(19)

Spectrum

Information Systems Industry

Decision Resources, Inc.

Bay Colony Corporate Center
1100 Winter Street
Waltham, Massachusetts 02154

Telephone 617.487.3700
Telefax 617.487.5750

Future Computing Environments: The Commodity Mainframe Era

Gordon Bell
Consultant to Decision Resources, Inc.

Business Implications

- The structure of mainframes renders them uncompetitive against micro-based multiprocessor computers. These micro-based systems also have cost advantages in manufacturing, software volumes, interactivity, and a base of trained users and programmers.
- The industry is transitioning from server-centric and client-centric environments to client-server, client-multiple-server, and network computing with multiple clients and multiple servers. This transition will take another decade because mission-critical legacy applications cannot readily be migrated to other environments.
- Resizing will continue to draw applications off mainframes onto networked PCs, workstations, and small servers; simple multiprocessors; large scalable multiprocessors; and massively parallel processing multicomputers. This trend will accelerate as network cost, performance, and maintainability improve.
- Simple multiprocessors will be the mainline of computing throughout this decade, to be replaced eventually by various forms of highly parallel processing computers. This environment will be based on standards, making components more commodity-like and the market more competitive. If this future environment is based on a few communications and computing standards, the market for computers will become vastly larger than it is today.

The success of large-scale, multimillion-dollar mainframes built from ECL technology and running proprietary operating systems from IBM and other hardware vendors (e.g., Fujitsu, Hitachi, NEC, and Unisys) has clearly peaked. While some analysts forecast the overall mainframe market to increase at 2-8% (less than half that of personal computers and workstations), others forecast a shrinking market for mainframes. We believe that the mainframe hardware market will be eaten away by alternative technologies, but large-scale transaction-oriented applications will grow moderately.

Mainframes operate as a central facility for batch processing, large transaction processing workloads, and general-purpose computing. Mainframes require a large, specialized staff to maintain both hardware and software. Created in the 1950s as the central computer for large enterprises, the mainframe now is following the same fateful path as proprietary minicomputers, which, ironically, were created in the 1970s for departmental applications and thrived on replacing mainframe functions. Minicomputer growth is even more anemic than that of mainframes because the minicomputer failed to attract enough mission-critical applications, which slow the migration to PCs and workstations.

With the advent of open systems[1] based on Unix, DOS, and industry standards, proprietary systems (particularly mainframes and traditional minicom-

1. The IEEE defines an open system as "a comprehensive and consistent set of international information technology standards and functional standards profiles that specify interfaces, services, and supporting formats to accomplish interoperability and portability of applications, data, and people."

puters) are out of favor. Unix became widely used in U.S. universities beginning in the early 1970s, resulting in a significant base of trained Unix programmers and users. Similarly, the wide-scale adoption of PCs has established the largest base of computer users. The mainframe has no such loyal following, except, perhaps, in MIS departments. Even mainframe programmers are dwindling as older programmers retire and new trainees gravitate toward Unix and DOS environments.

Beginning a little over 10 years ago, workstations became a natural choice for users hampered by overloaded minicomputers and mainframes. Enabled by fast, inexpensive, CMOS microprocessors, a bit-mapped graphic workstation provided high interactivity. Soon workstations and workstation servers connected to a local area network (LAN) replaced minicomputers for departmental and interactive computing. In the late 1980s, a LAN-based PC environment based on Novell NetWare servers came into prominence, connecting Microsoft DOS- and Windows-based PCs.

Today there are a multitude of computing alternatives to the original IBM 360 architecture—the basis for the majority of mainframes installed. (Mainframes from NEC and Unisys have architectures similar to the IBM 360.) Most of these mainframe alternatives fall into one of the following categories:

- *Multiprocessor.* A computer architecture in which two or more processors share common memory. (Most mainframes and supercomputers are also multiprocessors.)
- *Multi.* A microprocessor-based multiprocessor that typically supports fewer than 20 processors, where the microprocessors are organized around a single bus.
- *Computer Cluster.* A collection of independent computers (i.e., each computer runs its own copy of the operating system) that are interconnected to support a single application. The maximum number of computers supported is typically fewer than 16. (Most mainframes can be clustered.)
- *Scalable Multi.* A multi that has no theoretical limit on the number of microprocessors supported and to which additional microprocessors can be added incrementally.
- *Scalable Cluster.* A computer cluster that has no theoretical limit on the number of computers supported and to which additional computers can be added incrementally.

Why Today's IBM Mainframes Are Obsolete

To make the assertion that, like the dinosaur, the water-cooled ECL-technology mainframe is obsolete, we must first dissect it to examine the arcane structure required to support its software. Figure 1 shows the structure of the largest mainframe based on the IBM 360 architecture, the IBM ES 9000, which dictates the structure of IBM-compatible mainframes such as the Amdahl 5000 Series.

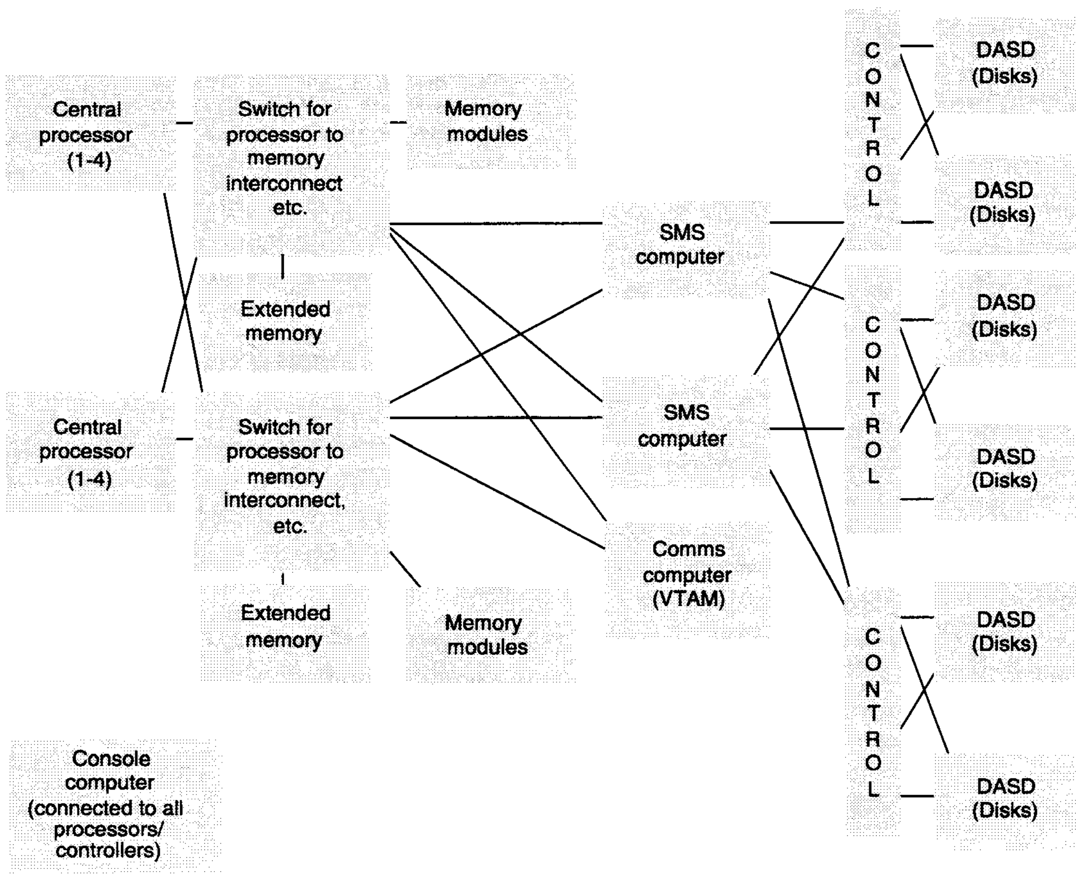
The mainframe's performance relies on two independent (but tightly connected) parts sharing a single operating system and clusters that have shared input/output (I/O). The ES 9000 has two quasi-independent subsystems, each having shared memory, up to four central processors (that may be added incrementally), and a switch to interconnect the parts. While this structure is essential for fault-tolerance, it may not meet the performance requirements for very large applications. Clustering technology is required to scale beyond a single computer with eight processors. The complexity of processor/memory interconnections limits mainframes to eight processors per computer; clustering is limited as well. For all practical purposes, about six mainframes can be clustered when they also have multiple processors.

