SRI VENKATESHWARAA COLLEGE OF ENGINEERING & TECHNOLOGY
(Approved by AICTE, New Delhi & Affiliated to Pondicherry University, Puducherry.)
13-A, Villupuram – Pondy Main road, Ariyur, Puducherry – 605 102.
Phone: 0413-2644426, Fax: 2644424 / Website: www.svcetpondy.com

Department of Computer Science and Engineering

Subject Name: **MOBILE COMPUTING** Subject Code: **CS E84**

UNIT I

Introduction: Wireless and Mobile Computing Architecture – Limitations of wireless and mobile communication – Wireless Telecommunication Networks: Digital cellular Systems, TDMA - CDMA – Wireless Networking Techniques –Mobility Bandwidth Tradeoffs – Portable Information Appliances.

---

2 Marks

1. What is mobile computing?

Mobile computing is a technology that allows transmission of data, via a computer, without having to be connected to a fixed physical link.

2. What is Mobility?

In mobile computing, mobility refers to characteristics of device to handle information access, communication and business transactions while in state of motion.

A person who moves

- ✓ Between different geographical locations
- ✓ Between different networks
- ✓ Between different communication devices
- ✓ Between different applications

A device that moves

- ✓ Between different geographical locations
- ✓ Between different networks

3.What are two different kinds of mobility?

User Mobility:

It refers to a user who has access to the same or similar telecommunication services at different places.

Device Portability:

Many mechanisms in the network and inside the device have to make sure that communication is still possible while the device is moving.

4. Find out the characteristics while device can thus exhibit during communication.

- Fixed and Wired
- Mobile and Wired
- Fixed and Wireless
- Mobile and Wireless

5. What are applications of Mobile Computing?

- Vehicles
- Emergencies
- Business
- Replacement of wired networks
- Infotainment
- Location dependent services

Follow-on services:

Location aware services
Privacy
Information services
Support services

Mobile and wireless devices:

Sensor
Embedded controllers
Pager
Mobile phones
Personal digital assistant
Pocket computer
Notebook/laptop

6. What are the obstacles in mobile communications?

- Interference
- Regulations and spectrum
- Low Bandwidth
- High delays, large delay variation
- Lower security, simpler to attack
- Shared Medium
- Adhoc-networks.

7. What is Wireless?

The term “wireless” refers, in the most basic and obvious sense, to communications sent without wires or cables. It is a broad term that encompasses all sorts of wireless technologies and devices, including cellular communications, networking between computers with wireless adapters, and wireless computer accessories. Wireless communications travel over the air via electromagnetic waves (radio frequencies, infrared, satellite, etc).

8. What is Wireless Networking

Networking technologies that connect multiple computers and devices together without wires -- i.e., in a wireless local area network or WLAN -- also fall under the wireless umbrella. Often, instead of referring to just "wireless" for these technologies, the term "wi-fi" or "wifi" will be used. Wi-fi covers technologies that incorporate 802.11 standards, such as 802.11g network cards and wireless routers.

9. List the types of wireless networks

Wireless networks use radio waves to connect devices such as laptops to the Internet, the business network and applications. When laptops are connected to Wi-Fi hot spots in public places, the connection is established to that business’s wireless network.

There are four main types of wireless networks:

- Wireless Local Area Network (LAN): Links two or more devices using a wireless distribution method, providing a connection through access points to the wider Internet.
- Wireless Metropolitan Area Networks (MAN): Connects several wireless LANs.
- Wireless Wide Area Network (WAN): Covers large areas such as neighboring towns and cities.
- Wireless Personal Area Network (PAN): Interconnects devices in a short span, generally within a person's reach.

10. Give Examples for wireless devices.

Examples of wireless devices include

Cell phones, PDAs, GPS systems, wireless mice, wireless keyboards, remote controls, wireless routers, wireless network cards, and pretty much anything else that doesn't use wires to transmit information**.**

11. What is meant by wireless and mobile computing architecture?

The architectural model of a mobile computing environment consists o f stationary and mobile component s . Fixed hosts are connected together via a fixed high - speed network (Mbps to Gbps). Some of the fixed hosts are special computers equipped with wireless interfaces, and are known as base (radio) stations (BS). They are also known as mobile support stations (MSS). Base stations, which are placed in the center of cellular coverage areas, act as access points between the mobile computers and the fixed network.

12. What are the limitations of wireless and mobile communication?

Frequent disconnection .
Limited communication bandwidth .
Heterogeneous and fragmented w ir el ess network infrastructure rapid and large fluctuations in the network QoS.
Security and anonymity
Service relocation
Support for location-sensitive applications

13. Definition of Digital cellular system

Digital Cellular System: Any cellular phone system that uses digital (e.g. TDMA, GSM, CDMA).

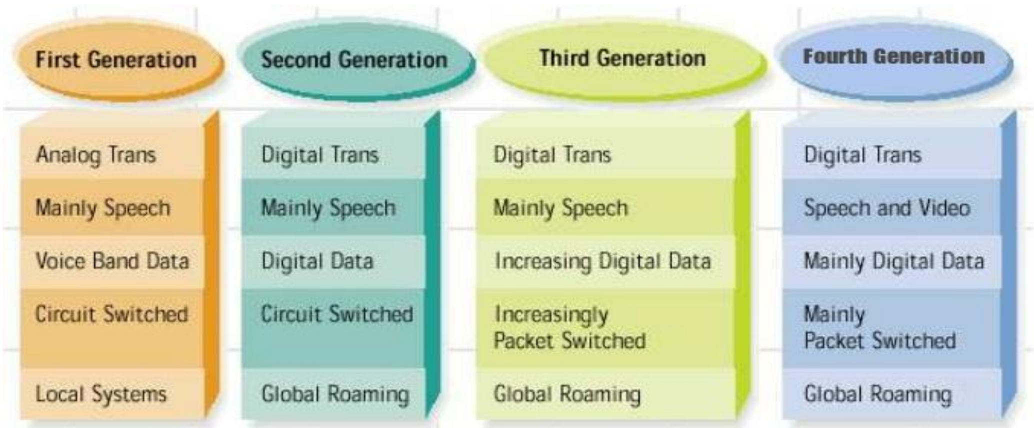
14.Write Cellular network generation

1G:First generation Analog cellular system
 -Analog voice
2G:First Digital Cellular System
 -Digital voice and messaging
2.5G: Digital Cellular System
 -Increase in digital data rates
3G: Digital Cellular System with increase in functionality -
 Broadband data and voice over IP
4G:Future re-architecting of digital cellular infrastructure -
 Increased data throughput

15. Write the Comparison of 1G to 4G systems.

Generation	1G	2G	3G	4G
Wireless Access	Analog	*Digital*	*Digital*	*Digital*
	FDMA	TDMA, *CDMA*	*CDMA*	*OFDMA, MC-CDMA*
Major Services	Voice	Voice	Voice	*Voice over IP*
		Internet (text only)	*Internet (text, images)*	*Rich Internet*
Core network	Circuit-based	Circuit-based	Circuit- and *Packet-based*	*Fully IP-based*

16. Write the Evolution of Cellular Networks



17. What is meant by TDMA?

TDMA is a digital transmission technology that allows a number of users to access a single radio frequency channel without interference, by allocating unique time slots to each user within each channel.

18. What is the purpose of TDMA?

One problem with TDMA is the wasted bandwidth of unused slots. Time slots are allocated to specific users whether or not they are using the slots (talking or transmitting data). Hughes Systems Network has contributed an enhancement of TDMA known as Enhanced TDMA (ETDMA) that attempts to correct this problem. Instead of waiting to determine whether a subscriber is transmitting, ETDMA assigns subscribers dynamically based on whether a user has voice/data to transmit.

19. What is meant by CDMA?

Code-Division Multiple Access (CDMA) offers a solution to the capacity limitation problem. It allows all mobile stations to concurrently use the entire spectrum (all channels) with much less interference. Instead of partitioning either spectrum or time into disjoint "slots", each subscriber is assigned a unique instance of a pseudo-noise digital signal. The transmission signal is "spread" over the entire spectrum, using the noise signal. CDMA is, therefore, known as a spread spectrum modulation scheme.

20. What are the groups of wireless technologies?

Wireless technologies can be grouped into at least six major categories: (1) in- room, point to point infrared, (2) in-room radio, (3) in-building radio frequency, (4) campus or metropolitan area packet networks, (5) wide-area packet/circuit switched data networks, and (6) regional-area Satellite Data Networks.

21. Define mobility-bandwidth tradeoffs.

In- building cellular offers the highest bandwidth (bi-directional), but very limited mobility. Micro-cellular offers lower bandwidth but allows for limited-speed mobility; macro-cellular offers much lower bandwidth but allows for the highest degrees of mobility. As can be noticed, in these networks, the larger the coverage area (the cell size), the higher the degree of mobility.

22. What is the use of portable information appliance?

It is used to record numeric information. This information was probably very important to the user of this appliance and in some way directly affected his livelihood. It may have actually provided the "function" of counting by allowing the user to create a mark corresponding to a piece of livestock. This would have been very useful to an individual whose society had not yet invented a system of numbers.

23 . Explain about PDA .

PDAs emerged in 1993 amid claims of single-point data organization, ubiquitous and instantaneous communications, and new operating paradigms using glitzy graphical user interfaces (GUI) and handwriting recognition.

24. Explain about palmtop computers.

It is likely that wireless network connectivity will trail wired connectivity in terms of performance for the foreseeable future. The best strategy for the developers of portable information appliance is to design products which either provides useful standalone functions such as an electronic still camera, or which complement wired network platforms. The emerging market of Palmtop Computers is a breakthrough in terms of the ability of the Palmtop to complement the desktop computer.

25. Define communicators.

The Communicator is a PDA concept that combines the benefits, portability and functionality of digital cellular phones and palmtop co mp u t er s . The idea is to stick a palmtop computer to a cell phone with data capabilities to provide remote access, in addition to the stand-alone form factor applications that can be found on palmtop computers. Internet access, telnet, email, and web browsing are all applications offered by communicators.

26. What is in building radio frequency?

This type of network, which is also known as Wireless LAN, expands the range of the infrared and the Bluetooth technologies by increasing the network diameter to about 200m. Unlike infrared and Bluetooth, in-building radio frequency is a cellular network, where mobile computers are allowed to roam within and across cells.

27. What is campus area packet network?

This network type encompasses the more traditional "cellular" networking paradigm. It is typified by a "pole top infrastructure" supporting network diameters of 0.2 to 5 miles with data rates of 20-128 kbps. Relay (or router) nodes a r e strategically placed to support the wider network diameter with a small price for increased latency.

28. Write some of the portable information appliances.

Two of the portable information appliances, the pocket watch and the printed book are relatively recent inventions which have transformed human society. The pocket watch enabled the level of logistical synchronization between individuals required for industrialization. Printed books, while not as interactive as paper and pencil, have also evolved as the preferred method for accessing standardized information in a portable format.

29. Explain satellite networks.

Satellite technology is still emerging. It is a downlink technology where mobile computers can only receive direct broadcast from a satellite. Outbound communication is initiated by the mobile computer through a modem DIAL-UP or other wireless technology. Hughes Network Systems pioneered the DirecPC net- work which uses the Galaxy satellite and which delivers 400 kbps downlink rate.

30. Write down the Nokia 9000 specifications (UQ April'13)

Table 3.4 The Nokia 9000 Specifications

Item	Specifications
Memory	8MB total: 4MB OS and applications, 2MB program execution, 2MB user data storage
Processor	embedded INTEL 386 processor
Operating System	GeOS TM3.0
E-mail protocols	SMTP, IMAP4, POP3 and MIME1
Weight	397g
Dimensions	173 x 64 x 38 mm
Displays	Grayscale 640x200 (illuminated) LCD

31.What are the limitations of Mobile Enivironment (UQ April'13)

limitations include:

▪ Frequent disconnection caused by one of the following events:

– handoff blank out in cellular networks; the problem is worse in micro-cellular networks

– long down time of the mobile computer due to limited battery lifetime

– roaming-off outside the geographical coverage area of the wireless service

▪ Limited communication bandwidth impacting the following:

– quality of service (QoS) and performance guarantees throughput and response time and their variances

▪ Heterogeneous and fragmented wireless network infrastructure leading to the following problems:

– rapid and large fluctuations in the network QoS ▪

Other problems include:

– security and anonymity

– service relocation

– support for location-sensitive applications

11 Marks

1. Write in detail about mobile computing (5)

Buzzwords such as *mobile, ubiquitous, nomadic, untethered, pervasive,* and *any time anywhere,* are used by different people to refer to the new breed of computing that utilizes small portable devices and wireless communication networks.

The difference between nomadic and mobile computing is particularly important to point out. Both nomadic and mobile computing require small portable devices. However, the kind of network used in nomadic computing does not allow mobility, or does so in the confines of a building, at pedestrian speed. Examples of such networks are DIAL-UP lines, which obviously do not allow any mobility, and Wireless Local Area Networks (W- LAN), which allow for limited mobility within a building facility.

Nomadic computing refers to the interleaved pattern of user relocation and "in-door" connection. Travelers carrying laptops with DIAL-UP modems are, therefore, nomadic users engaged in nomadic computing. Mobile computing, on the other hand, requires the availability of wireless networks that support "outdoor" mobility and handoff from one network to the next, at pedestrian or vehicular speeds.

A bus traveler with a laptop connected to a GSM phone or a CDPD modem is a mobile user engaged in mobile computing. Figure 1.1 depicts this taxonomy. It also shows ubiquitous computing to be the aggregate ability to compute in both the nomadic and the mobile modes. Mark Weiser, a pioneer and a visionary from Xerox PARC, had different view and definition for ubiquitous computing. The reader is referred to his famous 1991 article in Scientific American [91]. We caution the reader that, in this book, the term mobile computing is used to refer to both nomadic and mobile computing, to reduce the clutter.

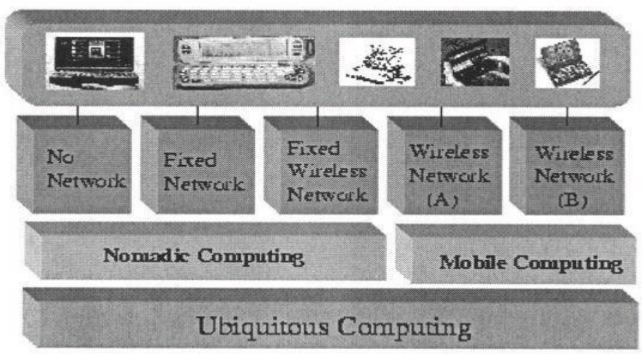


Figure 1.1 Ubiquitous = nomadic + mobile

2.Explain wireless and mobile computing architecture (6) APRIL/MAY2014

The architectural model of a mobile computing environment is shown in Figure 1.6 and consists of stationary and mobile components. Fixed hosts are connected together via a fixed high-speed network (Mbps to Gbps). Some of the fixed hosts are special computers equipped with wireless interfaces, and are known as base (radio) stations (BS). They are also known as mobile support stations (MSS). Base stations, which are placed in the center of a cellular coverage areas, act as access points between the mobile computers and the fixed network. Mobile computers can be in one of three states. The first state place

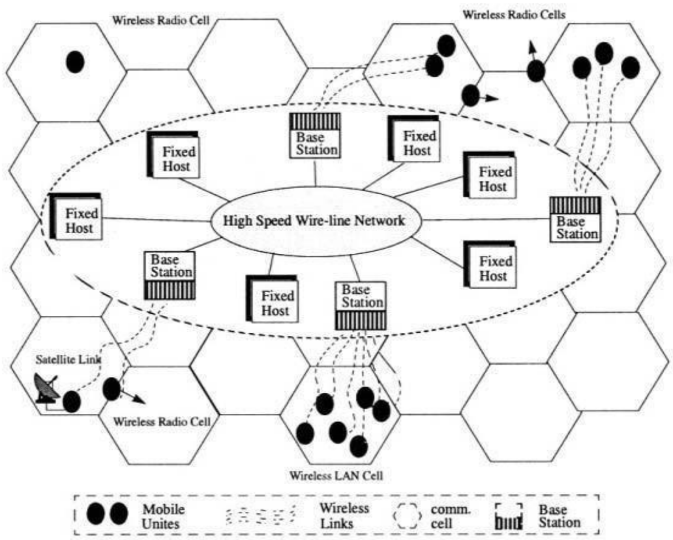


Figure 1.6 Mobile computing environment

a mobile computer within a cell and capable of communicating. The second state places the mobile computer out of range of any service cell and not ca- pable of communication. The third state places a mobile computer in a cell, communicating, but just ready to cross a cell boundary. These scenarios are depicted in Figure 1.6. Figure 1.6 is a generalized architectural overview of a typical wireless/nomadic system. Many such systems have been deployed both in the United States and Europe as well as in many other parts of the world. One such European sys- tem is the Global System for Mobile Communications (GSM). GSM, which is depicted in Figure 1.7, was originally developed by the European Institute for Research and Strategic Studies in Telecommunications (EURESCOM) as an advanced mobile communications technology. During early stages of deploy- ment, GSM was hailed as a superior wireless technology because the general architecture supported such features as roaming, minimum disruption when crossing cell boundaries, and connectivity to any number of public wired infras- GSM is gaining increased popularity in North America. Figure 1.8 quantifies GSM penetration in terms of number of states with GSM services in the US.

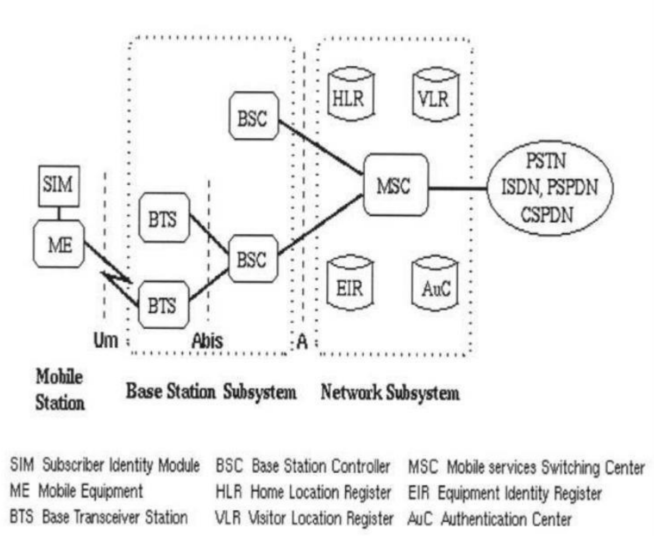


Figure 1.7 GSM architecture

The PCS system with AIN services outlined in Figure 1.9 is comparable to the overlay internetworking system described by Katz and Brewer [69]. The PCS/AIN system shown above is comprised of many different forms of communications (e.g. cellular, PCS, wired, POTS (Plain Old Telephone Service), etc.) with a centralized management scheme as defined by the Telcordia AIN standards. There exists interconnection across planes and between overlay planes to establish service attributes. One of the many issues to be addressed is how do wireless service providers and application developers create, deploy, and control applications support services given the systems described above.

3. Write the limitations of the wireless and mobile environment (5)

The limitations of wireless and mobile environment are as follows:

- Frequent disconnection caused by one of the following events:

– handoff blank out in cellular networks; the problem is worse in micro- cellular networks
– long down time of the mobile computer due to limited battery lifetime
– voluntary disconnection by the mobile user
– disconnection due to hostile events such as theft and destruction
– roaming-off outside the geographical coverage area of the wireless service

- Limited communication bandwidth impacting the following:

– quality of service (QoS) and performance guarantees
– throughput and response time and their variances
– efficient use of battery due to long communication delays (wireless interface requires battery energy during the slow send and receive)

- Heterogeneous and fragmented wireless network infrastructure leading to the following problems:

– rapid and large fluctuations in the network QoS

– mobility transparent applications perform poorly without some sort of mobility middleware or proxy.

– poor end-to-end performance of different transport protocols across networks of different parameters and transmission characteristics.

▪Other problems include:

– security and anonymity

– service relocation

– support for location-sensitive applications

There are other limitations related to platform and application development methodologies and languages. Operating systems for portable devices (other than laptops) are yet to reach maturity. Palm-OS, Windows-CE, EPOCH, and GeOS are the most significant operating systems developed for mobile computing. A version of Linux for hand-held devices is also being developed.

These operating systems are light weight with simplified, single-address space memory management. Application portability across these operating systems is currently a major problem. The use of Java is currently limited due to the in- adequate performance of JVM on most of these platforms. Development of mo- bile applications on these platforms is typically done through platform-specific SDKs supplied by the operating system vendors. Windows-CE development can also be done using Microsoft Visual C++.

4. Write in detail about wireless telecommunication networks (6)

Today, person to person voice communications, enabled by the telephone, is still perhaps the most powerful technology available to the average person. The benefit to cost ratio of this technology for the individual is enormous. An individual can use a telephone to conduct commerce, earn a paycheck in count- less ways, call for medical assistance, consult experts worldwide on any topic, and essentially obtain almost any critical information imaginable. The most sophisticated part of this technology is not in the telephone handset itself but in the enormous worldwide communications network to which the handset is attached.

The introduction of cellular telephones has certainly improved the individuals ability to access (or be accessed by) this voice network in any location. But the global network is now providing more than person to person voice communications. Data, images, and live video are now routinely transferred to the individual desktop computer. It is expected that these expanding capabilities will soon be available within some type of portable information appliance.

There are several well-established cellular infrastructures available today in different parts of the world. The European community has standardized largely on GSM. North America has broad AMPS coverage with a number of other standards competing in the PCS frequencies. Japan deployed the PHS infrastructure everywhere. A brief comparison of these predominant standards is shown in Table 2.1

Table 2.1 PHS, AMPS and GSM wireless technologies

	PHS	AMPS	GSM
Usage area	Cordless (in-home) Mobile/ In-building (Japan)	Mobile/ In-building (N. America)	Mobile/ In-building (Europe) (N. Africa/Asia)
Applicable travel speed	slow driving	driving speed	driving speed
Voice signal	Digital	Analog	Digital
Frequency band	1.9 GHz	900 MHz	900 MHz
Channel multiplex number	4	1	8
Radio wave coverage	Indoor: 50–100m Outdoor: 100–400m	1.5–10km	1.5–10Km
Terminal to terminal communication	Possible	Not possible	Not possible
Data communication	32 kbps (plan)	14 kbps	9.6 kbps
Standby time	around 200 hrs	up to 20 hrs	up to 40 hrs
Talk time	5 hrs	up to 150 min	up to 240 min
Terminal output	less than 10 mW	600 mW	800 mW
Modulation method	Shifted QPSK	FM	GMSK
Voice transmission speed	32 kbps, ADPCM	Analog	22.8 kbps

5.Explain in detail about D igital cellular systems (11) (UQ April'13)

Analog cellular systems such as North America's AMPS have the disadvantage that they are very expensive to expand and grow. Each mobile phone requires a dedicated channel to communicate in a cell site. The only way to expand in AMPS is to build additional cell sites which cost in the range of $500,000 to $1,000,000.

