Global Multi-Media Information Utilities

(CS Future Work Group Vision 1)[1]

August 1989

Ph. Janson

IBM Zurich Research Laboratory
8803 Ruschlikon - Switzerland

Abstract

The present report summarizes ideas that resulted from a joint effort by some 20 IBM Research Division staff members, who met regularly over a period of 18 months to discuss visions about the future of computing. The particular vision described here involves a hypothetical, global, multi-media, videotext-like information utility. The utility is assumed to have a geographical coverage similar to that of today's telephone system, over which it can deliver on a flexible pay-as-you-go basis the most varied, multi-media information services, such as multi-media electronic messaging, broadcasting, multi-media reference and encyclopedic database querying, rich transaction services, teleconferencing, on-line simulation, visualization, and exploration services etc.

[1] This footnote was added in 2010 to the original 1989 version of this paper. The paper was originally issued as a classified IBM internal report to document a vision developed by some 20 IBM researchers over a period of 18 months about how cyberspace – as we call it today – might look like in the 21st century. After Moshe Vardi, now at Rice University, dug it out from his files, the paper was declassified and is now made publicly available for its potential historical value. It may indeed be interesting to see retrospectively how computer scientists of the late '80s imagined cyberspace could be like 10-20 years later. It is interesting to see how generally accurate their vision was, albeit without today's "web" terminology and modulo a few wrinkles that still need some more work for ironing out.

1. Introduction

Visionary views of the future of computing can be triggered in many ways and based on many hypotheses: extrapolation of a technology, forecast of a market requirement, observation of an economic trend, fantasy of an inventor etc. This report discusses a vision that was motivated by several such hypotheses:

1. Application software will play a dominant role: it is a generally accepted observation that hardware technology is marching forward while the primary obstacle to pervasive availability and use of computing is software. Except for exotic hardware, MIPS, MBytes, Mpels, and Mbps are becoming commodities, whose availability, high performance, low cost, low power, and small size everybody takes for granted. On the contrary, software technology has traditionally trailed hardware technology: a lot of software still in use was designed to fit the limits of outdated hardware; it cannot exploit fully the potential of advanced hardware, fails to meet the needs of complex information systems, and often exhibits poor user-friendliness. One reason for this disparity is that advances in hardware technology are driven by clearly measurable cost/performance objectives, whereas metrics for rating software objectives and packages are not as clear, largely because it is hard to quantify the human factors involved. Thus, now that adequate hardware is affordable, we find that writing software to exploit it fully is hard because we do not understand user requirements, user interface design, and application specification as well as we understand corresponding hardware development processes.

   Based on these observations, one may forecast that the future evolution and expansion of computing in general and thus of equipment suppliers in particular will hinge much more on an ability to develop attractive software than on advances in hardware technology, which is becoming a commodity: software will count much more than hardware. The vision described in this report therefore puts a major emphasis on software in general and application software in particular.

2. A global multi-media information utility: given the predominant future role of application software, the vision suggests that massive R&D efforts must be devoted to understanding the requirements of applications and application substrates. Yet, from numerically-intensive computing to robotics, the spectrum and variety of possible applications domains is almost unlimited so that equipment suppliers even as large as IBM cannot possibly envision ever providing a complete palette of applications. Larger companies should focus first on those key application domains most likely to be in high demand. Beyond stressing application software in general, the vision discussed in this report chooses to focus on a set of applications, loosely called information systems, that it predicts will be particularly important and pervasively required.

In moving from the industrial age to the information age, the developed world has moved the focus of its activities from natural resources to information. More and more activities in today's economy are concerned with the management of vast amounts of information about every aspect of our lives. In fact, one observes dwindling supplies of conventional energy and certain raw materials and a tremendous waste of natural resources in general, often due to poor management and insufficient information. In the future, the wise management of these scarce or dwindling resources will require more and more sophisticated information systems for efficient creation, acquisition, analysis, sorting, exchange, or other processing of information.

The vision of future information systems discussed in this report revolves around the concept of a global "omniscient" and universally accessible multi-media information utility aimed at vastly improved information management capabilities, reduction of wasteful energy and resource consumption, and optimization of our use of time.

In brief, this utility would allow any consumer of information to locate, access, filter, restructure, classify, or in any other way process public or business information to suit his needs from anywhere in the world without having to leave his armchair. The vision addresses the belief that more and more people need more and more access to more and more information, and that that information should move to people, not the other way around. As much information as possible would be created, stored, transported, and processed electronically so that it would be available as rapidly, selectively, and universally as possible, and would have to be transported mechanically via envelopes, packages, or people traveling to convey it as seldom as possible. The information filtering functions are particularly important because, with the vast amounts and variety of information stored in the global system envisioned here, users would be overwhelmed and the system would hurt more than help, unless powerful tools allow users to select, extract, and control precisely what information they want to see.

In some sense, the proposed vision may not appear to be revolutionary: the envisioned system might be dismissed as a safely predictable extrapolation from, and merging of existing information tools that it may complement or even replace: mail, newspapers, catalogs, reference books, dictionaries, encyclopaedias, telephone, telegraph, telex, radio, television, VCRs, videotext, teletex, fax, electronic mail, EDI, on-line databases, bulletin boards, meetings, conferences etc. On the other hand, while such extrapolation may already have been predicted by many, the sheer size of the envisioned system and the idea that it may well completely supersede existing systems presents enormous technical, social, cultural, political, legal, and other challenges that we don't know how to address. Thus, rather than dismissing the vision as an evolutionary certainty, we should on the contrary study it carefully in order to be able to play a role in its evolution.

3. Capturing the consumer market: a last working hypothesis of the vision observes that, aside from a marginal home market, today's computing is confined mainly to the business and scientific worlds, while the need or desire for information certainly is not. Thus we suggest that the extension, in due time, of a global multi-media information utility to the home market could open the general consumer market to computing with unprecedented opportunities for the hardware and software suppliers. While the information utility discussed here finds its roots in the business and scientific worlds, it is eventually aimed at the broader consumer world as well.

   The rest of this report is organized in seven sections. Section 2 defines components of the information utility in some detail. Section 3 suggests specific information services that may be offered in such a system. Section 4 then investigates the extended distributed system services and application substrate required to support a global information utility. Section 5 discusses user interface requirements for the system. Section 6 looks at the requirements for user workstations. Section 7 addresses (very briefly) server issues. Finally, Section 8 tries to draw some preliminary lessons from the exercise, and to chart the course of future work in support of the vision.

2. System Components

Prior to discussing other aspects of a global information utility, it is desirable to identify the components of such a utility. These are not fundamentally different from the components of classical distributed systems, and are thus reviewed only briefly hereafter.

Network Suppliers vs. Subscribers

The envisioned system revolves around a global communication network through which the widest variety of information storage, transmission, processing, and management services can be accessed. Right away, we draw a line between the network itself and the information processing facilities attached to it, at essentially the same place and for the same reasons that the line is drawn today between "common carriers" (PTTs etc.) and their subscribers: inside the network boundary are communication links and switching nodes whose sole purpose is the transparent transport and routing of information between subscriber end nodes; outside the network perimeter are the subscriber nodes, which may both provide and/or use information services to/from one another. Separating the network from its subscribers is desirable, firstly for administrative reasons, and secondly, to allow independent evolution of network and subscriber systems.

