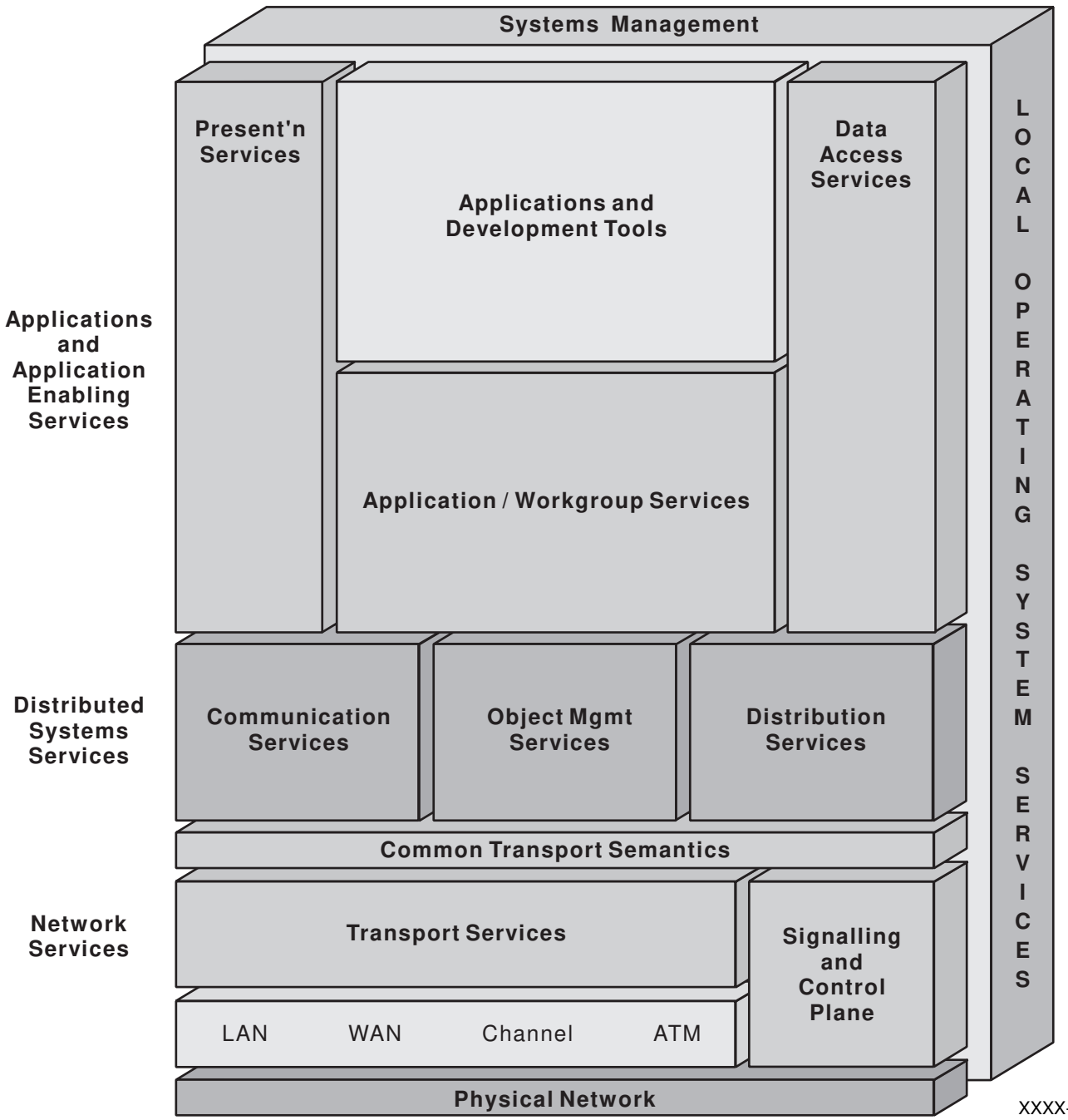
Open Blueprint Support for Mobile Computing



XXXX-XXXX-00

Open Blueprint Support for Mobile Computing

XXXX-XXXX-00

About This Paper

Open, distributed computing of all forms, including client/server and network computing, is the model that is driving the rapid evolution of information technology today. The Open Blueprint structure is IBM's industry-leading architectural framework for distributed computing in a multivendor, heterogeneous environment. This paper describes the Open Blueprint support for mobile computing.

The Open Blueprint structure continues to accommodate advances in technology and incorporate emerging standards and protocols as information technology needs and capabilities evolve. For example, the structure now incorporates digital library, object-oriented and mobile technologies, and support for internet-enabled applications. Thus, this document is a snapshot at a particular point in time. The Open Blueprint structure will continue to evolve as new technologies emerge.

This paper is a companion to the *Open Blueprint Technical Reference Library* collection, SBOF-8702 (hardcopy) or SK2T-2478 (CD-ROM). The intent of this technical library to provide detailed information about each Open Blueprint component. The authors of these papers are the developers and designers directly responsible for the components, so you might observe differences in style, scope, and format between this paper and others. For a general technical overview of the Open Blueprint, see the *Open Blueprint Technical Overview*, GC23-3808.

Who Should Read This Paper

This paper is intended for audiences requiring technical detail about Open Blueprint support of mobile computing. These include:

- Customers who are planning technology or architecture investments
- Software vendors who are developing products to interoperate with other products that support the Open Blueprint
- Consultants and service providers who offer integration services to customers

© **Copyright International Business Machines Corporation 1996. All rights reserved.**
Note to U.S. Government Users — Documentation related to restricted rights — Use, duplication or disclosure is subject to restrictions set forth in GSA ADP Schedule Contract with IBM Corp.

Contents

Preface 1

Overview of Mobile Computing 3
Business Justification for Mobile Computing 3
Mobile Computing Software Requirements 4
Client System Support for Mobile Computing 4
Server System Support for Mobile Computing 5
Network Support for Mobile Computing 6
User Access to Remote Locations 8

Mobile Computing Support in the Open Blueprint Structure 11
Using a Gateway Server 18

Appendix A. Mobile Device States 23

Appendix B. Notices 25
Trademarks and Service Marks 25

Appendix C. Communicating Your Comments to IBM 27

Figures

1. A Gateway Server 20

© Copyright IBM Corp. 1996

Preface

This publication is composed of two sections. The first section, "Overview of Mobile Computing", introduces mobile computing technology and includes a description of software requirements, systems support, and network support for laptop and notebook computers. The second section, "Mobile Computing Support in the Open Blueprint Structure", describes Open Blueprint resource manager support for mobile computing and provides an overview of application considerations.

This publication defines mobile computing as the ability to use or generate computer based information anytime and anywhere, without being tied to a specific physical location. These time and place attributes are not absolute and are a function of the requirements of a business problem and what an organization is willing to spend in return for a given solution.