In 1988, the Cellular Telecommunications Industry Association (CTIA) com- missioned a subcommittee called Advanced Radio Technology to define alter- native technologies that allows the cost effective cellular expansion in the US. Proposed technologies focused on Multiple Access network technologies. The first digital system accepted by CTIA is the TDMA system, which stands for Time Division Multiple Access and which allows users to share the radio channel through time division. The second digital system accepted by CTIA is CDMA, which stands for Code Division Multiple Access, and which allows users to share the entire radio spectrum through different, uniquely assigned codes for transmission and reception. In the next subsections, we briefly describe the TDMA and CDMA cellular systems.

Time-Division Multiple Access (TDMA) APRIL/MAY2014

TDMA is a digital transmission technology that allows a number of users to ac-cess a single radio frequency channel without interference, by allocating unique time slots to each user within each channel. Currently, a single channel is di- vided into six time slots, with each signal using two slots. This provides a 3 to 1 gain in capacity of AMPS. In dispatch systems (e.g. Motorola iDEN), a dispatch signal uses one time slot, thus providing a 6 to 1 gain in capacity. D-AMPS, GSM, iDEN and several PCS systems currently use TDMA. The Telecommunications Industry Association (TIA) provided an early standard for TDMA over AMPS, known as IS-54, which required digitizing the voice signal, compressing it and transmitting it in regular series of bursts, interspersed with other users' conversations. Second generation standard for TDMA by TIA is the IS-136 which uses TDMA on the control channel. TDMA is expected to be called TIA / EIA-136 once it becomes an ANSI standard.

One problem with TDMA is the wasted bandwidth of unused slots. Time slots are allocated to specific users whether or not they are using the slots (talking or transmitting data). Hughes Systems Network has contributed an enhancement of TDMA known as Enhanced TDMA (ETDMA) that attempts to correct this problem. Instead of waiting to determine whether a subscriber is transmitting, ETDMA assigns subscribers dynamically based on whether a user has voice/data to transmit. A phone conversation with long pauses will, therefore, not cause a loss of bandwidth, and will increase the spectral efficiency of TDMA. Today, TDMA is becoming a very popular air interface. Over 8 million digital subscribers worldwide utilize the IS-54 and IS-136 today. In the US alone, three of the top four carriers are deploying TDMA IS-136.

Code-Division Multiple Access (CDMA) APRIL/MAY2014

In frequency and time division multiplex systems, several hundred channels are available within the spectrum allocation of a carrier service. One channel of one base station is used for each conversation. Upon handoff, the subscriber station is directed via messaging to discontinue use of the old channel and tune to the new one. Without reusing the frequency assigned in the spectrum, the total number of cells that can be deployed can not exceed the available number of channels. Frequency reuse is very essential to the design of cellular systems that are based on frequency division multiplex.

Frequency reuse utilizes the fact that the attenuation of electromagnetic fields tends to increase with distance. Therefore, to reuse the frequency without incurring significant interference, only non-adjacent cells are assigned the same frequencies. Ideally, cellular frequency reuse is achieved by imposing a hexagonal array of cells in a service area. A seven cells hexagonal array is shown in Figure 2.1. Seven frequency channels represented by different gray levels are used, one for each cell. The hexagonal array can be replicated and connected, providing a larger coverage area, without using any but the seven frequency channels. Systems that use frequency reuse includes AMPS in North America, NMT in Scandinavia, and TACS in the United Kingdom.

In reality, cell coverage areas are highly irregular, and do not compare to the ideal hexagons shown in Figure 2.1. And even if ideal hexagons are possible, the frequency division approach offers limited capacity. Take AMPS as an example. Each AMPS operator in North America is allocated 416 channels (30KHz each). In a seven-way reuse hexagon, each cell will be allocated 416/7 = 59 channels. In this example, the capacity of cellular systems can not grow beyond the bandwidth offered by 59 channels, which is 1.8MHz.

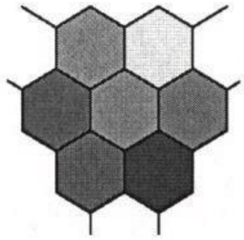


Figure 2.1 A hexagonal array of seven cells using seven different channels

Code-Division Multiple Access (CDMA) offers a solution to the capacity lim- itation problem. It allows all mobile stations to concurrently use the entire spectrum (all channels) with much less interference. Instead of partitioning ei- ther spectrum or time into disjoint "slots", each subscriber is assigned a unique instance of a pseudo-noise digital signal. The transmission signal is "spread" over the entire spectrum, using the noise signal. CDMA is, therefore, known as a spread spectrum modulation scheme. The spreading technique is also known as Direct Sequence scheme.

Frequency Hopping is another spreading technique, where the different segments of the subscriber conversation (or data) known as frames are transmitted on a sequence of randomly chosen frequencies within the spectrum. In either direct sequence or frequency hopping, the subscriber unit must communicate with the base station to agree on the direct sequence (the pseudo random digital code) or the sequence of frequencies to hop through. Signal interference in CDMA (between neighboring cells) is much less sensitive to most of the system parameters and is confined within a predictable average. This is one reason CDMA is attractive since it is easier to predict the achieved bandwidth based on the acceptable Noise to Signal Ratio (NSR) and the gain of signal spreading.

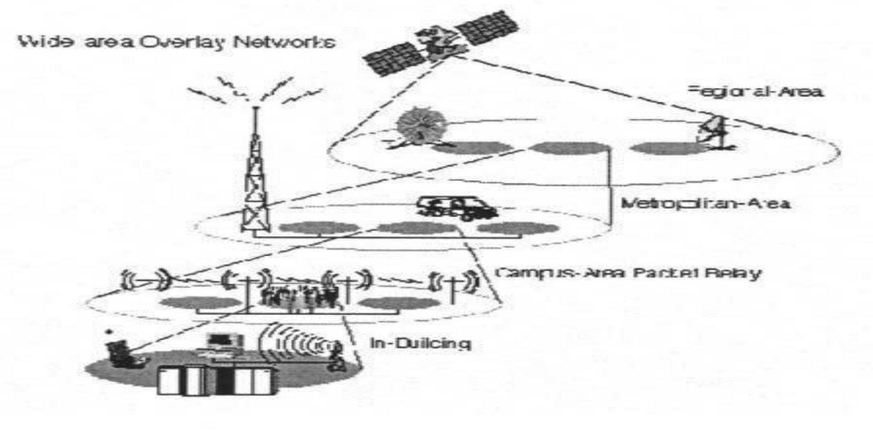


Figure 2.2 Wireless network overlay (Katz et. al)

Originally, CDMA was invented by Claude Shannon, who suggested that through noise-like carrier waves, bandwidth can be increased. Versions of CDMA has been in use for quite sometime by the military for the different reason of security. Transmitted signal is difficult to decode by an intercepting party due to the spreading and the unknown spreading noise signal. It is known by the military to be a Low Probability of Intercept (LPI) and Low Probability of Detection (LPD) air interface scheme. Since late 1980s, CDMA has been mi- grating into civilian applications and is now reaching maturity and impressive market penetration. Future wireless networks known as third and fourth generation wireless networks (based on where you are in the globe) are mostly based on CDMA .

6. Explain the wireless network technology (11)

Wireless technologies can be grouped into at least six major categories: (1) in- room, point to point infrared, (2) in-room radio, (3) in-building radio frequency, (4) campus or metropolitan area packet networks, (5) wide-area packet/circuit switched data networks, and (6) regional-area Satellite Data Networks. These six classes of networks have unique technologies which constrain the nature of the applications which can be supported by each of them. A similar taxonomy is provided in. Typically, an overlay of two or more network categories is used to provide continuous coverage in a mixed nomadic/mobile environment. Figure 2.2 shows an overlay of several network technologies. In the following subsections, we briefly summarize the characteristics and differences of these networks.

In-room Infrared

The in-room infrared class of networks generally has a network diameter of about 40–50m and supports bandwidths of about 1 Mbps. Applications supported by this type of infrastructure are limited to E-mail and collaborative- work applications due to the limited range of the system. The Infrared Data Association (IrDA) provides the most common standard used today for this network technology.

In-room Radio Frequency

The in-room radio frequency class of networks emerged in 1998 with the organized effort of the Bluetooth Special Interest Group. Bluetooth is a low-cost, short range radio that connects mobile PCs with other Bluetooth devices within a radius of about 10m. Very low energy consumption and about 1Mbps transmission speed makes this type of network attractive and suitable for inter-office device communication.

Hospital intensive care units, bank tellers, and desktop component inter connect may be example applications that could utilize in-room RF wireless technologies. The proliferation of portable devices such as 3COM's Palm Pilot, Windows-CE hand-held computers, and highly portable and powerful laptops such as the IBM Think Pads may incorporate Bluetooth transceivers to bridge the in-room wireless technology with fixed network infrastructures. The challenge laying ahead is to identify a suitable API for applications that will run a top this specific technology. Such API will allow for the design of ""infrastructure literate" applications that can accommodate the user expected performance levels while maintaining consistency across the infrastructure.

In-building Radio Frequency

This type of network, which is also known as Wireless LAN, expands the range of the infrared and the Bluetooth technologies by increasing the network diameter to about 200m. Unlike infrared and Bluetooth, in-building radio frequency is a cellular network, where mobile computers are allowed to roam within and across cells. Several standards are available today for this type of networks including the IEEE 802.11 and the Open-air interface.

Examples of Wireless LANs include Lucent/NCR WaveLAN and Proxim RangeLAN. Both ISA and PC Card interfaces are available with support or Windows and Linux. Proxim also provides additional support to a variety of Windows-CE devices. Wireless LANs can be used in both Infrastructure and Ad-Hoc Modes. In the former, Access Points are used and are connected to the fixed network through a dedicated router port. Wireless or nomadic devices with Wireless LAN interfaces access the network through the access point in the coverage area (cell). In this mode, the wireless LAN is used as a wireless extension of a fixed, high-speed network infrastructure (hence the name).

In the ad-hoc mode, several portable devices with wireless LAN interfaces are placed in the transmission range of each others. Each device is capable of communicating with any other device directly, without the help of any networking infrastructure. A private network is used to configure the network software (TCP/IP) among the ad-hoc group of devices. Ad-hoc networks is becoming increasingly important technology.

This technology, even though highly mature at this point in time, faces a few challenges. First, the IEEE 802.11 standard does not seem to be universally accepted (at least not yet). The OpenAir interface consortium, for instance, provides a competing proposal that is gaining popularity. Also, there is a lack of consensus on which air interface to use (direct sequence or the frequency hopping). Another challenge lies in the fact that wireless LANs are MAC-level networks that do not understand important features of IPv6 such as Multicast, RSVP, among other features. Unless, somehow, these features are implemented for wireless LANs, certain applications will be difficult to implement.

Campus/Metropolitan Area Packet Networks

This network type encompasses the more traditional “cellular” networking paradigm. It is typified by a “pole top infrastructure” supporting network diameters of 0.2 to 5 miles with data rates of 20-128 kbps. Relay (or router) nodes are strategically placed to support the wider network diameter with a small price for increased latency. For example, typical latency between a mo- bile device and the first relay node is about 40ms (assuming an uncongested network), and about 20ms between relay nodes.

Wide-Area Packet/Circuit Switched Data Networks

This network is comprised of a more familiar set of technologies and Regional Bell Operating Company (RBOC) services. One such offering is the Cellular Digital Packet Data (CDPD) service which is a packetized wireless transport that utilizes the unused channels of a cellular infrastructure. Motorola’s ARDIS and iDEN systems, Ericsson’s RAM (now called MobiTex), and the European GSM system are contained in this taxonomy. The iDEN network (Integrated Digital Enhanced Network) is a packet based voice/data network that uses the Mobile-IP networking protocol to route data packets.

Not only is this technology capable of supporting larger diameter networks, but they also tend to have lower bandwidths and higher latency effects than do the in-building networks. This tends to present a unique set of problems in application development. Significant body of research on network and system adaptation through infrastructure awareness components has been or is being conducted. However, the transformation of this research into commercially available "mobility middleware" is yet to occur.

Satellite Networks

Satellite technology is still emerging. It is a downlink technology where mobile computers can only receive direct broadcast from a satellite. Outbound communication is initiated by the mobile computer through a modem DIAL-UP or other wireless technology. Hughes Network Systems pioneered the DirecPC net- work which uses the Galaxy satellite and which delivers 400 kbps downlink rate. DirecPC also transmits continuous streams of multimedia information ranging from CNN broadcasts, to news, sports, and financial news feeds. Other Low Earth Orbit (LEO) systems are in planning and deployment phases including the Internet in the Sky project.

7. Explain mobility-bandwidth tradeoffs (5)

Another classification of the current wireless networking technology can be based on the "degree of mobility" offered by these networks. Multi-cellular wireless infrastructures range from in-building cells, to micro-cells (urban coverage), to macro-cells (suburban coverage), to satellite (global coverage). In- building cellular offers the highest bandwidth (bi-directional), but very limited mobility. Micro-cellular offers lower bandwidth but allows for limited-speed mobility; macro-cellular offers much lower bandwidth but allows for the high- est degrees of mobility.

As can be noticed, in these networks, the larger the coverage area (the cell size), the higher the degree of mobility. Satellite net- works are an exception and do not follow this trend. They offer the highest downlink bandwidth (no uplink possible with satellite networks), but they do not offer any mobility. Instead, they require a satellite dish to be stationed aiming at the satellite. Figure 2.3 shows a mapping of the mobility/bandwidth classification onto individual wireless networking technologies. In this mapping, mobility is further classified into indoor and outdoor, with outdoor mobility ranging from stationary, walking (pedestrian pace), and vehicular speed.

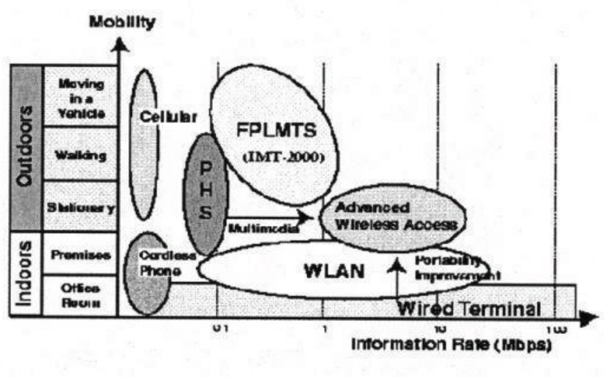


Figure 2.3 Mobility-bandwidth tradeoffs

The current mapping of wireless technology to the mobility/bandwidth classification is bound to change. At least this is ITU's and ETSI's vision and expectation of the third and fourth generation networks. For example, wireless LANs (an in-building technology) is expected to evolve into a network that allows for limited-speed mobility. Also, macro-cell networks are expected to improve on the bandwidth they offer. Figure 2.4 depicts this expected evolution.

8. Write about systems issues (6)

The rapid expansion of wireless Wide Area Network (WAN) services, wireless Local Area Networks (LANs), satellite services such as Hughes' Direct PC and the planned Low Earth Orbit (LEO) systems have created a large and fragmented wireless infrastructure. Given such a diverse set of technologies, the need to support mobile applications remains critical and even strategic to many industries.

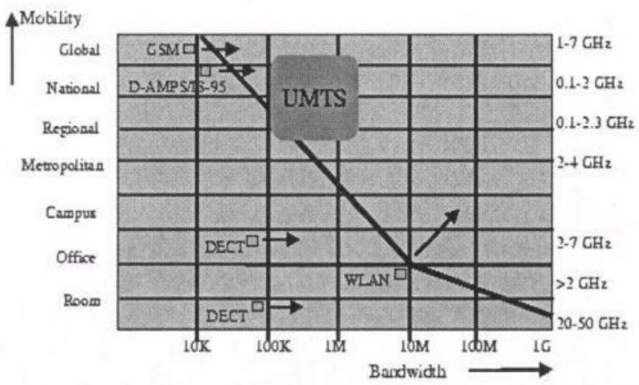


Figure 2.4 Expected mobility/bandwidth tradeoffs in 3G and 4G networks

Table 2.2 Application classes with examples

Application Class	Example Applications
Interactive Video	Video Conferencing, Distance Learning, etc.
Interactive Audio	Telephone, Digitized Voice over the Internet
Interactive Text/Data	Transaction Management, Credit Verification
Interactive Image	Teleconferencing, Collaborative Workgroups
Video Messaging	Multimedia E-mail
Audio Messaging	Voice Mail
Text/Data Messaging	E-mail, Telex, FAX
Image Messaging	High-Resolution FAX
Video Distribution	Television, VOD, PPV
Audio Distribution	Radio, Audio Feed, etc.
Text Distribution	News Feed, Netnews
Image Distribution	Weather Satellite Pictures
Video Retrieval	VOD
Audio Retrieval	Audio Library
Text/Data Retrieval	File Transfer
Image Retrieval	Library Browsing
Aggregate LAN	LAN Interconnection or Emulation
Remote Terminal	Tele–commuting, Telnet
Remote Procedure Call	Distributed Simulation

Table 2.3 Wideband CDMA Standard

Frequency band	2 GHz
Carrier bandwidth	5 MHz
Chip rate	5,115 Mcps
Frame length	10 ms
Voice	0.4-16 kbps
Video	128 kbps
Packet data	up to 128 kbps

The ability to scale performance and latency while accommodating an increasing user density is of paramount importance when designing and/or selecting a wireless infrastructure for a particular application. The choice of a wireless infrastructure must take into consideration the attributes of the application and the applications class of service requirements including bandwidth, network latency, service coverage, and general performance issues. Table 2.2 summarizes application classes as stringently defined by ITU-T Recommendation. These classifications have some loose definitions. For example, "interactive" usually means conversational, implying a person on either end of the application connection.

The term "messaging" generally refers to a person talking to a machine. An example would include leaving voice mail or sending a FAX. The term " retrieval" is generally thought of as a machine transferring information to a person. Also, the term "distribution" is typically thought of as a machine sending to people or machines who listen passively. The Client/Server architecture is a primary example of this application class. Application updates may include human intervention, but could be automated. The last five application classes listed in Table 2.2 are considered machine-to-machine interactions, al- though they may have to be "user" activated, while the actual transaction is between machines.

Multimedia Applications

As of today, there are limitations which prevent the effective exploitation of wireless networks by portable information appliances beyond the area of voice communications, text messaging, and limited data. While it is technically possible to transmit multimedia information such as a motion video clip from the internet into a portable wireless device, the standards, infrastructure bandwidth limitations, service costs, data compression technology, and power consumption considerations make this impractical at this time (1999). This limitations can be attributed to existing standards, which are limited in the level of service they can provide to the user of a portable information appliance. While they are effective for voice and text messaging, these standards do not support graphics intensive internet browsing or real time video at a high enough speed to make them practical.

In the case of video, the MPEG 1 standard provides for 352 X 240 pixel resolution, comparable to VCR quality video, and requires 1.14 Mbps data rate. This is well beyond any of the deployed wireless network standards. Today, consumer expectation is set by MPEG 2, which supports high resolution video of 1920 X 1080 pixels, which requires up to 80 Mbps of bandwidth (typical applications of this standard, however, may only require 6 to 8 Mbps).

To achieve wireless motion video data rates for portable devices, new wireless infrastructure standards will have to be deployed. One such standard is the Wideband CDMA approach proposed by Ericsson has the specifications shown in Table 2.3.

This standard has been adopted by the European community for the next generation of cellular service and could be implemented globally by 2002. The motion video quality enabled by such a service would, however, be less than MPEG 1 in terms of resolution and/or frame rate.

9.Write in detail about portable information appliances historical evolution (11)

The first portable information appliance was probably a piece of stone or clay with markings on it, used to record numeric information. This information was probably very important to the user of this appliance and in some way directly affected his livelihood. It may have actually provided the "function" of counting by allowing the user to create a mark corresponding to a piece of livestock. This would have been very useful to an individual whose society had not yet invented a system of numbers. Given the lifestyle of such an individual, ease of use, portability, durability, and reliability were all essential. Ease of use probably meant that the individual marks had to be deep enough in the appliance so as to be detectable by touching.

This would have been necessitated by the need to count goats in a heads-up mode while incrementing through the marks with the thumb. Once utilized, this appliance would have to be stowed in an extremely portable fashion so that it did not interfere with other activities such as attempting to frighten away predators, throwing sharpened sticks at predators and most importantly, running away from predators. Durability would have been important since the user did not have the means to protect the device from temperature variations, moisture, abrasion, and shock. To the user, this device may have played a very important role in establishing his credibility, accountability, and responsibility with respect to the rest of his community.

As the technology of mathematics and writing developed, human civilization progressed onward to the papyrus scroll (Figure 3.1) and ink pen. This appliance was highly portable and could convey very complex information. The user interface took a while to learn (reading and writing), and until relatively recently, only a limited number of individuals were able to use the technology. Still pen and paper persisted for several thousand years and is still the preferred portable information technology for most of the worlds population.

Two other portable information appliances, the pocket watch and the printed book (Figures 3.2 and 3.3) are relatively recent inventions which have transformed human society. The pocket watch enabled the level of logistical synchronization between individuals required for industrialization. Printed books, while not as interactive as paper and pencil, have also evolved as the preferred method for accessing standardized information in a portable format. Thus, pa-per and pencil, the printed book, and the pocket watch have been the dominant portable information appliances since the dawn of the industrial revolution.

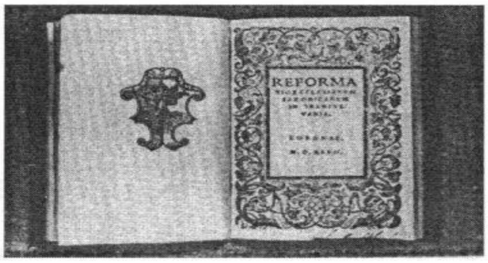


Figure 3.3 Printed book

The invention of the semiconductor technology in the Early 1960s began a transformation in portable information appliances, the full impact of which has yet to be realized. The first widely adopted electronic portable information appliance appeared in the early 1970s in the form of electronic calculators. Development work on these products began in the mid 60s and these designs exploited state of the art discrete transistor technology. By late 1960s, however, companies such as Texas Instruments, Rockwell and Intel had identified handheld calculators as a way to grow the market for Integrated Circuit technology. In 1970 there were several bulky hand-held calculators on the market at price points of around $300 and above.