The combination of multiprocessing and clustering creates additional complexity that increases the software burden of every large application. The IBM 360's complex operating system and arcane architecture are maintained by an ever-dwindling base of programmers who tend to resist new computing architectures. In this predictable scenario, the fewer the mainframe programmers, the less likely companies will commit to new mainframe purchases; and the fewer the mainframes purchased, the less likely new programmers will be trained on its architecture.

The duplexed mainframe comprises pairs of computers and software programs. The following describes some of their specialized control functions.

- The I/O control portion of the main computer operating system (including MVS, VM, AIX) executes channel commands. This structure is not suited to Unix, which is inherently small-file- and character I/O-based.

Figure 1
Duplexed IBM Mainframe Structure



Source: Gordon Bell.

- The Storage Management Subsystem (SMS) controls disks and archival storage. In addition to the main control computer, controllers and disks include computers not visible to users.
- The access interface (VTAM) allows the operating system to access the 3274 or 3745 communications computers and their operating systems.
- The console computer monitors the main computer and corrects errors.
- Computers within I/O devices and controllers that are not visible to users.

Computers and operating systems increase software size and complexity, automatically increase total memory, and cause data transmission delays. For example, disk caching and buffering occur within the disk, disk controller, storage computer, I/O channels, and main memory. (Some data transmission delays can be offset by increasing throughput.)

Multis, however, have a simple I/O structure whereby a disk (with a track buffer) reads directly into main memory under the control of an ordinary processor. Disk caches are located in main memory and can be traded off with other memory demands. Because all resources are available and interchangeable, any processor or memory can be used for application, control of communication, disk, and caching at any level. Compared with duplexed mainframe disks, redundant arrays of independent (or inexpensive) disks (RAID) reduces overhead and can increase performance by up to a factor of 4.

Whereas the relatively high cost of mainframes limits the number that are built, multis sell in an order-of-magnitude-higher volume, giving them a manufacturing learning curve cost advantage of approximately 50%.[2] Only a few thousand mainframe processors are produced every year; the key multi system component, CMOS microprocessors, can have unit volumes ranging from hundreds of thousands to tens of millions. Mainframes and multis use the same memory components; mainframes, however, traditionally have used special disks, giving multis the advantage with high-volume components such as commodity disks.

A market consisting of relatively expensive mainframes maintained by relatively large staffs results in a low-volume, high-price, custom software business. Table 1, based on the observations of Nathan Myhrvold of Microsoft, shows the economics of generating high-quality, profitable software: the value of each dollar of cost increases with the number of licenses sold. For example, a relatively complete, professional software package might cost $10 million to develop and maintain (enhance) on an annual basis. If 1,000 licenses are sold annually, the unit price would have to be $50,000, assuming R&D is 20% of net operating revenue. If 10,000 licenses are sold, then the price can drop to $5,000. With license sales of 100,000, the price can drop to $500, as low as high-end PC software is priced. Figure 2 illustrates the relationship between cost and volume.

Given the hardware, software, and market characteristics of mainframes based on the IBM 360 architecture and implemented with slow-to-evolve, expensive ECL technology, it becomes apparent why they are fundamentally obsolete. Mainframes will be replaced by the end of the century by any number of alternative computers and computing environments; however, centrally located enterprise servers that look similar to mainframes and are managed by both central MIS and functional groups will proliferate.

Despite its inevitable decline, the installed base of mainframes could linger for another 10-20 years because of the difficulty of porting large mission-critical applications to new computing architectures. The next generation of mainframes will use CMOS in place of ECL technology and could prolong the IBM 360 architecture if CMOS achieves price-performance objectives comparable to those of low-cost multi servers. However, even if CMOS technology does achieve these price-performance objectives, it will only slow the decline of the mainframe, not halt it.

Current Computing Environment

Today there are five distinct computing environments, differentiated by perceived purchase and operation costs, centrality of control, expandability or scalability, proprietariness, interactivity, and distributability. As shown in Figure 3, the five environments are as follows:

- Legacy mainframes with which users interact through 3270 terminals, or PCs and workstations emulating 3270 terminals.
- Legacy minicomputers running proprietary operating systems with which users interact through ASCII terminals, or PCs and workstations emulating ASCII terminals.
- Minicomputer-equivalent multis available from every Unix system vendor.
- LAN-based workstations and workstation servers running Unix.

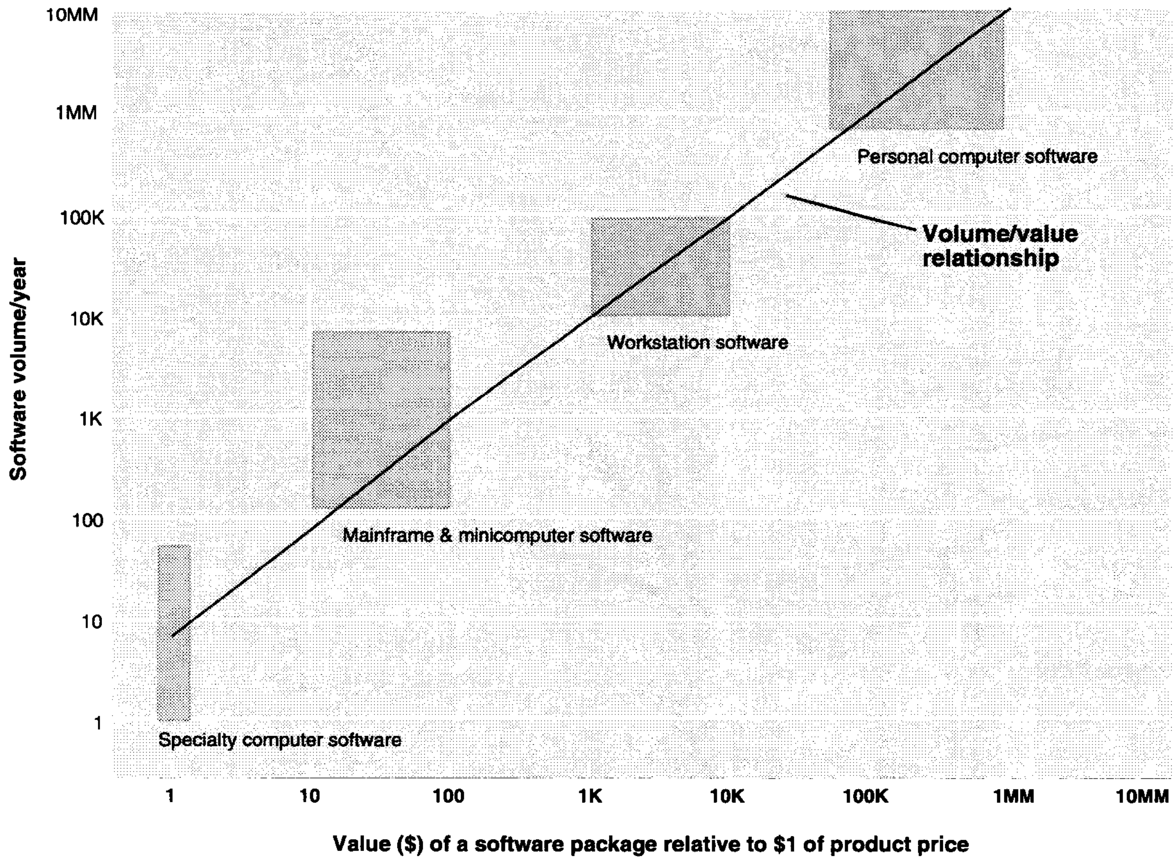
2. The rule of thumb for manufacturing learning curves is that each time the number of units produced is doubled, the per unit cost is reduced 10-20%.