Future networks will offer high-speed, low-cost communication (both point-to-point and broadcast) of any form of digitized information, text, data, graphics, image, voice, video etc. Alone, the construction of a world-wide network with such function, performance, and size requirements presents considerable technological, financial, and political challenges. While IBM may contribute some of its research and technical know-how towards solving the technological challenge, it is questionable whether it should, or even could, be a major player in solving the financial and political challenges. IBM undoubtedly plans to expand its existing Information Network into a full-fledged, global carrier facility. Yet there should be some doubt about how successful it will be in that arena. Installation and operation of plain transmission and switching facilities traditionally have been and, in some countries, are likely to remain the reserved playgrounds of telecommunication companies and state monopolies that are beyond IBM's control. And it is likely to be as hard for any computer company to penetrate the global telecommunication market successfully, as it is for traditional telecommunication suppliers to make significant inroads into computing markets.

Helped by the installation of fiber optic links, basic rate ISDN ports (2 x 64 Kbps), broad-band ISDN networks, and high speed trunk lines (T1/E1, T3 etc.) are well on their way; pilot metropolitan area networks (MANs) are being assembled in several countries; and it is quite reasonable to expect that common carriers will gradually phase out old iron and offer more and more high-performance leased and switched facilities able to meet the communication requirements of a global multi-media information utility. The process may be long, and will be evolutionary, but it leaves little room for doubt. Separating networking issues from subscriber and application issues allows the independent evolution of both aspects of the envisioned information utility, new information services becoming possible as new networking capacity is installed.

Server Subscribers vs. Client Subscribers

Having– informally – drawn a line between network suppliers and network subscribers, we now focus on the subscribers. As the functions and operation of subscriber equipment will call upon computing much more than upon communication technology, IBM could and should be a key player in this arena, as a supplier of information processing hardware and software, including extended distributed system services to be discussed later, such as system management, subscriber directories, authentication, access control, accounting and billing.

From the network supplier's point of view, all subscriber stations are technically indistinguishable (as in today's phone system), though they may differ widely in terms of their network usage characteristics and thus be subject to different tariffs.

Among subscriber stations, though, a clear distinction will exist between producers and consumers of information services, the former being referred to as server stations or servers, while the latter are called client stations or more loosely users. Of course, a subscriber station may offer services based on enhancing services offered by others: in

this case, it is simultaneously a client and server, but will be referred to as a server, because this quality is determinant in its mode of operation.

Servers will provide a variety of information services (to be discussed in the next section). Generally they will require relatively large information storage, retrieval, and processing capabilities, and will run on host-based systems, geared towards the back-end of computing, information storage, retrieval and processing; they may depend on moderate to substantial operator expertise or they may run "unattended"; they are generally not found lying around on desks. In this context, the term host is used to suggest a mode of operation in support of the information server role; it is not used to suggest any specific machine size or processing capacity, though many servers will be powerful. By contrast, client stations will be operated by users wanting access to the offered services. They will typically be workstations offering human interface support devices with varying degrees of data, graphics, image, video and audio capabilities. Again, the term workstation is used to suggest the provision of human interface support, not to suggest any specific machine size or processing capacity. Workstations are geared towards front-end computing, human interaction; they require little or no user expertise beyond common sense, and are as ubiquitous and inconspicuous as TVs or hi-fi sets. They will tend to be less powerful than server stations but this need not always be true.

This view thus foresees a polarization of computing resources into shared host installations on one hand and individual workstations on the other hand. This is not to say that minor services could not be offered by unattended personal computer-size machines, or that client users could not access the network through a private super-computer. Nor does the polarization rule out machines in the "mini" performance range between individual personal computers and large hosts. The argument bears on user interface considerations, not on computing power. Within the context of this document, client workstations do not necessarily differ from server hosts in performance, but mainly in environmental and user interface requirements. Selected client workstations may in fact offer more processing capacity than some service hosts: for instance, it may make sense both economically and functionally to locate processing power in professional workstations rather than in hosts to support 3-D design, real-time simulation and visualization functions.

3. Applications

Having briefly looked at the components involved in the contemplated global multi-media information utility, we now describe some of the applications that may be envisioned in such a future utility. We start with the simpler applications, which are already pre-figured today in some systems, albeit not on the scale and pervasiveness imagined here.

Multi-media Mail and Messaging

Under this generic label, we collect applications that include communication between two individuals. In our scenario, this would typically be individual clients, although intermediate mail servers may be involved in many cases.

Perhaps the oldest form of communication between people after face-to-face meetings is plain mail. However mail took away from communication an important aspect that face to-face meetings had: the real-time interaction. Thus early electronic communication systems such as telegraph, telex, and teletext were later developed to reduce the "transmission delay" of mail. Fax systems added the ability to transmit limited-quality pictures in addition to text. Nowadays, electronic mail and file transfer systems are common place within the world's research community, and are rapidly spreading through the intra-company business world, while EDI systems are taking off in the inter-company environment. However, such systems have not yet invaded the private home markets.

At the other end of the spectrum is the telephone system, which preserved the real-time features of face-to-face meetings, but missed some of the other nice features "discovered" with mailing systems, such as non-disruptive communication, transmission of non-voice information, delivery of information in bulk, in absentia, possibly with certified acknowledgement, random message retrieval etc.

Our vision suggests that a global multi-media information utility of the future will combine the best features of both of these worlds, and extend them to new types of information and new modes of communication to offer an integrated global multi-media phone-mail service not unlike the Pygmalion system developed by M.I.T. and DEC but on a much larger scale. This service would be available not just within isolated research communities, large corporations, or closed commercial partnerships (on a contractual basis), but to any individual, business employee or private citizen, from his desk, home, a public telecom booth (reminiscent of today's phone booths), or even a mobile terminal.

From the phone system, the utility will copy the ability to establish real-time communication between individuals, including call forwarding and rerouting, camp-on and call-back, automatic and abbreviated dialing, reverse and credit card charging etc. and to extend these functions to all forms of information, including whatever amount of hi-fi audio, data, graphic, picture, image, and animated video (i.e. videophone) can be transmitted in real-time with available technology at affordable rates. This will initially put restrictions on the amount of data, the resolution of images, and the speed of video information, but technology will keep pushing these limits up and costs down.

From the electronic mail and file transfer systems, the utility will copy the power to add all kinds of services to the basic in absentia bulk delivery capability, and extend it to all forms of information including voice (i.e. editable voicegrams), hi-fi quality sound, image and animated video, allowing even the combination of these into integrated documents.

Added services will include logging, tracing, certified acknowledgement, distribution lists etc. Possible enhancements include the concepts of active mail, animated messages etc., although these tend to be open doors for viruses.

Multi-media Broadcast Applications

Next the system will support the broadcasting of text, graphic, image, audio and video information in a way pre-figured today by bulk mailings, newspapers, radio, TV, and teletext systems.

All such broadcasting should support selective filtering mechanisms such as are offered today in teletext and some TV systems. Teletext systems continuously broadcast coded information during the fly-back time of the TV beam; this information is divided into numbered pages, and users can program their receiver to log and display selected pages, thus offering some filtering of the broadcast information. While a similar effect can be achieved for selectively taping a TV program on a VCR (by programming the channel and time for the recording), some TV systems actually include additional key-coded information in their broadcast, so that VCR's can be triggered to record a given program rather than blindly peeling whatever is being broadcast at a programmed time off the air. Such mechanisms will have to be generalized to allow users to retain only selected features or categories among all broadcast information. A small-scale precursor of such a system has been built by D. Gifford at M.I.T. in a project involving the broadcasting of New York Times excerpts.