By 1975, calculators had shrunk to pocket size and had fallen below the $20 price point. The age of portable electronic devices, enabled by the integrated circuit, was upon us. About this time, digital watches also began to replace mechanical watches which had been in place for hundreds of years.

By the early 1980s portable video camcorders had sold over 1 million units worldwide and penetration of portable electronics to the consumer had begun in earnest. This rapid penetration was driven by the compelling application of acquiring and storing motion video images. This trend was further accelerated by the introduction of 8mm format models which were highly miniaturized.

Personal organizers, such as the Sharp Wizard, were also introduced in this time frame and were most successful in Japan, where the use of personal computers was somewhat lagging that of North America. In North America, they were popular among technophiles but in general, these products tended to be a disappointment to individuals that had experienced desktop computing and found little compatibility between organizers and desktops.

Cellular phones have seen remarkable penetration worldwide . By the late1980s over 10 million units had been s old worldwide and the cell phone became a necessity for many and a status symbol for many others.

By the early 1990s, over one million Notebook computers had been sold world- wide as these products demonstrated their usefulness by turning spreadsheets and word processing into portable capabilities. Early models, in the late 1980s, from companies like Toshiba and Compaq, featured Monochromatic reflective LCDs. These systems were quite adequate for word processing and spread- sheets and were quickly adopted by traveling professionals. Transfer of data in and out of the notebook was achieved through magnetic disk. Prices remained relatively high ($2,000 +) due to two factors. First of all, manufacturers want to maintain high margins so the focus of the Notebook industry was on saturation of the business market, in effect, competing with desktop products. The second factor was the desire on the part of the user

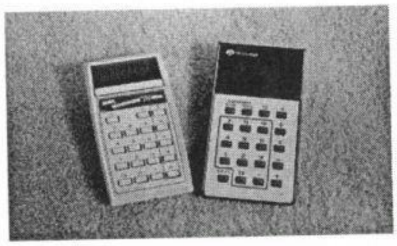


Figure 3.4 An early calculator

Figure 3.5 Portable video cameras (camcorders)

to have high performance which matched as nearly as possible that of a desktop system.

By the early to mid 1990s, several manufacturers were experimenting with the Personal Digital Assistant (PDA) product concept. These products attempted to span a gap between the personal organizer products and the notebook computer products. These products tended to compromise the miniaturization of organizers and lacked the full functionality of notebooks. Furthermore, they were typically crippled with an over sold and poorly performing handwriting recognition capability.

Most importantly, these early products tended to compete with, rather than complement the desktop or notebook computer. Several manufacturers attempted to add wireless communications to their PDA products to make them more appealing. Still, the lack of integration with the desktop PC and the bandwidth limitations of the wireless telecommunications infrastructure caused these products to fail. The telecommunications infrastructure in the mid 1990s offered only wire line and cellular modem capabilities with fairly low bandwidth (about 14.4 Kbs) for portable products.

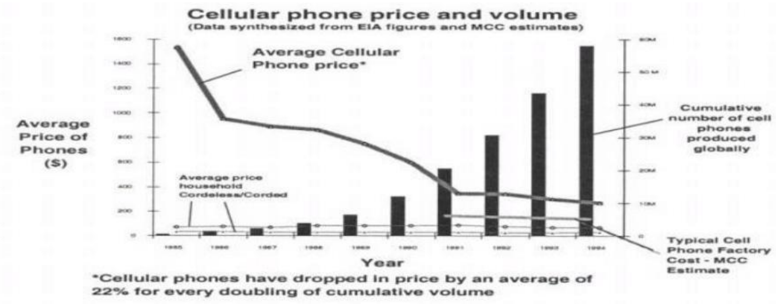


Figure 3.6 Cellular phones price and volume

10. Write about the advent of the PDA (11)

PDAs burst onto the scene in 1993 and mounted a headlong assault into the commercial market-place only to be quickly repulsed. When the initial exuberance subsided, the resulting carnage throughout the industry was both severe and widespread. Now, with forces re-marshaled and armed with a new generation of products, this same industry is attempting another assault, this time targeting the application specific vertical marketplace.

PDAs emerged in 1993 amid claims of single-point data organization, ubiquitous and instantaneous communications, and new operating paradigms using glitzy graphical user interfaces (GUI) and handwriting recognition. Most if not all of these claims fell short of consumer expectations. The reasons, while obvious in hindsight, lay hidden at the time. They were: high customer expectations, immature applications, and incompatible and unrealized infrastructures.

By 1993, the PC industry had introduced its most recent line of laptop computers which included computational and storage capacities that rivaled their most powerful desktop companions, even though computational and storage capacity had been doubling every 12 months in recent years. Grazing on these fertile fields had fattened the software industry and had bred a generation of software developers with inefficient development skills and tools. This in turn lead to unwieldy applications whose weaknesses were masked only by the raw computational and storage capacities of the hardware they ran on.

The result was that few wiry developers, and even fewer wiry applications existed that were capable of operating in the computational, power and storage barren environment of the PDA. Coincidentally, when the first PDAs appeared practically none were supported by third party software and embedded applications beyond the basic notepad, calendar, and calculator were virtually nonexistent.

Early on it was clear the success of the PDA rested heavily upon a variety of component and service infrastructures with the most critical of these enablers being wireless communications. In 1993, riding a sustained boom of 40% growth per year and giddy about recent cooperative initiatives, the cellular service providers boasted claims of a complete domestic wireless data infrastructure (CDPD) by the end of 1994. This effort, seemingly coordinated in its announcement, was enthusiastically received by the PDA industry but within a year the initiative would stall and lose much of its support. The breakdown came in the radio module that provides the link between the PDA and the wireless network. Initially predicted to be PCMCIA sized, it was soon realized that the requisite data radios would be both larger and more power hungry that anyone predicted. When they emerged, larger than some of the PDAs they were supposed to support, both industries recoiled under the letdown.

As if this was not enough, the whole industry was elevated to a high state of excitement, by a barrage of hype filled announcements, using phrases like *"Imagine if"* and *"Have you ever ... you will".* Every technology announcement from new processor architectures to handwriting recognition techniques added fuel to the flames. Claims like *"desktop performance in your palm", "time saving user interfaces", "ubiquitous communications", "transportable applications", "laptop functionality",* and more were touted loud and long. Market analysts and prognosticators joined in the frenzy, seemingly unable to separate future dreams from first article hardware.

As such, market expectations were set high, and high they stayed, as one product after another fell short and slammed into the reality wall. In fact, the ring of these claims still echoed in the ears of customers as they tried to use products that were expensive, bulky, fragile, unsupported, incompatible, uncooperative and unstable.

To make matters worse, costs were high and sales were low. The average price for a PDA in 1993 exceeded $750, some like the AT&T EO had prices that went as high as $2000– well outside the reach of many of the target customers. Consequently, in the first two years there were just 350,000 units sold. The volumes were so low in fact, that unlike most consumer electronics, they never crested the cost-experience wave which along with competition has the unrelenting ability to drive prices asymptotically toward the cost of the raw materials.

There were other problems as well with this initial surge of PDAs, but they served only to add to the mass confusion. The industry backlash, however, was both clear and severe. With hundreds of millions of dollars invested, two of the major players (AT&T EO, and IBM Simon) dropped out completely. The others fell back and re-grouped trying to understand what went wrong. What went wrong was equally as clear. Consumers were demanding usefulness and the first round of PDAs with limited applications and practically no communications simply did not fit the bill. Only a small percentage of the devices sold were ever really used. The vast majority were simply discarded amid the disappointment and frustration of the once excited user.

Today, the landscape has changed significantly. Unlike the excitement of the past, PDAs are now met with suspicion and skepticism. Regardless, a new battalion of products is moving into the fray. This time, however, there is an attempt to reduce the type, and in some cases manufacturers are trying to distance themselves from the past by avoiding the name PDA altogether, choosing instead names like pocket organizer and personal information manager (PIM).

Despite the reduction in hype, demands on and customer expectations of the PDA have continued to rise. One reason for this is that the laptop computer performance has continued to double every year. This, along with even higher resolution displays, improved ergonomics increased multimedia functions and more powerful applications has helped set a new standard for PDAs to meet. While PDAs have made progress in their operating systems, applications and third party support, they still fall woefully short of customer expectations.

Meanwhile, internet usage has erupted. Reliance on data stored in the myriad of html web sites, not to mention email services, has made wide area communications even more critical to the PDA paradigm than ever. Unfortunately, however, 18 months after ubiquitous wireless data services were promised by the cellular carriers, CDPD is in serious trouble. Southwestern Bell and Air touch have essentially stopped their CDPD deployment, leaving major holes like Los Angles, New Orleans and Atlanta in domestic coverage. This lack of clarity in the wireless infrastructure has caused confusion throughout the industry and continues to threaten the viability of the PDA.

This problem is compounded by the fact that the PDA manufacturers seem to rely on third parties to supply wireless modules for their products. This architectural approach results in a variety of inefficiencies and is due to a lack of expertise in wireless implementation. The integration of digital and RF circuitry at the semiconductor level will solve this problem in the future, but today, vendors that do not excel in both computing and communications design suffer a handicap.

Not surprisingly, a detailed look at the current offering of PDAs reveals that they are an outgrowth of PC concepts, utilizing the same worldwide components and manufacturing infrastructure that has been optimized to support desktop and laptop products. The silicon integration, displays, component size, software applications and substrate densities of this infrastructure has driven the PDA into one of 2 directions: either toward a fully functional product that is too large to be practical or toward a product that meets the ergonomics requirements of the paradigm, but that severely limits functionality and performance to fit.

The result is that while the average price of the new PDAs has dropped to $575.00, not much else has changed. Still starved for applications this new generation will not likely outsell its predecessors in the consumer marketplace, even though many industry projections say otherwise. Almost in recognition of this fact the strategists now say that the vertical market is the new focus of their attention but this is no panacea. The demands of the vertical market are many times more stringent than that of the consumer market and the procurement motivation is much less of an impulse. What is worse, the entrenched competitors like Symbol, Norand, and Telxon understand the operating environment and applications of the vertical market better, and will prove tenacious in their desire to maintain market share.

Prom the Early 1960s through the Mid 1990s, the advances in portable information appliances were impressive. Within the confines of a portable notebook, continuously increasing levels of computing power and display quality had been achieved. Ubiquitous, wireless voice communications via cellular phone had be- come common place. Consumers recorded hundreds of millions of hours of video data every year using hand-held camcorders. These products drove the development of important technologies. Silicon integration evolved from discrete transistor devices to single chips containing over 6 million transistors. Portable displays had evolved from simple numeric segment displays with less than 100 pixel elements into full color displays with over half a million pixels. Electronic and mechanical packaging technology was capable of connecting thousands of components in a compact volume compared to only a few tens of parts at the start of this period. Batteries in the early 60s could store no more than 100 watt-hours per litre. Their capacity today is up to 200 watt-hours per litre.

For all of these advances in hardware technology, however, many of these portable information appliances still seemed as static as the printed book. If they were connected to the outside world at all, it was through a low band- width wireless voice channel which was often unreliable for data transfer. Furthermore, the Internet appeared and created heightened expectations about information access. Without mobile access to the growing global information network, these portable devices would not live up to their

11. Explain the following in Detail. (UQ April'13)

(a)Explain palmtop computers (5)

(b) Hand-held computers (6)

It is likely that wireless network connectivity will trail wired connectivity in terms of performance for the foreseeable future. The best strategy for the developers of portable information appliance is to design products which either provide useful standalone functions such as an electronic still camera, or which complement wired network platforms. The emerging market of Palmtop Computers is a breakthrough in terms of the ability of the Palmtop to complement the desktop computer.

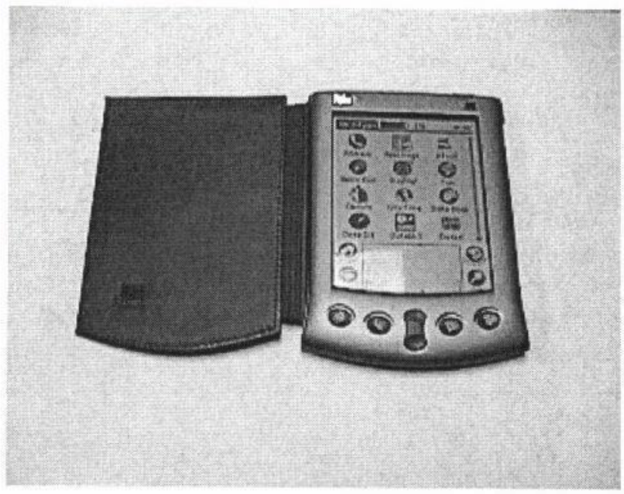


Figure 3.7 The Palm Pilot V

The Palm Pilot

The Pilot is a highly portable appliance which is the first truly viable substitute for traditional pencil and paper technology. With desktop synchronization, this device allows the desktop user to augment the networked desktop computing experience with a portable time management interface. While the Pilot is unlikely to provide services like high quality real-time video in the near future, this product concept has made important inroads into sensibly merging the interactions of portable and stationary information appliances.

Many other contemporary product designers have failed to take this approach by attempting to combine and therefore replace other devices. One example would be a smart phone that combines the functions of a cellular phone and a notebook computer. Such product concepts often end-up compromising the features which make the individual products appealing. For instance, may smart phones have poor display quality, unusable keypads, poor battery life, poor performance, and are much bulkier than most cellular phones. The result is a product that does not effectively replace either of the products that it is competing with.

Hand-held computers

The hand-held computer is another device that attempts to complement the desktop. It is much more capable than a Palm Computer, larger in size and weight, but can not be fitted in a pocket. Since their first emergence, hand-held computers have been competing with the Palm Computer market.

Table 3.1 The Palm Pilot V Specification

Item	Specifications
Size	4.7” x 3.2” x 0.4” (L x H x W)
Weight	4.0 oz. (including batteries)
Storage Capacity	2MB: 6000 addresses, 3000 appointments (approx. 5 years), 1500 to do items, 1500 memos, and 200 email messages.
Battery life	4-12 weeks (based on use) on 2 AAA batteries
Connectivity	RS-232C 9-Pin connector and 25-pin adapter; IR port; TCP/IP ready
Operating System	Palm OS
Applications	Date Book, Address Book, Mail, To Do List, Memo Pad, Expense, Calculator, Security, Games, HotSync, Others

Figure 3.8 Sharp Power Zaurus hand-held computer

Sharp Power Zaurus

Table 3.2 Sharp Power Zaurus specification

Item	Specifications
Processor	MIPS RISC Processor
Memory	16MB (ROM Upgradeable)
Display	6.5 High-Contrast Color LCD Touch Screen with Backlight (viewable area measured diagonally)
Colors	256
Resolution	640 x 240
Contrast control	Keyboard
Keyboard	64 Keys + 7 One Touch Application Keys
PC Card	one Type II slot
Audio	WAV file compatible with microphone, speaker, and external record button
Expansion Ports	Serial Port, PC Link, Printing
IR Port	IrDA 1.1 (115.2 kbps)compliant
Dimensions (w x d x h)	7.3 x 3.7 x 1.2 (186mm x 95mm x 29.6mm)
Weight	17.3 oz (490g)
Operating system	Windows CE

The Sharp Power Zaurus is a popular hand-held computer that competes with the Palm Computer market. The Zaurus which is depicted in Figure 3.8 is best described in terms of its specifications listed in Table 3.2.

Figure 3.9 The VADEM Clio tablet hand-held PC

Table 3.3 VADEM Clio specification

Item	Specifications
Processor	MIPS 4000
Storage	24MB ROM, 16MB RAM
Display	9.4" 640X480, 256 color
Operating System	Windows CE 2.1
Connectivity	IR port and built-in 33.6 kbps modem
I/O	keyboard, pen, and Type II PC card

VADEM Clio

Clio is a Windows CE based hand-held PC with a swing-top design that pro- vides three modes of interaction: keyboard, pen and tablet, and presentation modes. The three modes are achieved by swinging and/or folding the display around the keyboard base. The specifications of the Clio, which is shown in Figure 3.9 are listed in Table 3.3.

13. Explain communicators (11)

The Communicator is a PDA concept that combines the benefits, portability and functionality of digital cellular phones and palmtop computers. The idea is to stick a palmtop computer to a cell phone with data capabilities to provide remote access, in addition to the stand-alone form factor applications that can be found on palmtop computers. Internet access, telnet, email, and web browsing are all applications offered by communicators.

Table 3.4 The Nokia 9000 Specifications

Item	Specifications
Memory	8MB total: 4MB OS and applications, 2MB program execution, 2MB user data storage
Processor	embedded INTEL 386 processor
Operating System	GeOS TM3.0
E-mail protocols	SMTP, IMAP4, POP3 and MIME1
Weight	397g
Dimensions	173 x 64 x 38 mm
Displays	Grayscale 640x200 (illuminated) LCD

Nokia 9000

The Nokia 9000 is the most popular communicator, not only because of its appearance in the hands of Agent 007 in one of his recent movies (1997), but because of the unprecedented unique features and capabilities. The Nokia 9000 combined a compact personal organizer with Internet access and a versatile voice and text messaging system. The organizer includes: an address book, note editor, calendar with to-do list, calculator, and world clock.

A built-in browser, Telnet, and a VT100 Terminal emulation are built-in applications that bring the Internet to the mobile user anywhere GSM coverage is available. A multi-protocol email client, Short Message System (SMS) and a Fax application are also bundled to provide a wide spectrum of communication alternative, of course, in addition to the digital voice phone interface.

The specifications of the Nokia 9000 are listed in Table 3.4. Figure 3.10 depicts two pictures of the communicator. The picture to the right shows the communicator on a recharge base station and reveals the cell phone side of the device. The picture to the left shows an open communicator with a Web page on the backlit display.

Figure 3.10 Nokia 9000i Communicator

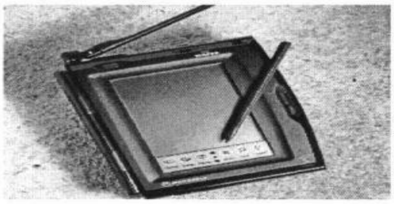


Figure 3.11 Motorola Marco hand-held computer

Motorola Marco

The Marco wireless communicator was introduced to the market one year be-fore the Nokia 9000 communicator (in 1995). It featured a built-in two-way wireless packet data modem allowing users to send and receive messages. The Marco Wireless Communicator, depicted in Figure 3.11, also included a fax and data modem, allowing information to be communicated through any telephone network. To augment its functionality, the Marco was equipped with two PCMCIA Type II slots to allow users to simultaneously operate third- party software applications and add memory to store more data. The Marco

Figure 3.12 Motorola Envoy hand-held computer

weighs 1.8 pounds and is 7.5 inches high, 5.8 inches wide, and 1.4 inches deep. The device features a bright portrait screen that allows easy reading in many lighting conditions.

At the time the Marco was introduced, Motorola had the vision of creating the first "wireless Newton". Newton OS 1.3 was therefore used. A similar product based on the Magic Cap operating system (from General Magic) was introduced in parallel. That was the Envoy depicted in Figure 3.12.

Unfortunately, the Apple Newton did not make it and despite all the software and personal information management tools loaded in the Marco, Motorola had only sold several thousand units before the device production was discontinued.

14.Write about sub-notebooks (micro-notebooks) (6)

As mobile users continue to demand lightweight, long battery life, and rugged portable computers, advances have been made in a number of diverse product concepts including what is now known as higher performance "micro- notebooks", or sub-notebooks.

Table 3.5 shows the specifications of the Sony PCG-707C sub-notebook that is depicted in Figure 3.13.

Table 3.5 Sony PCG-707C Sub-Notebook Specifications

Item	Specifications
CPU	233MHz Pentium Processor with MMX Technology and 256KB L2 cache
Memory	64MB EDO RAM standard
Hard drive	3.2GB, 2.5" (6.35 cm) disk
FDD	External with port replicator
Pointing device	Glide pad
PCMCIA card slot	Type II slot x 1, CardBus support
I/O ports	USB, VGA monitor, FDD, ASK and IrDA, PS/2, Modem (North America only), Mic-in, Audio
Dimensions	(W x D x H) 10.2" x 8.3" x 0.83"
Weight	1.40 kg (3.09 lbs)
Power supply	Li-ion battery pack (approx. 2.5 hours) With optional add-on battery (approx. 8.5 hours) Universal AC adaptor (100-240V AC, 50/60Hz)

Figure 3.14 HP Soujourn notebook

NOTEBOOKS

The notebook computer has enjoyed great success as the portable extension of the desktop computing environment. Notebooks are now starting to replace desktops for many users. Today the notebook market provides a most wanted portability by an increasing majority of users. We provide one example of notebooks which is the HP Soujourn. It weighs 3.2 pounds and is less than 0.71in thick. It uses an Intel Tillamook 233-MHz processor and comes with a 2.1 GB hard disk and a 64 MB of memory. Its display is limited though to only 12.1in SVGA. The HP Soujourn is shown in Figure 3.14.

15.Write about laptops (6)

Laptops are designed to replace the desktop. They can also be envisioned as nomadic desktops that can be easily moved from one place to another. The users of laptops require high performance, large high quality displays, and occasional portability. Such laptops may have maximum capabilities (as of 1999) such as up to 15.0in Color TFT (1024x768), integrated AC adapter, two battery support, up to 14GB disk storage, and 256MB memory. These capabilities come at the price of limited portability with these laptops weighing up to 8 Ibs.

Figure 3.15 Fujitsu Lifebook 900 laptop

Table 3.6 Fujitsu Lifebook 900

Item	Specifications
Model Number	985TX
Processor	Intel Pentium 233MHz with MMX Technology
System Memory	32 Megabyte SDRAM (included), Up to 160 Meg
System Cache	512k L2 Cache
Bus Architecture	PCI/CardBus
Display	13.3in LCD TFT Active Matrix Color
Video Memory	4 Megabytes SGRAM EDO
Maximum Resolution	1024 x 768/16M
Video	MPEG I/MPEG II
Zoomed Video	3D Graphics
TV Out	1280 x 1024/256 (External Monitor)
Hard Drive	5GB EIDE (Formated Size)
Floppy Drive	1.44 Meg Removable

It have only provide only one example of a laptop since almost all mobile users are familiar with laptops and their capabilities. The specifications of the Fujitsu Life book 900 laptop is summarized in Table 3.6. The laptop is depicted in Figure 3.15.