Table 1
Relationship of Volume, Leverage, and Software Price for Various Computer Sizes

Computer Type	Volume/Year	Price of Software ($)	Value of a Software Package Relative to $1 of Product Price ($)
Special computer (e.g., massively parallel)	One to a few copies	100K-10MM	1
Minicomputer or mainframe	100s-1,000s	10K-100K	10-100
Workstation	10K-100K	1K-10K	1K-10K
PC	1-10MM	30-500	50K-1MM

Source: Gordon Bell, as suggested by Nathan Myhrvold, Microsoft.

Figure 2
Software Volume/Value Relationship



Sources: Gordon Bell and Decision Resources, Inc., as suggested by Nathan Myhrvold, Microsoft.

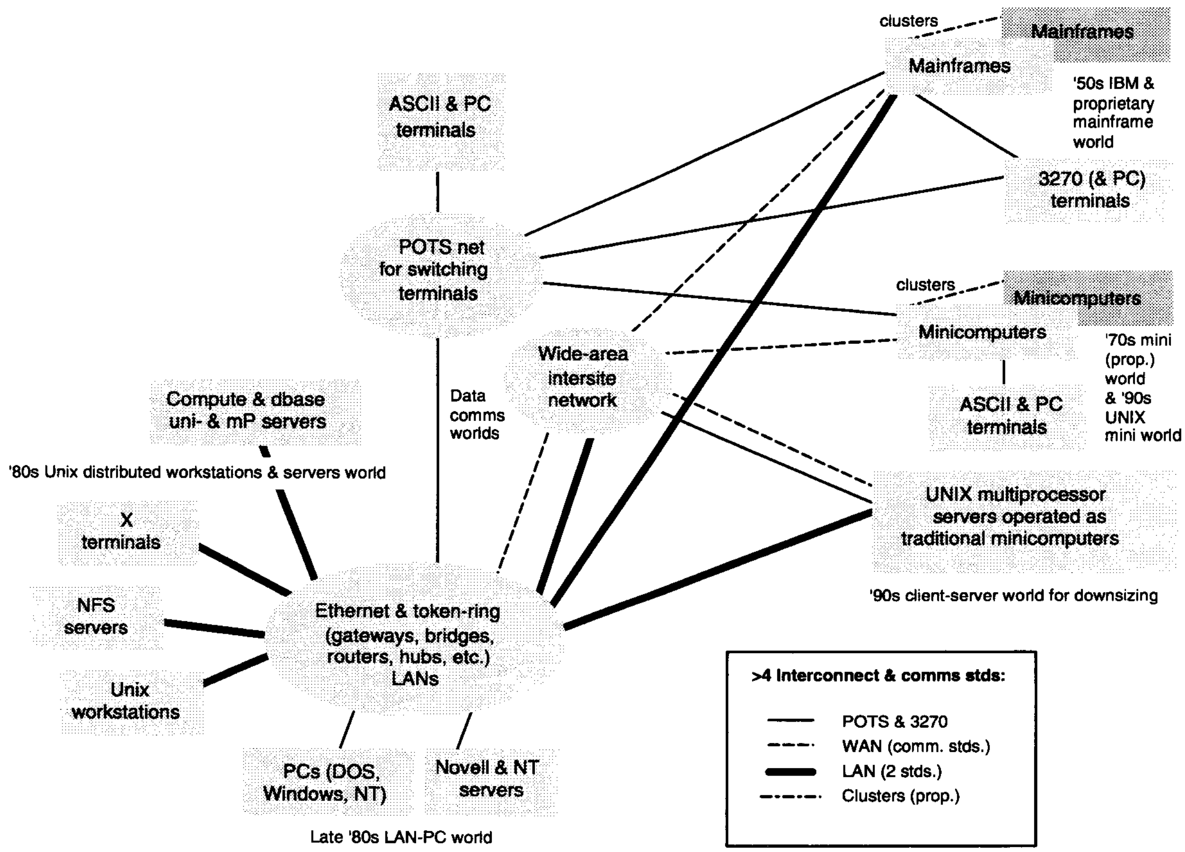
- LAN-based PCs and PC servers running Microsoft DOS, Windows (and, in the future, Windows NT), and Novell's Netware.

Table 2 describes some of the characteristics of these environments. Except for minicomputers, each has one dominant network to support it. Throughout the IS industry, legacy mainframes and minicomputers are being superseded by Unix uniprocessor and multiprocessor server platforms built from high-performance, low-cost, CMOS-based microprocessors. (The six major microprocessor architectures are Power PC [IBM and Motorola]; Alpha [Digital Equipment]; PA-RISC [Hewlett-Packard]; X86 [Intel]; R4X00 [Mips Technologies]; and SPARC [Sun Microsystems].)[3] Unix in any of its variants or dialects (Unicee) defines the open standard programming environment even though many dialects exist and every variant is unique. There is less portability among Unicee than there is among IBM-compatible mainframe platforms from IBM, Fujitsu/Amdahl, and Hitachi as the latter all run the IBM operating system. However, in Japan, Fujitsu and Hitachi supply proprietary operating systems and even unique applications, which lock in users to their platforms.

In the near future, Microsoft's Windows NT, which will run on the six major microprocessor architectures, could become the single, de facto industry standard Unicee promised but failed to provide. Almost every computer manufacturer today is building

3. For further discussion of microprocessors and their industry positions, see "Microprocessor Standards and Markets, Part II: Six Architectural Affiliations," *Spectrum, Information Systems Industry*, Issue 53, 1993.

Figure 3
Current Computing Environments



Source: Gordon Bell.

uniprocessor and multiprocessor computers around one of these powerful microprocessors.

Resizing

Today, computer buyers are considering microprocessor-based, open systems solutions—from PCs to workstations to minicomputer-sized servers—for a variety of reasons, including cost reduction, new applications, greater interactivity, autonomy from the central facility, and ability to distribute applications. This migration from mainframes to microprocessor-based systems often is described as downsizing and rightsizing; here we refer to it collectively as resizing.

Resizing of mainframe applications generally occurs in two ways. In *downsizing*, an application identical to the mainframe application is developed for a distributed set of microprocessor-based computers, for example, taking an electronic mail application run on a mainframe and distributing it to multiple PC mail servers. In *rightsizing*, an application is distributed on microprocessor-based computers to more closely fit the problem than did original mainframe application. Rightsizing, for example, is taking a mainframe application, such as transaction processing, and distributing the transaction portion of the application to PC clients and the database portion to PC servers. The mainframe may still do centralized processing. Rightsizing is sometimes called right-structuring or re-engineering, whereby the organizational structure, its processes, and the computing environment are all re-engineered for more efficient operation.

The arguments for resizing rest on the following benefits:

- Dramatically lower hardware costs (although when network costs are added, savings may not be so dramatic).

Table 2
Current Computing Environments

Characteristic	IBM or Other Legacy Mainframe	Legacy Minicomputer	Multicomputer	LAN Workstation	LAN PC
Introduction date	1950s; S/360=1964	Early to mid 1970s	Mid to late 1980s	1982	1986
Cost (system)	Most expensive	Expensive	Cost-effective	Expensive	Relatively expensive
Technology	Custom ECL	Custom bipolar	CMOS micro	CMOS micro	X86 micro
Reliability (1993)	Rarely fails	Solid	Very reliable	Workstations and networks fail	Reboots occur
Use	Large EDP and TP	Department and group	Department and group	Technical and high productivity	Office—desktop
Environment	Batch glass keypunch	Glass teletype	Interactive	Most interactive, WIMP[a]	Most interactive, WIMP[a]
Controlling organization	Central MIS	Departments and user groups or MIS	Departments and user groups or MIS	Departments and user groups or MIS	Departments and user groups or MIS
Staff	Large MIS staff	1-2 staff	1-2 staff	1-2 (plus network staff)	1-2 (plus network staff)
Location	Data center	Small room	Small room	Desktop	Desktop
Distributability	Central	Distributable	Distributable	Distributable	Distributable
Software type (availability)	Unique (small supply)	Unique (larger than mainframe)	Unix suppliers (larger than mini)	Unix suppliers (larger than mini)	PC and Unix suppliers (huge)
Principal programming environment	Proprietary IBM	Proprietary	Unix variant	Unix variant	Microsoft/Novell DOS/Netware
Network environment	SNA	SNA (IBM), DECnet (DEC)	TCP/IP & NFS	TCP/IP & NFS	Netware
Vendors	IBM and 360 clones, NEC, Unisys	IBM, Digital Equipment, DG, HP3000	DEC, HP, IBM, Pyramid, SGI, Sun, others	DEC, HP, IBM, Sun, SGI	Thousands of PC suppliers
Expandability (scalability)	Small scalability	Very small scalability	Limited scalability ≤30	Multiple workstations and servers on network	Multiple PCs and servers

a. Windows, Icons, Mouse, Pulldown menus.