© Copyright IBM Corp. 1996

Overview of Mobile Computing

Changing locations, whether walking from an office to a conference room or traveling around the world, has always been a part of the business environment. Yet, historically, the information technology (IT) infrastructure has lagged in its ability to support the user in a consistent manner, independent of location. With the advent of mobile computing technology, coupled with a business need, support for computing anywhere at anytime is fast becoming a reality. Communications and access to data, so familiar to the office or plant environment, is being extended to people while they travel.

Mobile solutions can support both mobile end-users and the management of mobile objects. This paper discusses mobility in terms of mobile end-user communications and access to computing and data located on server systems using notebook and laptop computers.[1]

The application of mobile technology spans most sectors of business and influences the activities of professionals and sales representatives alike. For example:

- A sales representative in a hotel room or at a customer location can obtain timely access to product information.
- Two consultants can collaborate on a report while they are traveling in different countries.
- An insurance representative can update a corporate database from a customer's home.
- An employee can have access to the same information at home, on the train and in the office.

Business Justification for Mobile Computing

Most organizations are under pressure to be more efficient and to deliver better service to their customers. Since its beginning, information technology has contributed to achieving both objectives. Mobile solutions extend the reach of information technology and offer an expanded set of alternatives which can directly contribute to an organization's efficiency. For example:

- Mobile computing provides additional ways for an organization to re-engineer its business and substantially cut costs. Several corporations, including IBM, have saved millions of dollars by mobile-enabling their sales forces, greatly reducing the cost of office buildings and furniture, and saving on the associated utilities. Savings such as this can increase corporate profits.
- By removing the need for an employee to remain at a fixed location, mobile computing can enable workers to spend a greater percentage of the work day doing their job. Said another way, workers can spend more time away from the office or plant and in the field performing their core job, whether it is working with customers or repairing equipment. Mobile computing can result in the need for fewer workers to provide the same level of service.
- Placing information in the hands of workers in the field improves the quality and speed of the decision making process. Workers can access current data such as the latest price list or the status of an order that is needed to make decisions in the field.
- Data obtained in the field can be captured and shared immediately with the entire organization. For example, new order information can be reflected in the corporate inventory database within minutes, not days.
- Mobile solutions improve customer service. For example, insurance customers can be told immediately what settlement they will receive as a result of a house fire.

© Copyright IBM Corp. 1996

Mobile Computing Software Requirements

There are many different ways to exploit mobile computing; there is no single mobile solution. At any point in time, an end user's mobile device can be connected or disconnected. When connected, the connection can be wireline or wireless; if the connection is wireless, the user can be moving within or across networks, or fixed in one place. When disconnected, a user can require access to data and programs that reside on a fixed server. An analysis of the possible device states and their values (as described in Appendix A, “Mobile Device States” on page 23), allows major software requirements for mobile computing to be organized into five categories:

- Location independence

  Mobile computing software must enable mobile users to move their computing devices freely and seamlessly among different networks with no loss of communication. The software must recognize each new location, adapt gracefully, and continue to support the network and the distributed and application services necessary for the task at hand.

- Network transitions

  As a device moves within a LAN or WAN, changes network types, or crosses network boundaries, the software must support each new network connection.

- Disconnected operations

  Mobile computing software must provide function that enables users to work when disconnected from the network.

- Optimization of low-bandwidth links

  Mobile computing software should support the efficient use of low bit-rate wireless and wireline networks. The software should minimize the total amount of data transferred and maximize the number of tasks a mobile user can perform once a connection is made.

- Ease of use

  Given the current state of mobile computing technology, transparency can not always be provided. Therefore, mobile computing software must present the user with an integrated view of the mobile environment, allowing the user to control and manipulate the environment in a consistent manner.

Client System Support for Mobile Computing

A mobile-enabled client system consists of software and hardware components that move with the mobile user and support general computing functions. The client system need not be physically remote from a corporate premise to be considered mobile. Mobile solutions can include client systems that move within a location such as the plant floor or an office complex.

To be mobile-enabled, the client system must be capable of being easily moved, but the device need not be capable of being used while actually moving. A mobile-enabled client system supports the essential elements of computing found in a fixed, desk-top environment, minimizing the differences when appropriate.

Portable and notebook computers contain specialized hardware such as smaller screens, specialized pointing devices and Personal Computer Memory Card International Association (PCMCIA) adapters. The client system can provide configuration and power-use information to systems management software for monitoring and control using Simple Network Management Protocol (SNMP) defined interfaces. Additionally, each element of software on the client system supports advanced power management, which conserves power and prolongs the life of the device's battery.

Mobile client software consists of an operating system such as OS/2, DOS, or Windows, general purpose applications, and the client components of resource managers and business solutions. Additionally:

- If the client system supports disconnected operations, the client software must contain a subset of the server function and the data (or just the data) found in either a server system or a business application or both. Disconnected mobile clients that contain server function and data enable users to perform tasks that, in a non-mobile environment, require a connection to a server.
- The client system must be able to reconfigure itself when it attaches to a new location. The client system must be capable of using newly attached hardware such as a CD-ROM drive in a docking station without restarting the client software.
- The client software must support network transitions such as movements from one subnetwork to a another, for example from a LAN to a WAN, or changes in media class, such as from wireline to wireless. Communications-aware resource managers accept real-time network configuration information and adjust as needed without restarting. Several transport protocols can be supported using the appropriate protocol stack as needed.
- Client software must support low-bandwidth links by limiting the amount of data transmitted between the client and server by using efficient protocols and data filtering.
- The client can contain a common dialer that applications and resource managers can use to establish communications links. The dialer uses telephone information that is maintained in a local directory on the client system.

Server System Support for Mobile Computing

A server system supports mobile solutions by supporting the activities of the mobile client. During a given period of time, the mobile client system can interact with several server programs running on one or more server systems. These server systems support the following requirements:

- Support of disconnected operations

  A client system can contain programs from one or more software servers that support a subset of hardware server functions. This allows the mobile user to perform activities while disconnected that normally require a connection to a server. When a connection exists between a client and a server system, the functions are supported by programs on the server system. The server also places data in a cache (or caches) on the client system. When disconnected, the mobile client software performs the subset of the server functions using cached data. When the client system connects, the server system assists the client in returning (reconciling) to the server any data that changed while the client was disconnected. An application can initiate, control and terminate caching and reconciliation using application programming interfaces (APIs). To further support disconnected operations, client/server flows can use asynchronous communications.

- Low band-width network support

  A mobile-enabled server supports low-bandwidth networks by limiting the amount of data sent to the client. The server optimizes client/server protocols to reduce chattiness, decrease the size of headers, and reduce the number of protocol exchanges used to establish and maintain a connection. In addition, the server filters data to prevent the transmission of unnecessary information and uses data compression techniques to reduce the amount of data that is transmitted.