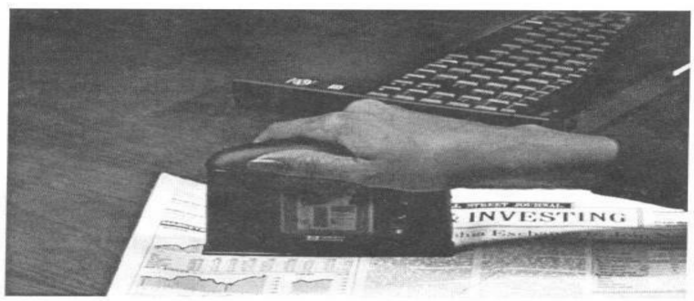


Figure 3.16 HP capture, store, communicate device

16.Write about other information appliances (5)

HP CapShare

HP's Capshare 910 is a hand-held portable device that allows mobile users to capture, store, communicate and print documents, The 5.5L x 4.1H x 1.5W (inches) device shown in Figure 3.16 weighs 12.5 oz and uses two AA NiMH rechargeable batteries that last for 100 document capture followed by a down- load. Typically, a mobile user capture documents from a newspaper or a magazine and then stores the document into his laptop or other portable device. Both PDF and TIFF formats are supported. The device has 4MB of memory and can capture from business cards and small receipts up to legal-size documents or 25in. newspaper columns. Maximum capture area of 119 square inches. A standard letter-size page takes about 6 seconds to capture.

Figure 3.17 First in-dash PDA

Clarion AutoPC

The Clarion AutoPC (depicted in Figure 3.17) is the first in-dash personal digital assistant. It integrates cellular telephony, Internet email, navigation software, GPS satellite tracking, contacts information and calendar, real-time information feeds (e.g. stock quotes and traffic information) in a single device. Hands-free interaction is possible through a speech recognition interface.

SRI VENKATESHWARAA COLLEGE OF ENGINEERING & TECHNOLOGY
(Approved by AICTE, New Delhi & Affiliated to Pondicherry University, Puducherry.)
13-A, Villupuram – Pondy Main road, Ariyur, Puducherry – 605 102.
Phone: 0413-2644426, Fax: 2644424 / Website: www.svcetpondy.com

Department of Computer Science and Engineering

Subject Name: **MOBILE COMPUTING**

Subject Code: **CS E84**

UNIT II

Emerging Wireless Network Standards: 3 G Wireless Networks – State of Industry – Mobility support Software – End User Client Application – Mobility Middleware –Middleware for Application Development - Adaptation and Agents - Service Discovery Middleware - Finding Needed Services - Interoperability and Standardization.

2 Marks

1. **Write some of the examples for emerging wireless networks.** Wearable Computing (MIT), Wearable Computer Systems (CMU), IBM Wearable PC BodyLAN: A Wearable RF Communications System

2. **Explain End-user client**

A flurry of activity appeared in the trade press in late 1995 describing the rush by vendors, both large and small, to market mobile client software packages. Some of those products are discussed in this section. Recent literature search suggests that many of these products never materialized, were re-targeted to wired networks, or in some cases, are still struggling with weak sales. However, there are some big players with deep enough pockets to continue to pursue this marketplace. The discussions here are restricted to those products and services that still appear to have a current or promised market presence.

3. What are the two key players in mobility middleware?

Two key players in the wired-network middleware market that provide sup-port for distributed are Novell's Netware and Microsoft's Remote Access. Neither of these products will be discussed further since neither has yet announced plans (that we have seen) for moving into the wireless middleware domain.

4. What are the limitations of mobility middleware?

Mobility-support software and products that are commercially available today leaves much to be desired in terms of functionality, performance, portability, and interoperability.

Translucent Overlays
Multi-Database Access
Multi-Database Access
Workflows
Location Dependent Services

5. What is meant by SDMA?

SDMA is a technology which enhances the quality and coverage of wireless communication systems. It uses a technique wherein the subscriber's access is via a narrow focused radio beam and the location of the subscriber is tracked adaptively by an intelligent antenna array system.

6. Define 3G network.

3G is the third generation of wireless technologies. It comes with enhancements over previous wireless technologies, like high-speed transmission, advanced multimedia access and global roaming. 3G is mostly used with mobile phones and handsets as a means to connect the phone to the Internet or other IP networks in order to make voice and video calls, to download and upload data and to surf the net.

7. How is 3G better?

3G has the following enhancements over 2.5G and previous networks:

- Several times higher data speed;
- Enhanced audio and video streaming;
- Video-conferencing support;
- Web and WAP browsing at higher speeds;
- IPTV (TV through the Internet) support.

8. Explain state of industry.

The suite of products that provide some support for mobile computing spans the technology space from end-user client applications, such as spreadsheets, Web browsers, through middleware, down to products implemented in hardware that provides cellular or other radio transmission-based communications services.

9. What is interoperability?

Interoperability is the ability of diverse systems and organizations to work together (inter-operate). The term is often used in a technical systems engineering sense, or alternatively in a broad sense, taking into account social, political, and organizational factors that impact system to system performance.

10. What is standardization?

Standardization or **standardization** is the process of developing and implementing standards. The goals of right standardization can be to help with independence of single suppliers (commoditization), compatibility, interoperability, safety, repeatability,
or quality.

11. What are the types of interoperability?

1) Syntactic interoperability
2) Semantic interoperability

12. What is the basic strategy of mobile ware office server?

The basic strategy that underlies Mobile Ware is to minimize mobile platform connect time by executing data transfers in a burst mode. The intent of this software is to make the mobile platform appear to the user as though it were actually a node connected into the wired network. The initial customer target focused on large sales staffs that were primarily mobile and who needed access on demand to sales support information that was too bulky and or volatile to carry on extended trips.

13. What is meant by WAP?

The wireless application protocol (WAP) standard currently being developed by the WAP forum group offers an OSI-like protocol stack for interoperability of different wireless networks. The WAP stack allows applications to register interest in quality of service events and thresholds (QoS). This, in turn, allows the application to be mobility-aware and adaptable to changes in the environment.

14. What is Shiva PPP (UQ: April'13)

Shiva's remote access client (known as PPP for Point-to-Point Protocol) enables mobile users to access servers embedded in either wireline or mobile servers almost seamlessly. For example, a client application that uses transaction processing services from BEA's Tuxedo can now access those services from a mobile platform using PPP. This software suite provides some limited security features such as limiting the number of login tries, or disconnecting a session and call- ing the user back at a pre-established number. However, it does not provide the rich collection of services available from Mobile Ware's Intelligent Transport Engine.

15. What is mobile ware office server?

Mobile Ware Office Server is an agent-based middleware for wireless or wire line access to application data. A service supported by Mobile Ware Office Server includes Lotus Notes, Web browsing, e-mail, and file transfer. A core component of the Mobile Ware Office Server is the Intelligent Transport Engine.

16. Explain Sybase SQL remote.

Unlike the Oracle Replication Manager, the Sybase product called SQL Remote has adopted a centralized model for managing replication. This product is a member of the Sybase SQL Any Where suite of tools (formerly called Watcom SQL).

17. What is Oracle replication manager?

Oracle has announced a version of its Replication Manager which will eventually support bi-directional replication among a collection of distributed and centralized server databases. The Oracle approach is based on a peer-to-peer model, much like Lotus Notes, in which a collection of distributed processes manage replication collectively.

18. What is Oracle lite?

This product is a cut-down version of the Oracle server that can run in a small portable system (or a desktop workstation). It can be used as a companion technology for the Oracle Agent Software to store local copies of subsets of corporate databases and can accumulate updates to the data that are generated locally at the mobile client.

19. What is UMTS (UQ: April'13)

UMTS, which stands for Universal Mobile Telecommunications System, is currently a project under the SMG (Special Mobile Group), a committee in ETSI. Decisions made in early 1998 by ETSI has given Europe a clear direction to- wards the realization of its third generation wireless communication system. ETSI agreed in January 1998 on two different UTRA methods: W-CDMA in the paired portion of the radio spectrum, and TD-CDMA in the unpaired portion.

11 Marks

1. **Write in detail about emerging wireless network standards (11)**

ITU (International Telecommunication Union) is a United Nation affiliated organization that oversees global telecommunication systems and standards. ETSI (European Telecommunications Standards Institute) is Europe's premier telecom standards organization well known for its development of the GSM standards. Both organizations are currently leading efforts to promote cooperation in the definition and development of future wireless networks. One goal common to both organizations is achieving seamless communication for the global consumer through cooperation on technical developments. Decisions made within these two organizations will have a dramatic effect on the future directions of wireless networks and services.

IMT-2000

The ITU (International Telecommunication Union), headquartered in Geneva, Switzerland, is an international organization within which governments and the private sector coordinate global telecom networks and services. IMT 2000 (International Mobile Telecommunications by the year 2000 project) is a project under the ITU that plans to facilitate cooperation in deciding global wireless access for the 21st century. Until recently, IMT 2000 was known as FPLMTS or Future Public Land Mobile Telephony System.

IMT 2000's vision is to "provide direction to the many related technological developments in the wireless industry to assist the convergence of these essentially competing wireless access technologies." IMT 2000 is expected to unify many different wireless systems, leading to the global offering of a wide range of portable services. It is expected that the IMT 2000 project will enable the merging of wireless services and Internet services, leading to the creation of a mobile multimedia technology and new modes of communication.

IMT 2000's vision of future wireless tele services is specified in terms of information rate, user delay sensitivity, and bit error rate requirements. Eight classes of services are specified as listed in Table 5.2.

Initial assignments of the IMT-2000 spectrum for Europe, US and Japan are shown in Figure 5.2. It shows how the dream of global roaming might be achieved in the future.

UMTS

UMTS, which stands for Universal Mobile Telecommunications System, is currently a project under the SMG (Special Mobile Group), a committee in ETSI.

Decisions made in early 1998 by ETSI has given Europe a clear direction to- wards the realization of its third generation wireless communication system. ETSI agreed in January 1998 on two different UTRA methods: W-CDMA in the paired portion of the radio spectrum, and TD-CDMA in the unpaired portion.

The main goals of the UMTS system can be summarized as follows:

Table 5.2 Third-generation requirements for wireless teleservices

Service Classification	Teleservices	Information Rate (kbps)	BER	Delay Sensitivity (ms)
Voice	Speech	8–32	10^{-3}	40
	Emergency Call	8–32	10^{-3}	40
	Teleconference	32–128	10^{-3}	40
Voice Band Audio	Facsimile	32–64	10^{-6}	100
	Telefax	64	10^{-6}	100
	Modems	32–64	10^{-6}	200
	Data Terminals	2.4–64	10^{-6}	200
Sound	Program Sound	128	10^{-6}	200
	High Quality Audio	940	10^{-5}	200
Video	Conferencing	384–768	10^{-7}	90
	Surveillance	64–768	10^{-7}	90
	Telephony	64-384	10^{-7}	40-90
Messaging	SMS & Paging	1.2–9.6	10^{-6}	100
	Voice Mail	8–32	10^{-4}	90
	Facsimile Mail	32–64	10^{-6}	90
	Video Mail	64	10^{-7}	90
	E-Mail	1.2–64	10^{-6}	100
Broadcast	Message	1.2–9.6	10^{-6}	100
	Multicast	1.2–9.6	10^{-6}	100
	SMS Cell	1.2–9.6	10^{-6}	100
	Public/Emergency Announcement (Voice)	8–32	10^{-4}	90
	Public/Emergency Announcement (Data)	1.2–9.6	10^{-6}	100
Data	Database Access	2.4–768	10^{-6}	200+
	Teleshopping	2.4–768	10^{-7}	90
	Newspapers	2.4–2000	10^{-6}	200
	GPS	64	10^{-6}	100
Teleaction	Remote Control	1.2–9.6	10^{-6}	100
	Remote Terminal	1.2–64	10^{-6}	100
	User Profile Editing	1.2–9.6	10^{-6}	200

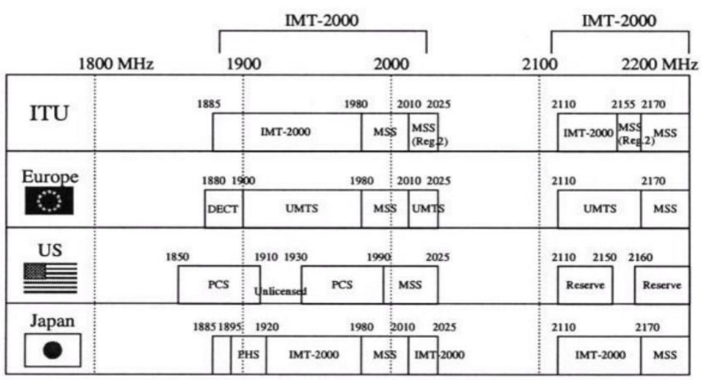


Figure 5.2 IMT-2000 spectrum assignment for Europe, US, and Japan

The accommodation of high speed, multimedia interfaces to support Internet applications at speeds of up to 2 Mbps, through a quantum leap in technology

- At least a 3-fold increase in spectral efficiency
- Support from an evolved GSM core network
- Compliance in meeting or exceeding ITU's Family Concept IMT 2000 System

The UMTS project schedule and milestones are shown in Figure 5.3. The first UMTS deployment is shown to be planned for the year 2002.

ACTS

Another organization that is greatly influencing the direction of wireless communications, particularly W-ATM, are the projects funded out of ACTS (the Advanced Communications Technologies and Services). ACTS is a group of European research projects with budget 50% funded by the European Economic Commission (EEC). The remaining 50% of the research funding is provided by those industry organizations involved in the research. ACTS broad objective is to develop advanced communications systems and services for economic

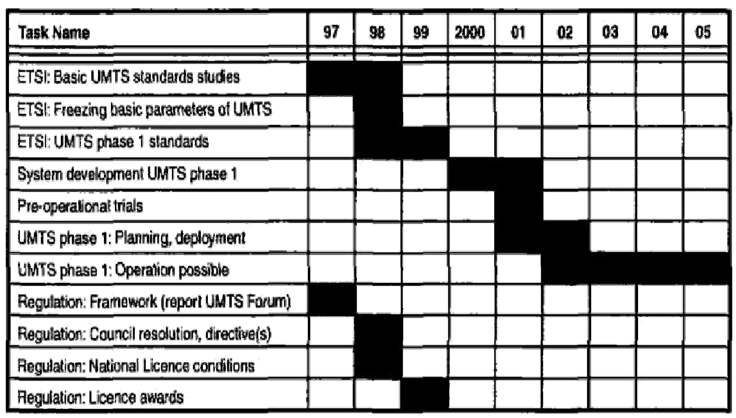


Figure 5.3 UMTS schedule and milestones

development and social cohesion in Europe. Research projects by ACTS include: multimedia, photonics, high speed networking, and mobile and portable communications.

2. Explain in detail about the Third generation wireless networks (11) (UQ April'13) APRIL/MAY2014

IMT 2000's original vision for third generation wireless networks was to create a single global communication system common to all countries and regions. This vision was too revolutionary to second generation wireless network providers who have invested heavily in current technology. To protect their investments, carriers requested ITU to consider a more evolutionary approach to third generation network standards. ITU, in turn, modified its vision into creating a "family of systems" that would converge and comply with a common set of requirements for third generation networks.

Following IMT 2000's vision, current research, development, and global standardization efforts are focused on upgrading second generation systems including GSM, CDMA, and TDMA. A major goal of this conversion is to upgrade these system evolutionary over time while maintaining the operation and profitability of the existing second generation network infrastructure.

TD-CDMA (Time Division, Code Division Multiple Access) and W-CDMA (Wideband Code Division Multiple Access) are the two major evolutionary network schemes currently under consideration by ITU. SDMA (Space Division Multiple Access) is also receiving attention as a network scheme with similar evolutionary nature.

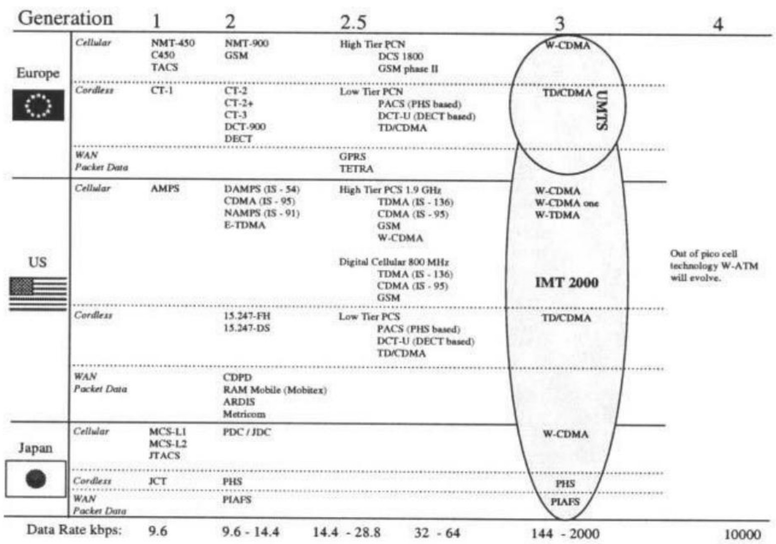


Figure 5.4 Evolution of wireless network technologies in Europe, US and Japan

Before summarizing the details of TD-CDMA, W-CDMA, and SDMA, we first describe the evolution of the wireless network technology in Japan, Europe, and the USA. Figure 5.4 helps clarify the alphabet soup in which the wireless industry swims. The figure defines the technology and generation in which a particular wireless system operates (or once operated). The data rates at the bottom of the chart apply only to the cellular technologies, not the cordless or WAN packet data rates. It is also important to note that at the time of publication, the third generation wireless technologies for the US had not been selected. Europe (presented in ETSI) just selected a combination of W-CDMA and TD-CDMA. Japan had chosen W-CDMA for their third generation wireless technology.

Time Division/Code Division Multiple Access

TD-CDMA is a proposed radio interface standard that uses CDMA signal spreading techniques to enhance the capacity offered by conventional TDMA system. Digitized voice and data would be transmitted on a 1.6 MHz wide channel using time-segmented TDMA technology. Each time slot of the TDMA channel would be individually coded using CDMA technology, thus supporting multiple users per time slot.

One design goal of TD-CDMA is to allow the CDMA technology to be smoothly integrated into the existing, second generation GSM TDMA structure world- wide. This will allow GSM operators to compete for wideband multimedia services while protecting their current and future investments. An important feature of TD-CDMA is its ability to adjust the ratio of spectrum allocated for the uplink and the downlink. The air interface can therefore be tuned to enhance the performance of certain applications such as Internet access and voice applications.

TD-CDMA uses the same frame structure as GSM 5.5. It has eight time slots with a burst duration of 577 micro seconds and a frame length of 4.616 m/sec as shown in Table 5.3. The TD-CDMA carrier bandwidth is eight times that of the 200 kHz GSM carrier equaling 1.6 MHz The compatibility between the TD-CDMA and GSM time bursts and frame structure permits the evolutionary step to third-generation systems. As many as 8 simultaneous CDMA codes are allowed in one time slot in TD- CDMA. This permits 8 users per time slot, or a larger combination of voice and data users to communicate without interference. For example, TD-

 CDMA can accommodate 11 voice users and 5 data users and still maintains the appropriate BER [83] (10^{-3} for voice and 10^{-6} for data). Eight users per time slot appears to be selected because it offers a happy medium between the number of voice and data calls that the system can accommodate.

Because TD-CDMA has the ability to assign multiple codes to one user, it permits broadband (or bandwidth on demand) transmission capabilities. Assuming a bandwidth of 1.6 MHz, a time slot with an information rate of 16 kbps using QPSK data modulation, eight possible users per time slot (eight CDMA codes per time slot) gives you an information rate of 128 kbps. If all eight time slots were allocated to a single subscriber in a pico cell environment or where mobility is restricted, 1024 kbps can be achieved. Changing to a different data

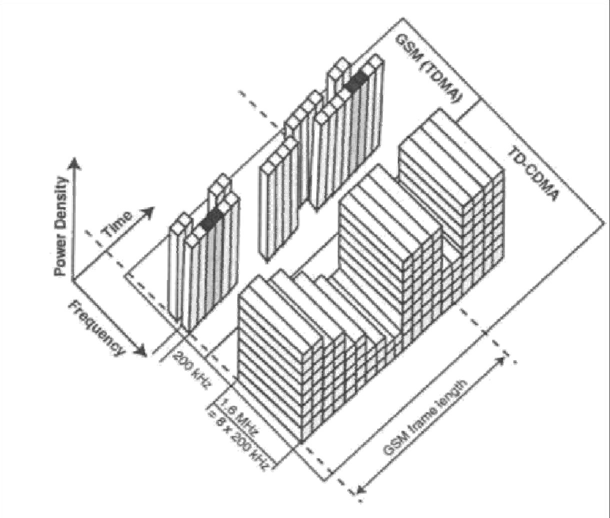


Figure 5.5 TD-CDMA vs. GSM(TDMA)

Table 5.3 TD-CDMA systems parameters

Bandwidth	1.6 or 3.2 MHz
Frequency band	2 GHz
Time slots per frame	8
Burst duration	577 micro sec
Frame length	4.616 msec
Full rate speech channels per carrier	64
Code slots per time slot	8
Chip rate	2.167 Mcps (million chips per sec)
Data modulation QPSK	(quadrature phase shift keying) or 16QAM (16 points quadrature amplitude modulation)
Spreading modulation	Linearized GMSK (gaussian minimum shift keying)

Modulation scheme like 16QAM instead of QPSK, an information rate of 2048 kbps is conceivable. Requirements for third-generation cellular systems are met by the TD-CDMA system.

Another advantage of TD-CDMA is the fact that intra-cell interference is orthogonal by time. This enables multiple subscriber signals to be received at differing power levels thereby eliminating the near-far effect and the need for a soft hand-off. The hand-off is conducted through a separate TD-CDMA or GSM carrier simplifying dual mode, dual band handsets. This is a divergence from GSM that conducts the voice and control channels on the same 200 kHz radio band. The sources studied on the TD-CDMA were not clear if mo- bile assisted hand-offs (MAHO), where the subscriber unit returns radio signal strength information back to the base station, is a feature in the TD-CDMA system.

The strategic importance of TD-CDMA can be summarized as follows:

1. The networks that TD-CDMA is catered towards are significantly deployed infrastructures, including:

- GSM: deployed in 74 countries, 200+ networks, and 20+ million subscribers and

- AMPS: deployed in 110 countries, 40+ million AMPS subscribers, 1.5+ million D-AMPS subscribers.