Source: Gordon Bell.

- Larger software market and more applications resulting from standards and lower cost platforms.

- Ability to geographically distribute data and applications.

- Elimination of slow response that comes with the batch nature of mainframe (record-based) and minicomputer (character-based) software.

- Greater interactivity and productivity brought about by screen-oriented PCs and workstations ("glass paper").

- Scalability—the ability to start with a single user application and obtain incremental resources.

- The promise of peer-to-peer computing, making way for future applications like video mail and videoconferencing.

- Empowerment of individual user departments to manage their own applications rather than depend on central MIS shops (which often have a large backlog of applications awaiting development). This advantage is made possible by the current generation of programmers brought up on Unix.

The downside of resizing is that fully distributed computing creates new information management problems, particularly with respect to network administration. Networking costs for distributed computing can be substantially higher than for central computers connected to dumb terminals (e.g., IBM 3270-series, ASCII, and eventually X terminals) using slow-speed communication lines. For large organizations, annual network costs for LANs can exceed $1,000 per node given the chaotic way that Ethernet and token ring networks are pieced together using repeaters, bridges, routers, brouters, gateways, hubs, and so on. Furthermore, because computing is separate from files, the networks are constantly changing, a result of network bottlenecks. In the future, a more centralized switching structure such as asynchronous transfer mode (ATM) will improve cost, performance, and maintainability.

To participate in the distributed, resizing market, nearly all computer companies are building similar, multiprocessor servers. These companies include the following:

- Mainframe companies that must respond with lower-cost, higher-performance, IBM-compatible CMOS-based systems.

- Traditional minicomputer companies (e.g., Bull, Digital Equipment, DG, HP, IBM, Motorola, Pyramid, Tandem) that provide Unix-based open systems.

- Workstation companies using the multi architecture (e.g., Sun, SGI).

- Decade-old start-ups that created the first multis (e.g., Encore, Sequent, Sequoia, Stratus).

- The hundreds of companies currently supplying Intel X86-based PCs as distributed multi servers (e.g., Acer, AST, Compaq, NCR, Tricord, Unisys, Wyse).

Multi servers are similar to the centralized mainframe in that they run a single copy of the operating system. However, unlike the mainframe, they usually are run by a department rather than central MIS. Servers can replace a portion of a single mainframe, but they are perceived as being unable to take on the entire task due to limited I/O capabilities and reliability. To handle the large, mostly disk I/O capacity required to fully replace a mainframe and induce users to switch to alternatives, multi servers must provide sufficient secondary memory, I/O throughput, and computational power at a lower cost. Today's large multi servers can support as much as two terabytes (trillion bytes) of disk space.

Transitioning to New Computing Environments

Computing environments can be defined on the following levels:

- *Server-centric.* Applications run entirely on mainframe, minicomputer, or multi and are accessed through terminal or terminal emulators running on PCs or workstations.

- *Client-centric.* Applications run entirely on PC or workstation, with occasional downloading of database information from mainframe, minicomputer, or multi. In these applications, the server is simply for files or databases.

- *Client-server.* Applications are divided to run equally in both mainframe/minicomputer/multi and PC/workstation environments; the graphical user interface is run entirely by PC or workstation.

- *Client-multiple-server.* PCs or workstations access multiple backends for critical information.

- *Network computing with multiple clients and multiple servers.* Clients and servers are peers in every respect with fully distributed computing, and the distinction between client and server disappears.

Today the IS industry is in transition from server-centric and client-centric computing to the latter three computing environments. The 30-year-old server-centric environment, where centralized computing is organized around a database and accessed via dumb terminals (i.e., record-oriented 3270 "glass keypunches" derived from the mainframe, and character-oriented "glass teletypes" derived from minicomputers), has been on the wane since the advent of the client-centric computing based on highly interactive, microprocessor-based personal computers and workstations. In turn, the client-centric application software environment, characterized by word processing, spreadsheet, and CAD software, is also losing favor as networks enable computing loads to be distributed among servers and clients.

Client-server computing is being driven by mainframe and minicomputer vendors striving to maintain their place in the computing budget and hierarchy. However, true client-server computing, where applications run equally well across all environments, is still in its infancy. Client-multiple-server and network computing with multiple clients and multiple servers are even less mature. (Video and other teleconferencing applications are likely to be the first and predominant use of these latter two computing environments.)

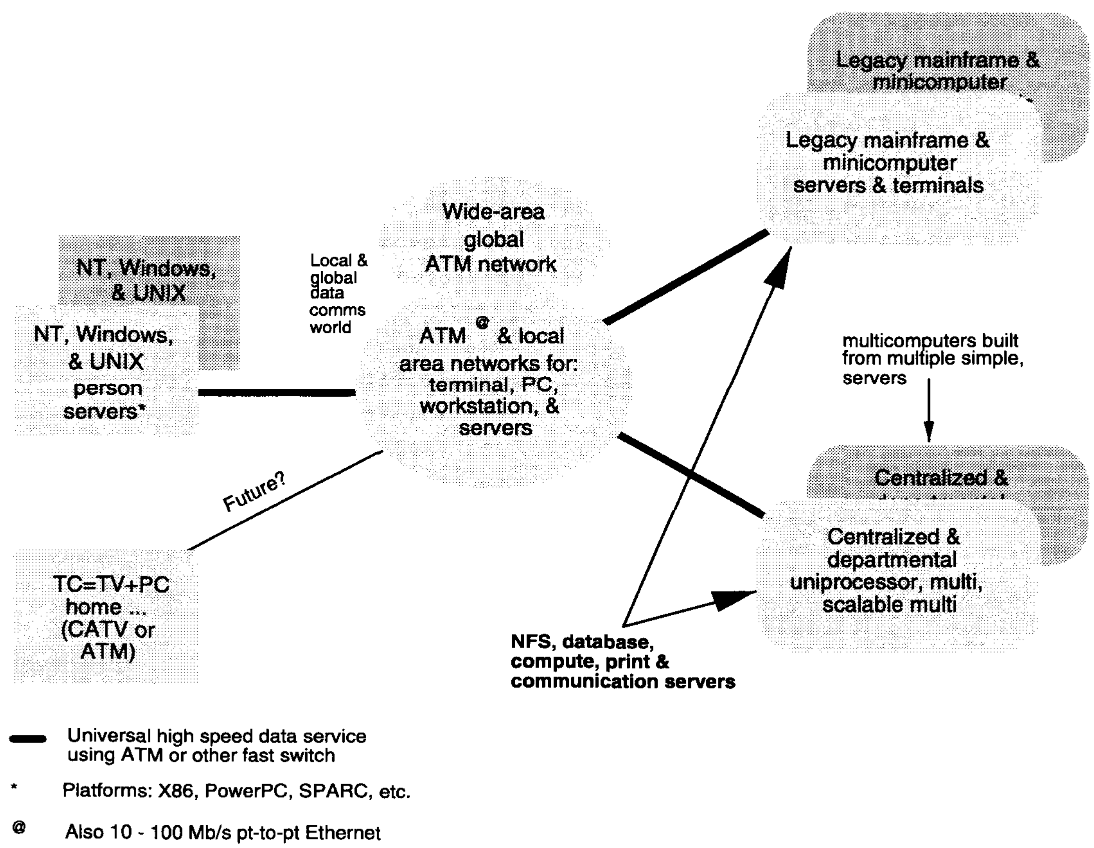
Future Computing Environment

We envision a very simple, 21st century computing environment (Figure 4) based on standards that allow complete interoperability among the three major components:

- A single network into which every computer is connected on a peer-to-peer basis.