Network Support for Mobile Computing

Communications networks support the wireline and wireless exchange of information between a mobile client, an enterprise, a public network such as the Internet, or a third-party network provider. The communications network used in a mobile solution reflects both task requirements (Is access to a host or LAN required? Does a user need e-mail?) and the prevailing network environment (What are the wireless usage rates? Is wireline access possible?).

There are three generic types of physical networks:

- Wide area networks
- Local area networks
- Personal area networks

Wide Area Networks

Wide area networks (WANs) support mobile client-to-client and client-to-enterprise communications and data exchange, frequently over large geographic areas. WANs can provide access to value added networks (VANs) and third-party information services such as America Online and CompuServe. WANs make use of public and private telephone and wireless networks.

Wireline: Mobile clients use dial communications to establish wireline connections to other clients and servers through a back-bone. Wireline connections are often implemented using either public switched telephone networks (PSTN) or faster networks such as integrated digital services networking (ISDN). Selection of a dial-up network is based on factors such as the availability of a network type, data throughput requirements, and cost. For example, in the United States, PSTN coverage is offered in every city but it delivers a slow bit-rate, normally in the 14.4k to 28k bits per second (kbps) range, requiring client and server software to be optimized for the low-bandwidth. ISDN is less prevalent and more expensive but delivers bit-rates of up to 128 kbps. Cable wireline networks are emerging and provides very fast data rates.

When a dial-up link is created, a circuit-oriented connection is established and maintained for the duration of a session. Overhead for creating the network connection and establishing any program-to-program communications such as logging into a LAN is often in the range of 20-40 seconds on PSTN networks.

Many applications that are designed for use in an enterprise by non-mobile client systems will work, unchanged, on mobile client systems that use wireline connections and the appropriate communication software (subject to the limitations imposed by the low transmission speed and the application's understanding of the network).

Wireless: Wireless WANs are implemented using radio frequencies (RF) to send and receive transmissions, so a client system can move from location-to-location while remaining connected to a network, or can quickly connect to the network. (How quickly a connection can be established varies by network type.) A mobile client system attaches to wireless networks using either a wireless modem or a wireless LAN adapter that is physically connected to the client system and is used to broadcast RF transmissions to and receive RF transmissions from the network.

RF support requires analog modem or digital radio frequency adapter support. Analog RF is a radio extension of PSTN. Digital networks provide functions such as polling that, in analog networks, must be performed by modems.

RF support can be subdivided into two groups: packet or circuit-switched. In packet wireless networks, the network is shared by many users and is used only when the client sends or receives data. To

complete a data transfer operation, information is broken into blocks of data that are combined with address information to form packets. The packets can be moved through multiple network paths. Software, which can be located on the client system, in the communications network or on an enterprise's base station, must break the data into multiple packets, then reassemble the packets into their original form. Examples of this software include MPT1 from Mobitex and ARTour from IBM. Network providers charge for the use of their packet networks using a per-packet or per-byte rate.

Existing applications must often be changed to support packet radio interfaces and to tolerate the propagation delays found in packet networks. Terminal emulation programs can have poor or inconsistent response times because of the propagation delay. File transfer programs require modifications to time-out values to ensure that the values are long enough to handle the delays imposed by the network. Messaging applications like paging, which transfer small amounts of data, can work without modification when an interface to the communications software is added.

Packet radio networks are often used for mobile solutions when:

- Messaging applications are used
- The network is shared among many users
- The volume of data transfer is small

Circuit-switched radio networks are based on standards such as Advanced Mobile Telephone System (AMPS) for analog support and Time Division Multiple Access (TDMA) for digital support. These networks have many of the same characteristics as wireline networks. Circuit-oriented connections are built on broadcast areas called *cells*. A cell is an area that is serviced by a single transmitter. Cells are connected together by mobile switching centers. Mobile users can move within cells or between cells while maintaining a connection. Circuit-switched network users are charged for the time that they are connected to the network.

Circuit-switched connections are frequently used by applications that require a persistent session such as terminal emulation and remote data access. Circuit-switched networks are used in mobile solutions when:

- The solution uses applications that require a network session
- Dedicated network connections are required
- The volume of data is high

Local Area Networks

Non-mobile client systems connect to LANs in the enterprise using a hardwired attachment such as a network adapter card and cable. Mobile users can access enterprise LANs from within a local geography such as a building or campus.

Local LAN access enables a mobile user to maintain an active network connection while moving within or between LANs. Local area networks can be extended using wireless and wireline technologies.

Wireline: Wireline LAN networks do not support continuous connectivity while moving. Their mobile support is focused on smoothing the network transition from LAN connected-to-disconnected-to-connected states. Relevant changes to resource configurations such as a new list of printers or file servers are reflected in the client system's configuration.

Wireless: An organization can use wireless local area networks to create a new, stand-alone LAN or an extension to an existing wireline LAN. Wireless LANs are similar to wireline LANs, but wireless LANs do not use cable connectors. Radio frequency (RF) and diffuse infrared (IR)[2] LAN adapter cards and their software device drivers permit client systems to remain connected to the network as they move throughout a location, but each of these adapters offers different bandwidth and flexibility. For example, IR adapters are limited to line-of-sight connectivity over distances of several feet. RF adapters permit movement throughout a large building or outdoors.

Personal Area Networks

A personal area network (PAN) can enable two portable computers to exchange data locally or a personal computer to use a printer in a guest office. Software is required to drive and manage the peer-to-peer connection, and to support specific functions such as file transfer. PAN networks can employ both wireless and wireline technologies.

Wireline: Wireline PANs require a physical connection. The most common implementation is the computer-to-computer PAN, which uses specialized cables attached to a parallel or serial port on the mobile device.

Wireless: Wireless PANs can operate across infrared (IR) and RF networks using either packet or circuit-switched radio networks. Direct IR is the predominate PAN type currently in use. For example, many portable computers provide IR adapters for point-to-point exchange of data. IR capability was recently added to selected desktop printers. Diffuse IR, though more expensive than direct IR, permits devices that are positioned anywhere in a room to exchange data. RF PANs are implemented using wireless modems and enable peer-to-peer networking. Both diffuse IR and RF PANs allow movement within a confined space.

User Access to Remote Locations

Mobile users use WANs to connect to systems, and they use LANs to obtain access to programs, data and resources. To best support mobile users, the working environment in the remote location should be equivalent to that of a local user. Many applications, such as TN3270, file transfer protocol (FTP), web browsing and fax, are supported using WAN access to both systems and LANs.

Remote System-to-System Access

Remote system-to-system access enables a mobile system to connect to another system. The connection is supported on any of the wireless or wireline WAN technologies using a variety of subnetworking protocols such as the point-to-point protocol (PPP), the serial line interface protocol (SLIP), and the integrated digital services network (ISDN) protocol. Mobile clients can connect to:

- Another client system

  One client system can connect to another to exchange data, to allow one mobile user to access files stored on the other client system, or to manipulate the desktop of another personal computer. An individual in an organization can connect to a mobile client system to perform diagnostics and add new programs.