Yet, while selective filtering is desirable for the users, to attract financing and services from information suppliers successfully, it will also be necessary to allow unfiltered broadcast services of the sort that dump advertisements in everybody's mailbox, radio waves, and TV cable. Additional mechanisms and rules may be implemented to confine advertising to certain times, limit it to certain volumes, flag it so it can be easily recognized and ignored or discarded, and ensure that users receiving it never have to pay for it.

Beyond limited client-side filtering techniques, the spatial scope of most broadcasting applications may have to be limited to less than the whole globe. While there may be a global market for certain TV (e.g. CNN) or radio (e.g. BBC World Service) programs, it is indeed hardly thinkable, technically, economically, and probably even socially, that everything that ever gets broadcast anywhere in the world should be broadcast to the entire planet. Thus broadcasting is one category of services where the "globality" of the information utility may be purposely curtailed.

Multi-media Browsing Applications

The above class of applications was based on wholesale broadcasting of information by a server with a limited passive filtering ability at the receiving client side. The present class of applications still assumes that clients have no control over the information stored in the server, but can selectively control what information is actually sent to them through active filtering at the server.

This generic class of applications includes the whole realm of browse-only on-line databases pre-figured by today's videotext systems in general, but extended to all forms of information and media, and accessible on a global scale.

The filtering techniques in this case include not only interactive user selection through menus and browsing, but also batch client selection through subscription: the user could obtain information either by interactively browsing through a database or by subscribing to it to be regularly notified of new information on topics he has selected, thus himself delimiting the content and organization of the information he wants to receive as part of his own personal multi-media electronic newspaper or magazine.

To the extent that many people enjoy the touch and feel of paper, and buy or subscribe to certain beautiful publications to collect them, newspapers and especially magazines and albums will continue to exist. However, for many other people swamped by today's information flood, the ability to filter what is available to them as narrowly as possible is becoming a necessity. For them, browsing services allowing selective access to a personalized multi-media electronic newspaper would be a welcome replacement to existing newspapers and trade journals.

The argument is even more compelling for information of other natures, such as reference books, dictionaries, encyclopedias, libraries, sales catalogs, record or video libraries etc. To be able to look-up the electronic Official Airlines Guide from one's desk or living room is already tremendously appealing to some people. In the same way, looking up phone books, record catalogs, sports calendars and results, weather forecasts from around the world, classified ads, city maps, hotel and restaurant guides, movie programs, and every other sort of information one uses daily would be very appealing to many. Users would not have to regularly re-order the latest (up-to-date) edition of whatever reference publication they need: the on-line version would always be up-to-date. And users would not waste time looking for the right reference and being distracted by other irrelevant material passing by their eyes: doesn't it sound like a familiar feeling to dive into encyclopedia articles from dams to dynamite when all you were looking for was dinosaurs? Several reference books such as the Oxford English Dictionary, the Encyclopedia Britannica, or the Grolier Encyclopedia are already available electronically on optical discs. Services could be built around these to offer on-line access to a continuously updated electronic edition, with fees being charged on a per volume or per query basis.

Such services should of course not be limited to textual information, and could be available on a global basis either through replication of databases, or through remote access depending on which is more economical. Record and movie companies could

offer excerpts of their productions as part of their on-line sales catalogs. The precious collection of manuscripts of the Getty Museum have been put on videodisc simply because they are too delicate to be manipulated by casual visitors, but such works could also be offered on-line, thus allowing remote access to a museum's collections.

Also, with electronic library search tools, such as hypertext systems, users could locate faster far more and/or better references to what they are interested in. For this to be possible, an essential feature that browse/subscribe databases should support is common data formats so that individual clients may extract related information, compute new views and relations, and build new networks of hypermedia documents from separately maintained databases, as dictated by their own needs and inclinations: for instance, a user could create cross-references or hypertext relations between entries in the Who's Who and the publications of a given author, or the biography of a composer and recordings of his works, as has been done for instance in the Music Department at U.C.L.A. In fact new services may be born from such cross-referencing or merging of pre-existing databases, libraries, discographies, and “videotheques”.

More sophisticated but still among interactive browse-only applications will be simulation applications, which will allow users to "browse" through (actually interact with) multi-media databases to experience the look and feel of virtual objects through sophisticated simulation packages. Flight simulators, both real ones as well as their more primitive though quite entertaining counterparts on today's personal computers fall within that category. Such applications consist of a simulation package embodying some physical concepts and laws, running on a workstation and executing against a database of information describing some virtual world. With tomorrow's database, communication, and processor technologies, one could easily envision a flight simulator database maintained on dedicated servers where it is constantly enhanced in geographical coverage and refined in level of detail so it can be used concurrently by multiple users flying or in fact sailing, driving or walking around the world, possibly even meeting one another on their respective screens. Precursors of such systems have already been built: the Aspen simulator designed at M.I.T. some years ago, the Penobscot Bay navigation simulator developed by DEC and M.I.T., or the Palenque DVI demonstration designed by Intel.

Interactive Transaction Services

Of course in many applications, once a user has obtained the information he was looking for, he will want to act on it, book, order, sell, bet, commit, or advertise something. To this end the system must also support a wide variety of on-line transaction services. It will allow a wide spectrum of applications from traditional on-line banking and shopping, to contracting, scheduling, ad placement, data archival etc., most of these coupled with automatic electronic funds transfer for billing.

Along these lines, the concept of a global hypermarket could progressively develop. Within the limits of administrative and legislative territories, it would be possible for

buyers to do comparative shopping and place calls for bids. In fact a whole new bargaining and dealing service may see the day through the system, which would collect and match buyers' calls and sellers' bids, and would manage and seal the contracts, all on-line.

Multi-media Conferences

A last class of applications that will be found on the system would be those involving communication among multiple clients, possibly through intermediate servers.

These include multi-media extensions to today's (store-and-forward) electronic conferences and bulletin boards, and (real-time) teleconferencing and "shared black board" systems. Such applications could help reduce business travel, which is consuming so much time and so many resources, yet remains essential today for lack of really practical and affordable technology offering the equivalent of face-to-face meetings.

Besides reducing the need for business travel, the system would also encourage such applications as teleworking: rather than wasting time and energy commuting, more and more of today's information workers, white collars, professionals, and managers, could carry out part of their work from home or more likely from small, local, distributed community work satellites providing just individually rented office space and convenience facilities, no meeting rooms, libraries, archive rooms, file cabinets, or computer centers.

Of course no electronic system is ever likely to replace completely the need (or desire) for business travel and commuting to work: communicating with effective realism and in real time all the subtle nuance of eye and body language involved in delicate business meetings is probably beyond technology. Furthermore, some people will always like business travel and prefer working away from home. However, where eye and body language don't matter and for people who like to minimize traveling, such applications could offer welcome alternatives.

Educational applications

The initial investment to get all the above applications off the ground will not be negligible, and thus most of them will be born within the context of the business world, which will take the burden of development and installation under the pressure of necessity and appeal. Yet, once pieces of the system start appearing and their cost starts diminishing through technological progress and economies of scale, there is no reason why its use should be limited to businesses.

The next sector which would likely join the system is educational institutions, starting with colleges and evening schools, and progressively reaching high and elementary

schools. Many of the services suggested above could be offered on the system with an educational slant to allow larger numbers of remote students to benefit from the expertise of selected teachers and pedagogical value of selected courses: individual and collective videophone and electronic mail between teachers and students, scheduled TV broadcasts of regular class material, on-line browsing and subscription to library services, interactive computer-aided education workshops, on-line exercises and tests etc.