2. The cost of changing the air interface in a cellular system is significant.

System design and setup with regards to the MSC (mobile switching center), the BSC (base station controller), the cell location, and frequency reuse are all based upon the characteristics of the access scheme. Selecting a revolutionary different access scheme is therefore more than just changing the air interface; it is a costly operation. The TD-CDMA system is de- signed to be an evolutionary –not a revolutionary– step from GSM second generation infrastructure to third generation infrastructure.

The benefits to taking a revolutionary step include the absence of a legacy system and a quantum leap in abilities. The risk, however, is shortening the return on investment on second generation infrastructure. But regardless of the philosophical underpinnings, keeping deployed infrastructure profitable is a concept well-embedded in the telecommunication industry.

3. TD-CDMA promises to be future proof:

▪ Spectral efficiency twice that of GSM

▪ Reuse of existing GSM network structure and principles: cell sites, planning, hierarchical cell structures

▪ Efficient interworking with GSM

▪ Inherent TDD (time division duplex) support for cordless operation

▪ Data rate up to 2 Mbps indoor, 1 Mbit in all environments

▪ No soft hand-off and fast power control

TD-CDMA has recently been agreed upon by ETSI as a third-generation solution for GSM service providers. Support from most major telecommunications equipment providers in Europe has played a role in ETSI's decision to adapt TD-CDMA.

Wideband Code Division Multiple Access

W-CDMA is a spread spectrum technology in which the entire band- width is shared by multiple subscribers for transmission. A subscriber's data is modulated with PN codes; the signal is then spread and transmitted across a wideband. The receiver is responsible for dispreading the desired signal from the wideband transmission and contending with interference. The dispreading process at the receiver shrinks the spread signal back down to the original signal and at the same time decreases the power spectral density of the interference . This inherent ability to manage interference is at the heart of W-CDMA.

In W-CDMA, there are four control channels: the pilot, synchronization, paging and access channels. The channels are identified in the transmission by using a specific PN code, either a Walsh or Hadamard function (see [48] for further explanation). In the 5 MHz W-CDMA forward link, there are two designated codes for possible assignment to two possible pilot channels, two codes for two possible synchronization channels, a maximum of seven inclusive sequential codes for paging channels, and the remainder codes are not assigned and are used for forward traffic channels (see Table 5.4). The traffic channels are assigned *n* channel code numbers based upon desired the data rate, *n* = 0 … 64 for 64 kbps, *n* = 0…127 for 32 kbps, and *n* = 0… 255 for 16 kbps. The re- verse link channels are the access and the reverse traffic channels. Codes remain unassigned on the reverse channels so that channel assignment can be done dynamically and in response to paging channels and to interference. The forward link and the reverse link are FDD (frequency division duplexed). Frequency separation depends on the countries frequency allocation scheme. The channels in both the forward and reverse links are frequency division multiplexed.

Table 5.4 W-CDMA channel usage [48]

Bandwidth (MHz)	Pilot Channel Number	Sync Channel Number	Paging Channel Number
1.25	0	32	1-7 (sequential)
5.0	0 and 64	32 and/or 96	1-7 (sequential)
10.0	0 and 128	64 and/or 192	1-7 (sequential)
15.0	0	384	1-7 (sequential)

W-CDMA's channel responsibilities can be described as follows:

- Pilot channel (Forward link)

The BTS (base transceiver station) transmits one or two pilot channels carrying a reference clock necessary for demodulation and the hand-off process. The pilot channel also carries information used in estimating BTS signal strength therein indicating the best communication link for the subscriber terminal. After deciding on the best pilot signal, the subscriber terminal demodulates the synchronization channel.

- Synchronization channel (Forward link)

The synchronization channel contains system parameters, offset time, access parameters, channel lists and neighboring radio channel lists, all necessary in synchronizing with the paging, access and voice channels. The synchronization channel always operates at 1200 bps.

- Paging channel (Forward link)

System parameters and paging information to groups or a single subscriber are continually sent on the paging channel. Pages are combined into groups permitting a sleep mode to be built into the subscriber terminal, extending battery life. Subscriber terminals can monitor multiple paging channels. Thus, when another cell's paging channel has a better signal, a hand-off is requested. The paging channel has a data rate of 9600 bps or 4800 bps.

- Access channel (Reverse link)

When a page is detected the terminal attempts to access the system through the access channel. The terminal increases signal strength sent to the BTS until the system responds, a random time limit has expired, or maximum power levels have been exceeded.

- Traffic channel (Both forward and reverse links)

Within the traffic channel, there are two types of in-band signaling used. Blank-and-burst in-band signaling where an entire 20 m/sec frame is re- placed with control information. Dim-and-burst is also in-band signaling but the control information is distributed throughout a variable number of 20 m/sec frames.

The forward and reverse channels are modulated differently, QPSK for the forward channel, and O-QPSK for the reverse channel. There are five steps to the modulation process:

1. A PN multiplier multiplies the user data by the Walsh or Hadamard function, uniquely identifying that information to a specific subscriber terminal. The functions are time-shifted so that the set of functions are orthogonal

2. The output of the multiplier is a code rate of 4.096 Mcps using 5 MHz bandwidth (see Table 5.5) that is split into two signals: in phase (I) signal and quadrature (Q) signal

3. The pulse shapes for the I and Q signals are smoothed, minimizing rapid signal transition that results in radio frequency emissions outside the allocated bandwidth

4. The balanced modulator multiplies the I and Q signals by two signals that are 90 degree phase-shifted. The number of bits per chip depends on the data rate supplied to the balanced modulator5. The output of the balanced modulator is then fed to a RF (radio frequency) amplifier

W-CDMA's 5 MHz bandwidth provides robust frequency diversity. Selective frequency fading usually affects only a 200 - 300 kHz range of the signal. Time diversity happens because of multipath fading channels. W-CDMA solves the problem in a couple of different ways. One way is the selection of only the strongest signal, a process similar to antenna diversity. Rake reception is another technique where weak signals are added together to build a strong signal. Inherent time diversity receiver provides robustness against fading.

Table 5.5 W-CDMA system characteristics

Bandwidth (MHz)	PN Rate (Mcps)	Code Rate (Mcps)	Data Rate (kbps)	Channel Selection
1.25	1.228	1.228	1.2-9.6	Walsh 64
5.0	4.096	4.096	2-64	Walsh 64X4
10.0	8.192	8.192	2-64	Walsh 64X8
15.0	12.288	12.288	2-64	Hadamard 48X4 & 96X8

Table 5.6 Maximum number of channels in W-CDMA

Bandwidth (MHz)	Total Number of Channels			
	64 kbps	32 kbps	16 kbps	Total
1.25	16	32	64	64
5.0	64	128	256	256
10.0	128	256	512	512
15.0	192	384	768	768

Exact time alignment in W-CDMA is not necessary. Time offset or variable offset is why W-CDMA is considered to be quasi-orthogonal by time. This characteristic is used to decrease the interference by shifting the time alignment by 1.25 msec.

W-CDMA employs ADPCM (adaptive differential pulse code modulation) as a speech coding method at a coding rate of 32 kbps. Voice is sampled and digitized at 64 kbps then supplied to the speech coder that characterizes and compresses the data at a rate of 16 - 32 kbps, depending on speech activity. The acceptable BER and the speech encoding are variable, which improves the capacity in a W-CDMA environment. Data is sent when speech is detected; there is no reservation of time slots or frequencies, thus you are achieving maximum utilization of the available bandwidth. To achieve the same effect in a TDMA or FDMA environment, statistical multiplexing must be employed with speech detection. Statistical multiplexing requires frequencies and time slots to be reassigned, thus complicating the system and raising the cost.

Power control in a W-CDMA environment is an open and closed loops. An open loop is a coarse adjustment of the signal strength. This means the sub scriber terminal continually receives from the BTS Radio Frequency amplifier adjustments measuring signal strength loss and the terminal reacts accordingly. A closed loop is fine adjustment of the signal strength, meaning in every 1.25 msec time slot from the BTS there is a power control bit indicating to the sub- scriber unit to increase or decrease transmission power. The end result is the signal received at the BTS is always at approximately the same power level.

There are two fundamental types of W-CDMA systems–synchronous and asynchronous. In synchronous operations, all symbol/chip transmissions of all subscribers are orthogonal by time eliminating co-subscriber interference. This decreases interference and increases channel capacity, but increases system complexity. Asynchronous operation, on the other hand, permits co-subscriber interference and allows more flexibility in system design, but lowers channel capacity. Synchronization on a system level is coordinated through the use of the synchronization channel. Synchronization on the subscriber level is coordinated via the pilot channel reference clock and is used in demodulating the received signal.

Hand-offs in a W-CDMA system are soft hand-offs. They are initiated by the subscriber terminal finding a better paging channel from a different BTS. The subscriber terminal communicates with both BTSs while the MSC coordinates the simultaneous communication. The call is then handed off from one BTS to the other, completing what is known as a soft hand-off. The hand-off is referred to as soft because the terminal is always in communication with a BTS, a condition that results in fewer dropped calls.

Japan’s jump into W-CDMA is encouraged by a lack of capacity in the presently deployed system. The population density is such that a third-generation system is needed immediately. NTT predicts that wireless subscribers will equal wireline subscribers in the year 2000 at 60 million.

The Japanese W-CDMA system will be connected to an advanced broadband digital wireline network. The wire line connections are to be as follows: ATM adaptation layer 2 (AAL2) to be used between the BTS and the MSC via the BSC, PSTN and ISDN from MSC to the central office, and TCP/IP for Internet connections. The experimental prototype includes three cell sites, seven mobile stations, and up to 2 Mbps transmission rate. Other W-CDMA system parameters for the NTT DoCoMo/Ericsson testbed are outlined in Table 5.7.The following are the features in the NTT DoCoMo/Ericsson W-CDMA experimental system:

1. Subscriber unit can receive multiple channels resulting in multimedia bandwidth. The NTT DoCoMo W-CDMA system can accommodate up to six

64 kbps channels simultaneously for a total bandwidth of 384 kbps per subscriber, enabling six different tele services at the one time. This bit rate achieves the NTT DoCoMo phase one testbed goal of 384 kbps per sub- scriber; phase two has the goal of achieving 2 Mbps per subscriber. For more details about W-CDMA channel information see Tables 5.5 and 5.6.

2. The system allows for future expansion with the aid of adaptive antennas.

Adaptive antennas use SDMA techniques. As explained in section (5.3.3), SDMA manages interference and thus increases the network capacity, im- proves link quality, increases signal range, reduces transmission power, and extends the life and profitability of the deployed infrastructure.

3. New random access procedure with fast synchronization that provides flexibility in user data rates

4. Protocol structure that is similar to the GSM protocol structure

5. Inter-Frequency Hand-off (IFHO)

6. Hierarchical Cell Structure (HCS), permitting hand-offs between different wireless systems, (i.e. a hand-off between PHS infrastructure to the *W*- CDMA infrastructure)

7. VOX - Voice activation silence suppression, does not send data when the audio level is below a threshold. VOX is also noted in the PHS ARIB standard as a low power consumption operation for the private system.

8. Speech coding Orthogonal Variable Spreading Factor codes (OVSF). Utilization of a speech detection tool and orthogonal speech codes provides maximum bandwidth utilization in the W-CDMA environment. The speech detection tool, as explained earlier, assists in transmitting only the necessary data by transmitting less when speech activity is low. The orthogonal speech codes prevent interference with other channels decreasing interference and increasing capacity.

Table 5.7 NIT DoCoMo/Ericsson W-CDMA experimental system specs

Frequency band	
Carrier bandwidth	5 MHz
Carrier frequency forward link	2175 MHz
Carrier frequency reverse link	1990.5 MHz
Number of carriers	2 per sector
Number of sectors	6 per base transceiver station
Chip rate	4.096 Mchips per second
Frame length	10 ms
Time slots per frame	16
Services	
Voice	8 kbps voice operated relay
Packet data	2.4 - 384 kbps
Video, UDI data	64 - 384 kbps
Exchange Terminals	
1.5, 2 and 6.3 Mbps	G.703/G.704/I.4331
Protocols	AAL2 & AAL5
Modulation/Demodulation	
Data (Downlink)	QPSK
Data (Uplink)	O-QPSK
Spread (Downlink)	BPSK
Spread (Uplink)	O-QPSK
Demodulation	RAKE
Coding	
Short code	256 to 16 chip layered orthogonal code
Long Code (forward link)	10ms $2^{18} - 1$ chip gold code cut into 10 ms lengths
Long Code (reverse link)	2^{18} *10ms $2^{41} - 1$ chip gold code cut into 2^{18} *10ms lengths
FEC Coding	
Inner coding for traffic channels	Convolutional (R=1/3, K=9) Soft decision Viterbi decoding
Inner coding for control channels	Convolutional (R=1/2, K=9) Soft decision Viterbi decoding
Outer coding for UDI data	Reed Solomon coding RS(36,32)
BTS diversity	RAKE + antenna, 1 to 8 chip time window, 2 branch
BTS synchronization	Asynchronous
Power control	Closed loop + Open loop

Space Division Multiple Access

SDMA is a technology which enhances the quality and coverage of wireless communication systems. It uses a technique wherein the subscriber’s access is via a narrow focused radio beam and the location of the subscriber is tracked adaptively by an intelligent antenna array system (see Figure 5.6). The name

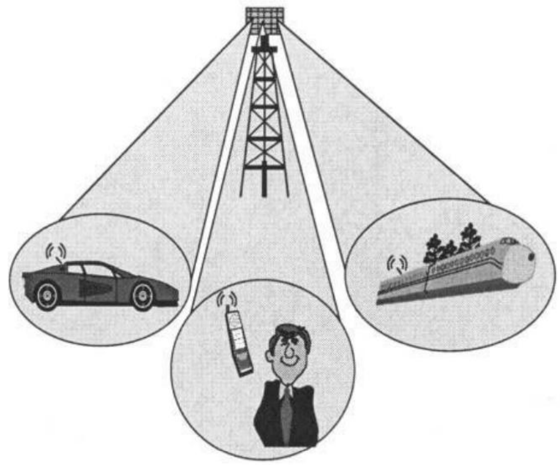


Figure 5.6 Space Division Multiple Access (SDMA)

SDMA is derived from the physical spatial characteristics between the focused radio beams. Spatial processing is not a new concept; it is used in presently deployed cellular infrastructures. For example, some cell sites are sectored at 120 degree. Also, most base station sites use two antennas for diversity reception regardless of whether they are sectored or not.

The most distinguishing aspect of SDMA is its management of interference. Reducing interference increases the effective network capacity, link quality and signal range. It also reduces the transmission power. Collectively, all the benefits brought by SDMA are expected to extend the life and profitability of second-generation network infrastructure.

SDMA is applied to the TDMA and CDMA systems differently because of the systems' basic differences. TDMA co-cell subscribers are orthogonal by time. Increasing the capacity in a TDMA environment by employing SDMA techniques requires multiple users on different radio beams to be assigned to the same carrier frequency and time slot. If the spatial component becomes insufficient between subscribers then an intra-sector hand-off is required to be initiated. The TDMA protocol needs to be expanded to permit these intra-sector handoffs.

CDMA subscribers use the same frequency and are quasi-orthogonal by time. Subscriber signals are distinguished by code filtering, not by time slots. Because of these CDMA characteristics, no intra-sector hand-offs are needed, keeping protocol overhead to a minimum.

Four types of interference concern cellular systems:

1. Background noise
2. External interference
3. Other cell interference, and
4. Other user noise

All systems deal with interference types 1 and 2. Interference types 3 and 4 are dealt with differently depending on the type of access method. In a TDMA system, interference types 3 and 4 are orthogonal either by frequency or time and do improve with frequency reuse planning. An interference signal from a neighboring cell base station is orthogonal by frequency to the desired signal. Also an interference signal from a co-subscriber within the same cell is orthogonal by time to the desired signal.

In a CDMA system, interference types 3 and 4 are spread across the same frequency and not necessarily orthogonal by time. In CDMA, all subscribers use the same frequency. The access method distinguishes between the desired signal and interference types 1 through 4 (which include co-subscriber interference) by using a Pseudo Noise (PN) code. The PN code is known by both the base station and the subscriber unit for spreading and dispreading the desired signal in the bandwidth. In a SD/CDMA environment, the spreading code acts like a direction estimator. The spreading code has the responsibility of locating the signal within the interference so the antenna array just has to establish an antenna beam in the direction of the user. In the TDMA environment, the antenna array has to distinguish between the interferer and the user whose signal structures are the same. The interferer signals have to be "nulled" before establishing a radio beam; the consequence of which might be the cancellation of all but one subscriber accessing that antenna array.

The employment of SDMA in a CDMA environment provides an easy increase in capacity. There are no additional protocols or controls that need to be implemented; only the deployment of an intelligent antenna array is required. Employment of SDMA in a TDMA environment, however, requires new frequency planning and alteration of protocols.

3. Explain fourth generation wireless research (6)

Beyond third generation wireless infrastructure is W-ATM (Wireless Asynchronous Transfer Mode). Wireless infrastructure is greatly influenced by the wire line infrastructure. Future directions in wire line infrastructure are towards ATM, and the wireless infrastructure will align itself appropriately. W-ATM is being researched because of the possibilities of providing high speed data transmissions with a low BER and high QoS in densely populated areas.

ATM is a protocol designed to accommodate multiple network services. The intent is to have one type of network for all types of data thereby increasing efficiency, services, and throughput, while decreasing costs and network complexity. ATM is an end-to-end communication system accommodating network services requirements for lossy or lossless data, bursty traffic with real-time requirements, or data with no time requirements.

W-ATM is a communication system for hybrid tethered/tetherless environments. It has challenges and obstacles at layers one, two, and three of the OSI (Open Systems Interconnection) model. Comments in this section will be focused on layer one issues and solutions. Many W-ATM layer one issues are mainly comprised of characteristics found in a mobile radio environment. This includes:

- Fading
- Multi-path propagation
- Signal attenuation, and
- Interference types, including inter-cell subscriber, intra-cell subscriber, background noise, and external or other noise.

Some significant research toward resolving W-ATM layer one issues is in the area of diversity reception like antenna arrays and SDMA. Diversity reception techniques solve some of the issues that W-ATM is facing, like fading and multi- path propagation. Multiple received signals caused by multipath propagation, received at antennas spaced at a distance of a fraction of the wavelength, allows the received signals to be treated as independent rays. Statistically, one of the received signals at a given point will not have faded, then by using diversity combining, the strongest portion of the two independent signals are used to create a third signal, partially eliminating the effects of multipath propagation and fading. Diversity combining also allows the mobile terminal to reduce the transmit power something battery researchers like to hear.

Higher frequencies and wide bands are capable of delivering the needed through- put rates expected of an ATM network. Interference and other layer one issues at higher frequencies are minimized by decreasing the transmission distance between the mobile terminal and the base station, thereby improving BER and QoS. Given the bandwidth, BER, and QoS requirements of an ATM network, W-ATM will be deployed first in the private and public pico-cell infrastructure or in highly populated areas, i.e. indoor wireless LANs or systems like Japan's PHS infrastructure.

Japan has the advantages of already having a publicly-deployed pico cell network and the population density to make a pico cell network financially possible. The PHS network is deployed on the island of Japan, where 125 million people are packed into an area slightly smaller than the size of California. With a publicly-deployed pico-cell network, the evolution to W-ATM will be swift and decisive. In the year 2002, Japan will be fulfilling the mobile multimedia dream with 10 Mbps wireless ATM links.

4. Explain in detail about state of industry and Mobility support software.(6)

The mobile client-application architectures which are emerging in commercial products can be roughly divided into three overlapping classes: Remote-Node,Client Proxy, and Replication. The basis for this subdivision is the need to address the problems associated with wireless bandwidth and battery limitations and the alternatives, that are commercially available today, for managing those problems. This classification is, therefore, different from the "research" classification. A brief description of each class follows.

- *Remote-Node* This approach attempts to create a facsimile of a fixed network client node by hiding all artifacts introduced by wireless communications. Under this model, all client software which run on a wired network platform would function without change on a mobile platform that includes a compatible OS and other library services. Accordingly, it places the most stringent demands on the middleware and other software (which supports the client application) to mediate the problems that arise as wire- less artifacts. As a result, this approach is most susceptible to failures in the wireless infrastructure. Software packages which adopt this approach may recognize some of the wireless limitations and adapt their behavior accordingly. For example, when response time is of concern, the limited bandwidth of wireless communications encourages the system to deliver records one at a time as they are retrieved from a database server rather than sending all record hits for some query. However, the ultimate goal is to provide an opaque overlay for the underlying ensemble of networks that shields the user from any concern for their interoperability. Remote-node applications can be realized by porting full clients (as used in the wireline network) to a mobile computer with compatible communication middle- ware. Shiva PPP is a famous middleware that supports most TCP/IP clients.

- *Client Proxy* This approach, characterized by products like Oracle Mo- bile Agents, attempts to minimize transmission costs and the impact of disconnects by buffering a client's requests, and/or the servers responses, and by resorting to batch transmissions. In this way, a user may select a variety of record types from several different tables, and then save battery power by disconnecting while the server processes the request. At some later time, the client can reconnect and receive a batch of records that satisfies all of the requests. The underlying assumption is that the end-user recognizes that periods of disconnect will occur, and that these periods will not impact the user's ability to perform useful work.

- *Replication* Clients which will be disconnected for extended periods of time, but which require immediate access to important data can satisfy those requests from locally cached replicas of key subsets of the database which are stored at some server site. Changes to the data that occur either at the client or the server must be reconciled through periodic client connects which may be initiated manually by the user, or automatically by the replication software. Some update conflicts may occur when multiple disconnected clients alter the same records. These collisions must be reconciled in some way.

5. Explain the End-user Applications (11) (UQ April'13)

A flurry of activity appeared in the trade press in late 1995 describing the rush by vendors, both large and small, to market mobile client software packages. Some of those products are discussed in this section. Recent literature search suggests that many of these products never materialized, were re-targeted to wired networks, or in some cases, are still struggling with weak sales. However, there are some big players with deep enough pockets to continue to pursue this marketplace. The discussions here are restricted to those products and services that still appear to have a current or promised market presence.

Oracle Mobile Agents

This product is a buffering and communications package for wireless platforms. A software agent that runs on the mobile client platform intercepts requests made by the client to the Oracle server and buffers them for a later transmission to the server. A companion Oracle agent runs on the Oracle server platform. That agent receives the buffered requests, submits them to the Oracle server, and buffers the responses for later transmission to the client. The server agent is capable of serving any number of mobile agents simultaneously. Conversely, a client agent can access any server agent that it knows about and for which it holds the appropriate DBA access privileges. Oracle agents can run on mobile platforms equipped with NT, Unix, or Windows and can communicate over TCP/IP using Shiva's PPP communications middleware. This product does not automatically support transactions or queries that span multiple Oracle servers.