- A server

  A client system can connect to a server to exchange data or to use a server resource such as a printer or fax. An application connection can be established to access support services such as mail, database or transaction. A client system can connect to a host system to log onto to a application such as Customer Information Control System (CICS) or Time Sharing Option (TSO).

Remote LAN Access

Remote LAN access enables a mobile client to log into a LAN segment as a full-function node with the capability to address any device on the LAN or to connect to a host system to log onto an application. The LAN session can be established over wireline and wireless connections such as X.25, PSTN and ISDN using specialized software.[3] A wireline, circuit-switched connection using PSTN or ISDN is preferable because many LAN applications are session-oriented and can transfer relatively large amounts of data. Circuit-switched wireless networks can be used to support remote access, but they provide a mobile user with the best service when the amount of data transferred is kept at a minimum.

Packet wireless networks are not often used to support remote LAN access, because packet networks are an expensive alternative for transferring any but the smallest amounts of data. In addition, the effective data rate of packet networks is low, because the packet network is shared by many mobile users.

On all networks, specialized software substitutes a WAN protocol such as PPP or SLIP for the LAN protocol (DLC and MAC). The specialized software allows the LAN requester function and transport services to work, relatively unchanged.

The major difference between local and remote LANs is the speed of the network. Although token ring LANs support a bit-rate of 4 to 16 megabits, and Ethernet LANS support a bit-rate of 10 to 100 megabits; remote LANs are limited by the speed of the WAN connection. Remote LAN access software must minimize the amount of data that flows between the mobile client system and local LAN nodes.

Mobile Computing Support in the Open Blueprint Structure

IBM's Open Blueprint structure is a standards-based architecture that defines the services required by applications in a distributed, multi-vendor environment. The Open Blueprint structure supports the core requirements for mobile computing in this environment. The following sections describe how each Open Blueprint component supports mobile computing.

Sub-Network

Sub-networks support the network types and wireless and wireline technologies described in “Network Support for Mobile Computing” on page 6.

WANs and LANs assist with supporting location independence and network transitions. Wired WANs enable mobile users to move from location to location, remotely accessing systems and LANs as needed. Wireless WAN technologies allow the mobile user to remain connected while moving. The Open Blueprint structure supports a number of digital wireless network technologies including Global System for Mobile Telecommunication (GSM) circuit cellular networks and Mobitex and DataTAC packet cellular networks. Packet exchange is supported using cellular digital packet data (CDPD).

The Open Blueprint structure supports the exchange of data using PANs. Open Blueprint PANs support IR and RF physical connections using standard interfaces such as the Networking Device Interface Standard (NDIS).

Transport Services

A transport protocol requires some form of connected operations. Mobile computing technical requirements include support of low-bandwidth links, location independence, and network transitions.

A mobile client system can move between LANs or across WANs. When a mobile client moves to a new network's domain, the client's IP address must change to reflect the new physical location. TCP/IP supports two services which enable location independence:

- Mobile IP

  Mobile IP, which is based on an Internet Engineering Task Force standard, enables a moving client to maintain its IP address and remain continuously connected. A new Internet routing element enables a server to track the location of a mobile client system and to forward packets to the client wherever it moves.

  For additional information on Mobile IP, see the Internet Engineering Task Force (IETF) home page at URL http://www.ietf.org.

- Dynamic Host Configuration Protocol

  The Dynamic Host Configuration Protocol (DHCP) supports mobility (but does not enable continuous connectivity). A mobile client system can physically connect to a new IP domain and, using DHCP client-server protocols, can obtain a new IP address and configuration parameters from a DHCP server. DHCP cooperates with the Directory resource manager service called Dynamic Domain Name System (DDNS) to ensure that the new name-to-IP-address mapping is reflected in the domain name system.

  For additional information on DHCP, see the IETF home page at URL http://www.ietf.org.

Support for network transitions extends the technical requirements for location independence to include movement between subnetwork types such as LAN to WAN or media types such as wireline to wireless.

© Copyright IBM Corp. 1996

When activated, most transport protocols form a binding with the underlying subnetwork. The transport and all programs that use it must be restarted each time the subnetwork binding changes. A client/server application and a set of network device drivers that are associated with IP support can eliminate these problems in selected wireless networks. For example, IBM's Advanced Radio Communications on Tour (ARTour) Emulator Express application does not need to be restarted when a connection is changed from wireline to wireless, or from one wireless network to another.

An efficient transport protocol minimizes the restrictions imposed by low-bandwidth networks in several ways:

- The transport protocol reduces the time needed to establish a session by minimizing the initial client to server protocol exchanges. The protocol continues to minimize hand-shaking once a connection becomes active.
- The transport protocol uses small, compressed headers to reduce network traffic.
- The transport protocol optimizes network throughput by transmitting multiple packets of data per acknowledgment and ensuring that the link is filled with data.

Signalling and Control Plane

Some remote users require dialer applications to establish network links. These dialer applications interface to the Signalling and Control Plane to manage connections.

Communication Services

Communications Services includes the Messaging and Queuing, Remote Procedure Call, Conversational and Hypertext Transport Protocol (HTTP) resource managers. The Messaging and Queuing resource manager supports disconnected operations. The Remote Procedure Call, Conversational, and HTTP resource managers can be used when the mobile client system is connected to the network. Additionally, HTTP has been extended to meet some of the demands of disconnected operations.

Additional information on the use of the Messaging and Queuing and HTTP resource managers follows.

Messaging and Queuing Resource Manager: The Messaging and Queuing resource manager supports asynchronous data transfer between a mobile client and a server across a wide variety of transport services. With queues on the mobile client, the Messaging and Queuing resource manager supports both disconnected and connected operations. For example, messaging on the client can be implemented using a local queue. The queue stores messages which enable the user to write and send information while disconnected. Messages sent to the mobile client can be queued at the server. When the mobile client reconnects to the server, messages on the client are transferred to the server, and messages on the server are transferred to the client.

To support low-bandwidth networks, Messaging and Queuing resource manager protocols reduce the size of header information, eliminate chattiness, and incorporate rules that prevent the accidental transfer of large data files.

Hypertext Transport Protocol Resource Manager: HTTP, which is defined by an IETF standard, is a connection-oriented protocol that supports the movement of World Wide Web (WWW) data that is composed of text, graphics and other multimedia formats. The mobile user's ability to obtain web information is subject to three major limitations.

- Mobile users cannot access HTTP protocols while disconnected, because HTTP requires a client/server connection.

- Low-bandwidth networks, particularly expensive wireless networks, are inefficient for downloading large files.
- The HTTP protocol is defined such that each time the user requests a web document by clicking on a hypertext link, the protocol invokes routines to create a connection to the document's URL, transfer the entire web document, and tear down the connection.

