Chapter 5: Internet and Intranet network technology

Computing is not about computers any more. It is about living. The digital planet will look and feel like the head of a pin. – Nicholas Negroponte'

In the spring of 1994, the Internet was cut free of American government support. What followed was continuous, almost explosive innovation and implementation of Internet and network technology. Internet and networks are now largely synonymous because internetwork connectivity can be accomplished with what seems to be increasingly easier tools. The issue becomes cost or, more often, cost control. The more an organisation spends on hardware, software and telecommunications, the more payoff management expects.

Less than one year ago, technology was the problem confronting an organisation that wanted to make the most of networking within and across departmental units. Products such as Lotus Notes provided a partial solution, but the product was costly, sometimes approaching thousands of dollars per user for certain commercial applications. The idea of seamless connections among distributed databases was often a powerful one. Employees could retrieve information from a common pool of data. The information could be integrated into desktop applications and be changed easily.

Intranets emerge

The public side of the Internet medium has dominated the popular and trade press. Another, possibly equally significant development, is taking place at the edge of the spotlight. Organisations are exploiting a number of advances in network technology to use the Internet in general and World Wide Web technology in particular, to create *Intranets*, an Internet/World Wide Web service for use within an organisation.[2]

An Internet server provides access to electronic mail, interactive services and video, sound and text such as that on the public Internet. The difference is not the look

---

[1] 'Being Digital'. New York: Alfred A. Knopf, 1995, p. 6.

[2] The intranet market is likely to become a fast-growing revenue stream for many information technology providers by mid-1996.

and feel of the information. It is that Internet technology is used to service users within an organisation or a non-public user group. Because software can be used to pass messages between different Intranet servers operated by a single organisation, the traffic is not confined to users at a single location. The Intranet provides a virtual value-added network at a fraction of the cost charged by commercial vendors such as General Electric Information Services, using software to create private tunnels through Internet traffic. [1]

According to Netscape, more than half of the firm's sales of its server software is for Intranet applications. The reasons include:

- Browser software has become the standard interface for accessing text and image, audio and video objects. Users are familiar with the interface or can learn it rapidly.

- Web technology and page building tools are readily available, stable and at low price.

- Internet technology provides a common platform for the viewing and sharing of information. Users on PCs, Macintoshes and UNIX workstations can view the same information and access the same image, audio and video objects.

- Software is available to permit automated Web page construction from material residing in such corporate databases as Oracle, Sybase and Informix. Oracle has developed a comprehensive line of Internet publishing tools that simplify the creation of customised Intranet services such as order entry and customer look up. The firm's new *Power Browser* software was designed with robust corporate applications in mind.

- Rapidly falling prices for higher capacity network hardware provide sufficient bandwidth for more advanced applications such as real-time video conferencing and shared applications or whiteboard functions.

Is the Intranet a Notes killer? The sharp cost difference between Notes and Intranet-compatible groupware such as Collabra, a unit of Netscape, is significant. Collabra is one of several software companies putting significant price pressure upon the Notes product. IBM spent millions to acquire a product that a savvy computer centre manager with Internet expertise can duplicate using lower cost Internet technology and now must reduce the price of Notes to remain competitive with the Internet groupware.

---

[1] General Electric Information Services, a unit of the General Electric Company, has forged a relationship with Netscape to link the GEIS VAN with the public Internet via technology that will be embedded in Netscape's next-generation browsers.

The outlook for Intranet applications is bright, possibly more extensible than growth of the public Internet. It is difficult to demonstrate the payoff from a public Web presence. It is much easier to illustrate cost savings from an Intranet effort. Typical Intranet applications include:

- *Human resource information.* Applicants' curriculum vitae, benefits information, application forms for educational reimbursement and dynamic information about pension and profit sharing accounts are among applications in this functional area. Human Resources software leader PeopleSoft provides tools that permit access to HR data from a Web browser.

- *Research.* For major chemical and pharmaceutical companies such as Pfizer, scientists at laboratories in different locations can retrieve, annotate and interact with data and colleagues elsewhere. The graphic capability of Web technology provides a low cost way to provide visual information to clarify numeric or chemical modelling data.

- *Internal communications.* Scott McNeely, founder of Sun Microsystems, uses an Intranet to communicate with employees. Each week he records an audio programme that employees can listen to in real time at their convenience. When a new programme is posted, employees see an icon on their workstations. A click jumps them to McNealy's home page and the programme plays. Other companies are using Web pages instead of traditional printed newsletters.

