Enterprise Information Architectures — They're Finally Changing

Wesley P. Melling
Gartner Group
56 Top Gallant Road, Stamford, Connecticut 06904
203-975-6533

1.0 Abstract

Substantive changes in the business environment — and aggressive initiatives in business process reengineering — are driving corresponding changes in the information technology architectures of large enterprises. Those changes are enabled by the convergence of a long list of maturing new technologies. As one of its many implications, the new IT architecture demands revised assumptions about the design and deployment of databases. This paper reviews the components of the architectural shift now in process, and offers strategic planning assumptions for database professionals.

2.0 Background

The Gartner Group is the computer industry's largest provider of strategic planning services for information technology executives. More than 120 analysts, averaging in excess of 20 years of experience in information systems, monitor developing technologies and vendor strategies on behalf of some 10,000 user clients (skewed heavily toward large enterprises). Because of the nature of the Gartner client base, vendors tend to be conscientious about keeping Gartner informed on future developments. In turn, a Gartner analyst will average five telephone contacts a day with clients seeking advice, making it easy to spot demand and usage patterns in the market.

As an analyst specializing in the management of architectural change and in implementation of high-volume transaction processing systems, the author has been in a position to observe first hand the recent acceleration in the rate of architectural change from traditional mainframes to open systems, discussed herein.

SIGMOD 94- 5/94 Minneapolis, Minnesota, USA
© 1994 ACM 0-89791-639-5/94/0005..$3.50

3.0 Introduction

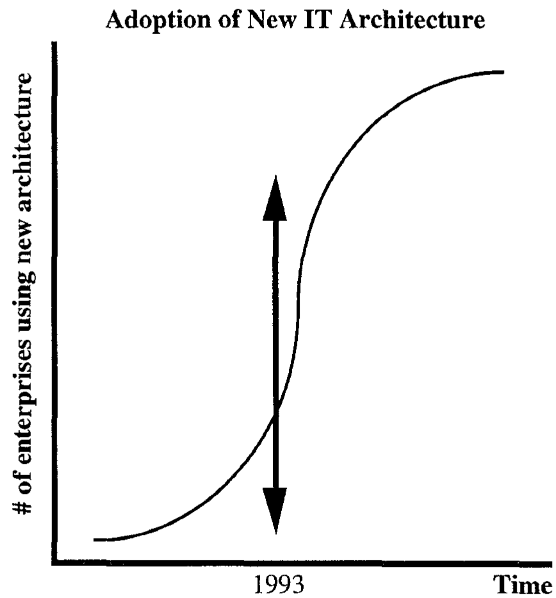
Information technology (IT) is changing eras. From 1964 to 1980, the information industry enjoyed a period of manageable, homogeneous computing on monolithic systems. By 1980, after many years of practice, IS professionals had come to understand how to implement application systems, buy from vendors and deal with end users. But during the 1980s, the industry entered a time of unmanageable, heterogeneous, networked systems, and by 1990, virtually all aspects of IT "standard operating procedure" had been overthrown, and computing architectures had spun out of control. LANs had experienced runaway growth, and desktops and servers of all sizes and persuasions had appeared like weeds. Multiple vendors, multiple hardware architectures, multiple operating systems and "islands of automation" were the norm in large organizations. Programming standards had eroded, and the notion of non-redundant, consistent, high-quality data seemed to be on the compost heap. IT spending had increased in the user community, while pressure to cut spending in the central IS department was common.

Against that backdrop of chaos, it is now clear that, in the 1990s, technology is converging to support an era of manageable, heterogeneous networked systems, and that synergy between user empowerment and coherent enterprise IT architecture is an imminent reality. The new synergy is based on a concept of virtual enterprise servers which we call the Enterprise Server Platform. The Enterprise Server Platform could be, *but is not*, envisioned as a centralized mega-server. Rather it is a collection of services, typically distributed across a variety of computers for price/performance, high availability and platform specialization reasons. Critical services include mail routing, file and print sharing, transaction and process management, decision support, network and systems management, resource brokering, software distribution, license management, security enforcement and inter-enterprise electronic data interchange (EDI). The location of the services, and the hardware/operating system architecture underlying the services, are transparent to the client. *Moreover, the server architecture that is visible to the application (and to the application programmer) is defined entirely in middleware.* (We use "middleware" to describe that set of systems software services that fits architecturally between the application program and the operating system.)

4.0 Why Now?

The notion that now might be the right time to move from traditional mainframes to a new architecture sounds like an idea whose time came ten years ago. Indeed, in industries which emerged after 1980 (e.g. cellular telephones, HMOs, industrial scale poultry production, medical testing laboratory chains, videotape distribution), the mainframe was rarely accepted in the first place. Yet it is evident that the bulk of the computing community is just now planning the move from mainframes. Three factors have slowed acceptance of a new architecture:

- A *user software investment* in MVS-based applications estimated at between 500 billion and one trillion dollars, which was viewed as "too large to walk away from" until it became clear that obsolescent applications had become impediments to corporate survival.
- *Human resistance to change.* This is hardly a new problem. In 1781, Benjamin Franklin observed that "to get the bad customs of a country changed and new ones, though better, introduced, it is necessary first to remove the prejudices of the people, enlighten their ignorance, and convince them that their interests will be promoted by the proposed changes, and this is not the work of a day." A Director of Technical Support who has invested 30 years of his life in MVS internals (and who is paid very well because of that expertise) has real difficulty understanding why a move away from that architecture will "promote his interests".
- The *nature of the new architecture* itself. Simple new ideas are accepted easily; the new architecture that is emerging is complex. New ideas that resemble old ideas are accepted quickly; the new architecture is drastically different both in concept and in semantics. A new idea should appear stable; the new architecture still looks painfully unsettled, since many of its components are emerging from the Unix community, which continues to display an almost complete inability to agree on anything at all. Finally, to be used in real businesses, a new idea must be of "industrial strength"; the new IT architecture, because of the number of new technologies involved, has until recently demonstrated a very short mean time to failure[1].