- Database and compute servers that hold corporate, department, and project data and that communicate with one another.

Figure 4
Twenty-First Century Computing Environment



Source: Gordon Bell.

- Client or people servers that access various data and computational processes in database and compute servers.

The future computing environment shown in Figure 4 is based on computing (operating system) and communications standards, which will enable any manufacturer to build a high-tech commodity-like component for the environment just as PC manufacturers do today. Power can be added incrementally, and data sharing is accomplished through common database servers or by using the network to aggregate data from various sources in parallel. Applications can run with equal facility on nearly any node, resulting in true application and data portability.

Ideally, a single data network will provide a single standard for communication. The current LAN-based infrastructure, which connects PC and workstation clients to various servers, will be replaced by a ubiquitous, reliable, high-speed ATM switch that connects distributed, commodity-like, microprocessor-chip-based PC and workstation clients to banks of centrally located and managed uniprocessor or multiprocessor servers. A user will be able to plug in a computer anywhere in the world as easily as connecting a telephone today. The user could then communicate with any other computer for mail, file transfer, video phone, video mail, client-server style work, software installation and update, automatic backup, and help or maintenance.

If such a computing environment will be based on a single communications and computing standard, the market for computers will be vastly larger than it is today. However, this scenario may doom some computer manufacturers, as the entire computing industry will take on the characteristics of today's PC industry. Manufacturers used to getting high margins on proprietary systems—including Unix platform suppliers who use a dialect of Unix to lock in user to platform—will have difficulty adjusting to a more competitive environment based on largely interchangeable computers. As in the past, they will resist adoption of standards and try to maintain nonportability of programs and data. If this future computing environment does not evolve, computing will continue in its current chaotic state, where everything from networks to computers and applications is customized.

Mainframe Alternatives: Niching Away at the Workload

Given a ubiquitous network, any computer can communicate with any other on a peer basis. Thus, computing types are, in principle, completely substitutable among the computer classes, from personal computers to mainframes. Scalable clusters and networked personal computers, workstations, and small servers allow users to add capacity incrementally in an arbitrary fashion—unlike traditional mainframes or multis, which expand only to a factor of 10. Over the long term, simple uniprocessors, multis, and scalable multis will become the main line of computing as shown in the computing environment model of Figure 4. Table 3 gives the advantages and limits of mainframe substitutes and the outlook for each type. The four key alternatives will be discussed in the following order:

- Networked uniprocessor personal computers, workstations, and small servers.
- Multis (up to 20 processors) that have been the main line of computing.
- Scalable multis (up to 1,000 processors).
- Scalable clusters.

Today, multis, which have limited scalability and cost \$40K to \$1 million (comparable to high-end PC and workstation servers and minicomputers), are the principal mainframe alternatives. However, the only alternative that can provide equivalent performance of one or more mainframes is a scalable multi. Scalable clusters are being used both for technical and commercial applications. Clusters used for technical computing and databases are inherently special purpose and will be superseded by networked uniprocessors and multis interconnected through ATM switches. Scalable multis and scalable clusters can be described as massively parallel processing (MPP) because an almost unlimited number of processors or computers (≥1000 by 1994) can be interconnected, and all processors work on the same application simultaneously.

Networked Uniprocessor PCs, Workstations, and Small Servers

Over the next decade, it should become possible to configure independent, distributed computers as a single system, given the development of a fast, low-

Table 3

Characteristics of Mainframes and Mainframe Substitutes

Type	Companies	Advantages over Mainframes	Limits	Short-Term Outlook	Long-Term Outlook
Mainframes and supercomputers (1-8 processors)	IBM, Fujitsu, Hitachi, NEC, Unisys	NA	Limited scalability to a few (20) processors; batch	Limited life	Replaced by ATM uni- and multiprocessor servers; supers may continue
Computer clusters	DEC, IBM, Tandem	Expansion for capacity	Small scalability (i.e., up to 10)	Exists; only way to add capacity	Replaced by ATM uni- and multiprocessor servers
Networked uniprocessor personal computers and workstations	Thousands of personal computer and workstation vendors	Cost, simple, effective, time to market, stability	Performance per node; distributing load is poor substitute	Main line for people servers and compute environments	ATM replaces LANs; a basis for scalable servers
Multi (1-20 processors)	Acer, Compaq, DEC, Encore, HP, IBM, NCR, Pyramid, Sequent, SGI, Sun, Stratus, Tandem, Tricord, Unisys	Nearly as simple as a uniprocessor	Slightly more expensive than uniprocessor, bus limits number of processors	Basis for servers to replace *partial* mainframe load	Multiple-P/chip = mainline; basis for ATM connected scalable multiprocessor; mainframe replacement
Scalable clusters	IBM, Meiko, Intel, Teradata	Cost; scalability to thousands	Limited functionality without shared memory	Used in parallel processing; nonexistent	Replaced by ATM uni- and multiprocessor servers
Scalable multis	KSR	Cost; scalability to thousands	Nonuniform memory access	Gain experience	Ultimate ATM-based servers

NA = Not applicable.

Source: Gordon Bell.

latency network that allows communication between each computer with minimal overhead. If ATM switches can provide this level of performance and reliability, thus replacing current relatively expensive and unreliable LANs, then the structure diagrammed in Figure 4 will evolve rapidly. It also will eliminate the inherent locality limitation of today's LANs. The natural evolution of current distributed client-server computing is toward networked uniprocessor personal computers, workstations, and small servers where any collection of computers can be used for arbitrary serial and parallel tasks. Proprietary scalable clusters, in which independent computers are interconnected through specialized fast switches, will be subsumed by the fast ATM switch. For example, Digital Equipment's recently announced ATM switch provides 64 ports operating at 156Mb/s, or roughly the equivalent capacity of a 128-node Thinking Machines CM5 massively parallel processing computer.

The bottom line is that, inevitably, all computers connected through a high-speed ATM switch will replace client-server workstation or PC environments using LANs, clusters using proprietary interconnects, and scalable clusters using ad hoc switches. Servers will comprise the simple uniprocessors and multiprocessors, and scalable multiprocessors. Clients then are personal computers and workstations. Workstations simply become large PCs that run Unix.

Multis

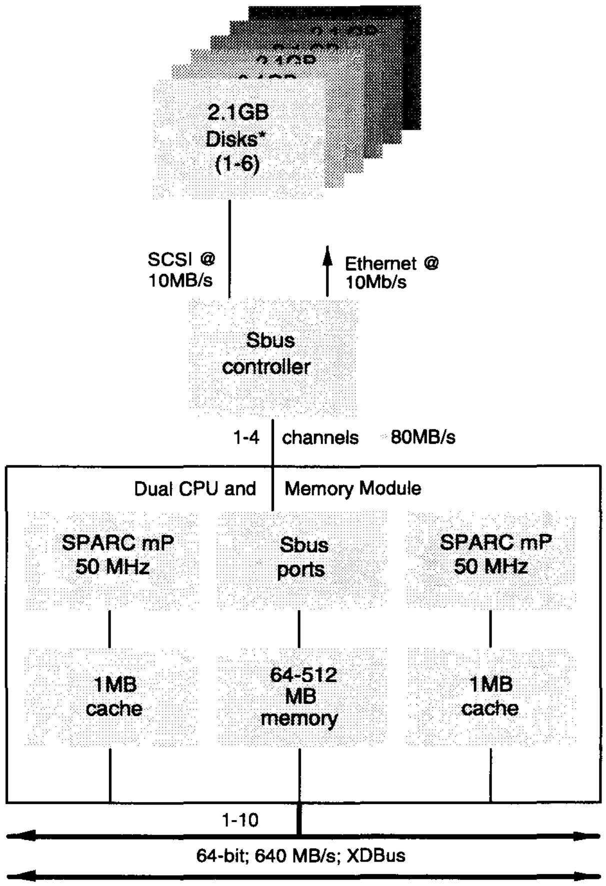
Since the 1960s, mainframe computing has been dominated by the simple multiprocessor—a computer organized so that every processor accesses a single, common memory. Each processor is connected to main memory through a centralized, cross-point switch. In the 1960s, mainframes with 1-4 processors were organized this way. In the early 1980s, supercomputers adopted the mainframe's multiprocessor structure for up to 4 processors. By 1992, Cray Research's C90 supercomputer could configure up to 16 processors. Currently, mainframes can configure up to 8 processors, a limit imposed by technology, operating system, and market price sensitivity. However, today's mainframes and supercomputers, which are organized around the central, cross-point memory to interconnect processors with memories, cannot practicably expand beyond 8 (mainframe) or 16 (supercomputer) processors.