Of course school is not just about learning from books and libraries. The whole social context is also essential and could not be replaced completely by an information system anymore than business travel could. However, a rich information system could be an outstanding complement to the school experience.

In certain aspects of education, well-guided use of computers has already proved dramatic benefits, and students seem to learn much more rapidly with the assistance of computers. For instance, in education perhaps more than in offices, simulation packages may play a key role: the ability to let students fly planes, watch the growth and reproduction of cells, model the propagation of an epidemic disease, visualize complex equations and mathematical models at play, perform chemical wet-lab experiments, smash particles together, travel the world etc. is unequaled in traditional education methods.

Home applications

Finally, the private home consumer market would probably be the last to join the system because of yet unclear needs for such services and the initial high costs involved. However, following our hypothesis that a global multi-media information utility could ultimately be attractive to the consumer market, the present section looks at what applications could see the light of day in that specific context.

In addition to being a window on the outside information world, a workstation in the home could become the central home switchboard, controlling telephone, answering machine, hi-fi, TV, VCR, door locks and openers, alarms, and control system, mostly through a universal IR or FM transmitter capable of interfacing to all existing equipment through their own remote control interfaces.

While previous attempts to push computers into homes have not had dramatic success, some ideas have worked. The main problem is a chicken-and-egg one: nobody will offer applications until consumers have workstations to access them, but customers won't buy workstations unless there are some attractive applications. Yet if properly designed, widespread, and affordable, an information utility could be appealing for home use, even if initially not quite viable commercially, as suggested by the Minitel system in France. The French Minitel system is based on very rudimentary, under $1000 workstations communicating with servers across a plain low-speed national X.25 network. The immediate success of the system resulted largely from the fact that the

workstations were given free of charge to millions of homes. Initial home use of the Minitel system was essentially only as an electronic phone book and bulletin board (aside from adult entertainment applications) because homes have relatively little use for other things such as airlines schedules and stock market quotes, and there weren't many other really "irresistible" services on Minitel at the time. However, the lesson seemed to be that home users are enthusiastic about Minitel, even based on the initially modest application offerings. At any rate Minitel was so successful in general that its use caused a major breakdown and called for immediate upgrade of the French Transpac X.25 network that carries Minitel traffic. Today, more and more services, mostly of the videotext kind are appearing. An overwhelming number of ads (on any media) and talk programs on French radio and TV advertise a Minitel code one can dial to get more information or contribute feedback on the subject at hand. Another example of a videotext system that seems reasonably successful in homes is IBM's joint venture in the Prodigy service.

An important vehicle to spur the development of home applications will be business applications. Indeed, home use of the system will be encouraged indirectly by its penetration of the business world as teleworking becomes more frequent. Most people today do not miss electronic mail because they don't even know what it is like to have on their desk or at home a service with almost the immediacy of the telephone at almost the cost of mail. However those of us who do know what it feels like, usually find it irresistible, have a terminal and modem at home, and wish we could use it to communicate with our families as well.

In order to really capture the home market, the system will have to offer more applications truly appealing to the home. As private citizens simply will not buy workstations without applications, application developers must take risks and develop attractive packages to break the chicken-and-egg problem. These applications include solo games (hopefully more intelligent than the average arcade blitz) as well as face your-neighbor or join-the-pack multi-user games, remote home control and alarm applications, on-line cook books, maybe church or community conferences for the elderly and handicapped, on-line shopping and banking, classified ads, enhanced answering services comparable to that which phonemail offers today to businesses, and general entertainment.

In the latter category, one may envision on-line ordering of recordings and movies, like radio and TV à la carte, the electronic equivalent of record and video rental stores: individual copies of audio or video tracks or just samples thereof could be sent to homes in real-time if bandwidth allows, during the workday otherwise, for viewing in the evening. The serious amateur could spend his evenings focusing on one composer, one conductor, or one actor without going down to the store every day, looking for things it may not carry or discovering those it does carry are already signed out.

Another gimmick that could ignite a home market is animated wall-painting art. After all, some art fans do spend thousands of dollars for paintings or other forms of home decoration. So why not offer large flat color graphics displays or screens onto which the

home workstation can project artful photographs, drawings, or even continuous high quality animated art pieces bought or rented over the network from computer art dealers? Let your fancier dreams come true ....

In the home as in education or in the office, we should expect simulations to be an overwhelmingly important set of applications. The Design&Decorate DVI demonstration by Intel is a fairly convincing precursor of things that might come in this arena. With sufficient bandwidth, one could even imagine entertainment packages based on simulations of places in the world: the user would be able to wander through the streets of a foreign city, explore national parks he does not know, fly over Caribbean beaches etc. based on libraries of street scenes, pictures of the world, or coded descriptions of places that he can call up.

4. System Substrate for Enabling Applications

Although the ultimate objective of the global multi-media information utility is clearly to deliver application services, no application will ever be possible without a substantial amount of underlying system services and building blocks to enable them. Research will be required not only to develop many of the applications suggested above but also for the development of the system substrate necessary to enable these applications. This section discusses a tentative (and by no means exhaustive) set of basic issues which must first be addressed by the system to allow the later development of specific services.

Directories

A first and essential function that will have to be provided in any global information utility is a universal directory service. Many directory designs exist today but none has ever been implemented and tested at the scale of a global network.

The OSI and CCITT standardization bodies are about to reach a fairly comprehensive and quite attractive standard (X.500) in terms of directories, which would, in principle, allow any user, anywhere in the world, to browse global hierarchically structured "yellow pages" listings of resources available anywhere else, then obtain from "white pages" all the information needed to reach and communicate with that resource. The scheme is promising in that it will be an international standard, has the full backing of CCITT, and will be the basic directory architecture for all OSI applications. This is a function for which IBM network architectures offer no equivalent. It is thus essential that we understand it and integrate it into all our systems.

That architecture supports the necessary decoupling between network providers and subscribers. The whole directory function executes itself as an application on the network, so that any commercial enterprise, not just the network supplier, could offer a

directory or piece thereof on one of its own servers. Furthermore, the architecture would allow entrepreneurs to open up new information services and advertise them in the yellow and white pages just as easily as they would today in the equivalent phone books, but with much greater timeliness.

Authentication and Access Control

Asecond fundamental prerequisite, directly related to directories, is authentication and access control support. In a global information system, the mere idea of every user having a registered userid with every server he might ever use is ludicrous and of course ruled out. In the days of time-sharing, a user's "home" system always used such an old scheme, but in a future global utility, most servers will not know about any users until these call in. In fact, authentication of users by servers will not be sufficient: every user has an equal right to be suspicious of any server sending it mail or advertisement for one of his services.

In the envisioned system, a single global sign on procedure, valid throughout the world and provided by legally bound and trusted registration systems supporting two-way authentication, is absolutely essential and should at the same time not put any visible burden on users. Such trusted registration systems should be operated by neutral, highly regulated organizations. These could be the network providers themselves, or more likely they could be outgrowths of banking and credit organizations, for reasons to become apparent soon.

Basically, at the time a client or server subscribes to the network, he should apply to the locally competent trusted registration system, and deliver his full identity, whereabouts, and credentials, to obtain what amounts to an electronic credit card, or in fact perhaps a really smart credit card embodying a secret DES or RSA key, whose matching key is safely kept by the registration server.