- *Sales and marketing,* Vantive Corporation uses an Intranet site to keep staff informed of prices, products and the latest marketing materials. Other firms use an Intranet for sales force co-ordination. A click allows a salesperson to send an order or electronic mail message to a colleague. Says one convert at American Multisystems in Milpitas, California: "Our Intranet pays for itself with faster turnaround on orders."

- *Technical support.* Many companies, including Microsoft, are using Intranet applications to provide employees with technical support.

- *Employee Web pages.* Staff use the Web to provide colleagues with information about projects, vendors and vacation schedules.

The benefit of Intranets is the promise of cost savings and improved operational efficiency. Long documents such as employee benefit plans can be made available on the Intranet saving the cost of printing and distributing thousands of documents. Engineering change orders, customer credit data and other important organisational information can be instantly available to staff anywhere in the world.

Key enabling technologies

High speed networking is becoming easier and cheaper. Several innovations are bringing the type of price performance gains associated with the personal computer market to Internet and intranet networks throughout the world.

ATM

Asynchronous Transfer Mode (ATM) is emerging as a major high speed networking technology. It is the telecom equivalent of a world-class gymnast who packs proven flexibility and intelligence into a small, compact package. ATM combines a robust header and a tight data packet that outperforms older technologies. In America, prices for ATM have fallen so that it is competitive with Frame Relay. By the end of 1996, ATM will be available throughout America. ATM is driven by the Regional Bell Operating companies. Ameritech, Bell Atlantic and SBC Communications now offer low-cost ATM services. The other Baby Bells are likely to follow suit. These developments have put pressure on backbone providers such as AT&T, MCI and WillTel to introduce expanded ATM services at competitive prices.

One incentive to move to an ATM and frame relay architecture is that network managers can move a central site to ATM and then migrate remote sites from frame relay links when demand **warrants.** A mid-term strategy for American telecommunication companies is the installation of ATM as carriers for their Frame Relay networks in order to maximise the ways to deliver ATM services. ATM provides an intermediate step between T1 and a 155 megabit per second Synchronous Optical Network (OC-3) link. ATM seems poised to grow more rapidly than T3 and OC-3 links. Revenue from the highest speed networks yields significant profits for carriers. As demand for high bandwidth Internet applications such as videoconferencing grows, demand for capacity will continue to grow rapidly. ATM is also attractive because it consolidates voice, video and data. Frame relay shines when only data are transported. However, ATM has an order of magnitude greater overhead than frame relay. But users want multimedia online services at a lower price despite **some** performance trade-offs.

Network managers will have more options at highly competitive price points. As user demand increases, the network infrastructure exists to provide capacity.' A new Microsoft Applications Programming Interface (API) gives developers a way to make use of this high speed switching technology. Programmers will be able to link existing software applications to ATM New software will be ATM -aware and be able to integrate easily with advanced network applications. Among the new desktop applications that the API is likely to drive are:

- Dynamic allocation of network resources to provide consistent, fast access to applications that require high bandwidth.
- Integrated videoconferencing and shared workspace applications; users would be able to see. talk and work on common documents.
- Deployment and access to distributed software and database applications without unacceptable delays or long response times.

---

[1] Information about ATM is available at http://www.nwfusion.com.

High speed links

High speed communication links exist in all developed countries. These optical networks provide dramatic capacity and a lower cost per bit than the older fibre networks they are replacing. Telecommunication companies are deploying high capacity channels because of the strong market demand for faster data transfer and bandwidth-hungry applications such as streaming video. The alphabet soup of communications options grows more complex each day. For specialists in the field, keeping pace with OC-3, ATM, 2D+B and other aracana is a big job. Suffice it to say that standard office networks are now able to run an order of magnitude faster over standard wire than was the case only twelve months ago. Backbone communications can support live video. The bottleneck for broadcast quality video at the desktop remains moving the data from the high speed backbone to the desktop. CompuServe has announced plans to spend more than $10 million on high speed ATM switches from StratCom, a carrier switch vendor. CompuServe's announcement signals a shift to ATM technology. ATM in addition to providing sufficient bandwidth to deliver video, offers a way to slash wide area networking costs by aggregating traffic from many links into a handful of high capacity trunk lines. The combination of ATM and new monitoring software from companies such as 3Com and Cisco Systems provides a way to manage traffic more efficiently, support Internet and frame relay access, and scale the network rapidly. Telecommunications giants such as AT&T, Nippon Telegraph & Telephone and others have used these technologies for more than a year, but the diffusion of high capacity lines and advanced technology to other types of companies marks a dramatic broadening of the high speed communications technology that is likely to continue for the foreseeable future.