Source: Gartner Group

Figure 1: Contagion Phase for the New IT Architecture

Nevertheless, it can now be clearly observed that an emotional breakthrough has occurred and that a *majority* of large enterprises are currently in the process of deploying major applications on new architectures. In short, we have entered what is variously described as the "contagion" or "conflagration" phase of acceptance of a new concept. Fuel for the trend is provided by the ease of finding successful examples of the new architectures at work. Enterprises of the size of Texaco, Mobil, The Fritz Companies, Hyatt and Zale's have successfully moved major (and mission-critical) applications off traditional mainframes. Companies like Tyson, Blue Cross of Texas and Bow Valley Energy have replaced mainframes altogether. Proof of concept exists.

---

[1] For a rigorous articulation of the factors that speed up or slow down acceptance of a new idea in a society, see *Diffusion of Innovation*, Everett M. Rogers, Free Press, 1982

5.0 Business Process Reengineering

Business Process Reengineering (BPR) is a key element in the factors driving IT architectural change. BPR is defined as "the fundamental rethinking and radical redesign of business processes to achieve dramatic improvements in critical, contemporary measures of performance, such as cost, quality, service, and speed"[2].

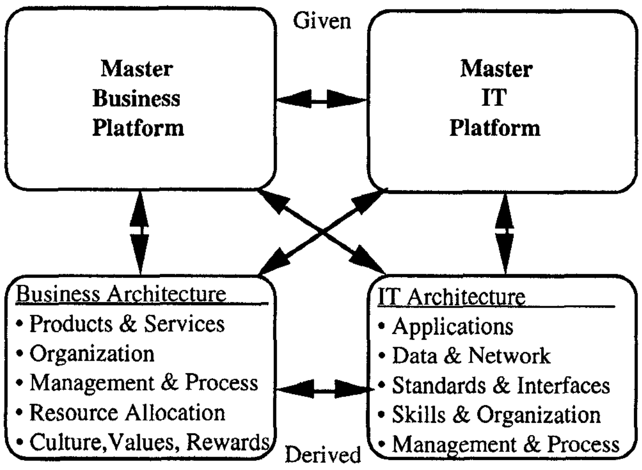


Figure 2: Henderson Master Platform Model

John Henderson of M.I.T. provides an excellent model (see Figure 2) for understanding the interaction of business strategies and IT architectures. In that model, a Business Master Platform refers to the set of strategies, markets, regulations, product technologies and resources chosen by an organization as relevant to its mission. From that is derived a Business Architecture, that set of products and services, organizational structures, management processes, resource allocations, and values and rewards which are necessary to implement the Business Master Platform. The corresponding IT Master Platform is understood as the range of relevant computing technologies available to the enterprise, and the ways in which those technologies can be used for competitive advantage. The IT Architecture is that set of specific computing architectures and products that are chosen to implement the IT Master Platform, and the support infrastructures, skill sets, decision processes and administrative mechanisms used to deploy those architectures. The model has several corollaries:

- There is bi-directional impact between the business and technology master platforms.
- If the business or technology Master Platform changes, it is highly unlikely that the derivative IT architecture can survive.
- Alignment between the business and technology architectures is crucial for success, but may take a decade or longer to achieve.

This model bears directly on the shift in IT architecture, since it suggests strongly that when Business Process Reengineering causes radical redesign of business processes, it will also cause a radical redesign of information technology processes — and indeed that is exactly what is happening. Figure 3 shows some of the IT architecture implications of the new business architectures.

Business Architecture	IT Architecture
Business unit autonomy	Multiple vendors, networks, platforms, operating systems Buy versus make
Fewer levels of management	Ubiquitous mail, notes, image management, teleconferencing
Reorganization of work from task-centric to process-centric	Shift from OLTP monitors to process managers
Integration of the supply chain	Cross-vendor C/S applications Multiprotocol routing Trustworthy messaging
Globalization	Multivendor application portability. Global networks No-excuses 24x365 operation.
Intense focus on customer service	Fast application development Cross-vendor C/S applications Trustworthy messaging 24x365 operation
Increased worker mobility. Growth in telecommuting.	Wireless communications Asynchronous messaging Database replication 24x365 operation
Intense focus on cost	Use of latest technology

Source: Gartner Group

Figure 3: Changes in Business/IT Architecture

[2] Michael Hammer and James Champy, *Reengineering the Corporation*, HarperBusiness, 1993

5.1 Business Unit Autonomy

Enterprises are reorganizing as sets of small, autonomous strategic business units to achieve the speed and flexibility required to survive under the new rules of competition and geopolitics. The shift downward in authority has these implications:

- Since IT will be increasingly central to business unit success, and since business unit managers know it, multiple business units, each with a strong CEO, will inevitably create multiple IT strategies, using multiple vendors. Vendor-transcendent standards will be critical for tying the enterprise back together.