After initial registration, subscribers could establish any communication among themselves without any further advance notice according to one of several possible triangular authentication schemes. For instance, a client may contact a server and pass it a piece of encrypted information whose authenticity the server can unmistakably verify by obtaining a digitally signed guarantee from the trusted system with which the client is registered. Conversely, by authenticating himself to his own registration system, the contacted server can obtain a digitally signed certificate of identity that he can send back to the client who contacted him. Many schemes exist to protect certificates against recording and replaying by masqueraders.

Once authentication is completed, the client knows that he is dealing with the server he wanted, while the server can decide whether to honor the request or not, based on geographical location, organization, function, role, credit line, or whatever other attributes of the client are relevant. If service is granted, the server must also know where to send any bills for his services. This is the advantage of letting credit or

banking organizations operate registration systems: the authentication certificates issued for each service transaction, in addition to authenticating the involved parties, could include an account number and key that servers can use to issue certified electronic charge slips for expended services.

A related security issue is that of electronically notarizing documents like contracts so that neither parties can deny having respectively signed or received a document. This is necessary for instance to allow registration systems to secure the electronic charge slip transactions mentioned above.

Many solutions exist for notarization as well as for authentication, usually based on cryptography and registration systems acting as legally bound and trusted authorities. However, in this arena as in the directory arena, IBM's technology and products are lagging behind compared to many competitors.

Accounting and Billing

Next, as suggested in the foregoing section, accounting and billing issues will play an essential role in the envisioned system. In many cases, the procedure may be sketched as above, involving only one server and his client through the registration service. However in some cases, complex models of charging may be involved: all charges may go to the server (800 service), service charges may go to the server while network charges may go to the client, or vice-versa, the service may involve sub-services, with a cascade of charges landing directly or indirectly on the client's account etc. All such situations will require detailed analysis and variable amounts of cryptographic support from registration and/or billing services.

A related important accounting issue is that of copyright protection and royalty payments. In a system where just about any information that people write, compose, record, film, stage, assemble, compute, or otherwise offer for consumption, could easily find its way from some initially authorized server owning it, through intermediate servers who may have bought varying rights to use and/or resell it, to any number of client, there is a clear need to separate the rights of the various servers involved from those of the original information creator or owner. Every time a radio or TV station broadcasts a program, or a publisher prints a work, royalties are paid, sometimes wholesale, to the author or a royalty clearinghouse. The same must hold for the electronic world described here but is compounded by the fact that one author's information may be considered embedded in re-seller's service through more or less explicit hyperlinks, which will complicate the royalty payment issues and computation. These essential problems need to be understood, and research should determine how much automated support could be provided by new algorithms and techniques to help account for royalties due as the system operates.

As for directory and security services, IBM products offerings are very limited when it comes to accounting. IBM's own network service (IIN) has had to invent its own

solutions to accounting because SNA offers essentially no architectural support to such functions. Even so, the various solutions and mechanisms used in the various IIN domains and territories are not always homogeneous, and certainly would not satisfy some of the more advanced requirements put forth in this section.

Software Distribution

A fourth important issue that the system will have to deal with is software distribution, especially its standardization, compatibility, and automation. Specific application software distribution may be an interactive information service like any other. However, some software, more precisely that dealing with network and related system functions (directories, authentication, accounting, etc.) will need to be highly standardized, to evolve with guaranteed compatibility, and be distributed automatically, at least within consistent geographic areas. In small-scale and experimental distributed systems for professionals, this is no real issue. However, in large scale systems for the mass market, one cannot afford to install new software releases everywhere at once and one cannot depend on untrained users to install it. Already in relatively 'local' distributed systems such as Athena at M.I.T. and Andrew at C.M.U., software maintenance poses challenging problems. Therefore in future systems that would offer national or even global coverage, the standardization, compatibility, migration, and ease of installation challenges are bound to be of a magnitude comparable to those found in today's phone, radio, hi-fi, TV, or video systems, where the 'software' distribution issues are in fact still minor and even so not always solved.

RAS

Moving from software distribution to maintenance, the problems are compounded by issues of reliability and availability. Once an information utility is meant to be used and depended on as a 'multi-media telephone' system, it must live up to the telephone system RAS requirements, which go far beyond most of today's information systems. That issue concerns mainly the network supplier, but also the suppliers of key system services like directories, authentication, and accounting, as well as the organizations responsible for the maintenance of client workstations. Without 24-hour availability and low MTTR figures, no subscriber will want to rely on such a utility.

Communication Paradigms

Now moving down to more specific issues related to network communication, the system will have to support a variety of communication paradigms. While the old flavor of terminal-host logon session may remain necessary for some transaction applications, and the use of peer-to-peer sessions between intelligent stations may be the most widespread communication paradigm, the need will increase for new modes of interaction such as short-hold file transfer connections, single-shot remote procedure

calls for such things as short queries, reliable mailgram communication, cooperative processing between client and server through message exchange or virtually shared storage, connectionless and multi-cast communication, many-to-many conference connections etc. Distributed operating systems and access methods should progressively incorporate standardized stubs supporting such paradigms. IBM's network architectures are far behind many competitors' with respect to the variety of communication paradigms supported.

One important class of paradigms will be needed to support indirect communication, call rerouting, call forwarding, using one's office workstation from one's home station to access third party services etc.

High-speed Communication Protocols

Another communication feature that server and client stations will have to support is high-performance protocols. Anything designed and tuned for less than 64 Kbps is of no use because it would be too restrictive in the proposed scenario. In fact, for many applications involving images and animated video, several Mbps are more on target, and that is per "connection", meaning that servers must be capable of much greater rates to accommodate multiple simultaneous clients. Switches will even have to support Gbps data rates, thus completely ruling out the present hop-by-hop approach to error and flow control. WANs will have to be regarded as larger LANs, using essentially connectionless protocols among switching nodes, leaving it up to the ends to control communications - if necessary, using concepts such as rate-based flow control, non-sequential transmission, check-point mode error control, etc. For many forms of traffic, error control is not even needed, and thus should neither be forced by the network, nor implemented by the end stations.

Standards

An overwhelmingly important aspect of the system will be its uncompromising reliance on standards, standards for communication protocols, for data representation, for electronic mail, for file transfer and database access, for database structures, for user interfaces, and for hardware and software in general. Standards per se are not a technology or a field of research; however, they are not born out of nothing: IBM should strive to support standards, common subassemblies, unified choices, and homogeneous designs in every aspect of all its products to be a major player in a global information utility. It could create and promote its own standards whenever possible, but should support others where they exist and are successful.

5. User Interface

Having discussed possible applications and their support services, the next natural step is to deal with user interfaces. Together with applications, it is the user interface to applications and workstations that will make the global multi-media information utility happen ... or fail.

Standardized Direct Manipulation

In spite of the progress in user interface design over the history of computing, from JCL to today's mouse-and-menu driven models, human interface design remains an ill-understood technology. In the old days, the problem clearly had to do with the limitations of the technology. Nowadays, that can hardly be true. On the contrary, the problem seems to be that we do not know how to harness the variety and power that some interface technologies offers. As a result, the interface designer tends to overwhelm the user with too many bells and whistles: for instance, why should a VCR user be forced to program his VCR to tape what is on channel 31 between 18:30 and 21:30 on day 3 (converting from GMT to local time if he gets a foreign satellite channel), when all he knows is that he wants to tape "Gone with the Wind" regardless of which channel and time it comes on within the coming week? As suggested earlier, proper keying of shows on the cable could trigger the VCR according to the latter paradigm.