In the mid 1980s the limited scalable, multiple microprocessor computer (multi) was introduced by Encore, Sequent, Stratus, and several (now defunct) computer companies including Synapse, which built the first multi in 1982 using up to 20 Motorola 68000 microprocessors. The multi consists of multiple CMOS microprocessors organized around a single, shared bus, which allows any processor with a cache to access any memory module. The system is simple, economical, and effective because the cache provides fast access, limits memory traffic, and is able to "snoop" the bus such that a single, coherent memory is maintained. The performance of a multiprocessor is determined by its bus speed, microprocessor speed, and cache size. Today, all minicomputer and workstation manufacturers use this multiprocessor structure to build computers that scale from 1-4 and in some cases 1-20 processors, depending on the microprocessor and bus speed and cache size.

The Sun SPARCcenter 2000 is a large multi, noteworthy in terms of processing power and primary and secondary memory, and the buses that support them. Figure 5 shows a block diagram of the SPARCcenter 2000, which supports up to 20 processors, 240 2.1GB disks, and 5GB of primary memory. Each processing module supports 4 Sbus I/O cards. An Sbus card supports a variety of I/O devices including 10Mb/s SCSI and Ethernet cards and a variety of communications interfaces. The SPARCcenter 2000 is managed by Sun's two operating systems. Solaris 2 is designed for multi-threading. It can host centralized, relational databases (e.g., Ingres, Informix, and Oracle) accessed through SQL as a replacement for mainframe database applications. In this application, the ability to run multiple disks in parallel and handle client requests coming through its multiple Ethernet connections is important.

Table 4 shows the performance characteristics of the IBM ES 9000 mainframe versus the SPARCcenter 2000 (and other multis that can take over some fraction of a mainframe's load) as well as those of the most cost-effective mainframe alternative, a single personal computer. Few of these alternatives have demonstrated that they have the capacity of a large mainframe; however, as noted in the table, several alternatives (e.g., SPARCcenter 2000 and Power Challenge XL) have raw processing power equal to or greater than a mainframe, while prices are over an order of magnitude less.

Figure 5
SPARCcenter 2000 Block Diagram



*max. disk 10 x 4 x 2.1GB; max. memory = 5,120MB

Source: Gordon Bell.

Table 5 shows the incremental price and price performance characteristics for a single, generic IBM PC, several multis, and an IBM mainframe multiprocessor. These price data for the basic system and incremental processor, memory, and disk are the basis of Figure 6, which shows the performance for these computers plotted against system price. Performance is measured in SPECint92 and is simplistically derived by multiplying the individual processor power times the number of processors. For each SPECint92, one megabyte of memory is added to the configuration.

We can draw several conclusions from Figure 6:

- Several systems that sell for less than $1 million are capable of running many of the same applications as a $10 million mainframe.
- Ultimately, the most cost-effective system is simply the clustering of ordinary PCs. This latter structure is, of course, the base component of the scalable computing environment described in Figure 4.

Multis will continue to be the mainline and foundation of computing at least until 1995 and quite likely until 2000, when scalable multis will take over. Quite likely, scalable multis will be built as a hierarchy of ordinary multis.

Table 4
Mainframe and Mainframe Alternatives for Off-Loading

Computer	Number of Processors	Bus Speed MB/s	Performance (SPECint92)	Maximum Memory (MB)	Maximum Disk (GB)	I/O Rate (MB/s)	Processor Architecture	O/S
IBM ES9000	1-8	NA	8x50 (MIPS)	9216	NA	800	IBM 360	AIX, MVS, VM
IBM PC and Compatibles	1	NA	1x32	32	05	NA	Intel X86	DOS, NT, Unicee
Compaq	1-4	NA	4x65	512	140	NA	Intel X86	NT, Unicee
HP 9000	1-12	1,000	12x80	2,048	1,900	256	PA-RISC	HP/UX
Tricord ES5000	1-6	267	6x65	64M-1G	488	28x5/10	Intel X86	NT, Unicee
Sequent Symmetry	1-8 [a]	NA	8x65	960	NA	NA	Intel X86	NT, SVID
NCR 3500	1-16	400	16x65	2,048	3,000	28x5/10	Intel X86	NT, Unicee
SPARCcenter 2000	2-20	640	20x66	5,120	6x48x2.1	36x10	SPARC	Solaris, Next O/S
SGI Challenge XL	1-36	1,200	36x94	16,000	960-3,000	1,280	R4X00	Unix

Note: Multis not included: Acer, Compaq, Data General, Digital, Encore, IBM, Pyramid, Sequoia, Stratus, Tandem. Multicomputer clusters not included: Auspex, IBM, NetFrame, Parallan, SPARCclusters.

NA = Not available
a. System is capable of more processors, but unannounced.

Sources: Vendor literature and Gordon Bell.

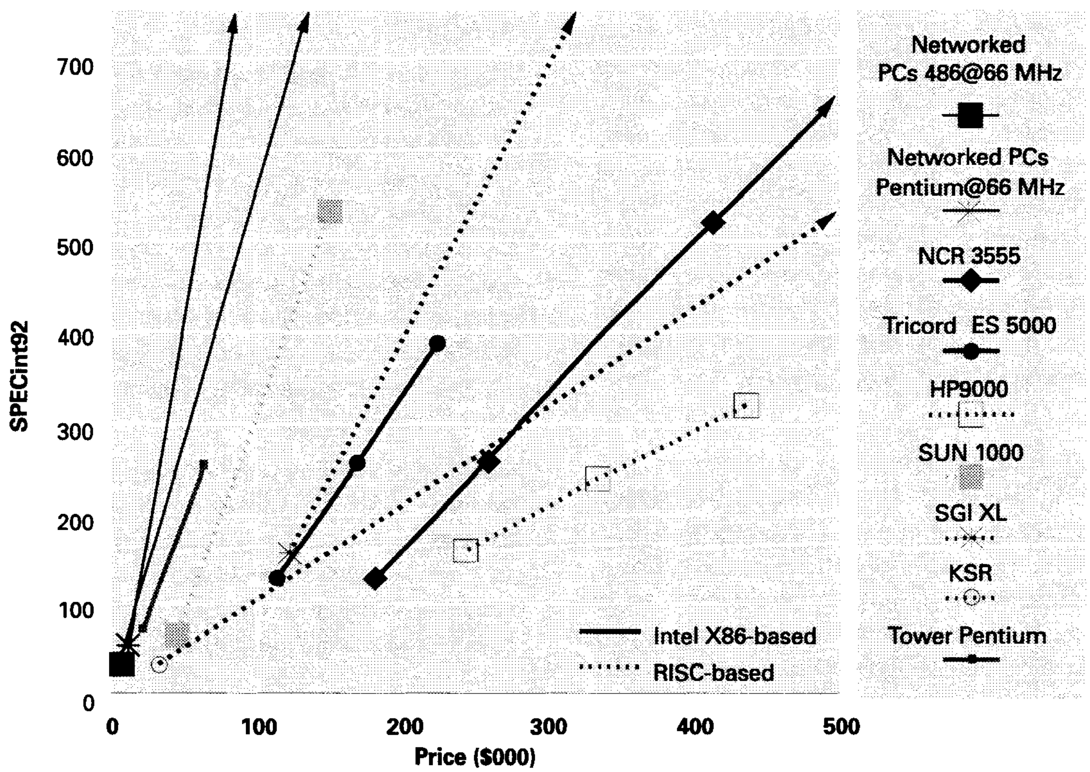
Table 5
Incremental Prices for Various Computing Alternatives, 1993

Computer	$Base (K)	$K/Processor	Δ $/SPECint92 Processor	Δ $/MB (memory)	$/MB (disk)
IBM mainframe	16,000 for 8	NA	40,000(MIPS)	NA	3-6
IBM PC	1	2-4:Pentium	30-60	40-60	.75-1.5
HP 9000	100	65	814	235-390	4.8-11
Pentium multis	15-100	10-25	150-380	164-220	2.5-4.5
SGI multi	30-80	7.5-20	80-212	100-170	3.75
Sun multi	35-80	10	152	65	2-2.6

NA = Not applicable.