Application function can minimize these limitations.[4] This function on the client creates local caches that store web documents based on a most-recently-used algorithm and explicit file names, for example, *always cache these 20 text files and three graphics objects*. The caches permit the client's web browser to use the data while disconnected.

The extensions on the server support proxy functions that allow important web data from a variety of servers to be stored in a single location. The proxy functions reduce the time it takes a mobile client to obtain the web data.

The extensions enable the mobile client and server to work in a cooperative manner. First, the extensions enable a single HTTP connection to support many URL request/data transfer iterations. Second, HTTP header information passed between the client and server is substantially reduced once the initial connection is established. And, third, the server maintains information about the state of the data in the client's cache, which permits the server to limit the data transfer to just the sections of a document that have changed since the document was last accessed by the client.

The IETF is currently extending HTTP in the form of draft standards called HTTP/1.1 and HTTP/NG to support persistent connections, cache-control, and proxy behavior. These changes will minimize the three limitations described in the beginning of this section.

Distribution Services

Distribution services support mobile client-to-server exchanges of data using the Directory, Security and Transaction resource managers.

Security Resource Managers: Security requirements in the mobile environment are not entirely unique, but are very important considering the characteristics of a mobile client system and a network. Mobile computing utilizes public networks such as air-waves and PSTN back-bones to transport private data. Mobile computing offers new entry points into the enterprise such as wireless access or dial-up connections. Mobile client systems are less secure than desk-top units because they are comparatively easy to lose or steal.

The Security resource managers provide the basis for protecting the mobile environment. Client cryptographic services are used to protect data that is stored on the client system and that is transmitted over the network. Authentication controls such as passwords are used to restrict the use of the client system, and to control access to programs and data. Servers located in the enterprise also make use of cryptographic services to protect data transmitted over the network. Firewalls are used to prevent intrusion into the enterprise.

Directory Resource Manager: Applications and resource managers on the mobile client system make use of directory services provided by the Directory resource manager server, which is located on a server system. To support location independence for mobile users, the mobile client system can contain a subset of the directory information and function in a local store. This information is used to support the client dialer functions.

In cooperation with the TCP/IP resource manager, the Directory resource manager supports DDNS. When DHCP establishes a new IP address for a mobile client (see “Transport Services” on page 11), the

Directory resource manager dynamically updates the name-to-IP-address mapping in the domain name system to reflect the client's new IP address.

In the future, when a mobile client is a guest in a network, the Directory resource manager will help to support a service called *resource discovery* which assists client applications with learning about the existence and capabilities of devices, applications and services in a network to which the client system has not been previously configured to connect. Additional information about resource discovery can be obtained from the Salutation Consortium home page at URL http//www.salutation.org.

Transaction Manager Resource Manager

The Transaction Manager resource manager supports two-phase commit processing. Mobile client users can be involved in the commit process. If a participating disconnected user delays the completion of a transaction, the server will either hold the resource or assume a back-out. These are important considerations when designing applications for disconnected users.

Application Services

Application services support several core application functions required by mobile users.

Mail Resource Manager: The Mail resource manager enables the client to cache mail documents and act (read, delete, respond) on them while disconnected from the network. The Mail resource manager enables the user to create and store new mail locally. The Mail resource manager services on the client provide the appearance that the mail has been fully processed and that the results have been transferred to the appropriate server.[5] In reality, the actions for each piece of mail are queued on the disconnected client and sent to the server the next time the client's mail service is connected to the network.

Collaboration Resource Manager: The Collaboration resource manager supports asynchronous and real-time collaboration. Collaboration services are used to enable a mobile user to participate in joint activities such as sharing information, document creation and audio/video conferencing.

Asynchronous collaboration supports disconnected operations. This support includes the ability to cache and use documents on the mobile client system. Documents that have been changed can be returned (reconciled) the next time the client connects to the server. The act of caching and reconciling collaboration documents requires interaction between the Collaboration resource manager and an application. The Collaboration resource manager supports low-bandwidth networks by providing the option to transmit only the pieces of a document that have been modified.

The Collaboration resource manager enables mobile client users to participate in synchronous, real-time collaboration with either mobile or non-mobile users. Bandwidth constraints limit a mobile client's participation in asynchronous collaboration and can prevent the use of most wireless networks and many of the slower wireline media for audio and video conferencing.

To help support location independence, the Collaboration resource manager contains a set of named profiles that correspond to specific locations. When an application or user selects a profile, the Collaboration resource manager creates a network environment for the client system. The user can select a profile that invokes a set of changes based on named locations such as "Home", "Office", or "Hotel". A network environment consists of a communications port, subnetwork, transport, and dial string (if required).

Workflow Resource Manager: Workflow is a method of modeling and implementing business processes. The Workflow resource manager supports asynchronous collaboration and supports assigning work items to mobile users. The Workflow resource manager runtime client can be located on a mobile client system. The Workflow runtime client can, for example, enable a sales representative in the field to approve an invoice for a printer, and release the printer for shipment to the customer. The Workflow resource manager uses the Messaging and Queuing resource manager to support disconnected operations.

A mobile user can check-out work items from the server, complete the work items while disconnected, and queue them on the client so they will be transmitted to the server when the user reconnects. Work items that are checked-out to a mobile user for later processing are removed from the work lists of all other users.

A disconnected mobile user can start a process instance on the mobile client system. The process instance information is kept on the mobile client and transmitted to the server at the next connect opportunity.

Event Services Resource Manager: The Event Services resource manager can be used by an organization to support the asynchronous distribution of information from a server system to a mobile client system. For example, Event Services can be used to broadcast an updated price list to a sales force that uses mobile devices. Clients that are connected to the network receive the updates immediately. Disconnected clients will receive the updates the next time the mobile client system reconnects to the network.

Data Access Services

Data Access Services support the storage, retrieval and management of data. Mobile-enabled data access services allow the mobile client to use these services when the mobile client is disconnected from the network or connected over low-bandwidth networks. Three Data Access resource managers are mobile-enabled.

Storage Management Resource Manager: The Storage Management resource manager is responsible for making back-up copies of client data and storing the copies on a server system. The Storage Management resource manager overcomes two difficulties with performing back-up of data located on a mobile client system. First, it supports the ability to conduct a back-up operation that spans multiple connect-disconnect-connect cycles. Second, it can perform back-up over low-bandwidth networks by performing incremental back-ups on the smallest possible unit of changed data such as a file, record, or field (but never on an entire disk drive), limiting the amount of data that must be moved through the network.

Relational Database Resource Manager: To support disconnected operations, the Relational Database resource manager copies data from a server system to a compatible store on the client system. Applications on the client can use this data while the client is disconnected. The Relational Database resource manager can also support returning (reconciling) the changed data to the server. Caching and reconciliation are performed jointly by the Relational Database resource manager and an application.