Future user interfaces should be as standardized and easy to memorize, natural to use, or even guess, as screwdrivers or bicycles, so that any user, not just a hi-tech fan can drive them. They should make extensive use of direct manipulation, be self-explanatory, with on-line help, identical prompts and commands across applications. They should be user-oriented rather than machine oriented. The user should not have to know anything about the processor, memory size, disks, I/O, files, processes, pipes, or other arcane concepts, anymore that a car driver needs to know anything about carburetors or catalytic converters today.

In some sense, application designers should not be free to design their own user interfaces any more than they should be free to define their own byte ordering, their own record format, or their own communication protocols: they must conform to established conventions and use common subsystems to access the user interface as they do to communicate across networks or access databases. Users must be able to hand pick a selection of separately designed applications and fuse them together easily into one synergistic package, where such operations as cut-and-paste across application boundaries are automatically possible where they make sense.

Several promising window-menu-icon-mouse-based user interface packages, such as IBM's PM, MIT's X-windows, Microsoft's Windows, and several others from HP, NeXt etc. have been defined in the past few years following basic concepts set forth in Xerox's original Smalltalk system. However, more experience with them will be required to determine their best features and make them converge towards common standards. This standardization process will pose a special challenge in a global system as it must

take into account national languages, cultural biases, and graphic idiosyncrasies of different people.

Data Navigation through Hyperlinks

A potentially significant and promising user interface paradigm, which has only recently started to receive wide attention, is hypertext (hypermedia). It allows a user glancing at a screen or listening to an audio piece to move the focus of attention to another related screen or audio excerpt simply by selecting words, phrases, icons, or other objects on the screen and following the logical thread it represents through one or more databases. The specific feature of hyperlinks that make them different from everything else is their "branchiness", their ability to confer to the objects they connect a dimensionality much greater than that of classical, sequential, tabular, relational, or hierarchical media and databases. While some see this dimensionality as a drawback as it tends to lose the user, the technology has many supporters. It has already appeared on the market with Apple's Hypercard. It should be understood, tried, and probably offered by IBM in a more substantial form than a simple local workstation package.

In the context of a global multi-media information utility, the hypermedia concept takes on an enhanced significance in that global hypermedia links may be created to allow users to navigate through and create new views and relations from separate, distributed databases. A professional composing a hyper-document would imbed in it direct hyperlinks to the works he means to cite, rather than painfully typing in references. "Readers" would then be able to directly select these links and see the real things instead of having to chase them through references. The set of all databases maintained on-line would thus form a hypernet of information on which the user's workstation would be a powerful window.

Voice

Another much touted aspect of user interfaces is their support of voice synthesis and understanding. While voice synthesis seems like a desirable, and indeed, for some people is an essential feature, the desirability of speech understanding is more questionable. There is a clear need to support voice input for plain digital recording of audio documents. There are even reasons to support word recognition for a limited command vocabulary or for automated typing. What is more arguable is whether it makes any sense to support full-fledged speech understanding.

Aside from the fact that full-fledged speech understanding is very MIPS intensive, and that the technology is not very advanced yet, there are serious questions about the need and suitability of such interfaces. Full speech understanding must cope with natural language expressions and such expressions tend to use concepts that are not just hard to grasp for a computer, but in fact often vague, imprecise, inaccurate, or fuzzy even by any human measure. It is questionable whether even the most "intelligent"

user interface could ever make any sense out of some of the fuzzy utterances that people sometimes make.

This does not mean that voice digitizing and word recognition (as opposed to speech understanding) are completely out: they make much sense as a means to submit simple commands within a limited "canned" language (e.g., "print", "call (nickname)", "search on (keyword)" etc.), merely as a substitute for the sort of commands we are allowed to **type** today, but not for free speech.

One may also envision the need for free-speech input without semantic recognition as a substitute for power typing but hardly as a basic paradigm for man-machine interaction. Even so, free-speech input in some cases would not even need conversion to written text. Written text has been preferred to voice in the past only because technology was not able to store and forward voicegrams. But as soon as that capability is available, many people may prefer it to written text for informal correspondence as long as they don't need any editing capability. Much of today's informal correspondence using handwritten memos or even e-mail could be replaced in the future by informal voicegrams. Even a system as primitive as Rolm's existing PhoneMail could enjoy fairly heavy use if all businesses and homes had access to it.

Similarly, some written – or in fact even spoken – information may be replaced in the future by visual information of various kinds. In many cases, written or spoken information was used in the past simply because the production, storage and transport of visual information was costly, if possible at all. With an information utility providing graphics and image editors, and plenty of bandwidth and MB to store and forward visual data, people may truly apply the view that "one picture is worth a thousand words" for certain kinds of messages. In such cases, the use of written words could be restricted to bullet-style phrases, headlines, keywords and captions meant to comment or explain the documents they accompany concisely. However, actual documents could be mostly voice/picture ones, with text added strictly where necessary and more convenient.

Handwriting

Handwritten character recognition is another user interface form that will be useful for certain applications beyond simple signature recognition for authentication purposes: for instance, for many people who never did or never will learn typing, the recognizable hand-written entry of commands as well as bulk text will be a tremendous asset. Also for mobile portable workstations, the ability for an LCD or other flat screen to double up as a hand-written text input device would save the need to lug a keyboard device around, which would be a significant volume and weight saving.

Eyeball Tracking

Another potentially useful mechanism which should be explored for enhancing the friendliness of future user interfaces is eyeball tracking devices to be used as substitute for pointing devices.

6. Workstations

Having discussed various aspects of the user interface to the system, it is appropriate to ponder next about the kind of workstation technology and workstation features that will be required to support such user interface characteristics.

Standardization

One essential concern is the design of workstations will be their standardization. Standardization has already been underlined in the context of the user interface itself, as well as in the context of data formats, protocols, and software for the system substrate. It will play a key role in workstations because they must be able to communicate with vast numbers of other workstations and servers around the world, which will necessitate that they be roughly about as standardized as telephones are today. The lack of standardization comparable to that of telephones in end stations such as TV receivers, for instance, already poses certain problems: analog broadcasts in one standard need to pass through converters to be viewed by receivers of another standard; but the exchange of digital video documents between different countries may become even more complicated unless common digital formats are adopted. Unfortunately, there is presently little reason to expect that HDTV standards will be any more unified than present TV standards.

Standardization does however not mean that all workstations would be the same, much less that their manufacturing would be the monopoly of any company. Workstations would clearly come in several versions, as telephones do today, from inexpensive versions with elementary functions to powerful versions with sophisticated options.

Types

Workstations would probably come in four different flavors: professional, office, home, and portable.

By the year 2000, it seems reasonable to assume that high-end, professional stations could include multiple processors of super-mini power, perhaps even with an array processing unit, 100 MB of memory, an 4 Mpel-16 bit-A3 color screen with graphics processor (A3 is two side-by-side A4, which is European lingo for roughly 8 1/2" by 11"), and 1 GB of – possibly optical– disk storage. In the more distant future, such

workstations could even have holographic output devices to design or simulate 3-D objects in 3-D rather than 2-D space.

Office workstations would be in a mid-range, including perhaps a single main processor of 10-20 MIPS, 10-20 Mbytes of memory, a 1 Mpel-16 bit-A4 A4 or A3 APA color display and come with a 100 Mpbs LAN or MAN adapter. As suggested earlier, when teleworking becomes common practice, the office workstation may often double up as home workstation, thus spurring the demand for home applications, which in turn will foster a market for lower cost home workstations.