Oracle Lite

This product is a cut-down version of the Oracle server that can run in a small portable system (or a desktop workstation). It can be used as a companion technology for the Oracle Agent Software to store local copies of subsets of corporate databases and can accumulate updates to the data that are generated locally at the mobile client. Oracle may provide the Oracle Lite server with "two way" replication which could automatically propagate updates either to the client from the central server site, or vice versa. Recently, Oracle and Palm Computing (a 3Com company) announced an alliance to integrate the Oracle Lite client database and the 3Com Palm III and Palm Pilot organizers, allowing new and existing Palm Computing platform applications and data to be replicated, synchronized, and shared with an Oracle 8 database server.

Oracle Software Manager

This product is intended for a database administrator who needs to propagate software updates to remote copies of the Oracle server. It is capable of performing the distribution via hardwired networks or through wireless connections. It is not clear whether this package is versatile enough to accomplish distributed software update to a collection of mobile devices as though the entire operation were a distributed transaction. For example, if the DBA needs to update the mobile Oracle-Lite server software for entire sales staff, the updates may have to be performed individually by the DBA.

Oracle Replication Manager

Oracle has announced a version of its Replication Manager which will eventually support bi-directional replication among a collection of distributed and centralized server databases. The Oracle approach is based on a peer-to-peer model, much like Lotus Notes, in which a collection of distributed processes manage replication collectively.

Sybase SQL Remote

Unlike the Oracle Replication Manager, the Sybase product called SQL Remote has adopted a centralized model for managing replication. This product is a member of the Sybase SQL Any Where suite of tools (formerly called Watcom SQL). Also, Sybase has optimized its replication server to accommodate users that are only occasionally connected. So while this product has been developed with wired network users as a primary target, the software does include a component that recognizes the frequent disconnects that typify mobile users.

6. Explain mobility middleware (6) APRIL/MAY2014

The majority of products targeted for the middleware market rely on TCP/IP and socket-like connections for the client server interface whether they are in- tended to be deployed in the wireline network arena or the wireless domain. Variants of TCP have been proposed to circumvent the problems that plague TCP for some wireless applications. By choosing to adopt this defacto standard transport protocol, vendors are positioning their products for deployment in a large existing infrastructure. As a result, it is already possible to surf the Internet using a Netscape interface on many wireless platforms and a simple cellular phone connection.

Two key players in the wired-network middleware market that provide sup- port for distributed users are Novell's Netware and Microsoft's Remote Access. Neither of these products will be discussed further since neither has yet announced plans (that we have seen) for moving into the wireless middleware domain. However, Microsoft Exchange has been integrated with Shiva's PPP software that allows communication of clients to servers through the cellular phone network.

MobileWare Office Server

This suite of products was introduced in 1995 as a solution to managing mobile access to corporate data. The basic strategy that underlies MobileWare is to minimize mobile platform connect time by executing data transfers in a burst mode. The intent of this software is to make the mobile platform appear to the user as though it were actually a node connected into the wired network. The initial customer target focused on large sales staffs that were primarily mobile and who needed access on demand to sales support information that was too bulky and/or volatile to carry on extended trips.

The current flagship product, MobileWare Office Server, includes a native Lotus Notes mail and database replication support. MobileWare Office Server is an agent-based middleware for wireless or wireline access to application data services supported by Mobile Ware Office Server includes Lotus Notes, Web browsing, e-mail, and file transfer. A core component of the MobileWare Office Server is the Intelligent Transport Engine. The transport engine provides several features including:

- Connection Profiles: The user chooses from a collection of profiles based on current working environment (LAN, Dial-up, or wireless connections, and TCP/IP, NetBios, etc.). Each profile contains a set of tuned parameters that optimize the communication between the clients and the servers.

- Data check pointing: This ensures efficient recovery and fast reconnection after failures and involuntary disconnections.

- Automatic reconnection in response to involuntary lost connections Data compression.

-

- Dynamic Packet-Scaling: Based on the current connection quality and capacity, data packets are dynamically re-sized to minimize connection time.

- Encryption and authentication. Uses DES encryption for per-connection authentication.

- Security. Forces re-authentication from the client upon receipt of any unregistered packet.

- Queuing. Application data is stored on both the client and the server in a client assigned outbox until a connection is made to transfer the data.

- Follow-Me Server. Uses a notification and delivery mechanism for events such as arrival of data to the client's outboxing on the server. The user's mobile computer is notified if connected, and alternative notification procedures (such as paging) are allowed. MobileWare Corporation was founded in 1991, and is now a private subsidiary of Itochu Japan.

Shiva PPP

Shiva's remote access client (known as PPP for Point-to-Point Protocol) enables mobile users to access servers embedded in either wireline or mobile servers almost seamlessly. For example, a client application that uses transaction processing services from BEA's Tuxedo can now access those services from a mobile platform using PPP. This software suite provides some limited security features such as limiting the number of login tries, or disconnecting a session and calling the user back at a pre-established number. However, it does not provide the rich collection of services available from Mobile Ware's Intelligent Transport Engine described above

7. Explain Adaptation (11)

Mobile computers must execute user- and system-level applications subject to a variety of resource constraints that generally can be ignored in modern desktop environments. The most important of these constraints are power, volatile and nonvolatile memory, and network band-width, although other physical limitations such as screen resolution are also important. In order to provide users with a reasonable computing environment, which approaches the best that currently available resources will allow, applications and/or system software must adapt to limited or fluctuating resource levels. For example, given a sudden severe constraint on available bandwidth, a mobile audio application might stop delivering a high-bit-rate audio stream and substitute a lower-quality stream. The user is likely to object less to the lower- quality delivery than to the significant dropouts and stuttering if the application attempted to continue delivering the high-quality stream. Similarly, a video application might adjust dynamically to fluctuations in bandwidth, switching from high-quality, high-frame-rate color video to black-and-white video to color still images to black-and-white still images as appropriate. A third example is a mobile videogame application adjusting to decreased battery levels by modifying resolution or disabling three-dimensional (3D) features to conserve power.

The spectrum of adaptation

At one end of the spectrum, adaptation may be entirely the responsibility of the mobile computer's operating system (OS); that is, the software for handling adaptation essentially is tucked under the OS hood, invisible to applications. At the other end, adaptation may be entirely the responsibility of individual applications; that is, each application must address all the issues of detecting and dealing with varying resource levels. Between these extremes, a number of application-aware strategies are possible, where the OS and individual application each share some of the burden of adaptation. While applications are involved in adaptation decisions, the middleware and/or OS provides support for resource monitoring and other low-level adaptation functions. The spectrum of adaptation is depicted in Fig. 6.1. we are concerned primarily with middleware for adaptation, that is, software interfaces that allow applications to take part in the adaptation process. Pure system-level adaptation strategies, those which take place in a mobile-aware file system such as Coda (e.g., caching and hoarding), are covered elsewhere in this book.

Resource monitoring

All adaptation strategies must measure available resources so that adaptation policies can be carried out. For some types of resources— cash, for example—monitoring is not so difficult. The user simply sets limits and appropriate accounts. For others, more elaborate approaches are required. The Advanced Configuration and Power Interface (ACPI) provides developers with a standardized interface to power-level information on modern devices equipped with "smart" batteries. Accurately

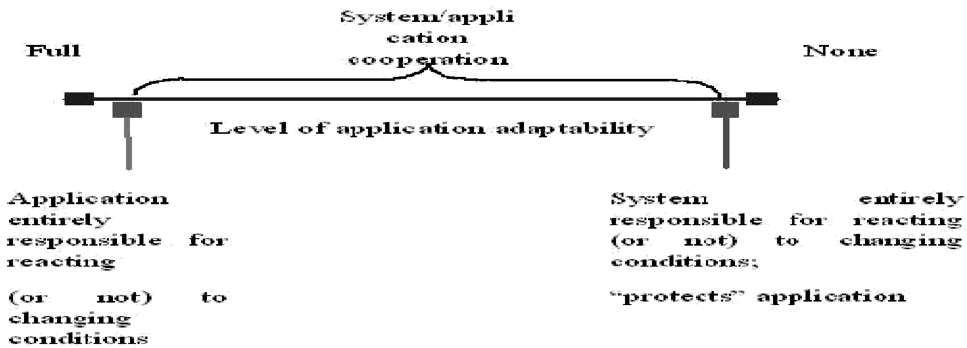


Figure 6.1 At one end of the spectrum of adaptation, applications are entirely responsible for reacting to changing resource levels. At the other end of the spectrum, the operating system reacts to changing resource levels without the interaction of individual applications. measuring network bandwidth over multi hop networks is more difficult. Some approaches are described in Lai and Baker (1999) for the interested reader. Whatever methods are used to measure resource levels have a direct impact on the effectiveness of the entire adaptation process because accurate measurement of resource levels is critical to making proper adaptation decisions.

Characterizing adaptation strategies

The Odyssey project (Noble et al., 1997; Noble, 2000) at Carnegie Mellon University was one of the first application-aware middleware systems, and it serves as a good model for understanding application-aware adaptation. In describing the Odyssey system, Satyanarayanan proposed several measures that are useful for classifying the goodness of an adaptation strategy. We describe these—fidelity, agility, and concurrency below.

Fidelity measures the degree to which a data item available to an application matches a reference copy. The reference copy for a data item is considered the exemplar, the ideal for that data item—essentially, the version of the data that a mobile computer would prefer given no resource constraints. Fidelity spans many dimensions, including perceived quality and consistency. For example, a server might store a 30- frame-per-second (fps), 24-bit color depth video at 1600 × 1200 resolution in its original form as shot by a digital video camera. This reference copy of the video is considered to have 100 percent fidelity. Owing to resource constraints such as limited network bandwidth, a mobile host may have to settle for a version of this video that is substantially reduced in quality (assigned a lower fidelity measure, perhaps 50 percent) or even for a sequence of individual black-and-white still frames (with a fidelity measure of 1 percent). If the video file on the server is replaced periodically with a newer version and a mobile host experiences complete disconnection, then an older, cached version of the video may be supplied to an application by adaptation middleware. Even if this cached version is of the same visual quality as the current, up-to-date copy, its fidelity may be considered lower because it is not the most recent copy (i.e., it is stale).

While some data-dependent dimensions of fidelity, such as the frame rate of a video or the recording quality of audio, are easily characterized, others, such as the extent to which a database table is out of date or a video is not the most current version available, do not map easily to a 0 to 100 percent fidelity scale. In cases where there is no obvious map- ping, a user's needs must be taken into account carefully when assigning fidelity levels. More problematic is the fact that fidelity levels are in general type-dependent—there are as many different types of fidelity-related adaptations as there are types of data streams; for example, image compression schemes are quite different from audio compression schemes. Generally, an adaptation strategy should provide the highest fidelity possible given current and projected resource levels. Current adaptation middleware tends to concentrate on the present. Factoring projected resource levels into the equation is an area for future research.

Agility measures an adaptation middleware's responsiveness to changes in resource levels. For example, a highly agile system will deter- mine quickly and accurately that network bandwidth has increased substantially or that a fresh battery has been inserted. An adaptation middleware's agility directly limits the range of fidelity levels that can be accommodated. This is best illustrated with several examples, which show the importance of both speed and accuracy. For example, if the middleware is very slow to respond to a large increase in network bandwidth over a moderate time frame (perhaps induced by a user resting in an area with 802.11 WLAN connectivity), then chances to perform opportunistic caching, where a large amount of data are transferred and hoarded in response to high bandwidth, may be lost. Similarly, an adaptation middleware should notice that power levels have dropped substantially before critical levels are reached. Otherwise, a user enjoying a high-quality (and power-expensive) audio stream may be left with nothing, rather than a lower-quality audio stream that is sustainable

Agility, however, is not simply a measure of the speed with which resource levels are measured; accuracy is also extremely important. For example, consider an 802.11a wireless network, which is much more sensitive to line-of-sight issues than 802.11b or 802.11g networks. A momentary upward spike in available bandwidth, caused by a mobile host connected to an 802.11a network momentarily having perfect line of sight with an access point, should not necessarily result in adjustments to fidelity level. If such highly transient bandwidth increases result in a substantial increase in fidelity level of a streaming video, for example, many frames may be dropped when bandwidth suddenly returns to a lower level.

The last measure for adaptation middleware that we will discuss is con- currency. Although the last generation of PDAs (such as the original Pilot by Palm, Inc.) used single-threaded operating systems, capable of executing only one application at a time; newer PDAs, running newer versions of Palm OS, variants of Microsoft Windows, and Linux, run full- featured multitasking OSs. Thus it is reasonable to expect that even the least powerful of mobile devices, not to mention laptops that run desk- top operating systems, will execute many concurrent applications, all of which compete for limited resources such as power and network band- width. This expectation has a very important implication for adaptation: Handling adaptation at the left end of the spectrum (as depicted in Fig. 6.1), where individual applications assume full responsibility for adapting to resource levels, is probably not a good idea. To make intelligent decisions, each application would need to monitor available resources, be aware of the resource requirements of all other applications, and know about the adaptation decisions being made by the other applications. Thus some system-level support for resource monitoring, where the OS can maintain the "big picture" about available resources needs and resource levels, is important.

An application-aware adaptation architecture: Odyssey

The Odyssey architecture in greater detail. In the spectrum of adaptation, Odyssey sits in the middle—applications are assisted by the Odyssey middleware in making decisions concerning fidelity levels. Odyssey provides a good model for understanding the issues in application-aware adaptation because the high-level architecture is clean, and the components for supporting adaptation are clearly delineated. The Odyssey architecture consists of several high-level components: the interceptor, which resides in the OS kernel, the viceroy, and one or more wardens. These are depicted in Fig. 6.2. The version of Odyssey described in Nobel and colleagues (1997) runs under NetBSD; more recent versions also support Linux and FreeBSD. To minimize changes to the OS kernel, Odyssey is implemented using the Virtual File System (VFS) interface, which is described in great detail for kernel hacker types in Bovet and Cesati (2002). Applications interact with Odyssey

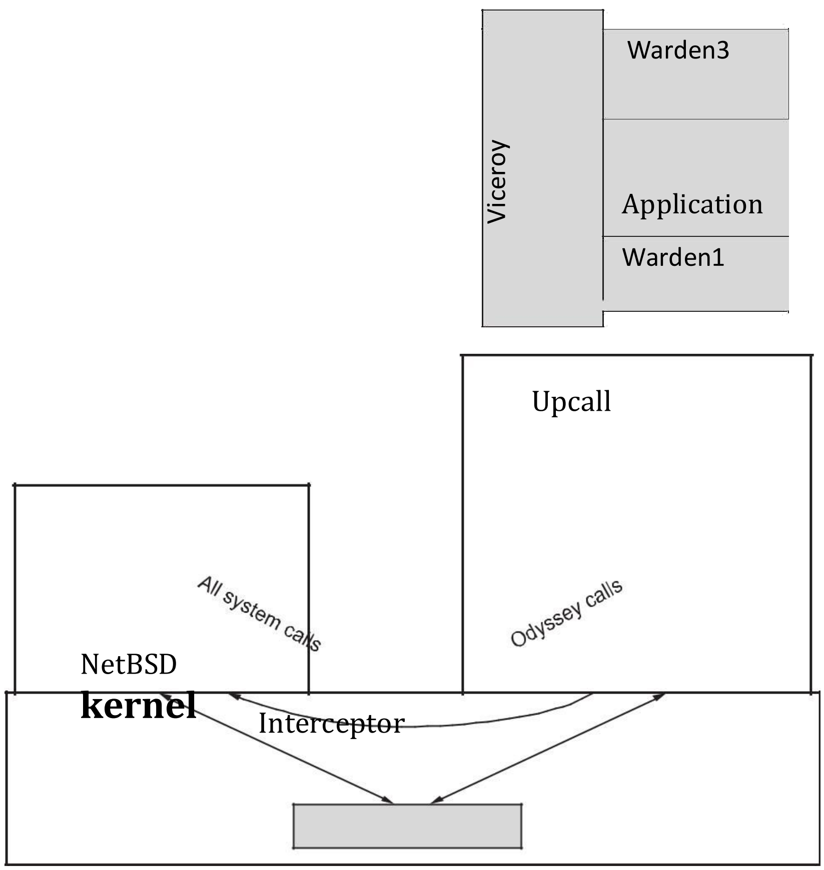


Figure 6.2 The Odyssey architecture consists of a type-independent viceroy and a number of type-specific wardens. Applications register windows of acceptable resource levels for particular types of data streams and receive notifications is when current resource levels fall outside the windows. Odyssey using (mostly) file system calls, and the interceptor, which resides in the kernel, performs redirection of Odyssey-specific system calls to the other Odyssey components.

The basic Odyssey model is for an application to choose a fidelity level for each data type that will be delivered—e.g., 320 × 240 color video at 15 fps. The application then computes resource needs for delivery of each stream and registers these needs with Odyssey in the form of a "window" specifying minimum and maximum need. The viceroy monitors avail- able resources and generates a callback to the application when available resources fall outside registered resource-level window. The application then chooses a new fidelity level, computes resource needs, and registers these needs, as before. Thus applications are responsible for deciding fidelity levels and computing resource requirements—the primary contribution that Odyssey makes is to monitor resources and to notify applications when available resources fall outside constraints set by the application. Before describing a sample Odyssey application, the war- dens and viceroy are discussed in detail below.

Wardens. A warden is a type-specific component responsible for handling all adaptation-related operations for a particular sort of data stream (e.g., a source of digital images, audio, or video). Wardens sit between an application and a data source, handling caching and arranging for delivery of data of appropriate fidelity levels to the application. A warden must be written for each type of data source. An application typically must be partially rewritten (or an appropriate proxy installed) to accept data through a warden rather than through a direct connec- tion to a data source, such as a streaming video server.

Viceroy. In Odyssey, the viceroy is a type-independent component that is responsible for global resource control. All the wardens are statically compiled with the viceroy. The viceroy monitors resource levels (e.g., available network bandwidth) and initiates callbacks to an application when current resource levels fall outside a range registered by the application. The types of resources to be monitored by the viceroy in Odyssey include network bandwidth, cache, battery power, and CPU, although the initial implementations of the Odyssey architecture did not support all these resource types.

A sample Odyssey application

We now turn to one of the sample applications discussed in Nobel and colleagues (1997): the xanim video player. The xanim video player was modified to use Odyssey to adapt to varying network conditions, with three fidelity levels available—two levels of JPEG compression and black-and-white frames. The JPEG compression frames are labeled 99 and 50 percent fidelity, whereas the black-and-white content is labeled

1 percent fidelity. Integration of xanim with Odyssey is illustrated in Fig. 6.3. A "video warden" prefetches frames from a video server with the appropriate fidelity and supplies the application with metadata for the video being played and with individual frames of the video.

The performance of the modified xanim application was tested using simulated bandwidths of 140 kB/s for "high" bandwidth and 40 kB/s for "low" bandwidth. A number of strategies were used to vary bandwidth: step up, which holds bandwidth at the low level for 30 seconds, followed by an abrupt increase to high bandwidth for 30 seconds; step down, which reverses the bandwidth levels of step up but maintains the same time periods; impulse up, which maintains a low bandwidth over a 60-second period with a single 2-second spike of high bandwidth in the middle; and impulse down, which maintains high bandwidth for 60 seconds with a single 2-second spike of low bandwidth in the middle. Both high and low bandwidth levels are able to support black-and-white video and the lower-quality (50 percent fidelity) JPEG video. Only the high bandwidth level is sufficient for the 99 percent fidelity JPEG frames to be delivered without substantial numbers of dropped frames. In the tests, Odyssey maintained average fidelities of 73, 76, 50, and 98 percent for step up, step down, impulse up, and impulse down, respectively, all with less than 5 percent dropped frames. In contrast, trying to maintain the 99 percent fidelity rate by transferring high-quality video at all times, ignoring available network bandwidth, resulted in losses of 28 percent of the frames for step up and step down and 58 per- cent of the frames for impulse up. Several other adapted applications are discussed in the Odyssey publications.

MORE ADAPTATION MIDDLEWARE (APRIL 2013)

Puppeteer. For applications with well-defined, published interfaces, it is possible to provide adaptation support without modifying the applications directly. The Puppeteer architecture allows component-based applications with published interfaces to be adapted to environments with poor network bandwidth (a typical situation for mobile hosts) with- out modifying the application . This is accomplished by outfitting applications and data servers with custom proxies that support the adaptation process. A typical application adaptation under Puppeteer is a retrofit of Microsoft PowerPoint to support incremental loading of slides from a large presentation or sup- port for progressive JPEG format to speed image loading. Both these adaptations presumably would enhance a user's experience when handling a large PowerPoint presentation over a slow network link.

The Puppeteer architecture is depicted in Fig. 6.4. The Puppeteer provides a kernel that executes on both the client and server side proxies, supporting a document type called the Puppeteer Intermediate Format (PIF), a hierarchical, format-neutral format. The kernel also handles all communication between client and server sides. To adapt a document, the server and client side proxies communicate to establish a high-level PIF skeleton of the document. Adaptation policies control which portions of the document will be transferred and which fidelities will be chosen for the transmitted portions. For example, for a Microsoft PowerPoint document, selected slides may transferred, with images rendered at a lower fidelity than in the original presentation. The import driver and export driver parse native document format to PIF and PIF to native document format, respectively. Transcoders in Puppeteer per- form transformations on data items to support the adaptation policies. For example, a Puppeteer transcoder may reduce the quality of JPEG images or support downloading only a subset of a document's data. A typical Puppeteer-adapted application operates as follows:

■ When the user opens a document, the Puppeteer kernel instantiates an appropriate import driver on the server side.

■ The import driver parses the native document format and creates a PIF format document. The skeleton of the PIF is transmitted by the kernel to the client-side proxy.

■ On the client side, policies available to the client-side proxy result in requests to transfer selected portions of the PIF (at selected fidelities) from the server side. These items are rendered by the export driver into native format and supplied to the application through its well- known interface.

■ At this point, the user regains control of the application. If specified by the policy, additional portions of the requested document can be transferred by Puppeteer in the background and supplied to the application as they arrive.

Coordinating adaptation for multiple mobile applications The adaptation middleware architectures is coordination among adaptive applications. Odyssey and Puppeteer, for example, support sets of independently adapting applications but do not currently assist multiple applications in coordinating their adaptation strategies. When multiple applications are competing for shared resources, individual applications may make decisions that are suboptimal. At least three issues are introduced when multiple applications attempt to adapt to limited resources—conflicting adaptation, suboptimal system operation, and suboptimal user experience.