Sources: Vendor literature and Gordon Bell.

Figure 6
Price/Performance of Selected Multis



Note: Configured systems include a 500MB disk and 1MB of memory/SPECint92.

Source: Gordon Bell.

Scalable Computers

The High Performance Computing and Communications Initiative[4] stimulated the idea that computers can scale (add components incrementally) over an almost unlimited range of processors, memory size, secondary memory size, and I/O. Teradata, now part of AT&T (NCR), developed a proprietary, large-scalable computer for database applications. To make such a large-scalable computer for general-purpose applications,[5] computer modules (i.e., processor-memory pairs) must be interconnected by a high-performance, low-latency switch (i.e., network). Performance, using an appropriate measure, then grows in proportion to the number of resources implemented.

The ideal, scalable computer should be size scalable; that is, it should be useful as a single processor and extend to thousands of processors, with correspondingly scalable I/O. It should also be able to handle a wide range of scalable parallel applications, including general workload applications. The interconnect network must be generation scalable; that is, it can be implemented over a range of technology generations and provide a compatible program environment for at least a decade to support binary compatibility of applications among generations. Furthermore, since the processor-memory pairs are independent, the ideal scalable computer should be geographically scalable, that is, distributable beyond a single room to include a campus. By solving many problems in security and fault-tolerance, a distributed computer that would occupy a building or even a large campus might be designed.

In a race to provide a peak teraflop of computer power by 1995,[6] two MPP computer structures (Figure 7) have been introduced: the scalable multi and the scalable cluster. Table 6 shows differences between the two. Scalable multis (Figure 7a) communicate by accessing a single, shared common memory and have a single copy of the operating system. Scalable clusters (Figure 7b) are independent computers, having independent copies of an identical operating system, which communicate by passing messages to one another through a central switch. Current proprietary scalable clusters are not general purpose and will likely disappear within a few years to be replaced by a collection of PCs, workstations, small servers, and multis connected through a high-speed switch, as described previously.

Scalable Multis. In 1993 Kendall Square Research introduced the KSR-1 scalable multi, which was upgraded to the KSR-2 in November 1993. This system is a true alternative to a mainframe because it (1) is general purpose, (2) runs a single operating system, and (3) can allocate any of its resources to a common workload. Unlike a mainframe that is limited to an order of magnitude scaling range, the KSR-1/2 can grow over a dynamic range of 1,088 processors. We anticipate that within the next decade all companies will evolve their multis to be scalable.

The KSR-1/2 is a size and generation scalable, shared-memory multiprocessor computer (as shown in Figure 7a). Its structure and programming model (as shown in Figure 8) is simply 1 to 1,088 multiprocessors nodes that access a common memory. Each node comprises 32MB of primary memory and a 64-bit superscalar processor. A 1,000-node system provides almost 30 times more processing power, primary memory, I/O bandwidth, and mass storage capacity than a mainframe.

The KSR-1/2 is significant because it provides (1) size (including I/O) and generation scalable shared memory multiprocessing where every node is identical; (2) an efficient environment for both arbitrary workloads (from transaction processing to timesharing and batch processing); and (3) sequential to parallel processing through a large, hardware-supported address space with an unlimited number of processors, a strictly sequential consistent programming model, and dynamic management of memory using its "ALL-CACHE" mechanism.

The KSR-1/2 supports every form of parallelism: multiple users may run multiple sessions comprising multiple applications, comprising multiple processes (each with independent address spaces), each of which may comprise multiple threads of control running simultaneously sharing a common address space. The

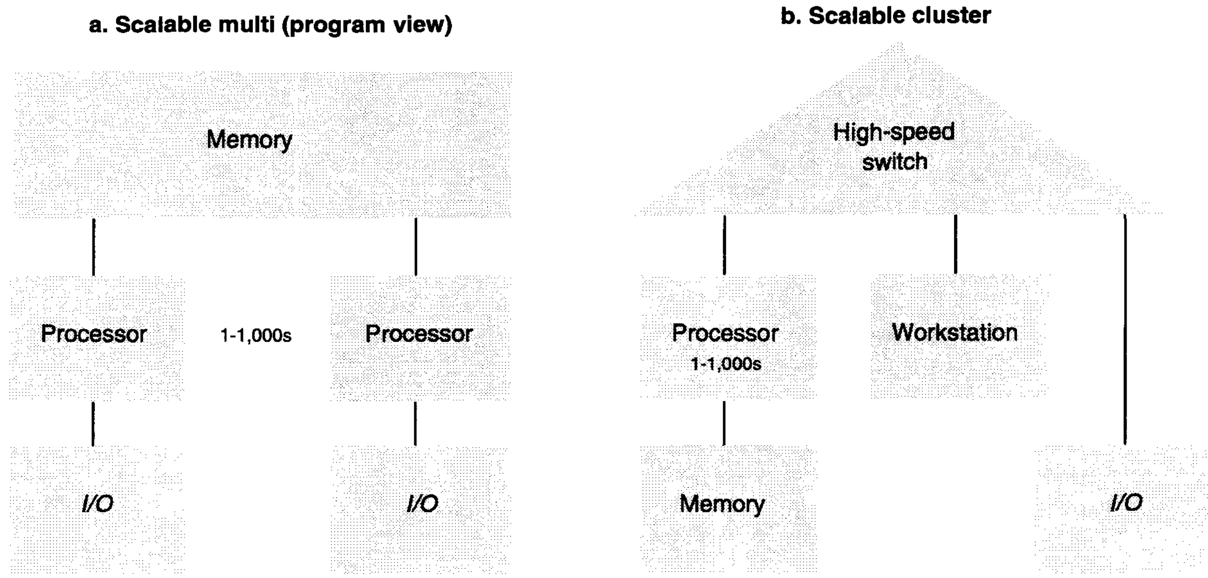
---

4. The High Performance Computing and Communications Initiative is one of six initiatives within the U.S. Federal Coordinating Council for Science, Engineering and Technology.

5. Test for general purposeness: Can the computer efficiently process a wide range of jobs (including a workload consisting of sequential to parallel, small to large job sizes, short to long run-times, and interactive to batch response times) requiring a variety of processing, memory, database, and I/O resources?