File Resource Manager: The File resource manager enables disconnected operations by supporting data caching. Data reconciliation must be done by the application.

Systems Management

Resource managers jointly support core systems management functions such as logging, notification, and change management, and they have extensions to enable the management of remote client access and disconnected operations.

- Remote access hardware such as modem pools and software employed in low-bandwidth networks can be monitored and managed using SNMP-compliant applications. The systems management function supports capacity planning, work-load balancing, and availability management. API's enable system management applications to receive this information and respond as needed.
- Mobile clients used in both disconnected and connected modes of operation are supported in a number of ways. For example, software distribution from a server to a mobile client can span multiple connect-disconnect cycles.
- Client applications can use Desktop Management Task Force (DMTF) interfaces to obtain information about the state of a disconnected device's power supply and configuration. A client system can be monitored on an exception basis. When the client is disconnected, error alerts and other recordable events are recorded in a local store on the client. When the client is connected to a server, the log and the new events can be routed to an application on the server.

Presentation Services

Open Blueprint Presentation Services provide the functions needed to display information to a user. The Web Browser and View resource managers enable the mobile client system to display information in a consistent manner.

Web Browser Resource Manager: The mobile user's ability to navigate and view WWW data is limited by the Web Browser resource manager's dependence on the HTTP resource manager. The Web Browser resource manager can be enabled for low-bandwidth networks by using the functions described in “Hypertext Transport Protocol Resource Manager” on page 12.

There are two approaches to enabling the browser for disconnected operations.

- Use HTTP protocols to access data in a cache on the mobile client. This might require the use of the functions described in“Hypertext Transport Protocol Resource Manager” on page 12 .
- Use file I/O to access web data in a byte file system on the client system. This requires an application to identify the data to the File resource manager before being disconnected.

View Resource Manager: Documents that are formatted for printers such as PostScript and Mixed Object: Document Content Architecture-Presentation (MO:DCA-P) documents can be a problem for a disconnected mobile user who does not have access to a printer. The View resource manager can be used to view these documents on the client system while either connected or disconnected.

Applications and Development Tools

The Open Blueprint structure supports the development of distributed mobile applications using a variety of services and tools. Lap-top and notebook computers do an excellent job today of supporting many of the stand-alone, personal-productivity applications used by mobile professionals and sales people. Even with smaller displays and keyboards, it is relatively easy to manipulate local data using a word processor, spread sheet or graphics application. Difficulty can arise when a client application uses data that is normally stored on another device or that interacts with applications and resources at remote locations.

A specific mobile solution design dictates the appropriate selection of a tool or service.

Application Design: A set of application design criteria is defined in the Open Blueprint structure that must be situationally applied. These design criteria are offered as general guidelines and should be used in conjunction with organizational programming practices. The application design criteria cover the topics of:

- Access to data
- Ease of use
- Asynchronous support
- Network independence

Access to Data: The application designer can provide access to data using one of two methods: either give the mobile client on-demand access to the data where it normally resides, or cache the data on the client system. The selection of either or both methods is dictated by the answers to several questions: What are the sources of the data? How much data does the client need to access? Where is the data currently stored? Which applications share and update the data? How frequently is the data accessed? How up-to-date must the data be?

The two methods for accessing data are:

- On-demand access

  On-demand access to enterprise data means that the client can read or write the data at one or more remote locations within the time constraints dictated by the solution and subject to the availability of an appropriate network. For a traveling business executive, it might be sufficient to offer nightly server access using a hotel phone line. Conversely, a sales representative might require very quick, on-demand wireless access to a customer database.

- Caching data

  Caching data is the process of storing data on the mobile device so client applications can use the data while disconnected. If the enterprise data stored on the client is read only, that is, never changed while on the client, the client application need only identify what data will be needed and issue the appropriate requests to have it copied. If the data is read/write, in addition to the first step, the application must support the return (reconciliation) of the changed data to the correct enterprise store.

The two models can be combined so the client application uses a data cache for a period of time or until a predetermined event occurs, and then connects back to the enterprise to either refresh the cache or reconcile with the appropriate data store, or both.

To support mobile computing, the application can use cryptographic services or perform data compression.

Ease of Use: The mobile environment can be complex. Special dialing instructions, multiple access numbers, and network incompatibilities can cause ease-of-use difficulties with connecting to and accessing a remote system. For the most part, mobile users are without the immediate advice of a friend or the support of a central HELP Desk. These factors increase the importance of simplifying end-user tasks.

Collectively, the mobile applications on a client system should have a consistent look and feel. This is more than simply selecting a single desk-top shell such as OS/2 Presentation Manager or the Windows 95 Desk-top. An application platform must also be selected that supports the mobile user in a consistent and natural way. The collaboration model is an example of an application platform that naturally supports mobile users (see “Support for Mobile Applications” on page 18).

The mobile client system must be able to control and adjust its configuration based on the current location. The ideal solution is one in which applications sense the current location of the client system and adopt the configuration and behavior that is most appropriate (use of a specific dial prefix, file transfer program, or terminal emulator, disabling the download of graphics, and so on).

Asynchronous Support: A fundamental means of supporting disconnected operations is to use the messaging and queuing model. In addition to using asynchronous data flows, this model supports three important services: client side queuing, assured delivery of messages, and support for a variety of transport protocols.

Network Independence: To facilitate mobility, an application must work within a variety of networks that change depending on the location.

- To support network transitions, application design dictates as much separation as possible between the application and its dependence on any one transport protocol or subnetwork.
- A single connection between the client and a server can be shared between two or more applications. An example is the simultaneous use of a dial connection by mail, LAN requester and database applications using multiple protocols (TCP/IP, NETBIOS). Each application must be designed to use minimal network resources and to not depend on exclusive access to networking facilities.
- To support low-bandwidth networks, application client/server protocols should minimize the flow of data through the network. In addition, an application might be required to support increases and greater variations in network latency.

Support for Mobile Applications

The Open Blueprint structure supports an asynchronous collaborative model of mobile computing, based on the technologies supported by Lotus Notes. By integrating and exploiting the mobile capabilities of several resource managers, this model provides a set of core mobile services and tools that provide the basis for developing and deploying mobile applications. The collaborative model provides services that support the following:

- A mobile-enabled mail service that supports disconnected operations and low-bandwidth networks.
- The re-creation of parts of the collaborative server on the client system with the ability of an application or user to enable or disable the server functions.
- The replication and resynchronization of document server data on the mobile client. The collaborative model supports indirect caching of transaction, relational, and hierarchical data if it has been converted to a collaborative format.
- A set of named profiles that can be used to reconfigure a mobile client for a specific location.
- Mobile-enabled workflow services that support disconnected user participation in business processes.
- Agent scripts that reside on servers and clients and act on behalf of the disconnected mobile client. Agents are application logic that allow developers to automate application tasks on both servers and clients. Such tasks include data handling, replication, and messaging services.
- Authentication, authorization, and cryptographic services that support the mobile client and the data flows between the mobile client and server.
- A client/server protocol that is optimized for low-bandwidth connections.