- Projects and problem solution will be increasingly addressed by cross-functional and cross-business-unit teams, which will be heterogeneously equipped and geographically dispersed. Open standards will permit workgroup application integration despite the heterogeneous base. Global networks and effective groupware will vitiate the problem of geographic dispersion.

- There will be a natural tendency to distribute the operations of a modern enterprise a) to get response close to the customer and b) to preserve the agility of small units. That distribution of function will push IT departments toward client/server architectures and distributed databases. When compounded by the inverse (or at best linear) economies of scale now demonstrable in hardware, software and data management technology, justification of monolithic centralized systems will be difficult, if not impossible.

Strategic Planning Assumption: Database administrators should plan for an environment in which distributed databases, with substantial subset replication, are deployed across heterogeneous platforms, with update at the lowest possible organizational level.

5.2 Fewer Levels of Management

As levels of management are removed in corporate restructurings, staff to manager ratios are moving from 7:1 up to 15:1 or even higher. Managing fifteen geographically dispersed subordinates is well beyond the capabilities of the "Come to my office!" management style.

Enterprises are rapidly turning to "groupware" — computer-based application support for team communication and team work. Global networks are obvious underpinnings for these systems. Sometimes overlooked is the data impact of groupware. Data types (e.g. image, voice) and data methods not supported by today's relational technology are proliferating, and database sizes are exploding as these systems are deployed.

Strategic Planning Assumption: Database administrators must develop strategies for inclusion of non-coded information and for the resulting explosion (2-4x) of database size.

5.3 Reorganization of Work from Task-centric to Process-centric

Transaction processing systems institutionalize and calcify an "assembly line" approach to the organization of work, in which an enterprise's activities are decomposed into simple tasks, and then the simple tasks are linked together with [often-complex] work-flow procedures to achieve an output. Each task is energetically optimized, but the overall process is typically not optimized. (This partially explains why an insurance company whose systems provide subsecond response time for *each task* still takes five weeks to pay a simple, unchallenged health claim.)

Such "task-oriented jobs in today's world of customers, competition, and change are obsolete. Instead, companies must organize work around process."[3] From a systems perspective, that means moving from quickly processable, highly stereotyped input, the class of application a transaction processing monitor is built to deal with, to multi-dimensional, flexible-path, long-duration processes, which will require a new infrastructure "that combines transactions, distributed tasks, message queuing, batch scheduling and the kind of operational control offered by monitors"[4].

On-line transaction processing (OLTP) using the traditional assembly line model is so important in most commercial enterprises that the software that controls it — the transaction processing monitor — has become the core defining element in most computer architectures. However, to

---

[3] Ibid.

[4] David Vaskevitch, *The Bigger the Application, the Smaller the Database*, Proceedings of the High Performance Transaction Workshop, September 1993

support process-centric approaches to organizing work, a new class of infrastructure software must and will emerge by 1997 (0.8 probability), combining elements of transaction management, workflow management, message queuing and software fault tolerance. "Process managers" will go beyond the functionality of "message managers" now found in the three level client/server model. Transaction processing monitors will still exist in high volume environments, but will no longer be the enterprise's core architectural foundation. By 2000, a TP monitor will be no more than a supportive (and subservient) performance technology found inside the encapsulated transactional steps invoked by a process manager.

Process managers will display several general characteristics of interest to database administrators:

- Timeframes will be extended (days and weeks instead of seconds), and most communication will be asynchronous. This implies a shift in basic exception-handling constructs. Most of the transaction processing systems in use are biased toward flushing exceptions out of the system for subsequent human resolution. That is appropriate for a "transaction", where system and human are working synchronously, but not for a "case" that takes days or weeks, where humans may not be present, and where process steps are asynchronous and even parallel. Process managers will have to try repeatedly to complete failed process steps. Unfortunately, few extant database managers have ACIDity/recovery mechanisms built for extended time horizons or for transaction workarounds, and emerging process manager technology will have to over-ride (or trick) the database managers in the short term.

- The possibility of midpath content change, along with business rule exceptions, application software failures and hardware failures, all imply journalling and roll-back capabilities analogous to those found in transaction processing systems. Rollback may be required to a state which existed days or weeks in the past, and may be partial. The database ACIDity/recovery issues raised are non-trivial.

- If we look at examples of reengineered business processes, we find that they tend to cross what were previously departmental boundaries, to span multiple days, and to be accomplished by staff members with comprehensive and varied job descriptions. In crossing departmental boundaries, a work item being processed may well encounter not only different hardware/software platforms, *but also different styles of computing*. A sales order, wending its way through the sales fulfillment process in a "make-to-order" environment, will encounter transaction processing systems, electronic mail systems, material requirement planning systems, manufacturing process control and shop floor management systems, warehouse material handling systems, shipping optimization systems and financial systems. Despite the heterogeneity in both platforms and styles, data will still have to move through the process without loss of integrity or control. Consistent definition of data across the organization will be imperative. On the other hand, the notion of maintaining consistency by consolidating all data for the process into a single database on a single centralized system will be:
  - unresponsive and unacceptably expensive in a geographically distributed environment,
  - impossible in an enterprise biased toward buying applications instead of building them (the majority),
  - too slow to develop and too cumbersome to maintain in today's marketplace, where speed of response to customers and time-to-market for new products and services is the dominant survival issue.