Several sample scenarios illustrate these concerns. First, consider a situation where a number of applications executing on a mobile host with limited power periodically write data to disk. This would occur, for example, if two or more applications with automatic backup features were executing. Imagine that the mobile host maintains a powered-down state for its hard drives to conserve energy. Then, each time one of the automatic backup facilities executes, a hard disk on the system must be spun up. If the various applications perform automatic backups at uncoordinated times, then the disk likely will spin up quite frequently, wasting a significant amount of energy. If the applications coordinated to perform automatic backups, on the other hand, then disk writes could be performed "in bulk," maximizing the amount of time that the disk could remain powered down. This example illustrates suboptimal system operation despite adaptation.

Another issue when multiple applications adapt independently is conflicting adaptation. Imagine that one application is adapting to varying power, whereas another application is adapting to varying network bandwidth. When the battery level in the mobile device becomes a concern, then the power-conscious application might throttle its use of the network interface. This, in turn, makes more bandwidth available, which might trigger the bandwidth-conscious application to raise fidelity levels for a data stream, defeating the other application's attempt to save energy.

A third issue is that in the face of limited resources, a user's needs can be exceedingly difficult to predict. Thus some user participation in the adaptation process probably is necessary. To see this, imagine that a user is enjoying a high-bandwidth audio stream (Miles Davis, Kind of Blue?) while downloading a presentation she needs to review in 1 hour. With abundant bandwidth, both applications can be well served. However, if available bandwidth decreases sharply (because an 802.11 access point has gone down, for example, and the mobile host has fallen back to a 3G connection), should a lower-quality stream be chosen and the presentation download delayed because Miles is chewing up a few tens of kilobits per second? Or should the fun stop completely and the work take precedence? Efstatiou and colleagues propose using an adaptation policy language based on the event calculus (Kowalsky, 1986) to specify global adaptation policies. The requirements for their architecture are that a set of extensible adaptation attributes be sharable among applications, that the architecture be able to centrally control adaptation behavior, and that flexible, system-wide adaptation policies, depending on a variety of issues, be expressible in a policy language. Their architecture also allows human interaction in the adaptation process both to provide feedback to the user and to engage the user in resolving conflicts (e.g., Miles Davis meets downloading PowerPoint). Applications are required to register with the system, providing
a set of adaptation policies and modes of adaptation supported by the application. In addition, the application must expose a set of state variables that define the current state of the application. Each application generates events when its state variables change in meaningful ways so that the adaptation architecture can determine if adaptive actions need to be taken; for example, when a certain application is minimized, a global adaptation policy may cause that application to minimize its use of system resources. A registry in the architecture stores information about each application, and an adaptation controller monitors the state of the system, determining when adaptation is necessary and which applications should adapt. Another policy language–driven architecture advocating user involvement is described in Keeney and Cahill .

8. Write about Mobile Agents (11)

We now turn to another type of mobile middleware, mobile agent systems. Almost all computer users have used mobile code, whether they realize it or not—modern browsers support Javascript, Java applets, and other executable content, and simply viewing Web pages results in execution of the associated mobile code. Applets and their brethren are mostly static, in that code travels from one or more servers to a client and is executed on the client. For security reasons, the mobile code often is prevented from touching nonlocal resources. Mobile agents are a significant step for- ward in sophistication, supporting the migration of not only code but also state. Unlike applets, whose code typically travels (at an abstract level at least) one "hop" from server to client, mobile agents move freely about a network, making autonomous decisions on where to travel next. Mobile agents have a mission and move about the network extracting data and communicating with other agents in order to meet the mission goals.

Like adaptation middleware, mobile agent systems (e.g., Cabri, Leonardi, and Zambonelli, 2000; Gray, 1996, 1997; Gray et al., 1998, 2000; Bradshaw et al., 1999; Lange and Oshima, 1998; Peine and Stoplmann,1997; Wong, Paciorek, and Moore, 1999; Wong et al., 1997) support execution of mobile applications in resource-limited environments, but mobile agent systems go far beyond allowing local applications to respond to fluctuating resource levels. A mobile agent system is a dynamic client-server (CS) architecture that supports the migration of mobile applications (agents) to and from remote servers. An agent can migrate whenever it chooses either because it has accomplished its task completely or because it needs to travel to another location to obtain additional data. An alter- native to migration that an agent might exercise is to create one or more new agents dynamically and allow these to migrate. The main idea behind mobile agents is to get mobile code as close to the action as possible—mobile agents migrate to remote machines to perform computations and then return home with the goods.

For example, if a mobile user needs to search a set of databases, a traditional CS approach may perform remote procedure calls against the database servers. On the other hand, a mobile agents approach would dispatch one or more applications (agents) either directly to the data- base servers or to machines close to the servers. The agents then per- form queries against the database servers, sifting the results to formulate a suitable solution to the mobile user's problem. Finally, the mobile agents return home and deliver the results.

The advantages of this approach are obvious. First, if bandwidth available to the mobile user is limited and the database queries are complicated, then performing a series of remote queries against the servers might be prohibitively expensive. Since the agents can execute a number of queries much closer to the database servers in order to extract the desired information, a substantial amount of bandwidth might be saved (of course, transmission of the agent code must be taken into account). Second, continuous network connectivity is not required.

The mobile user might connect to the network, dispatch the agent, and then disconnect. When the mobile user connects to the network again later, the agent is able to return home and present its results. Finally, the agents are not only closer to the action, but they also can be executed on much more powerful computers, potentially speeding up the mining of the desired information. Of course, there are substantial difficulties in designing and implementing mobile agent systems.

Why mobile agents? And why not?

We first discuss the advantages of mobile agents at a conversational level, and then we look at the technical advantages and disadvantages in detail. First, a wide variety of applications can be supported by mobile agent systems, covering electronic commerce (sending an agent shop- ping), network resource management (an agent might traverse the net- work, checking versions of installed applications and initiating upgrades where necessary), and information retrieval (an agent might be dis- patched to learn everything it can about Thelonious Monk).

An interesting observation made by Gray and colleagues (2000) is worth keeping in mind when thinking about agent-based applications: While particular applications may not make a strong case for deployment of mobile agent technology, sets of applications may make such a case. To see this point, consider the database query example discussed in the preceding section. Rather than using mobile agents, a custom application could be deployed (statically) on the database servers. This application accepts jobs (expressing the type of information required) from a mobile user, performs a sequence of appropriate queries, and then returns the results. Since most of the processing is done off the mobile host, the resource savings would be comparable to a mobile agents solution.

Similarly, little computational power on the mobile host is required because much of the processing can be offloaded onto the machine hosting the custom application. However, what if a slightly different application is desired by a mobile user? Then the server configuration must be changed. Like service discovery protocols, covered in mobile agent systems foster creation of powerful, personalized mobile applications based on common frameworks. While individual mobile applications can be written entirely without the use of agent technologies, the amount of effort to support a changing set of customized applications may be substantially higher than if mobile agents were used.

Mobile agent systems provide the following set of technical advantages

■ The limitations of a single client computer are reduced. Rather than being constrained by resource limitations such as local processor power, storage space, and particularly network bandwidth, applications can send agents "into the field" to gather data and perform computations using the resources of larger, well-connected servers.

■ The ability to customize applications easily is greatly improved. Unlike traditional CS applications, servers in an agent system merely pro- vide an execution environment for agents rather than running customized server applications. Agents can be freely customized (within the bounds of security restrictions imposed by servers) as the user's needs evolve.

■ Flexible, disconnected operation is supported. Once dispatched, a mobile agent is largely independent of its "home" computer. It can per- form tasks in the field and return home when connectivity to the home computer is restored. Survivability is enhanced in this way, especially when the home computer is a resource-constrained device such as a PDA. With a traditional CS architecture, loss of power on a PDA might result in an abnormal termination of a user's application.

Despite these advantages, mobile agent architectures have several significant disadvantages or, if that is too strong a word, disincentives. One is that neither a killer application nor a pressing need to deploy mobile agent technology has been identified. Despite their sexiness, mobile agents do not provide solutions to problems that are otherwise unsolvable; rather, they simply seem to provide a good framework in which to solve certain problems. In reflections on the Tacoma project (Milojicic, Douglis, and Wheel, 1998), Johansen, Schneider, and van Renesse note that while agents potentially reduce bandwidth and tolerate intermittent communication well, bandwidth is becoming ever more plentiful, and communication is becoming more reliable. As wire- less networking improves and mobile devices become more powerful and more prevalent, will mobile agents technologies become less relevant? Further, while a number of systems exist, they are largely living in research laboratories. For mobile agent systems to meet even some of their potential, widespread deployment of agent environments is required so that agents may travel freely about the Internet.

A related problem is a lack of standardization. Most mobile agent systems are not interoperable. Some effort has gone into interoperability for agent systems, but currently, there seem to be no substantial market pressures forcing the formation of a single (or even several) standards for mobile agent systems. The Mobile Agents System Interoperability Facility (MASIF; Milojicic et al., 1999) is one early attempt at fostering agent interoperability for Java-based agent systems.

All the disadvantages just discussed are surmountable with a little technical effort—apply a good dose of marketing, and most disappear. There is a killer disadvantage, however, and that is security. Even applets and client-side scripting languages (such as Javascript), which make only a single hop, scare security-conscious users to death, and many users turn off Java, Javascript, and related technologies in their Web browsers. Such users maintain this security-conscious stance even when interacting with Web sites in which they place significant trust because the potential for serious damage is high should the sandbox leak. Security for mobile agent systems is far more problematic than simple mobile code systems such as Java applets because agents move autonomously.

There are at least two broad areas of concern. First, agents must be prevented from performing either unintentional or malicious damage as they travel about the network. Could an agent have been tampered with at its previous stop? Is it carrying a malicious logic payload? Does it contain contraband that might be deposited on a machine? Will the agent use local resources to launch a denial-of-service attack against another machine? Essentially, if agents are to be allowed to get "close to the action," then the "action" must be convinced (and not just with some marketing) that the agents will not destroy important data or abuse resources. Second, the agents themselves must be protected from tampering by malicious servers. For example, an agent carrying credit card information to make purchases on behalf of its owner should be able to control access to the credit card number. Similarly, an agent equipped with a proprietary data-mining algorithm should be able to resist reverse engineering attacks as it traverses the network.

Agent architectures

To illustrate the basic components of mobile agent architectures, a high- level view of Telescript (White in Milojicic, Douglic, and Wheel, 1998) works well. Telescript was one of the first mobile agent systems, and while it is no longer under development, many subsequent systems bor- rowed ideas from Telescript. There are a number of important components in the Telescript architecture: agents, places, travel, meetings, connections, authorities, and permits. These are depicted in Fig. 6.5. Each of these components is described in detail below.

Places. In a mobile agent system, a network is composed of a set of places—each place is a location in the network where agents may visit. Each place is hosted by a server (or perhaps a user's personal computer) and provides appropriate infrastructure to support a mobile agent migrating to and from that location. Servers in a network that do not

Figure 6.5 The major Telescript components are illustrated above. Tom has just dispatched an agent which has not yet arrived at the theater server. When Tom's agent arrives, it will interact with the static agent in the box office place to arrange for theater tickets. Daryl previously dispatched an agent to purchase tickets and has a connection with her agent in the box office place, so she can actively negotiate prices. Daryl's agent and the box office agent have identified each other through their respective authorities and permits associated with Daryl's agent have been evaluated to see what actions are permitted. The static agents in the drugstore and music store places, which both reside on a shopping center server, are currently idle. To interact with the drugstore or music store agents, Daryl or Tom's agents will have to travel to the drugstore and music store places .offer a "place" generally will not be visitable by agents. Places offer agents a resting spot in which they can access resources local to that place through a stationary agent that "lives" there, interacting with other agents currently visiting that place.

Travel. Travel allows agents to move closer to or to colocate with needed resources. For example, an agent dispatched by a user to obtain tickets to a jazz concert and reservations at one of several restaurants (depending on availability) might travel from its home place to the place hosted by the jazz club's box office before traveling to the places hosted by the restaurants. The primary difference between mobile code strategies such as Java applets and agents is that agents travel with at least part of their state intact—after travel, an agent can continue the computation it was engaged in at the instant that travel was initiated. Migration is studied in further detail below in the section entitled, "Migration Strategies."

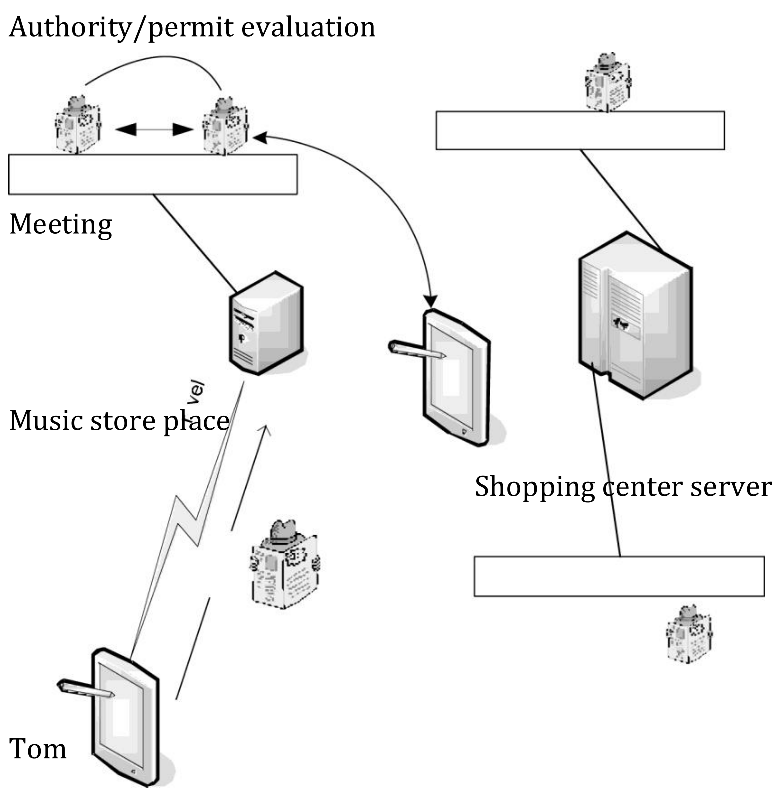
Meetings. Meetings are local interactions between two or more agents in the same place. In Telescript, this means that the agents can invoke each other's procedures. The agent in search of jazz tickets and a restaurant reservation (discussed under "Travel" above) would engage in meetings with appropriate agents at the ticket office and at the restaurant's reservation office to perform its duty.

Connections. Connections allow agents at different places to communicate and allow agents to communicate with human users or other applications over a network. An agent in search of jazz tickets, for example, might contact the human who dispatched it to indicate that an additional show has been added, although the desired show was sold out (e.g., "Is the 11 P.M. show OK?"). Connections in Telescript require an agent to identify the name and location of the remote agent, along with some other information, such as required quality of service. This remote communication method, which tightly binds two communicating agents (since both name and location are required for communication), is the most restrictive of the mechanisms discussed in further detail below in the section entitled, "Communication Strategies."

Authorities. An agent's or place's authority is the person or organization (in the real world) that it represents. In Telescript, agents may not with- hold their authority; that is, anonymous operation is not allowed—the primary justification for this limitation is to deter malicious agent activity. When an agent wishes to migrate to another location, the destination can check the authority to determine if migration will be permitted. Similarly, an agent may examine the authority of a potential destination t determine if it wishes to migrate there. Implementation of authorities in an un trusted network is nontrivial and requires strong cryptographic methods because an agent's authority must be un forgeable.

Permits. Permits determine what agents and places can do—they are sets of capabilities. In general, these capabilities may have virtually any form, but in Telescript they come in two flavors. The first type of capability determines whether an agent or place may execute certain types of instructions, such as instructions that create new agents. The second type of capability places resource limits on agents, such as a maximum number of bytes of network traffic that may be generated or a maximum lifetime in seconds. If an agent attempts to exceed the limitations imposed by its permits, it is destroyed. The actions permitted an agent are those which are allowed by both its internal permits and the place(s) it visits.

Other issues. A number of details must be taken into account when designing an architecture to support mobile agents, but one of the fundamental issues is the choice of language for implementation of the agents (which might differ from the language used to implement the agent architecture). To support migration of agents, all computers to which an agent may migrate must share a common execution language. While it is possible to restrict agents to a particular computer architecture and OS (e.g., Intel 80 × 86 running Linux 2.4), clearly, agent sys- tems that can operate in heterogeneous computer environments are the most powerful. Compiled languages such as C and C++are problematic because agent executables must be available for every binary architecture on which agents will execute. Currently, interpreted languages such as Java, TCL, and Scheme are the most popular choices because many problems with code mobility are alleviated by interpreted languages. In cases where traditionally compiled languages such as C++ are used for implementation of agents, a portable, interpreted byte code typically is emitted by a custom compiler to enable portability. Java is particularly popular for mobile agents because Java has native support for multithreading, object serialization (which allows the state of arbitrary objects to be captured and transmitted), and remote procedure calls. Other factors, aside from the implementation language for agents, include migration strategies, communication, and security. Migration and communication strategies are discussed in detail below.



Migration strategies

To support the migration of agents, it must be possible to either capture the state of an agent or to spawn an additional process that captures the state of the agent. This process state must then be transmitted to the remote machine to which the agent (or its child, in the case of spawn- ing an additional process) will migrate. This is equivalent to process check pointing, where the state of a process, including the stack, heap, program code, static variables, etc., is captured and stored for a later resuscitation of the process. Process check pointing is a very difficult problem that has been studied in the operating systems and distributed systems communities for a number of years, primarily to support fault tolerance and load balancing (Jul et al., 1988; Douglis and Ousterhout,1991; Plank, 1995). In general, commodity operating systems do not provide adequate support for check pointing of processes, and add-on solutions [e.g., in the form of libraries such as libckpt (Plant, 1995)] are nonportable and impose significant restrictions, such as an inability to reconstitute network connections transparently. A number of research operating systems have been designed that better support process migration, but since none of these is viable commercially (even in the slightest sense), they are not currently appropriate platforms. Check pointing processes executing inside a virtual machine, such as Java processes, are a bit easier, but currently most of these solutions (Richard and Tu, 1998; Sakamoto, Sekiguchi, and Yonezawa, 2000; Truyen et al., 2000) also impose limitations, such as restrictions on the use of call- backs, network connections, and file activity. The virtual machine itself can be check pointed, but then the issues of portability discussed earlier reemerge, and network connections and file access will still pose problems. So where is this going? The punch line is that if commodity operating systems are to be targeted by agent systems—and for wide-scale deployment, this must be the case—then completely capturing the state of general processes to support migration is rife with problems.

One solution is to impose strong restrictions on the programming model used for mobile agents. Essentially, this entails capturing only the essential internal state of an agent, i.e., sufficient information about its current execution state to continue the computation on reconstitution, combined with a local cleanup policy. This means that an agent might perform a local cleanup, including tearing down communication connections and closing local files, before requesting that the agent middleware perform a migration operation. For example, in Aglets (Lange and Oshima, 1998), which is a Java-based mobile agents system, agents are notified at the beginning of a migration operation. It is the responsibility of an individual agent, on receiving such a notification, to save any significant state in local variables so that the agent can be properly "reconstituted" at the new location. Such a state may include the names of communication peers, loop indices, etc. Agent migration in Aglets begins with an agent initiating a migration (its own or that of another agent) by invoking dispatch(). A callback, onDispatch(), will be triggered subsequently, notifying the agent that it must save its state. After the migration, the agent's onArrival()callback will be invoked so that the agent can complete its state restoration.

Communication strategies

Communication among agents in a mobile agent system can take many forms, including the use of traditional CS techniques, remote procedure call, remote method invocation (e.g., using Java's RMI), mailboxes, meeting places, and coordination languages. Each of these communication strategies has advantages and disadvantages, some of which are exacerbated in mobile agent systems. One consideration is the degree of temporal and spatial locality exhibited by a communication scheme (Cabri, Leonardi, and Zambonelli, 2000).Temporal locality means that communication among two or more agents must take place at the same physical time, like a traditional telephone conversation. Interagent communication mechanisms exhibiting temporal locality are limiting in a mobile agent's architecture because all agents participating in a communication must have network connectivity at the time the communication occurs. If an agent is in transit, then attempts to communicate with that agent typically will fail.

Spatial locality means that the participants must be able to name each other for communication to be possible—in other words, unique names must be associated with agents, and their names must be sufficient for determining their current location. Some of the possible communication mechanisms for interagent communication are discussed below.

Traditional CS communication. Advantages of traditional CS mechanisms such as sockets-based communication, Remote Method Invocation (RMI) in Java, and CORBA include a familiar programming model for soft- ware developers and compatibility with existing applications. Significant drawbacks include strong temporal and spatial locality—for communication to be possible, agents must be able to name their communication peers and initiate communication when their peers are also connected. RMI and other communication mechanisms built on the Transmission Control Protocol/Internet Protocol (TCP/IP) also require stable network connectivity; otherwise, timeouts and subsequent connection reestablishments will diminish performance significantly. Examples of agents systems that use traditional CS mechanisms are D'Agents (Gray et al.,

1998) and Aglets (Lange and Oshima, 1998). In Aglets, an agent first must obtain another agent's proxy object (of type AgletProxy) before communication can take place. This proxy allows the holder to transmit arbitrary messages to the target and to request that the target agent per- form operations such as migration and cloning (which creates an identical agent). To obtain a proxy object for a target agent, an agent typically must provide both the name of the target agent and its current location. If either agent moves, then the proxy must be reacquired.

Meeting places. Meeting places are specific places where agents can congregate in order to exchange messages and typically are defined statically, avoiding problems with spatial locality but not temporal locality. In Ara (Peine and Stolpmann, 1997), meeting places are called service points and provide a mechanism for agents to perform local communication. Messages are directed to a service point rather than to a specific agent, eliminating the need to know the names of colocated agents.

Tuple spaces. Linda-like tuple spaces are also appropriate for interagent communication. Linda provides global repositories for tuples (essentially lists of values), and processes communicate and coordinate by inserting tuples into the tuple space, reading tuples that have been placed into the tuple space, and removing tuples from the tuple space. Tuple spaces eliminate temporal and spatial bindings between communicating processes because communication is anonymous and asynchronous.

9. Explain interoperability and standadization (11)

The wired infrastructure has been designed and deployed around a rich set of international standards. For example, the legacy local area network (LAN) consists of such technologies as Ethernet, Token-Ring, and Token-Bus which were defined in precise details by the IEEE 802 committees. Moreover, newer network technologies like FDDI, HIPPI, and Fibre Channel have been defined by the ANSI X3T working groups with a mature set of approved specifications. Both the IEEE and the ANSI bodies add further credibility to their work by helping international organizations like ISO and ITU to easily migrate the specifications into international standards bodies for worldwide acceptance. Similar efforts are underway in the ATM Forum to create a set of implementation agreements which should permit interoperability between different vendor implementations and products.