Using a Gateway Server

Enterprises need to extend non-mobile applications to support mobile users. However, extending these applications presents some challenges. Many existing applications are designed for a fully connected, non-mobile environment. Existing software does not enable easy replication of resources on mobile clients. Existing software assumes that communications links are stable, always connected, and have high bit-rates. Existing software is not designed to maintain application integrity during changes in the

underlying communications support. It can be impractical to move much of the existing software to a mobile, disconnected environment and expect to maintain the semantics of the software.

One approach to overcoming these challenges is to configure a *gateway server*. A gateway server (shown in Figure 1 on page 20) is a logical entity located between the application and network environments and the enterprise. The gateway server is a logical entity because the functions represented by the gateway server are implemented using multiple Open Blueprint resource managers and applications. The functions contained in the gateway server extend the enterprise information and communications structure to the mobile entity, while separating the two environments and providing a single point of control. The gateway server can provide:

- Remote client access to the enterprise
- Control of information leaving the enterprise
- Separation of the enterprise and the external application and network environments.
- Systems management support
- Consolidation of access to supporting networks

Mobile Client Access to the Enterprise

The gateway server supports the mobile client's access to the enterprise as follows:

- The gateway server owns the subnetworking, transport and communications hardware and software that provide the connectivity between the client system and the enterprise. For example, the server controls the wireline modem pools and the software required to support 3270 emulation or remote LAN access. Additionally, the server is responsible for network address administration.
- The gateway server contains directory information that the mobile client uses to locate services within the enterprise such as a transaction server. The client uses either an explicit name or network address, or in the future, resource discovery.
- The gateway server can be configured to contain a complete set of Open Blueprint security functions, including cryptography, which is used to protect data transmitted between the server and client, and firewalls, which are used to authenticate and authorize access to the enterprise.

Control of Information Leaving the Enterprise

The gateway server is the exit point for information leaving the enterprise and going to the mobile client. The gateway server provides two key services:

- It stores information for the mobile client until the client connects. For example, the results of a query to multiple databases can be queued at the gateway server until the mobile client connects, at which time the information is moved to the client. Similarly, workflow navigator function can queue workflow data and work items at the gateway.
- The gateway server contains the outbound dialing functions such as a dialer and a directory containing client phone numbers that are used to initiate a connection to the client. This is an important set of services that are used when a non-mobile client within the enterprise must communicate with a moving client, or when there is information that must be "pushed" from the enterprise to the mobile client.

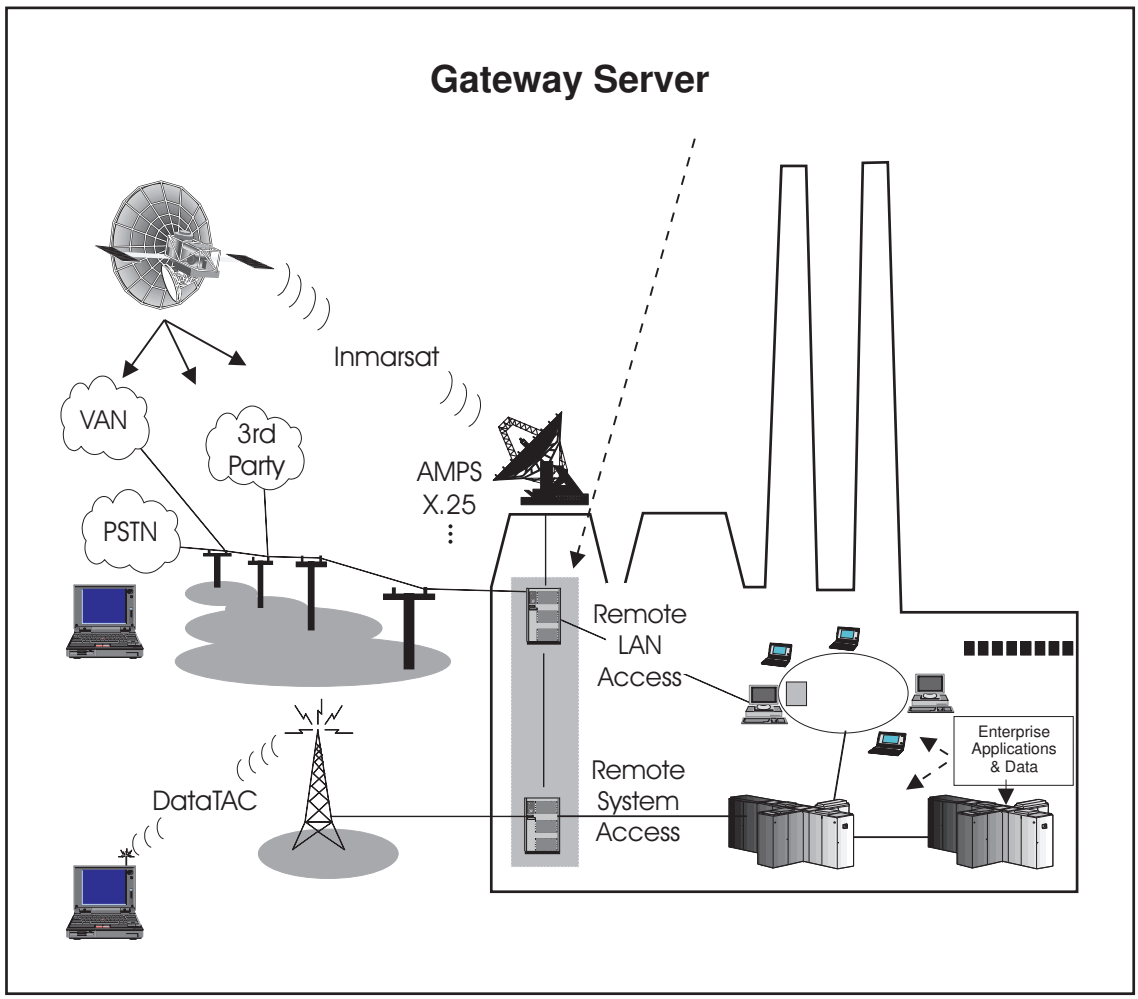


Figure 1. A Gateway Server

Separation of the Enterprise and the Mobile Client

Significant differences can exist between a network used by the mobile client and the communications infrastructure used in the enterprise. The differences include bandwidth, error conditions, protocols and the availability of a persistent connection. The gateway server helps to mask these differences.