6. See *Ultracomputers, A Teraflop Before Its Time*, Gordon Bell, Communications of the ACM, August 1992.

Figure 7
Massively Parallel Processing Computer Structures



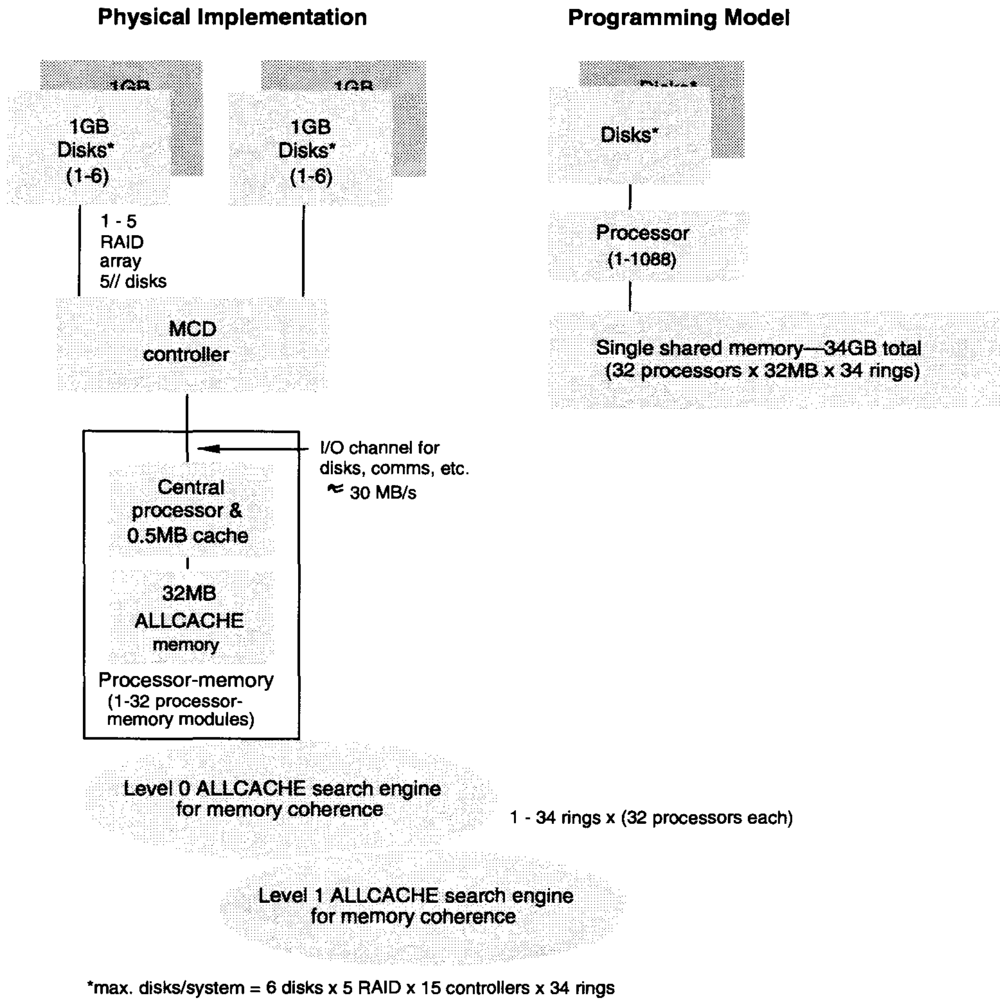
Source: Gordon Bell.

Table 6
Scalable Multi and Scalable Cluster Characteristics

	Multiprocessors	Computer Clusters
Memory model	Single address space in a single coherent memory	n-independent address spaces in n-independent computers
Resources	One fungible set	Bound to each computer
O/S structure	One copy; one copy of work queue	n-copies of a distributed O/S work are bound to a fixed set of computers; resources may be idle
Programming	Transparent; tasks automatically run on idle processors	Reprogramming required to allocate work and task for particular computers
Computer types (today)	Supercomputers; mainframes; "multi-based" minicomputer and workstation servers; and scalable multiprocessor servers	Networked workstations; computer clusters; and scalable clusters (e.g., IBM, Intel, Meiko, Teradata)
Servers (future)	Supercomputers; "multi-based" servers; and scalable multiprocessors (for next mainframe)	"Multi-based" servers; PCs and workstations connected via ATM network

Source: Gordon Bell.

Figure 8
KSR-2 Scalable, Shared Memory Multiprocessor



Source: Gordon Bell.

KSR-1/2 is used both for supercomputing and large-scale, commercial applications. Its commercial programming environment for transaction processing can access relational databases in parallel with unlimited scalability as an alternative to multiple mainframes. The 32- and 320-node systems are estimated to deliver more than 1,000 and 10,000 transactions per second, respectively, giving them 5-10 times the throughput of the largest mainframes.[7]

Proprietary Scalable Clusters. Proprietary scalable clusters, offered by Intel, Meiko, NCube, and Thinking Machines, are being used for both technical and commercial applications. Technical applications are run massively parallel; that is, all processors work on an application simultaneously and communicate with one another through message passing. Two programming models are used:

- *Single Program, Multiple Data (SPMD).* A program is written in a supercomputer style whereby a sequential program operates in vectors and arrays.
- *Message Passing.* Multiple, identical programs operate on independent data elements and then communicate results with one another by passing messages.

7. As this analysis was being finalized, Kendall Square Research announced a large reduction in previously reported revenues for 1993 and 1992, due to accounting irregularities. Our use of the KSR-1/2 as an example of a scalable multi represents our assessment of the viability of this computer architecture type, not of Kendall Square Research itself.

Scalable clusters have been used to provide parallel access to a common database. Similarly, transaction processing should be possible using computers that access a common database. The oldest and most successful scalable cluster, measured by installed systems, was built by Teradata. The Teradata computer was organized to be a SQL database machine accessed through a direct connection to an IBM mainframe.

Since proprietary scalable clusters fail the general purpose computing test, they will evolve either to scalable multis or to scalable clusters based on standards similar to the ATM networks described previously.

The Competitive Picture

Mainframes are in for the fight of their lives. By not adopting CMOS technology, mainframe designers allowed microprocessor-based uniprocessor and multiprocessor systems to replace many of the functions of mainframes for all but the largest applications. The move to downsizing/rightsizing is a natural result of users' gravitation to less expensive computing solutions and more interactivity. With the advent of open systems and standards, and a large base of users and programmers trained on microprocessor-based systems, the mainframe's ability to retain users with legacy software is doubtful. Implementing CMOS technology into mainframes to reduce costs is too-little too-late strategizing.

We believe that distributed, client-server computing will evolve naturally to replace uncompetitive, unresponsive central mainframes. This evolution will be driven by lower manufacturing costs due to high-volume components, more adherence to standards, increased competition, and most importantly, development of a fully connected, peer-to-peer network where any computing resource or set of resources can easily carry out any function or application.

A high-volume, competitive personal computer environment will mean that a software industry has formed to supply a plethora of low-cost, high-quality, applications. The low cost of personal computers and the ability to interconnect them through a ubiquitous ATM switched network will eliminate the seam between hodgepodge LAN and WAN environments.

The likely result is a very simple structure based on just a few standards. This structure would handle virtually all computing, relegating the mainframe to a few lingering legacy applications. It will also create a market where minicomputers and workstations will have to compete directly with PCs. In anticipation of this development, most minicomputer and workstation vendors are building multis in order to capture higher margin sales now going to mainframes and, to a lesser extent, large minicomputers. However, with everyone chasing the same market niche with similar products, the result will be a highly competitive environment—not unlike today's PC market—in which margins are razor thin and only a few high-volume, low-cost producers are profitable. In effect, we are entering the era of the commodity mainframe.

About the Author

***Gordon Bell** is a computer industry consultant at large. He spent 23 years at Digital Equipment Corporation as vice president of research and development, where he was the architect of various minicomputers and time-sharing computers and led the development of Digital's VAX and the VAX environment. Mr. Bell has been involved in, or responsible for, the design of many products at Digital, Encore, Ardent, and a score of other companies. He is on the boards and technical advisory boards of Adaptive Solutions, Chronologic Simulation, Cirrus Logic, Kendall Square Research, Microsoft, Visix Software, University Video Communications, Sun Microsystems, and other firms.*

Mr. Bell is a former professor of Computer Science and Electrical Engineering at Carnegie-Mellon University. His awards include the IEEE Von Neumann Medal, the AEA Inventor Award, and the 1991 National Medal of Technology for his "continuing intellectual and industrial achievements in the field of computer design." He has authored numerous books and papers, including High Tech Ventures: The Guide to Entrepreneurial Success, *published in 1991 by Addison-Wesley. Mr. Bell is a founder and director of The Computer Museum in Boston, Massachusetts, and a member of many professional organizations, including AAAS (Fellow), ACM, IEEE (Fellow), and the National Academy of Engineering.*

Eric P. Blum, Research Program Manager

93-11-55

About Decision Resources, Inc.

Decision Resources is an international publishing and consulting firm that evaluates worldwide markets, emerging technologies, and competitive forces in the information technology, life sciences, and process industries. Decision Resources links client companies with an extensive network of technology and business experts through consulting, subscription services, and reports. For additional information, please contact Marcia Falzone by phone at (617) 487-3749 or by fax at (617) 487-5750.