- Three aspects of the process-centric environment point inexorably to the use of powerful workstations with a client/server application model:
  - The varied nature and constant interruptability of each employee's new job essentially demands a windowing interface.
  - The multimedia component of most new processes exceeds the capability of non-programmable terminals,
  - The need to integrate multiple incompatible legacy systems is most conveniently addressed with intelligence on the desk.

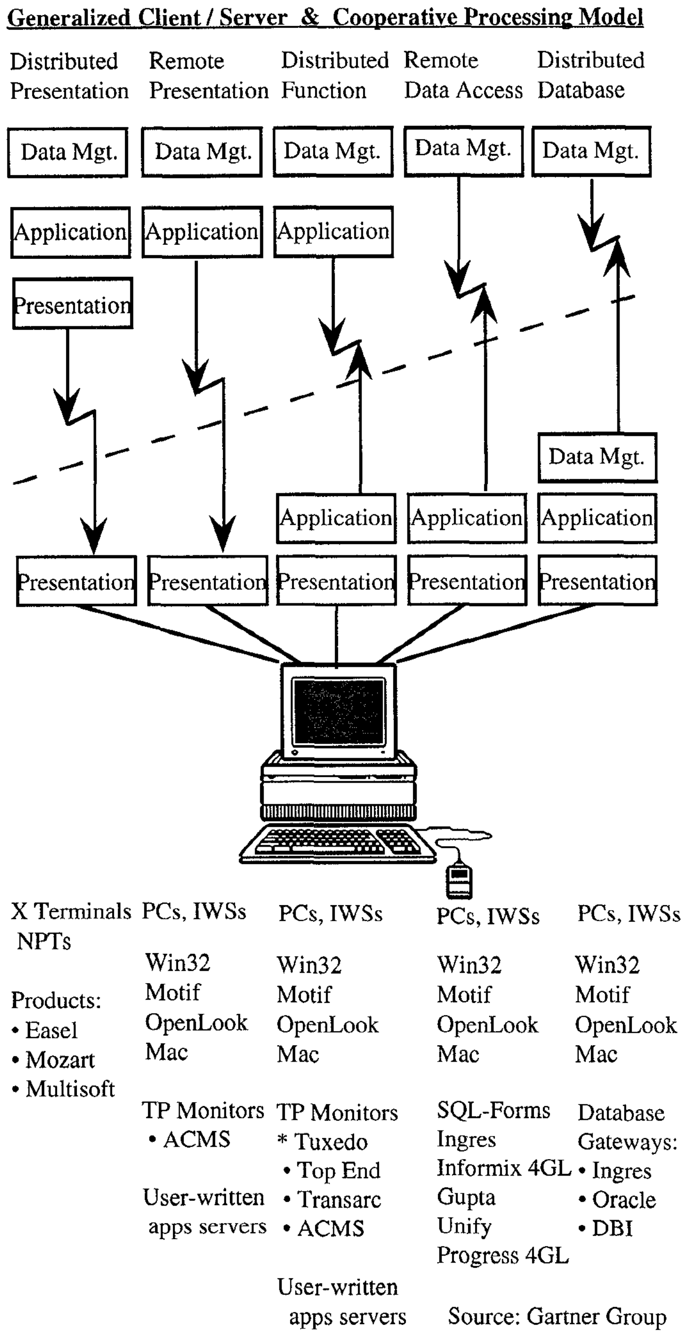


Figure 4: Client/Server Spectrum

Multiple models exist for the client/server application architecture (see Figure 4). From left to right in Figure 4's spectrum, they move from "fat server" approaches, where almost all the work is done on the server, to "fat client" approaches, where the entire application is on the client, and only an SQL call moves across the network to the server.

The simplistic approach of performing most of the application on the workstation and reserving the server for SQL services ("fat client") is appealing because it exploits the power of the chip on the desk and because the paradigm is architecturally simple. However:

- Keeping the processing program close to the data (i.e. on the *server*) substantially reduces transmission volumes.
- Exploiting the database's stored procedures facility raises the integrity and security of the data, while at the same time reducing application programmer load. (This is particularly appropriate when dispersed server applications are being built as encapsulated objects, as they would be in a process manager environment.)
- Permitting client programs to deal directly with database views exponentiates version control administration effort when data is reconfigured.
- In practice, "fat server" applications have tended to scale better in performance than "fat client" applications.

We believe, therefore, that in environments where volume scaling and data security are important issues, a more balanced application architecture will prevail (something roughly equivalent to what we have labeled "distributed function" in Figure 4) .

• Process managers will effectively force an object-oriented systems implementation. The most stable components of the new systems will be individual tasks and the software components that support those tasks. The most fluid will be the routes through those tasks. Ideally, the new systems will permit user managers (what Hammer/Champy call "process owners") to design processes simply by linking together existing tasks in new ways. The concepts of software reusability and of dynamic linkage found in object-oriented software, then, become critical to meeting the business requirement, as do high-level (object-oriented) graphical tools that permit end-user managers to script new processes on a point-and-click basis. In addition, encapsulating methods and rules with the data will provide a necessary element of data protection in this access-biased, flexibility-biased environment.

• It must be assumed that as an activity moves through a process, crossing departmental and functional lines, it will encounter workers equipped with dumb terminals, PCs, Macintoshes, Unix workstations, laptops of every type, PDAs[5], RF[6] data entry devices and voice input devices. Database administrators can assume no limitations on the types of workstations that must be supported or on the types of networks that will be used.