The gateway server supports the subnetwork protocols and transport services that are used by both the mobile client and the stationary enterprise. This support relieves the mobile client system and the enterprise server programs from having to understand each others' protocols. It also ensures that the programs in the enterprise are not affected by network transitions and common wireless network events such as signal fading. Separation of the enterprise and the mobile client is supported as follows:

- At the transport level, the gateway server contains a pass-through function that routes a mobile client's request or session to other stationary servers. Depending on which protocols are in use, the gateway server can provide protocol conversion, for example TCP/IP to SNA, eliminating the need for the client to support all enterprise transport protocols.
- The gateway server minimizes the difference between the low-bandwidth, mobile client network and the high-speed enterprise network by filtering and compressing information to reduce the amount of data that is transmitted to the client.
- To support disconnected operations, the gateway server is used as single point of implementation for applications which provide a persistent presence on behalf of the mobile client. A simple example is

the placement of a mail server on the gateway server. A more complex example consists of the use of intelligent agents or proxies that respond to requests from clients or servers, or perform work on behalf of the mobile user while the client is disconnected from the network.

Systems Management Support

The gateway server is a critical point for obtaining the information needed to manage the mobile client and supporting networks. This does not imply that it contains a systems management console. Rather, as a single point of access, the gateway server is well positioned to coordinate the collection of the raw data used in the essential systems management disciplines such as availability management and capacity planning.

The gateway server can be used to stage the distribution of software updates and the collection storage management data.

Consolidating Access to Supporting Networks

The gateway server can consolidate access to the network service providers such as America Online and Compuserve. Although the client can access network services directly (most third-party providers support direct access to their services), it is sometimes more efficient to centralize and route access to save costs or to exercise control over service use.

---

[1] Low function devices such as personal information managers (PIMs) are another class of device that have their own, unique set of attributes. This paper excludes them for the purpose of simplifying the more general support issues, although pagers and smartphones are supported by telephony applications.

[2] IR communications modulates invisible light waves to transmit data. Like visible light, IR is blocked by walls. Diffuse IR spreads out and can bounce off objects before being received. It can be contrasted with direct IR, which must be transmitted from point-to-point (direct line-of-sight), but has a greater range and a higher bit-rate.

[3] This type of software is represented by products such as IBM's LAN Distance and Shiva's LAN Rover family.

[4] IBM's ARTour Wireless Web Express product is an example of an application that contains this function.

[5] In some cases, mobile users can reprocess the mail anytime before they connect to the network.

Appendix A. Mobile Device States

This paper describes the possible states of a mobile device. At any given time, each state has one of two values. These states and values are:

State	Value
Mode of operation	Connected or disconnected
Network media	Wireline or wireless
State of mobility	Moving or fixed

These states and values are described in the next several sections.

Mode of Operation

A mobile system's mode of operation is either connected or disconnected. The mode of operation determines whether a system can communicate with a network or another device.

- Connected:

  When a mobile system is connected, it is linked to a network or another device. When the connection is active, data transfer and communication flow can occur.

- Disconnected:

  When a mobile system is disconnected, there is no working connection between the system and a network or another device, so communication or data exchange is not possible. The user can only use data and programs that are stored on the mobile system.

Network Media

Network media defines the class of media that is used to make the connection between a mobile device and a network or another device. Network media is either wireline or wireless.

- Wireline network media links a device to another device or to a network using a physical connection such as a phone circuit.
- Wireless network media links a device to another device or to a network through open-air using technologies such as radio frequency and infrared. There is no physical connection that the device user can see or feel.

State of Mobility

State of mobility describes how a portable device is used; either while moving or while in one of many fixed locations.

- *Moving while computing* describes portable devices that change location while performing computing functions.
- *Fixed while computing* describes devices that are always in some fixed location while performing computing functions (fixed in an office, then moved to the home where the device is fixed again).

© Copyright IBM Corp. 1996

A Day in the Life of a Mobile User

The state associated with a mobile device can shift as the mobile user moves around. At any point in time, the mobile system can exist in one of the possible combinations of operations, media and mobility, described previously in this section. During the course of a day, a mobile client system can exist in several of these combinations. For example,

- A user uses a notebook computer as a stand-alone PC (stationary and disconnected) on a conference room table.
- Later, the user moves the computer to an office and connects to a wireline LAN using a docking station. Later that day, the user takes the PC across the site to attend a meeting. The user connects to a new LAN, and the application and communications software make the appropriate network transition from the first LAN to the second LAN.
- At the end of the day, the user boards a train and connects to a wireless LAN over a low-bandwidth, wireless connection to download e-mail.
- When the e-mail is downloaded, the user disconnects from the network and reads the mail while traveling home.

Appendix B. Notices

References in this publication to IBM products, programs, or services do not imply that IBM intends to make these available in all countries in which IBM operates. Any reference to an IBM product, program, or service is not intended to state or imply that only IBM's product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any of IBM's intellectual property rights may be used instead of the IBM product, program, or service. Evaluation and verification of operation in conjunction with other products, except those expressly designated by IBM, is the user's responsibility.

IBM may have patents or pending patent applications covering subject matter in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing
IBM Corporation
500 Columbus Avenue
Thornwood, NY 10594
USA

Trademarks and Service Marks

The following terms are trademarks or service marks of the IBM Corporation in the United States and/or other countries:

ARTour
CICS
IBM
IBMLink
MO: DCA
Open Blueprint
OS/2
Presentation Manager

The following terms are trademarks or service marks of other companies:

ARDIS	Ardis Company
America Online	America Online, Incorporated
CompuServe	CompuServe Incorporated and H&R Block, Incorporated
DCA	Digital Communications Associates, Incorporated
Lotus Notes	Lotus Development Corporation
MCI	MCI Corporation
Mobitex	Televerket
NDIS	3Com Corporation and Microsoft Corporation
Prodigy	Prodigy Services Company

Microsoft, Windows, and the Windows 95 logo are trademarks or registered trademarks of Microsoft Corporation.

© Copyright IBM Corp. 1996

Appendix C. Communicating Your Comments to IBM

If you especially like or dislike anything about this paper, please use one of the methods listed below to send your comments to IBM. Whichever method you choose, make sure you send your name, address, and telephone number if you would like a reply. Feel free to comment on specific error or omissions, accuracy, organization, subject matter, or completeness of this paper.

When you send comments to IBM, you grant IBM a nonexclusive right to use or distribute your comments in any way it believes appropriate without incurring any obligation to you.

- If you prefer to send comments by FAX, use this number:

  United States and Canada: 1-800-227-5088.

- If you prefer to send comments electronically, use one of these ID's:
  - Internet: **USIB2HPD@VNET.IBM.COM**
  - IBM Mail Exchange: **USIB2HPD at IBMMAIL**
  - IBMLink: **CIBMORCF at RALVM13**

Make sure to include the following in your note:

- Title of this paper
- Page number or topic to which your comment applies

© Copyright IBM Corp. 1996

Printed in the United States of America
on recycled paper containing 10%
recovered post-consumer fiber.

XXXX-XXXX-00