Faster-switching hardware

As noted above, companies such as Stratcom, Cascade Communications and Cisco Systems have become drivers of high speed network technology, particularly in regard to ATM High speed routers, devices which create bottlenecks when networks become large, are being replaced by switching technology. Surging demand for higher speed access and better control of traffic consolidation can be met by installing advanced switches. Internet providers UUNET Technologies, Netcom On-Line Services and Performance Systems International Network (PSINet) have launched programmes to install switches that will support multiple T3 lines. [1] Cascade switches, for example, can support 1.55 megabit per second links with upgrades to 622 megabits per second by 1997. MCI Communications and other major telecommunications companies are also moving to more advanced switching technology. To support the American supercomputer network, MCI has adopted ATM technology.

---

[1] PSI has become one of the first American vendors to offer dial-up ISDN connectivity across the United States.

As recently as 1994, advanced switching technology was deployed by the major telecommunication companies and in specialised telecommunication applications for government and industry. Advanced switching technology is being deployed in an increasingly wider range of network environments. UUNET, a company in which Microsoft has a 20% stake, plans to offer high speed local access in 300 cities worldwide by 1997. The higher speed switched network architecture poses a significant challenge in countries where the local telecommunication companies do not provide this type of service at competitive rates.

One-stop solutions

What about companies or organisations without the in-house expertise to launch into high speed networking? The solution to advanced networking problems has been reduced to a decision about which integrator to call. AT&T, bolstered by its stake in Bolt, Beranek and Newman, provides comprehensive high speed networking services. In addition to working out the details of connecting local area networks to high speed links, AT&T/BBN offers a comprehensive monitoring service. For firms concerned about security, AT&T/BBN will monitor traffic to and from designated sites. The service provides additional measures of security with such features as automatic site shut down and tracking of the intrusion back to its source. AT&T's Internet strategy is to provide similar services throughout the world, either directly or in partnership with other firms. An example of AT&T's capabilities is the installation of a multimedia servicer (Multimedia Communications Exchange Server) that links a company's voice and data networks and supports videoconferencing. AT&T can configure the hardware and network to support Intranet or Internet applications.

Consider the proprietary IBM SNA (System Network Architecture). For years, this network environment followed the drum beats of IBM. Unique software ran on proprietary hardware to give a customer speed, reliability, security and the confidence that it was an IBM product. The reality was that SNA was sufficiently difficult to link to the Internet that other IBM divisions developed Internet connectivity solutions for the mid-size AS 400 and RS/6000 range of machines. The lustre of SNA faded and many customers moved to less costly and somewhat less technically complex network architectures.

Today, the 'new' IBM will not only provide SNA- to- Internet connectivity solutions, IBM will provide cradle-to-grave Internet services regardless of the customer's network architecture. The meshing of the proprietary SNA protocols with the TCP/IP of the Internet remains technically on a par with winning a stage of the Tour de France. The problem is no longer the customer's. IBM can, and will, make the networks 'talk'.

A similar change has taken place at the leader in organisational networking. Novell has watched the market change in the last two years. Novell's proprietary, often unstable and almost inexplicably popular NOS or Network Operating System has enjoyed as much as two-thirds of the market for networks. Small companies with five employees to behemoths with 1,000 or more users in one department rely on Novell. Certified Network Engineers and Certified Novell Resellers keep the faith

burning. A network engineer with 'CNE' after his name is almost certain of employment anywhere in the world.

Linking Novell networks to the Internet was not impossible. Third-party products from Cisco Systems and others eased the burden. But configuring servers and workstations to load the additional software necessary to make the link was tedious work. The result was often unstable and prone to crashes. Novell, however, has not been blind to the Internet's evolution into a new medium. Consequently the firm provides a wealth of free technical information about linking Novell networks to the Internet. [1] Novell now provides software to facilitate loading the TCP/IP protocol stack on servers connected to the Internet and offers different options for linking workstations to the Internet.

Problems of protocol conversion, security and adequate bandwidth can now be solved more easily and more economically than ever before. The benefits of high speed networking, however, are available to those countries, organisations and individuals with the resources to purchase them. Use of the Internet as a medium is confined to those locations, organisations and individuals with resources. More than half of the traffic on the Internet originates in North America. The centres of greatest usage correspond to areas where infrastructure and necessary resources are available.