To the extent that most homes today can afford telephone, hi-fi, TV, and VCR, it should be possible to initially complement then replace these devices by an integrated workstation in a comparable price range. Home workstations could initially consist of hardware not much more sophisticated than today's PCs. They could just use the TV as a display unit and be connected to existing telephone, answering machine, hi-fi and VCR. The communication adapter might be a combination of a basic-rate ISDN plug with a video cable plug. While such hardware would clearly restrict the type and sophistication of possible applications somewhat, home workstations could still be fairly successful, as suggested by the French Minitel experience, which is based on sub-$1000 stations.

The last type of workstation would be portable. It would be a watered-down version of the home machine, where every effort is made to keep it light and compact by including in the basic package only essential features, and allowing additional peripherals to be plugged into the back-panel or connected through an infra-red link. The communication adapter could be a (limited-capacity) cellular phone link or a basic-rate adapter to be plugged into ISDN-like sockets in public facilities like planes, trains, rental vehicles, hotels, exhibition halls, or public "phone" booths. The unit itself could be a flat panel the size of an A4 pad, on which the user could either directly draw or call a soft keyboard, all packaged in a way that would make it convenient for use away from the office or home.

Peripherals

Workstations would generally come with a "disk" but not necessarily in the traditional sense. First, these disks may be RAM disks to avoid the noise problems. Then, at least in professional and office workstations, they may be organized as plain caches to avoid the back-up and information management issues. Such a disk would contain all the elementary IPL and system software to join the system, as well as frequently used utilities (e.g. multi-media editor) and most recently used data files. However, it would be treated simply as a medium term cache, all other information being stored for the long term in organization or community supplied, secured, backup data servers. Based on the Andrew file server experience at CMU, the whole-file caching approach minimizes the load on servers while allowing clients to work for a while even if their file server is down. The whole disk could even be read into memory at IPL time, and then hardly

used during normal work. Writable data files could be automatically saved to the disk and to the data server over the network at regular intervals or at the end of the work session without the user even having to think about it.

An important office workstation option, especially in the early stages, is a video digitizer or scanner to absorb documents originally produced on paper. The requirement includes not only plain scanning but in some cases even visual recognition (e.g. OCR) to enable further electronic processing of information created by people not yet equipped with their own workstation.

Similarly, support will be required for I/O devices providing the special user interface function suggested earlier: voice input, handwriting recognition, eyeball tracking etc.

Personalization and Mobility

An important feature that workstations, especially office and home ones, should have is their ability to operate in two modes: starting as impersonal devices, like telephones or TVs today, then potentially switching over to a custom-tailored alter-ego mode. Clearly an office worker or "house-person" will want to tailor his workstation to personal tastes and needs. Yet the workstation should still be usable by other colleagues, family members or visitors of his office or home, as is his telephone today. The solution to this dual aspect of workstations requires that they start as impersonal devices so they can always be powered on and used by anyone, leaving personal profiles and information inaccessible until the proper password or key is supplied.

At the same time, the system as a whole should not uniquely associate any network port or workstation with any client, so that a user may request any information service, from anywhere in the world just as well as he could from his home or office. In fact, he should be able, as much as possible, to cause any workstation in the world to behave like his usual one by retrieving profile information and personal data over the network from some server or indeed from his usual station. This again speaks for functional sub-setting across workstation types and standardization across workstation manufacturers.

7. Servers

The support of very large scale, cooperating but autonomous, integrated but heterogeneous, distributed databases offering extended data types, data compression and decompression, high performance parallel search algorithms, flexible filtering mechanisms, multi-media hyperlink navigation etc. will undoubtedly pose many challenges to server machine design and programming. Unfortunately, the work group devoted too little time to identifying, let alone discussing such server design issues. They are therefore not further addressed here.

8. Assessment

So far, this report has done little more than suggest ideas, concepts, and scenarios to set the stage for subsequent necessary investigations. Along these lines, the work group produced some initial reflections on the basic ideas, trying to draw preliminary conclusions, issue a few caveats, and hint at where and how work should proceed. It is our view that the present report is only a 1 pel/sq.inch view of the list of research topics that need to be addressed. Future work must review each topic individually in great detail, study corresponding state-of-the-art, identify issues to be explored, and then put the picture back on the loom and re-weave it with a progressively finer resolution.

Opportunities and Risks

While the future may come sooner or later and never quite as suggested in our vision, many of the ideas proposed are likely to see the light of day in some form at some not-so-distant point in time. Much of the necessary technology and know-how is already here or around the corner, and pilot projects reminiscent of subsets of our system are already operational or under way, often at competitors' facilities whose ability to visualize the future is less impaired than our own by past history, a large installed customer base, and a somewhat one-sided host-biased view of the world.

Yet the future we described holds many promises and opportunities for IBM if we can marshal our research activities and focus them on some vision, this one just as an example. Precisely because of its past history and strengths, IBM could play a significant role and reap much profit in the server arena. Yet much of the action in the future will revolve around workstations and user interfaces, an arena in which we are not famous. There is some concern that we must move fast if we want to be a non-negligible player there. Manufacturers like Xerox, Apple, Sun, Apollo, NeXt and the like have an edge over us at least in experience if not in products that are more powerful, more user-friendly, generally better accepted, and often less expensive than ours, while they are having less and less difficulty interfacing to hosts. While our PC's and PS's are doing relatively well in today's office market, the line between the office and professional markets is blurring, and they will not stand the competition of future low-end professional workstations. At the high-end, our RT is no match for most competitors' offerings, though the chances of its successor appear to be good. Finally, in the home market, IBM is essentially absent, and our various attempts to make portable computers have been failures compared to Compaq's and Toshiba's.

Remembering our conjecture that the office market will spur the home one, we should focus first on professional and office segments of our scenario, later trying our hand at the low-cost home segment, seeing how much power per dollar we can pack and finding the optimal point for the home workstation.

Towards the Global Utility

The global information utility of the future will not happen overnight. There are many chicken-and-egg problems on the way. Introduction of the system must be encouraged by designing its pieces so that they can be installed progressively, one at a time, in an evolutionary way. For instance, initial advanced applications may develop around isolated embryos of high-speed networks, pilot links, and MAN's; they may provide less than global coverage but should be so attractive that they will motivate more and more communities and areas to install the necessary high-speed networks. At the same time, many other applications may be accommodated by and live with modems and phone lines as that will remain the most widely available network for a while. At some point in time, the digital upgrading of the phone system and the expansion of high-speed networks will meet, thus opening opportunities for more sophisticated services, which will in turn open the way to even more widespread broadband ISDN and MAN technology.

Similarly, early workstations should be designed to operate with existing telephones, TV's, VCR's, hi-fi etc. rather than forcing the user to buy all new devices and discard the one he already has.

It is essential that IBM come up with attractive applications suited to different levels of network performance. It should engage in more commercial experiments like Prodigy, putting together public e-mail systems, bulletin boards, databases, dictionaries, encyclopedias, and doing its own PC maintenance and software distribution over the phone hand-in-hand with some credit card company.

Needed Research

This section is a bird's eye view of some of the technical issues where better understanding is needed to support our scenario.

- Applications: first and foremost, much work is called for on possible innovative and attractive services and applications for a global information utility. All along, this report has suggested several: they all revolved around extensions to existing services, e.g. extensions to multiple media, extensions through unification and integration of interfaces, extensions through customizing and automation of otherwise known processes etc. There are bound to be many more applications that could be built upon the basic extensions of today's information systems and services.