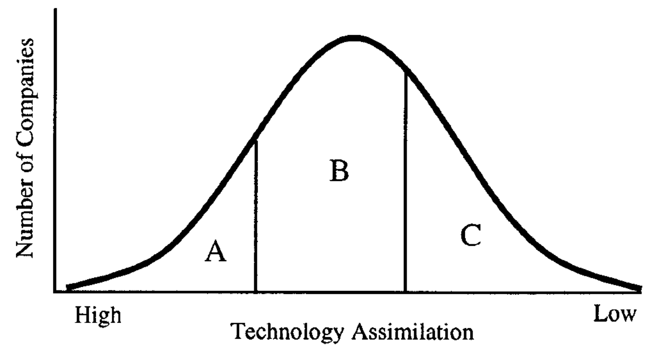
While we can describe the characteristics we expect to see in mature process managers of the future, there are no mature process managers available today. Embryonic process managers (e.g. ATT/GIS' ProcessIT, Digital's Electronic Case Handling Office [ECHO], Early Cloud's Message Driven Processor [MDp], ICL's ProcessWise, Intelus' ProcessFlow) are moving in the right direction, but are hardly complete. *Type A users* (see Figure 5) implementing new systems to support BPR processes should consider coding transactional steps as encapsulated server objects (using either TP Lite or TP Heavy tools), and using an object-oriented workflow manager or an embryonic process manager to control the whole process, invoking the encapsulated OLTP steps along the way. *Type B and C users* should write clean client/server OLTP applications, and encapsulate them later when they implement then-mature process managers (1997 is a reasonable planning target for process manager maturation, 0.8 probability).

Strategic Planning Assumption: Database administrators can no longer rely on a simple pairing of DBMS and TP monitor as a foundation for IT architecture, but must become adept at using combinations of process managers, TP monitors and asynchronous messaging systems..

Strategic Planning Assumption: Database administrators must assume that work items which pass from department to department will also pass from one platform/database pair to another, with attendant integrity and synchronization issues to be resolved.

Strategic Planning Assumption: In the short term, "TP Lite" — the use of the functionality found in client/server and relational database technology to implement a transaction processing system without a TP monitor — will be the most commonly used solution for transaction processing (0.7 probability), and this approach will become even more prevalent as more enterprises engage in business process reengineering (0.8 probability).

Information Technology Cultures



Culture	Type A	Type B	Type C
Known As	Pioneers	Moderates	Followers
Approach	Aggressive (High Risk)	Balanced (Low Risk)	Cautious (Risk Averse)
CEO's Vision of IT	Change Agent Compet've Edge	Productivity	Cost Efficiency
IT Technical Sophistication	High	Moderate-High	Low-Moderate
Funding	Flexible	Variable	Constrained

Figure 5: IT Cultural Comparison

5.4 Integration of the Supply Chain

Perhaps even more important than business unit autonomy in driving the shift to heterogeneous architectures will be the integration of enterprises with their suppliers and customers. Those working relationships will often be consummated by establishing linkages between computer systems. This increasing linkage of enterprises has several effects on IT architecture:

- Even when a large enterprise has constrained itself to one computer vendor internally (rare) and to one computer architecture (even rarer), the external linkages instantly create a multivendor, multi-architecture integration issue. Because the integration of an enterprise's supply chain can not depend on

[5] Personal Data Assistants
[6] Radio Frequency

standardization on any one chip or operating system or database manager, since few enterprises have any control over their partners' platforms, enterprises whose business architecture includes a serious program for value chain integration should consider Gartner's First Law of Open Systems — *The employment of portability at any one layer of a computer system architecture enables substitution in all the underlying layers.* Moving responsibility for portability to the application development environment level, for example, avoids the continuing chaos we anticipate over the next five years in chips, operating systems and database management systems.

- Because all computer vendors will be forced by the market to support open standards, open standards become the most likely mechanism for inter-enterprise exchange. Where open standards are also adopted internally, they minimize the number of technologies an IT department will have to support.

- As client/server application architectures increasingly cross enterprise boundaries, the use of remote procedure call (RPC) mechanisms proprietary to any one database management system will become increasingly awkward.

Strategic Planning Assumption: Database administrators can no longer rely on a monolithic or even homogeneous hardware and operating system environment, even for the most critical applications.

Strategic Planning Assumption: Databases management systems will be a key component of the integrating middleware that defines the new architecture, but prudent users will stay independent of any one database manager by using development tools and standards that span multiple DBMSs.

5.5 The Trend to 24x365 Operation

There is an observable difference in user (and IT professional) expectations for availability in recent customer projects, with an increasing insistence on essentially non-stop computing. Part of the trend is driven from IT back into the business architecture — in a year when a fully disaster tolerant system (of the new architecture) costs less than a disaster-vulnerable mainframe, it becomes tempting to specify disaster-tolerance. Part is driven by the competitive climate, and by the increasing intolerance of customers for less-than-perfect service. Where users once accepted "Computers are just like that" as a reasonable explanation for scheduled (and unscheduled) outages, they now know better. The availability standard for the 1990s is 24 hours a day, 7 days a week, 52 weeks a year[7] *with availability defined as that period of time during which users can perform interactive update*, and with no excuses accepted.

The factors forcing 24x365 operation include:

- A continued trend toward globalization of enterprises, with operations spanning increasing numbers of time zones.

- An ever-more-demanding customer base, trained to expect 24-hour service by ATMs and L. L. Bean, and quick to move to competitors.