For countries without these resources, the benefits of high speed networking are beyond reach. Consider the Ivory Coast. The country is without a high speed Internet connection. The ambitious plans of AT&Tto lay a ring of high speed fibre around Africa will not be completed until 1999 or later. In Africa, twelve countries are linked to the Internet; 32 are not. Most nations, not just those in Africa, still have state controlled telecommunication systems. These systems lack the funds to build a high speed infrastructure. In some instances, the authorities fear that access to the Internet will jeopardise the comfortable tradition of centralised information control.

Advocates of advanced wireless technologies believe that areas such as Africa and part of the Asian continent can deploy newer technologies. Much of North America and large portions of Western Europe have copper wire networks. These can be costly and difficult to upgrade. Nations such as the Ivory Coast can build a state-of-the-art system without having to upgrade an outmoded system (as is happening in Russia and parts of Eastern Europe, for example). If resources can be marshalled to build advanced networking and telecommunication infrastructures, the benefits to these nations can be significant. Barriers are likely to remain for the foreseeable future. These include political resistance, the global squeeze on available cash and a lack of expertise. Paradoxically, as high speed internetworking becomes cheaper and easier to deploy, many parts of the world are destined to remain outside the new medium's environment. Without access to the global

---

[1] Novell's Internet site is heavily used. Technical information, product information, documentation and software updates may be found at http://www.novell.com

network, the opportunities to reverse lacklustre economic and educational performance remain bleak.

Web server technology

Developments are entering the market at a staggering rate. UNIX remains the Internet server operating system of choice.[1] Two products warrant brief comment. First, is O'Reilly & Associates Web Server for Windows NT and Web Server for Windows 95. The price on the carton is $500. However, mail order prices for this software are less than $300. This price point will be under severe market pressure as other firms exploit the large and growing market for Windows-equipped PCs, workstations and servers.

The second product is Windows NT itself. The publicity surrounding Windows 95 has largely obscured the significant strides made in Microsoft's back office software products. Beginning with Release 3.51, Windows NT has been gaining momentum for certain types of server applications. Microsoft has found the established network server market ripe for attack for three reasons. Although robust, support of UNIX-based systems is expensive. Staff require specialised skills, programming expertise and time to fine-tune the UNIX implementation at a particular site. Another issue is the cost of UNIX software. The large number of UNIX installations, (estimates run as high as 75% of academic, research and large commercial enterprises having one or more UNIX networks), means many different versions of UNIX which translates into small market slices for many platform-dependent applications. Even basic software tools for word processing cost one or more orders of magnitude more than a comparable PC application. Finally, the technical headaches associated with managing a fast-growing network application are difficult to relieve in today's business climate. Cost control, rapid response and simplified maintenance are gaining importance.

Windows NT, although arguably less powerful for certain applications than a comparable machine equipped with a current release of a UNIX variant from Sun, Hewlett Packhard or IBM, has several appealing advantages:

- Interaction with the operating system, basic system functions and application software is through the familiar point-and-click interface. Setting up a Sun Microsystems Sparc 20 as an Internet server can take as little as four hours. Set-up time for a comparable RISC machine running NT is about one hour.

- NT runs on a range of PC and RISC hardware. A site can begin with the NT server software available in special promotions for nothing for sites with 10 or more users and a spare PC. Entry-level cost is essentially zero. Prices for a full network installation of the software for Internet access are

[1] Sun Microsystems is the most popular hardware platform for Internet services. It has a 40% market share.

less than $1,000, thousands less than comparable licence fees for many competitive products.

- Set up of hardware is simplified with the integration of plug-and-play technology. When a device is added, the system recognises the new device and handles many configuration steps automatically.

- Performance on such platforms as the Digital Equipment Alpha is comparable to performance of optimised UNIX running on high-end machines from other vendors. As the performance of the hardware increases, costs fall. In this environment, NT offers a Web site a way to increase response time at lower overall costs.

- The 70 million users of Windows provide a large installed base of customers for applications software. With Windows 95, Microsoft has introduced an intermediate step between 'old' Windows and NT. Windows 95 compliant software should run on NT-equipped machines. In practical terms, the cost of robust multi-tasking, 32-bit software will remain low. For all practical purposes, the price threshold for server software is now in the sub-$500 category and falling.