- Data Bases and Server Technology: a second essential ingredient for the envisioned utility, and one where IBM should have an edge over many competitors, is multi-media databases supporting navigation along hyperlinks

across possibly heterogeneous partitions. Much work is needed here along the lines of scaling and standardization for globally distributed but cooperating multi-media databases, passive and active filtering techniques, parallelism for high-performance access, data compression and decompression techniques etc.

- User Interface and Client Technology: a third important research area where investigations are necessary is user behavioral models, user interface support tool kits, user interface standards, multi-media interfaces etc.

- Distributed Computing: a fourth subject where work is needed is paradigms for distributing applications across servers and between servers and workstations: how much of the work can or should be performed on the servers in the background vs. how much can or should be performed in real-time on the workstations? How could or should the work be divided among servers for availability, performance, or other reasons? What are good guidelines that could generally be applied to organize and distribute most applications? What are good models for communication among servers and between servers and clients? What are consequences for operating system structures? Etc.

- System Management: in the area of system management, more work is required on inter-network directories, authentication, access control, accounting, billing, and related control and management mechanisms for global-scale inter-networks.

- Communication: two particular aspects of communication systems that need further work are high-speed protocol design, their implementation and support in operating systems, and methods for multi-media integration both on common network links and in common electronic documents. A further important issue, though not directly a research topic, is standardization of the protocols.

Other concerns

Beyond business and technical concerns, societal concerns are also raised by the envisioned utility and will have to be explored.

Consider legal concerns first:

In most countries, the legal system has traditionally had difficulty in coping with technological innovations, and the delay in adapting is only increasing as technology progresses faster and is getting more sophisticated: for instance, while the ubiquitous copier, the tape recorder, then the cassette recorder and the VCR already posed challenges to most copyright laws, the advent of the digital audio tape and computers with powerful graphics and image copying, editing, and processing capabilities is stressing intellectual property laws to their limits. Computer software itself is a form of intellectual property that seems to have fallen in a grey zone between copyright and

patent laws. With its powerful communication, integration, retrieval and processing abilities, a global information utility is certain to raise critical new legal issues.

Next, conducting business through a global information utility will immediately raise issues about the legal value of contracts sealed through the system, and the liability of the parties as well as perhaps the network itself in case of litigation.

This in turn raises the question of which court should be competent in litigations opposing parties located in different political entities, who conducted business through a server located in yet another different political entity. And in which language and according to which Constitution should the applicable laws be written.

Then there are socio-economical concerns:

One may immediately think about the future of printed works. On one hand, there is little chance that people will completely abandon paper as a medium for novels, magazines or other material one typically reads in waiting rooms or in bed: paper remains after all more attractive than anything else for entertaining reading. On the other hand, however, outside the realm of entertaining reading, conducting any amount of research, study, analysis, thinking, correlation, annotation or synthesis on serious material is far more efficient with the help of a computer hooked into a vast information network offering efficient search and query capabilities. Thus paper publishing may lose when it comes to non-entertainment writing such as reference material etc.

As another example, one may wonder about the future role of existing institutions such as travel agents or stockbrokers who, today, are the paper equivalent of some of the information services one would probably automate first in the future. What purpose would travel agencies serve if customers could browse through video catalogs of exotic destinations and on-line databases integrating the OAG with car rental and hotel listings, all of which would be presented through travel-marketing expert systems who -sorry, which- would conclude the deal by an instant booking. There is no more room for other intermediaries between seller and buyer in such a system. The same would hold for stockbrokers if investors could directly bid for stock on-line and the system would set prices by tracking offer and demand. The "who, sorry, which" quip may however suggest a partial refutation to the above argument: while alert business users may prefer the cold efficiency of an on-line system, the overwhelming majority of people for whom traveling or buying stock is a rare experience could not be bothered to learn how to deal with the complexity of automated travel or stockbroker service and will always prefer relying on a friendly agent who can "hold them by the hand" and in whose advice they have confidence to settle their travel plans or buy stock.

Another concern deals with the scope and meaning of a "market". As long as customers must take their car and rely on impressions to shop around for good deals, the cost-quality range within which individual shop keepers can afford to play is loosely defined, of limited geographical scope, and subject to such non-quantifiable values as shop ambiance, friendliness, and convenience. Once customers can consult indexed consumers reports, get system assistance to locate the best deals in the country

(packaging and delivery included), or even try clothes on-line, there is no longer so much of a margin for individual stores to set their style.

On another tack yet, the fast dissemination of information through a global information utility may dramatically affect certain societal behavior. It is a fact that the computerization of the world's stock exchanges has taken much of the float away from the world's money supply and made the financial business much more volatile and subject to instant rumors. The same could be said of modern politics and diplomacy to some extent. A global utility is bound to extend such volatility to other aspects of life as well.

Finally there are cultural concerns:

While the technology is there to start building a global information utility, natural language translation techniques are not yet up to the task of providing everything automatically in all languages. Some languages, most notably English are certain to play a primary role initially. And just as TV progressively reduced cultural differences among American people and is now doing the same through invasion of satellite TV in Europe and around the world, there is little doubt that a global information utility will accelerate that phenomenon, which may be a source of concern to some.

Also among cultural concerns is the fact that every major invention related to information handling has had profound effects on culture in the widest sense: the printing press took away the need to memorize; television degraded reading skills because watching it passively demands less effort than reading. Thus the potential value of a global information utility should be weighed against its potential cultural drawbacks.

Finally, the advent of global information utilities will raise interesting creativity questions. In all existing forms of information publishing, from books to movies through musical compositions, some fraction of what gets published is worth reading, watching, or listening to, while most of the material is somewhere between boring, poor, abysmal and tasteless. There is little doubt that things will be the same in an electronic multi-media information world. The question is whether it might even be worse. It seems that writing good books or composing good music is within the reach of more people with smaller budgets than producing good movies or good TV shows, therefore one can rightfully wonder how many people will have the talent and the money to produce good "info-ware" for global electronic distribution.

Acknowledgements

This report would never have been written without the support of A. Peled, Research Division Vice-President for Systems and Director of Computer Science, and the initiative of J. King, then of the Almaden Research Center. In early May 1988, A. Peled accepted to sponsor an effort proposed by J. King to organize a series of regular

meetings to discuss future directions in Computer Science Research among Division staff members.

Future research directions may consist simply of a list of hard or interesting technical issues to be investigated separately. However, the participants in the meetings readily agreed that the Division should have a visionary plan of where Computer Science should aim as a whole, rather than a list of isolated problems to be worked on. Thus the group started discussions aimed at defining visionary scenarios about the future of computing.

The mere fact that the author proposed a discussion and volunteered to write a report on one such vision in no way entitles him to exclusive ownership of that vision: many people have dreamt about similar visions. In particular the vision discussed hereafter is the cross-pollinated product of the foresight and imagination of 19 more people who participated in the brainstorming effort on the subject:

Tsahi Birk, Roger Haskin, Jim Kasson, Jim King, Bruce Lindsay, Jorge Sanz, Bob Selinger and Moshe Vardi of the Almaden Research Center, Tom Kamimura of the Tokyo Research Lab, and Sanjay Addanki, Steve Boies, Larry Koved, Mark Laff, Dan Ling, Kevin McAuliffe, Dan Prener, Basil Smith, and Mark Wegman of the T.J.Watson Research Center.

They should be thanked for their contributions, while the author shall bear the blame for any shortcoming of this report.

The author also thanks Liba Svobodova for her insightful comments, thought-provoking suggestions, and stylistic remarks on successive drafts of the report.