- Increased mobility of the workforce.

The next three components of change in business architecture share the implication of 24x365 operation; those components have other implications on IT architecture as well.

5.6 Globalization

As enterprises of all sizes become increasingly multi-national, the effects on IT architecture can be drastic. Governmental favoritism toward local computer manufacturers, governmental pressure toward certain systems standards, multiplicity of regulatory environments and inconsistency in national infrastructures all play a role.

With transactions and processes spanning time zones half a world apart (see Figure 6), one of the most important of the pressures on the old IT architectures is the disappearance of "night-time", that convenient bloc of time when batch could be run, and databases backed up, and errors corrected, and with it, the disappearance of "week-end", when new processors could be

[7] For convenience, this non-stop style of operation is often referred to as "7x24" or "24x365" operation. We prefer the "24x365" usage, since it is deliberately explicit in giving no relief for holidays.

installed, and databases reorganized and new system and application software installed.

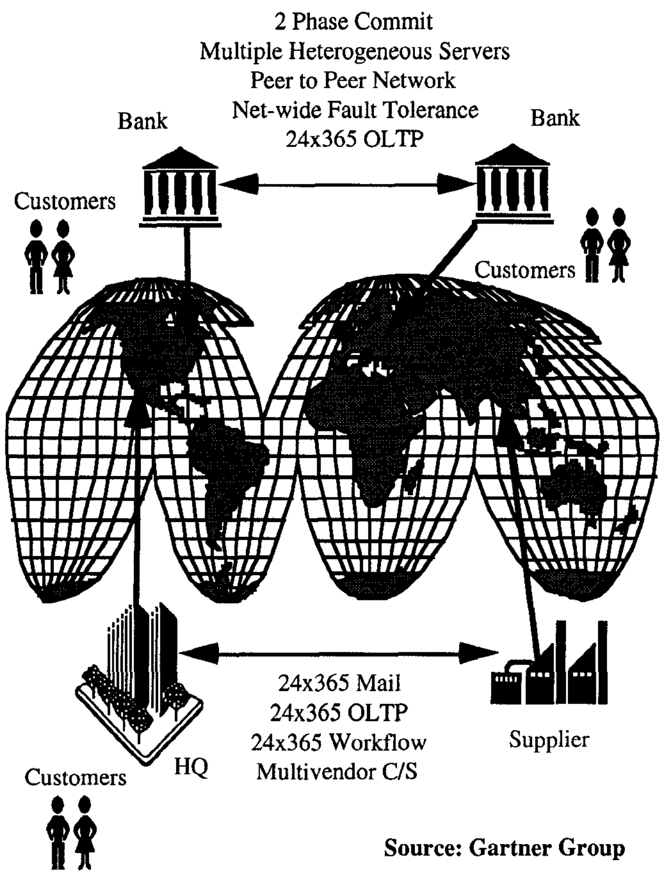


Figure 6: 24x365 Interactions in a Global Network

Strategic Planning Assumption: Database administrators can no longer rely on "batch windows" or "back-up windows" or "database reorganization windows" or even "Software Release Level upgrade windows". All these activities must be accomplished "hot".

5.7 Intense Focus on Customer Service

Consumers (and commercial customers) worldwide are now accustomed to receiving service at any time — 24-hour banking at ATMs, 24-hour grocery stores, 24-hour catalog (and cable TV) shopping services, 24-hour airline reservations. Those consumer have little patience with, and no loyalty for, enterprises that do not meet their standards of convenience. While they may "understand" that a data center has been flooded, or damaged by a bomb, or buried in the rubble of an earthquake, they will also place their order somewhere else.

The good news is that high-availability systems are now easy to find (and to afford) as a wave of clustering technology sweeps across Unix platforms. While truly disaster tolerant systems are still only available from Digital, Tandem and Unisys, we expect even disaster tolerance to be a common systems characteristic in 1996.

The bad news is that operations managers no longer have any excuse for "being down", not even for "scheduled downtime".

While hardware fault tolerance mechanisms will continue to be applied to high availability situations, *network-wide fault tolerance and disaster tolerance require software solutions*, usually involving the deployment and synchronization of loosely coupled clusters and of geographically dispersed redundant resources.

Strategic Planning Assumption: Database administrators must provide fault tolerant and disaster tolerant data resources, which implies expertise in managing and synchronizing geographically dispersed shadow and standby databases.

5.8 Increased Worker Mobility

Database update has historically been founded in an assumption that the source of data was either a) batch input, or b) an on-line terminal [effectively] hard-wired to the host. In the new model, we will increasingly deal with a worker who needs to down-load a subset of data to a powerful, portable system, update the data locally and return it at some convenient time (convenient to the worker) for reintegration.

Strategic Planning Assumption: In most commercial enterprises, database administrators should begin to consider the mobile computing style as the standard, and fixed-site computing as a degenerate case.

5.9 Intense Focus on Cost

It is no secret that all enterprises — in both the private and public sectors — are under intense cost pressure, with lay-offs and corporate losses common news items. In that bleak context, users are demanding IT solutions that are more accessible, flexible and productive — and dramatically less expensive. In theory, many of the technologies (e.g., GUIs, network computing, object-oriented programming and distributed system management) could be supported by MVS/ESA or by other traditional mainframes. But in practice, the new synthesis casts mainframes in a minor role. Users are voting with their pocketbooks, emphasizing commodity microprocessors and portable software. We believe that this is occurring for three fundamental reasons:

- Traditional mainframes still cost too much. Five key deterrents to mainframe use — processor cost, disk cost, system software cost, application package cost and technical support cost — will all have to be resolved before the mainframe can compete in objective terms.