The choices in Web server software are exploding. Netscape offers its software and security tools for $5,000. Open Market (Cambridge, Massachusetts) has matched the Netscape price point. Similar products with a wide range of additional features and enhancements are available from such firms as America Online's NaviSoft unit. CompuServe provides a nearly identical suite of server software through its Spry Communications subsidiary.

Web servers are, for all practical purposes, available like pizza (ready to eat) or in low cost do-it-yourself kits (easy to make). UNIX is likely to remain a strong force in the market. Windows NT is a technology to reckon with over the next 12 months.

Database connections

Rapid advances are being made made in linking industrial-strength relational databases to Internet and Intranet applications. Without tools designed to interact with client-server databases, the browser technology puts a pretty face on static or difficult to manage data. With database connectivity, the browser can access an application that can be refreshed and upgraded.

The most dramatic developments have come from the big three in the relational database world: Oracle, Sybase and Informix.[1] By the end of 1995, Oracle had allied itself with Spyglass Technologies, the firm that is the authorised distributor and licensing agent for the Mosaic browser technology. Oracle introduced a complete

---

[I] Informix acquired the database technology of Illustra (Oakland, California) in December 1995 to strengthen its Internet technology.

suite of Internet publishing tools plus tools to permit online interaction with data stored in Oracle tables. Oracle's Power Browser will support its own scripting language and can be integrated with Oracle's proprietary Power Objects technology. The toolset supports a local Structured Query Language database and server-based tables.

The Oracle tools allow Web pages to be stored in an Oracle database. Alternatively, the tools can generate HTML tagged output in real time. The Oracle Web tools will be integrated with text retrieval tools built upon the Open Text (Waterloo, Ontario) technology. Any client capable of communicating using hypertext transfer protocol can use Oracle Web resources. Oracle has designed an extensible and open architecture. Users gain access to multimedia Web applications built upon Network Loadable Objects. The NLOs will support applications such as Adobe PDF viewers, Java applets and VRML scripts.

Other database developments include:

- IBM's introduction of a Web interface to DB2 databases.
- Computer Associates' support for Web access its Ingres product.
- Sybase's Power Builder tools that allow the development of Web-enabled databases.

One of the more interesting network technology thrusts has been the partnership between Netscape and Informix. These two firms have joined with Spider Technologies to create front-ends to Informix, Sybase and Oracle databases.

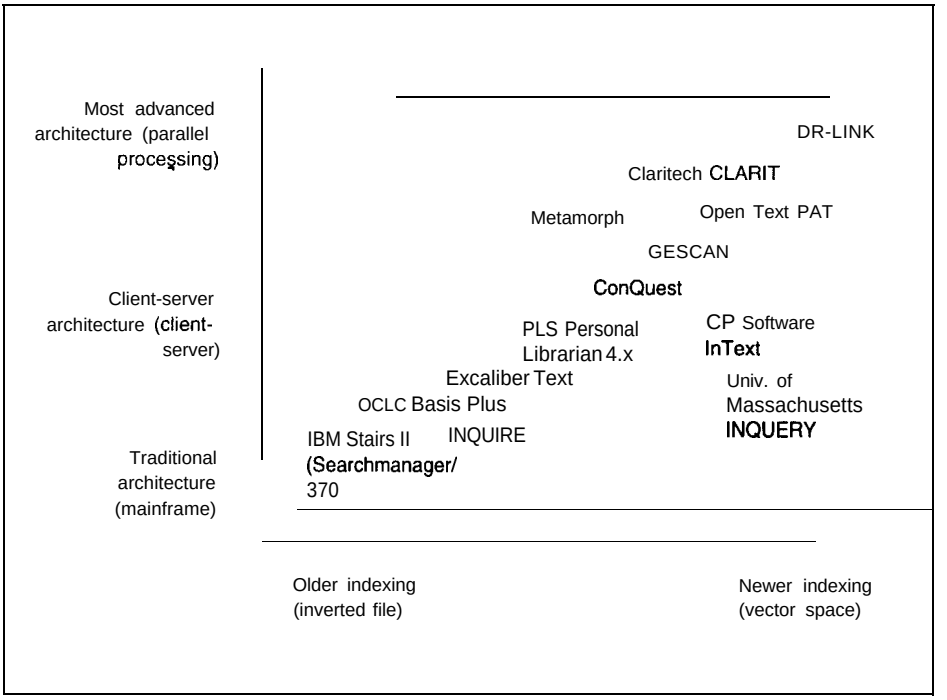
Spider is a programming tool that allows users to create applications built upon a client-server database. Spider is taking an open architecture approach. One challenge facing Spider and other tools that link client-server databases to the Internet is that the Internet is inherently asynchronous. A robust client-server database application is inherently synchronous. If Spider can resolve the problems associated with Internet access to legacy data, its integrated browsers will provide a platform for a new range of Internet and Intranet applications. The $500 Web Whacker from The Forefront Group indicates that powerful, low-cost tools will be available for a range of client-server platforms.'

Finding information

Looking up information in an SQL database is becoming easier, but what about full-text queries? Many high-value Internet sites contain large amounts of textual data. The legacy data of organisations with Intranets is largely text. Much of the text is not structured with HTML or SGML tags, yet users want to examine archives of news groups, examine the full text of technical papers and gain access to the growing body of full-text commercial information stored at Web sites.

---

[1] The Spider tools are available at http://www.w3spider.com. Price ranges from $5,000 for a single user, to $60,000 for an unlimited number of users.



The y axis represents the processing environment and the x axis the hardware environment for the text retrieval packages. The most advanced text retrieval software exploits client-server architecture and parallel processing environments.

Access to full text has been becoming easier. Only within the last 12 months has full text software exploded as a business segment. More than 40 companies are shipping full text retrieval software. Many of the most interesting have been in existence less than one year. The table above arrays the major providers of text retrieval software.

The rush to full-text retrieval began in 1993 with the introduction of WAIS (Wide Area Information Service technology). WAIS technology uses a book metaphor for organising its information. Encyclopaedia Britannica selected the WAIS text retrieval tools for its Internet version of its information. WAIS, Inc is now an operating unit of America Online.

Available toolkits

Another leader in full-text retrieval example is Fulcrum Technology (Ottawa, Canada). FullText, the text-retrieval engine, has been licensed by Microsoft for use in its Network. FullText supports natural language queries. A user enters a word or series of words in standard English. No special commands or Boolean logic are required. The Fulcrum technology is robust and supports relevance ranking, a process that shows the documents that most closely match the query at the head of the list. The traditional online systems typically display documents in reverse

chronological order, not by their relevance to the user's query. The Fulcrum full-text software is a tool kit. A licensee of Fulcrum uses the tools to build a full-text retrieval application. The cost of the full suite of software can easily reach $20,000 per year or more depending upon the application.

The strong demand for full-text software has triggered rapid innovation and price pressure in this Internet software segment. Open Text has introduced a version of its natural language, full-text software for Internet applications. Oracle Software has licensed the OpenText technology for use in its database software. Personal Library Software (Rockville, Maryland) has followed the same path and added database look-up technology in order to give its licensees access to large domains of unstructured full text and SQL databases. Even the ageing Basis Plus software has been tweaked to support Internet queries, SGML and HTML and quasi-natural language functionality.

Prices for these software suites are falling. Vendors are reluctant to quote prices without knowing the licensee's application. Text retrieval software, which once cost upwards of $20,000, can now be obtained as shareware, a component of an SQL database engine, or on a stand-alone basis for as little as $2,000.

Commercial database operations such as Knight Ridder Information Services, STN and Lexis-Nexis have been slow to introduce natural language tools. Only West Publishing's WIN (West Is Natural) product has had significant impact on the high-end online market. Among the second-tier online providers, DataTimes (Oklahoma City, Oklahoma) has licensed the Personal Library Software technology and taken a stake in Intell.X, a Virginia-based software firm with text summarising technology. Innovation in full-text online access has gone to relative newcomers such as InfoSeek, a company discussed elsewhere in this Briefing.

Architext has enjoyed a rapid rise since its introduction in late 1995. Engineering Information has licensed the Architext software. The firm uses its tools to create an index of the World Wide Web. The abstracting tools used to build a brief description of a Web site's content are more robust than those used by other Web indexing companies such as Lycos and Yahoo.

The Architext software supports natural language queries and interactive browsing. The query-by-example feature allows users to instruct the software to "find more documents like this one" with a mouse click. The results of a query can be display in a group by topic and subject. A search of the Architext Web index can be displayed by topic, relevance or site. Automatic abstracting routines are included in the Architext toolkit. Hypertext links are automatically recognised and converted to HTML format. One of the most interesting features of the software is its ability to generate indexes that comprise only 15% of the source file's size. Most indexing schemes require indexes equal to or greater than the source file's size. Filtering and routing tools are included in the Architext toolkit.