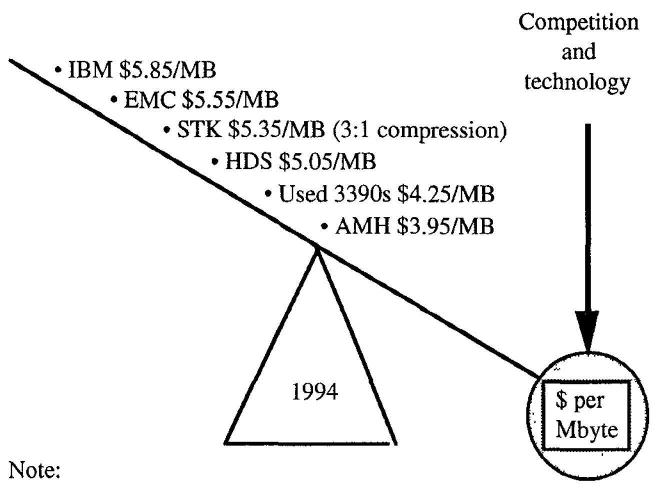
- Traditional mainframes are perceived to be more closed and proprietary than other platforms.

- The pace of MVS software evolution is slow. GUIs, object-oriented programming, and network processing including client/server, are happening faster on non-MVS systems.

The higher unit-of-work cost of the mainframe used to be justified in terms of its absolute performance. That is no longer possible. In 1992, "midrange" uniprocessors at the component level matched mainframe uniprocessors in absolute performance. In 1993, both Tandem and Digital offered systems configurations that could outperform the largest traditional mainframes in absolute OLTP performance, and Digital's Alpha posted batch sort benchmarks faster than Crays or ES/9000s. Besides falling behind in absolute performance, the traditional mainframe has lost the edge it once had in functionality, availability, integrity and security.

We believe that over a five-year planning horizon, platforms that are POSIX-compliant[8] and XPG3-branded[9] will continue to be the most aggressively priced (0.8 probability). The vendors currently at the most competitive levels of price/performance include both RISC and CISC systems and both Unix and non-Unix systems. (We expect Windows NT servers to join the price/performance leaders in 1995, adding to the heterogeneity of the group.) What the price/performers share is XPG3 compliance. The new governing dynamic in systems pricing is that once a vendor gives up a software lock-in strategy and ships truly open (XPG3-branded) systems, then that vendor has no choice but to stay on the commodity price/performance curve.

Projected 1994 Mainframe DASD Average Selling Price



Note:
- ASPs are configured prices
- 360 Gbyte acquisition
- 3990-3s or equivalent
- Controllers with minimum of 64 Mbyte cache
- Most advanced vendor offering

Source: Gartner Group

Figure 5: 1994 Mainframe DASD Selling Prices

From a database administrator's perspective, one of the key economies in the architectural shift is the cost of storage itself. The disks which are commonly used on mainframes had 1993 list prices as high as $10.53/MB, and street prices in the $7.00/MB range. Those prices are falling rapidly, but in 1994 will still range from $4.00/MB to $6.00/MB (street price, see Figure 5), at a time when the [faster] disks used on alternate architectures are selling in the $1.00/MB to $2.00/MB range. Given the future projected price trends for both architectures, we see mainframe disk continuing to be at a 2:1 (or even 3:1) cost disadvantage for at least three years (0.7 probability). One might expect that the mainframe disk would make up for higher pricing with superior performance, but in fact, the inverse is true (see Figure 6).

---

8 POSIX is an industry-standard application programming interface (API) used for interaction with a broad range of operating systems.

9 X/Open, a standards-setting industry consortium, publishes the X/Open Portability Guide (XPG), now available in Versions 3 and 4. These guides contain standards that go well beyond the operating system interface. X/Open administers a branding process for complying platforms, and has to date branded most Unix and several non-Unix platforms.

1993 Disk Performance Comparison

Manufacturer	IBM	IBM	DEC
Storage type	3390-3	3390-9	Storageworks
Model	A38/B3C	A98/B9C	SW812
Capacity	90	271.68	108
Transfer rate	4.2 MB/sec	3.9 MB/sec	4.9 MB/sec
Mean access time	19.6 ms	45.3 ms	15.1 ms
List price	$954,850*	$1,289,050*	$226,438**
MMC 2nd year on	$1,175	$880	$2,252
Raw $/MB	$10.53*	$4.74*	$2.25**

* Requires 3990 controller ** Includes two controllers

Source: Gartner Group

Figure 6: 1993 Disk Performance Comparison

Henderson's Master Platform model postulates that there is bi-directional impact between Business and IT Master Platforms and between Business and IT Architectures (see Figure 1). Disk storage is a case where a sudden change in technology and price is driving a change in the business architecture. While the drop in price has resulted in lower spending in some cases, in others it has merely unleashed pent-up demand. As a example, a multi-billion-dollar U.S. distribution company processes about 400,000 sales order line items a day as input to its sales order fulfillment activity. An equivalent number of records are consequently inserted into a decision support archive for use by the marketing and inventory management functions. Those corporate functions have always wanted five years of detail on-line. The decision, made when storage was $15.00/MB, was to provide 13 months of detail and five years of summary data, a less-than-satisfactory compromise. The same decision, made now when storage is $2.00/MB, would result in a database five times the size, and would meet the real user requirement.