The real innovation, however, is going explode on the scene in the next 12 months. Funding by the American government has given a number of universities and defence contractors an opportunity to develop radical new technology that will

change full-text retrieval more significantly than the products of the companies now shipping product.

Next-generation technology

The next-generation technology – flowing from Carnegie-Mellon University, Syracuse University and the University of Massachusetts – are built upon the premise that a person seeking information does not necessarily know what information is needed to answer a specific question. The advanced text retrieval engines have been built to combine several different processes so that the user can ask a question and obtain only the most relevant information; specifically:

- *Linguistic analysis.* The new software examines the source documents in terms of the information contained in the language, syntax and structure of the document itself. The index of pointers to the content of the document and the domain of documents reflects this sensitivity to language.

- *Statistical analysis.* The technology examines the frequency with which words and phrases occur in the document and in the context of the domain of documents of which the single instance is a part.

- *Vector space indices.* The data generated by these algorithms are analysed and mapped into an imaginary hypercube where each concept or information element is given a weight and a direction. The resulting index structure has a number of remarkable properties that can be exploited when a user wants to extract information from the domain of information.

Before information can be extracted, the system must examine the user's query or statement of information need. The same technology that processed the information stored in the index is used to analyse what the user needs. The user can provide information in virtually any form appropriate. For instance, a skilled online searcher might formulate a Boolean query. A business person looking for marketing information might enter several sentences or fragments saying what was needed.

These advanced systems then match the query with the data in vector space and display the information according to the requirements of the user; for example, bibliographic entry only, bibliographic entry plus machine-generated extract of the most important information in the document or documents, or the full text of the documents.

Another such system is DR-LINK, developed by TextWise (Syracuse, New York), DR-LINKrepresents one benchmark in intelligent text processing and retrieval. The software includes automatic indexing classification, term expansion, proper noun extraction, and real-time updating of index data. The system has been in development for over five years at a cost of millions of dollars in both government and private funding. The software entered beta test in 199.5 and will be commercialised in 1996.

Like the other advanced retrieval software, a single person has been the driving force behind the technology. Elizabeth Liddy (MLS and PhD), a former research

librarian and winner of the ASIS 1992 Doctoral Dissertation of the Year Award, designed the system.

DR-LINK combines multiple software modules to analyse, categorise and essentially 'understand' full text. The algorithms permit DR-LINK to begin to compete with human indexers and researchers. The automation allows more robust document analysis without the large staff of professional indexers needed to cope with the exploding volume of textual information.

A user may ask the system a complex question in natural language; for example: "What pending or proposed legislation might adversely affect banks in France? What are banking officials saying about it?" The DR-LINK technology has the sophistication to find all documents responsive to this type of query.

Matching documents do not necessarily have matching words in them. The user can use relevance feedback – marking a portion of text or a complete document to say, "Find me more documents like these" – as a way of refining the query.

DR-LINK like most of the advanced text processing systems, runs on RISC processors. The RISC platforms support parallel processing: that is, many functions are carried out simultaneously. Parallel processing is almost mandatory for advanced text processing for several reasons. First, the computational requirements of the linguistic and statistic algorithms are needed to speed throughput. Second, the mapping of vector space is a computational-intensive process that must be carried out continuously in order to ensure that the index is up-to-date. Finally, the processing of the user's queries must occur in real time against the vector space indices.

With more than 40 companies marketing advanced text processing tools in North America alone, prices will be forced downward as these firms seek to expand their market share. The first-generation services such as West's WIN have limits imposed by their architectures. Second-generation services such as the widely-acclaimed Individual Inc and Desktop Data real-time filtering services are often less powerful and flexible than the marketers' brochures advertise. The third-generation systems now on the market from Open Text, ConQuest Software (now a unit of the document imaging specialists Excaliber Technologies, San Diego, California) and other companies, run on existing computing platforms.

The fourth-generation systems such as DR-LINK and Claritech (Carnegie-Mellon University) exploit the newest computing platforms. As the computational power of the new RISC architectures falls in price, the application of these advanced engines to network publishing will increase rapidly.

The death of intermediated online services will occur. But only individual companies will die. The business of providing value-added, online retrieval of information will grow more rapidly than many analysts believe.