Strategic Planning Assumption: Database administrators should treat open systems as the norm for investment and training, and should assume that traditional mainframes are exceptions.

Strategic Planning Assumption: Database administrators should plan for decision support archives to triple in size at a minimum as falling prices release pent-up user demand.

6.0 Remaining Obstacles

Three elements (besides human resistance to change) continue to impede the transition to the new architectural base:

- A lag in the availability of industrial-strength systems management tools for many of the platforms (especially the Unix platforms) which are used in the new architecture
- Continued distrust of Unix file systems
- A lag in high availability Unix solutions

6.1 Systems Management

[Note: This issue will be ameliorated in 1994 and essentially eliminated in 1995 (0.7 probability).]

If a new architecture is to be the platform for the mission-critical applications that used to run on mainframes, then users expect to have systems management tools like those on the mainframe. While Digital's VMS systems have for some time offered reasonable tools, most Unix systems have only in the past year begun to catch up, and then only for homogeneous environments. But in this brave new world, few user environments are homogeneous, and even the relatively "simple" case of finding a set of systems management products that can manage both Hewlett-Packard UX and IBM Aix systems, for example, is far from easy. The cross-vendor management task becomes even more daunting when more than two vendors' systems are required targets, or where OpenVMS or OS/400 or NT are added to the equation.

Confusion in the Unix systems management market is running at an all-time high. The "disintegration" of the Open Software Foundation's Distributed Management Environment has created a strategic planning vacuum for many users. Unix System Laboratories' (USL's) Distributed Manager (DM) also appears to be stillborn, as Novell has apparently decided to stop developing and licensing the Tivoli-based framework and SVR4 vendor-supplied management applications to SVR4 system vendors.

Users should not wait for standards efforts (such as COSE or X/Open) to address the practical requirements of Unix (or multiple operating system) systems management. The good news is that technologies to address cross-vendor systems management tools are emerging from a handful of system vendors and ISVs. HP's Operations Center now supports HP, IBM and Sun Microsystems Inc. Unix platforms, while IBM's System Monitor/6000 agent technology is shipping on IBM and HP platforms, with Sun and NCR Corp. support scheduled for 1Q94. Computer Associates Inc.'s Unicenter for Unix product should be available across HP, Sun, Sequent IBM and DG platforms in 1Q94. In 1994, we also expect to see cross-vendor Unix management products delivered by OpenVision Technologies Inc., Legent Corp. and Candle Corp., as well as a more complete suite of management applications from Tivoli.

6.2 Reliable Unix File Systems

Much of the investment that vendors like HP and DG have put into their Unix implementations has been in the area of adding robustness to Unix file systems, a mandatory investment, since file system fragility was a major obstacle to Unix penetration in the commercial market. 1994 will see the emergence of new file systems for Unix that are both journalled and high-volume-capable, and the issue will be essentially eliminated by 1996 (0.7 probability).

6.3 High Availability for Unix

High availability concerns for Unix are being eliminated in 1994 as the majority of Unix vendors announce clustering technologies aimed at the kind of availability and functionality previously demonstrated in the VAXcluster.

A cluster is a loosely coupled arrangement of independent systems that 1) share a common storage resource, 2) appear to users as a single system and 3) execute a common application The goals of "clustering" are twofold — scalability and availability. "Good" scalability is either the ability to run a given job twice as fast with twice as much hardware, or the ability to run twice as big a job in the same elapsed time with twice as much hardware. Availability is the ability to continue operation after failure of any system component. In 1994, cluster technology already exists to continue interactive processing through hardware failure/maintenance, operating systems change, database reorganization, and natural and military disasters.

7.0 Summary

In *Future Perfect*, Stanley Davis characterizes the deficiencies of traditional organizations this way: "In the industrial context, organization always lags behind strategy. Because of the assumption that you have to know what you want to do before you can know how to do it, all organizations based on the industrial model are created for businesses that either no longer exist or are in the process of going out of existence."[10]

In sharp contrast is this description of a modern enterprise taken from Marilyn Ferguson's *The Acquarian Conspiracy:* "The network is the institution of our time: an open system, a dissipative structure so richly coherent that it is in constant flux, poised for reordering, capable of endless transformation. This organic mode of social organization is more biologically adaptive, more efficient, and more "conscious" than the hierarchical structures of modern civilization."[11]

The 1990s will be a decade of rapid change in the economic, political and competitive environment. Enterprises will adjust by changing their "business architectures" — organizations, policies, practices and decision methodology — to something approaching Ferguson's optimistic, even idealistic, vision. Information technology (IT) organizations will realign their IT architectures to support the new business architectures, and will select from a long list of new technologies to do so, effectively discarding the most fundamental computing paradigms of the last two decades. There will be new ways of defining and organizing data, new ways of designing programs, new tools for code construction, new ways of deploying hardware, new ways of interacting with the end user and with other enterprises, new ways of managing business processes. The new environment changes the fundamental assumptions used by database designers and administrators for the past two decades.

---

[10] Stanley M. Davis, *Future Perfect,* Addison-Wesley, 1987

[11] Marilyn Ferguson, *The Acquarian Conspiracy: Personal and Social Transformation in the 1980s*, Tarcher, 1980