Automating information retrieval and processing

Automation of online information retrieval will drive network publishing. In the business climate that has emerged since the end of the Cold War, pockets of growth and depression will remain a permanent part of the economic landscape. Furthermore, competition will require that organisations will have to reduce costs in order to remain competitive. Even firms selling products to a narrow, price-insensitive niche will find themselves under relentless attack from competitors who can apply technology more cleverly.

Automation of many network publishing functions is underway. The most common examples are at hand, but they are usually seen as low-value utilities or more often overlooked completely. One example is the automation built into FrameMaker. All versions of FrameMaker can be extended with C programs to permit automatic generation of documents from databases or spreadsheets. Text is captured, tagged and poured into frames. The person setting up the system determines whether the output is in printed form or in World Wide Web pages. The barrier to wider use of this technology remains complexity and the need for programming.

New tools are beginning to flow into the market. IBM, now owner of Lotus Notes, will include scripting tools in the next release of Notes to automate monitoring, routing and output of information in Notes databases. Extensions written by users of Notes are available without charge from many user groups. Commercial extensions of Notes that add additional automation functionality are readily available. Cuadra Associates has added a web front-end to its powerful STAR information management software. STARweb makes a range of advanced search and retrieval functions available to any user of web browsing software.

A more important software development has been the blending of filtering tools with routing software. The software monitors a stream of information. The stream may be commercial news feeds or internal documents in word processing format. When documents match a particular set of criteria or specific search criteria, they are routed to the person or department requesting the information.

One company has extended filtering and routing technology by developing some powerful post-processing tools. InText Systems (San Francisco, California) is a unit of the Australian CP Software Group (Canberra). The InText software allows a site operator to monitor information; for example, Internet news groups. When information matches a query, the data are written to the system.

InText software then creates an extract of the original document. The length of the extract can be set by the system administrator or the person formulating the query. InText software then routes a message to the person posting the query informing him that a match has been located. A token of the document, not the actual document itself, and the extract are sent by electronic mail to the user. If the user wants to examine the full-text document, the original file is opened and may be copied to the user's workstation.

The complete suite of InText filtering, extracting and routing tools is priced depending upon the user's application. A version of the extraction program is available for about $100 from a licensee of InText technology.

Summarising, abstracting and extracting technology is advancing at a rapid pace. DataTimes, producer of full-text databases of business news, has developed the Intell.X summariser. The software automatically reduces a newspaper article or an e-mail message to a quick read summary of the headline and its main points. Unlike the InText approach, Intell.X displays the main points as bulleted items. The user can then display an enriched 20-30 word summary of the article containing themes and topics. The user may examine the full text at any point. Similar routing functions are available without charge from Pegasus Software, also located in Australia.

Outlook

Innovation will continue unabated for the the next three to five years. The Internet has become a full-scale applications development platform. New needs are discovered at the same time that users are demanding greater bandwidth to gain access to traditional data and video. The foundation for continued innovation is in place. The Web technology provides a stimulus in diverse network technology disciplines. The reason is that everyone can get into an interactive, positive feedback loop. When an application is built for the network, the distribution channel blends with the development environment. Innovation can diffuse rapidly, thus triggering another round of innovation. With the network becoming the next-generation application platform, applications are platform independent. More people can then make use of the application environment. As usage expands, new demands are placed upon network technology, reinforcing the drive for innovative solutions

The technology environment will evolve rapidly. Network technology will make more widely available at a lower cost such developments as:

- Provide sufficient bandwidth for simultaneous audio and video real-time applications to be used from properly equipped desktops.
- Groupware will become a standard component of a software application.
- Bandwidth will continue to expand, although the deployment of the most advanced architectures will be limited to those nations, organisations and individuals who have the resources to invest in the technology.
- Local Internet service providers in North America and Western Europe will be squeezed between users who demand more bandwidth and the large telecommunication companies who will provide 'Internet dialtones'.
- Local area network technology will be in flux with battles raging among industry leader Novell, newcomer Microsoft, and established providers such as IBM.

In short, network technology will demonstrate the type of change and capability expansion that characterised PC software in the late 1980s. Technology is no longer a barrier. The principal obstacles that must be overcome have to do with making a decision about what supplier's technology and software to acquire. As a result, the emphasis is shifting rapidly from bits and bytes and packets to financial and marketing issues.

The proliferation of powerful, lowest cost technologies will permit small businesses and other organisations to 'Internetise' their businesses. The lustre will fade from the Internet as it is described in the popular media. The gap will be filled with intranet applications that will have profound business consequences worldwide.