The wireless industry currently embraces a small number of standards. The closest effort is within the IEEE 802 working group which recently completed the IEEE 802.11 Wireless MAC (media access control) standard. The primary objective of the IEEE 802.11 effort is to permit wireless LANs from different vendors to interoperate.

IEEE 802.11 does not, however, address the needs of the wide area wireless networking industry which currently deploys various packetized protocols (e.g. CDPD, GPRS) across unused cellular channels. Each network type is based on its own set of assumptions about the kinds of service the customers are willing to purchase. Service providers for each of these types of networks have different goals and strategies and do not seem likely to provide interoperability among the other classes of service.

Other mobile infrastructures are also lacking in internationally recognized standards. This is evident in the cellular telephone industry: a PHS telephone will not function in a cell serviced by a GSM or PCS infrastructure. The same is true for any combination of the aforementioned technologies. Moreover, cordless telephones, infrared transmission, satellite channels, and most mobile communication systems are either based on proprietary data interfaces, or have implemented selected parts of existing and/or emerging deployment agreements. Such key attributes as Quality of Service, Location Register con- tents, Database formats, update policies, and data exchange rates are left to the equipment providers and service providers which may be based more on deployment schedules than on availability of standards and interoperability guarantees. The emerging UMTS system standard which is expected to be deployed by the year 2002, will provide a golden opportunity for interoperability of data links interfaces, digital voice, and wireless data and services.

Many of the client-application products, or the communications substrate that they rely on, are recognizing that several competing wireless transmission pro- tocols exist with each network type. They also recognize that the number of such protocols may grow or shrink. As a result, these client-level packages are adapted to use the popular underlying protocols. This limited form of inter- operability appears to meet the needs for developers of client software. As an example, the Oracle Mobile Agents product discussed previously supports both CDPD and Shiva PPP. However, no client software we have seen claims to migrate seamlessly among the different wireless network classes.

At some level, interoperability among the various network classes can be pro- vided by adopting popular communications standards. For example, those client applications developed to exploit TCP/IP in wired networks can inter- operate without change in the wireless domain if some variant of TCP/IP is offered as a service, e.g. lETF's Mobile IP. However, the quality of service that is provided by this approach may not be transparent, or even acceptable. In addition, it remains the client's responsibility to transfer among the various competing network services.

The heterogeneity of the existing and emerging wireless network protocols poses not only a need for interoperability, but also a stringent quality of service requirements. This is because the inherent unreliability and bandwidth limitation largely varies from one network to the other, leading to rapid fluctuations in the quality of the provided services. Recent research efforts proposed extensions to formal open systems standards.

The wireless application protocol (WAP) standard currently being developed by the WAP forum group offers an OSI-like protocol stack for inter operability of different wireless networks. The WAP stack allows applications to register interest in quality of service events and thresholds (QoS). This, in turn, allows the application to be mobility-aware and adaptable to changes in the environment. The WAP stack also provides negotiation protocols between producers and consumers of data to optimize the necessary level of data presentation based on the nature of data, the current wireless network between the source and the destination, and the capabilities of the destination device. Content negotiation should play a major role in maintaining QoS across heterogeneous networks.

Another proposal in extends the ISO Reference Model for Open Distributed Processing (RM-ODP) so that clients are able to adapt to the variation of the network service they encounter. Under this proposal, applications will have to be mobility-aware, but nonetheless will be able to maintain the required QoS.

10. Explain shortcomings and limitations (6)

Mobility-support software and products that are commercially available today leaves much to be desired in terms of functionality, performance, portability, and interoperability. The continuous decline in wireless communication cost and the recent convergence towards a truly global and standard communication system will help reduce the business risk associated with investments in this software market. This will be true for giant software vendors as well as small startup companies. Several limitations and shortcomings of existing products are summarized below.

- **Translucent Overlays.** None of the product offerings or announced plans for products that we have seen have included the vision of a translucent client context for exploiting nomadic applications. Each client or middle- ware offering is tailored to a specific kind of network service and assumes the client will manage its transition from one network class to another as the mobile platform roams about.

- **Multi-Database Access.** DBMS vendors such as Oracle and Sybase which are announcing mobile client products are providing connection ser- vices only to their proprietary DBMS product. This trend mirrors the path that these vendors have adopted for the fixed network environment. Products which have emerged for the wired environment which make it possible or a client to interact with a variety of vendor DBMS, e.g. ODBC inter-faces, will not automatically extend service into the wireless domain. This failure occurs because current solutions that couple mobile clients with server DBMS include insertion of vendor-specific agents on both ends of the wireless connection to mediate wireless artifacts. Thus, a Mobile Oracle client can only talk to those servers that are serviced by the Mobile Oracle server agents .

- **Mobile Transactions.** We have seen no product that addresses issues of transaction management in the mobile environment. The concept of a mo- bile transaction, where the locus of control of the transaction is maintained by the mobile user, remains as a research idea needy of commercialization. BEA Systems markets a variety of products that couple wired clients with a variety of TP monitors, and implicitly to any of the DBMS products those TP monitors serve. If mobile transaction products will ever be made available, they will be offered by companies such as BEA Systems and Transarc.
- **Workflows.** Workflow products lag database access products in their migration to wireless and mobile environments. We have seen no workflow products targeted for the mobile domain in our literature searches.
- **Location Dependent Services.** None of the server or client software packages targeted for the wireless domain claim to offer location dependent services. This important class of services will be essential at both ends of the wireless client-server communications link.

SRI VENKATESHWARAA COLLEGE OF ENGINEERING & TECHNOLOGY
(Approved by AICTE, New Delhi & Affiliated to Pondicherry University, Puducherry.)
13-A, Villupuram – Pondy Main road, Ariyur, Puducherry – 605 102.
Phone: 0413-2644426, Fax: 2644424 / Website: www.svcetpondy.com

Department of Computer Science and Engineering

Subject Name: **MOBILE COMPUTING** Subject Code: **CS E84**

UNIT III

Mobile Networking: Virtual IP Protocols - Loose Source Routing Protocols - Mobile IP – CDPD – GPRS – UMTS - Security and Authentication – Quality of Service – Mobile Access to the World Wide Web.

2 Marks

1. What is mobile networking?

The mobile networking protocol should also be transparent to the hosts and routers which do not understand or support mobility. Thus, the mobility unaware routers should be able to route packets destined to a mobile host as normal IP data packets. Security is another important concern in internetworking. In mobile networking it is more so, since the mobile nodes will be visiting foreign networks, requesting services, and accessing data. Thus, it is important that the security of the visiting network is not breached due to the presence of a foreign node in its network.

2. Write Short notes on virtual IP protocols.

A mobile network is a virtual network with a virtual address space. A mapping is maintained between the physical or actual IP addresses and the Virtual IP addresses. This mapping is performed by the mobile host which obtains a *care of address* from the local network being visited using either the Dynamic Host Configuration Protocol (DHCP) or the BOOTP protocols or by any of the link layer protocols. Below, we describe two Virtual IP protocols.

3. What is loose source routing protocol?

This approach was proposed by David Johnson of CMU in 1993. It uses the Loose Source Route option available in the IPv4 for routing packet data. The option allows the source to specify the intermediate gateways in the IP packet. Thus, the source can control the route the IP packet takes. At each destination, the gateway picks up the next IP address from the IP packet, sets it as the destination, and advances a pointer stored in the IP packet header.

4. Define mobile internet protocol.

The Mobile Internet Protocol (Mobile IP) [88, 80, 19, 7, 59, 76] defines en-hancements to the Internet Protocol to allow routing of IP packets to mobile nodes in the internet. The IP version 4 assumes that the Internet Protocol of a node uniquely identifies the point of attachment of the node to the internet- work. Packets are routed based on the IP address.

5. What are the components of MIP?

The major architecture components of the mobile IP protocol are:

Mobile N o de (MN)
Home Agent (HA)
Foreign Agent (FA)

6. What is DHCP?

The mobile node obtains a temporary IP address on the foreign agent network to be used for forwarding. This can be done using the IETF Dynamic Host Configuration Protocol (DHCP).

7.Explain the routing path of a datagram.

The routing path of a datagram sent from a fixed host to a mobile node is as follows: (1) the datagram is sent from the fixed host to the home agent using standard IP routing; (2) the home agent encapsulates the received datagram inside another datagram and sends it to the foreign agent (IP-in-IP tunneling); (3) the encapsulated IP packet is received by the foreign agent, decapsu- lasted, and forwarded to the mobile node; (4) the mobile node replies by sending a datagram to the fixed host through the foreign agent.

8.Define cellular digital packet data.

CDPD is a connectionless multi-network protocol, proposed originally by the CDPD Forum (now called the WDF Forum). It is based on the early versions of Mobile-IP . The idea behind CDPD is to share unused channels in existing Advanced Mobile Phone Systems (AMPS) to provide up to 19.2 kbps data channel.

9.Explain GPRS.

GPRS is a GSM packet data service developed by the European Telecommunication Standards Institute (ETSI) as part of GSM phase 2+ developments. The goal of GPRS was to support data transfer rates higher than the 9.6 kbps achieved through GSM’s circuit switching technology. Unlike Mobile-IP, GPRS is not restricted to IP packet data protocols, and offers connection to standard protocols (such as TCP/IP, X.25, and CLNP) as well as specialized data packet protocols.

10.Explain briefly about security and authentication.

In a mobile computing environment, it is desirable to protect information about the movements and activities of mobile users from onlookers. In addition to the basic security concerns in wire line systems (authentication, confidentiality, and key distribution), a new issue is the privacy and anonymity of the user’s movement and identity. In fact, a typical situation arises when a mobile user registers in one domain (home domain) and appears in a different foreign do- main; the user must be authenticated and his solvency must be confirmed. Usually during this process the user has to provide a non-ambiguous identity to his home domain and has to verify it. If no care is taken, this identity can be tapped on the air interface in a cellular environment or through the signaling protocols exchanged on the registered wired network.

11.What should be the quality of service in networking.

Mobile network protocols such as Mobile-IP and GPRS provide mobility trans- parency at the network layer level. This allows the higher layers of the protocol stack to be used unchanged. Unfortunately, there are ill consequences to this transparency that are mostly attributed to the constraints of the wireless and mobile environment. For example, transport layer protocols that rely heavily on timeout mechanisms for re-transmission, if used unchanged, will perform poorly under variable delays and limited bandwidth.

12.Explain about mobile access to the world.

More and more users are becoming increasingly dependent on information they obtain from the World Wide Web. Users are also demanding ubiquitous access, any time, anywhere, to the information they rely on. Several research efforts ex- plored the problems associated with wireless access to the Web. Most solutions used a Web proxy that enabled Web browsing applications to function over wireless links without imposing changes on browsers and servers. Web proxies are also used to prefetch and cache Web pages to the mobile client's machine, to compress and transform image pages for transmission over low-bandwidth links, and to support disconnected and asynchronous browsing operations.

13.What is wireless WWW?

A prototype consisting of commercially available PDAs and a wireless LAN has been used to provide a "proof of concept" for the Wireless World Wide Web (W4). A simplified version of Mosaic was ported to the PDA for the purpose of experimenting with response time performance and to sort out design choices. A PDA cache was used to improve the performance.

14.Explain the design of MOWSER..

The design is based on a mediator server that filters retrieved information according to the limitations of the mobile unit. Color, resolution, display mode, sound capability, and maximum file size are among the factors considered. The browser, called MOWSER, connects to two servers in the fixed network. The first is the preference server that maintains the user profile; the second is a proxy server that implements all the filtering indicated by the preference server.

15.What are the methods of web access optimizations?

- Web Express optimization methods are summarized below:

  Caching
  Differencing
  Protocol reduction
  Header reduction

16.What is dynamic URL?

The Mobisaic project at the University of Washington extends standard client browsers to support dynamic URLs and active documents. The Mosaic Web client and the URL syntax are modified so that when the user traverses a dynamic URL, the client resolves any references to dynamic information it may contain and sends the result back to the server. This is helpful in defining location-sensitive resources.

17. Draw the WAP protocol stack diagram.

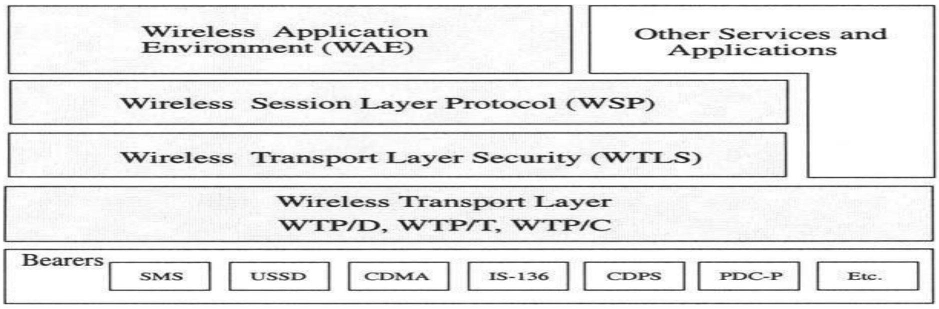


Figure 7.5 The WAP protocol stack

18.Explain the Loss profile approach.

A Loss Profile is proposed and is defined to be a description, provided by the application, of an "acceptable" manner in which data for its connection may be discarded. The loss profile is used in the event of bandwidth reduction at the wireless end of the connection. An elaborate example of a loss profile is given on viewer perception of a video clip under data loss. The loss profile is used by a specialized session layer which is transparent to the application.

19.What are the research efforts that address QoS concerns in the wireless and mobile environment.

- ✓ Optimizing TCP/IP for Mobile Networks
- ✓ QoS driven, full protocol stacks

20.Draw the diagram of GPRS.

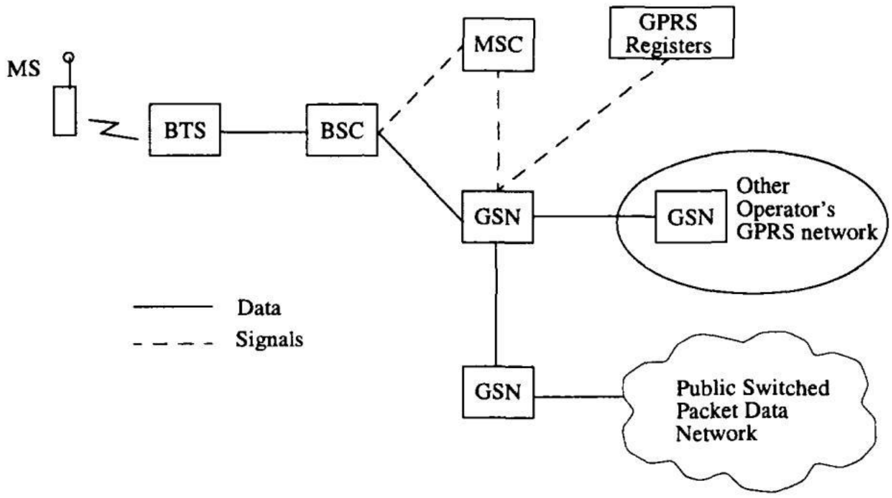


Figure 7.2 General Packet Radio Service

21.What is IETF?

The Internet Engineering Task Force (IETF) is a forum of working groups responsible for identifying operational problems and proposing solutions to these problems.

22. What is IRTF?

The Internet Research Task Force (IRTF) is a forum of working groups focusing on long-term research topics related to Internet protocols, applications, architecture, and technology.

11 Marks

1. EXPLAIN MOBILE NETWORKING

Internetworking mobile computers with the fixed-network raises the additional requirements of mobility transparency and mobility and location management. The mobility behavior of a node should be transparent to a peer node. A peer node should be able to communicate with a mobile node using some fixed IP address irrespective of the current point of attachment. The mobile net- working protocol should also be transparent to the hosts and routers which do not understand or support mobility. Thus, the mobility unaware routers should be able to route packets destined to a mobile host as normal IP data packets. Security is another important concern in internetworking.

In mo- bile networking it is more so, since the mobile nodes will be visiting foreign networks, requesting services, and accessing data. Thus, it is important that the security of the visiting network is not breached due to the presence of a foreign node in its network. Authentication of the mobile nodes and foreign networks is also important. Thus, at a minimum, mobile networking protocols should provide authentication and security features comparable to those found in fixed-network IP protocols such as IPv4 and IPv6.

In this section, various approaches and protocols for mobile internetworking

are examined, including:

- Early approaches: virtual IP mechanisms
- Loose source routing protocol
- The Mobile Internet Protocol (Mobile-IP) Cellular Digital Packet Data (CDPD)
- The General Packet Radio Service protocol (GPRS)

Emphasis is placed on protocol mechanisms, leaving out the details which can be obtained by following cited work and web resources.

1.1.1 Early Approaches: Virtual IP Protocols

In this approach, a mobile network is a virtual network with a virtual address space. A mapping is maintained between the physical or actual IP addresses and the Virtual IP addresses. This mapping is performed by the mobile host which obtains a *care of address* from the local network being visited using either the Dynamic Host Configuration Protocol (DHCP) [38] or the BOOTP protocols or by any of the link layer protocols. Below, we describe two Virtual IP protocols.

1.1.1.1 Sunshine And Postal

The earliest solution for managing mobile hosts was proposed by Sunshine and Postal in 1980. They proposed that the mobile hosts be assigned a virtual IP address which can be used to identify them. A mobile host in the foreign network is required to obtain a care-of-address, and to update its location in a mapping database. When a packet has to be routed to the mobile host, its current location is looked up in the database and the packet is transmitted to that location.

1.1.1.2 The SONY Protocol

This protocol was proposed in 1992 by F.Teraoka et al. of Sony Laboratories. In this scheme, a mobile host has two IP addresses associated with it. A virtual address, which is immutable and by which it is known to the outside world, and a physical address, which is acquired from the local network. Two sub layers are introduced in the network layer and are used to map the physical address to the virtual address. The transport layer interfaces with the network layer through the virtual layer interface and addresses its packets to the virtual address of a mobile host. A cache called the Address Mapping Table (AMT) is used for fast address resolution. A copy of this cache is maintained at each host/router. The VIP (Virtual IP) is implemented as an IP option. A set of packet types is also defined for host communication.

On entering a foreign network, the mobile host obtains an IP address and informs its home network of its current location. The home network broadcasts this information so the AMT cache gets updated. A stationary host, when required to communicate with a mobile host, looks up its cache. If the mapping is available, the packet is transmitted in the normal fashion by appending the VIP header. If the cache entry is not available, the packet is addressed to the VIP address. A set of connection gateways are required for the co-existence of mobility aware and mobility unaware hosts on the network.

1.1.2 Loose Source Routing Protocol

This approach was proposed by David Johnson of CMU in 1993. It uses the Loose Source Route option available in the IPv4 for routing packet data. The option allows the source to specify the intermediate gateways in the IP packet. Thus, the source can control the route the IP packet takes. At each destination, the gateway picks up the next IP address from the IP packet, sets it as the destination, and advances a pointer stored in the IP packet header.

The home network maintains a database of all mobile hosts native to its net- work. When a mobile host changes location, it informs its home network of its new location. When an IP packet destined to the mobile host arrives at the home network, the packet is forwarded to the mobile host at the current location address, and the corresponding source host is informed of the current location of the mobile host. The corresponding host can use this information to cache the location, thus avoiding communication with the home network until the mobile host changes its location again. Source route set up is done by the corresponding host.

Loose Source Routing is an IP option which can be used for address translation. LSR is also used to implement mobility in IP networks.

Loose source routing uses a source routing option in IP to record the set of routers a packet must visit. The destination of the packet is replaced with the next router the packet must visit. By setting the forwarding agent (FA) to one of the routers that the packet must visit, LSR is equivalent to tunneling. If the corresponding node stores the LSR options and reverses it, it is equivalent to the functionality in mobile IPv6.

The name loose source routing comes from the fact that only part of the path is set in advance. This is in contrast with strict source routing, in which every step of the route is decided in advance when the packet is sent.

1.1.3 The Mobile Internet Protocol (Mobile-IP)

The Mobile Internet Protocol (Mobile IP) [88, 80, 19, 7, 59, 76] defines en-hancements to the Internet Protocol to allow routing of IP packets to mobile nodes in the internet. The IP version 4 assumes that the Internet Protocol of a node uniquely identifies the point of attachment of the node to the internet- work. Packets are routed based on the IP address. In a mobile environment, the point of attachment of the mobile node will be different from time to time, and the mobile nodes could be attached to different networks. For IPv4 to work correctly in the mobile environment, the mobile node will either have to be assigned a new IP address every time it changes its point of attachment, or the host specific routing information has to be supplied throughout the network. Both of these alternatives result in scalability and connection management problems. The mobile IP protocol describes a mechanism which allows nodes to change their point of attachment on the Internet.

The major architecture components of the mobile IP protocol are:

- Mobile Node (MN): is a host or a router that changes its point of attachment to the network from one sub network to another. The MN is known throughout the network by an IP address assigned to it in the home net- work. The mobile node can communicate from any location as long as the link layer connectivity to the internetwork is established.
- Home Agent (HA): is a mobile-IP capable router on the mobile node’s home network. The HA maintains the location information for the mobile node. It also acts as the tunneling agent for packets destined to the mobile node. The HA manages the registration and authorization information of all the mobile modes belonging to its network.
- Foreign Agent (FA): is a mobile-IP capable router that the mobile node has visited. After attaching to the foreign network, the mobile node is required to register itself with the FA. The FA detunnels and routes the packets destined to the mobile node. The FA may also act as a default router for mobile nodes registered with it.

The mobile IP protocol can be summarized as follows:

1. The mobility agents (HA and FA) in the network broadcast their avail-ability through agent advertisement packets.
2. The mobile node, after connecting to a network, receives information about the mobility agents through the agent advertisement broadcasts. Alternatively, the mobile nodes can solicit the agent information if no broadcasts have been received.
3. The mobile node determines the network it is attached to. If it is connected to the home network, it operates without mobility services. If it is returning back to the home network, the mobile node deregisters itself with the HA and operates without mobility services.
4. If the mobile node is attached to a foreign network, a care-of-address is obtained from the FA.

5. The mobile node operating from a foreign network registers itself with its home agent. The foreign agent then acts as a relay in this registration process.

6. When the mobile node is away from its home network, datagrams destined to the mobile node are intercepted by the home agent, which then tunnels these datagrams to the mobile node's care-of-address. The tunneled pack-ets destined to the mobile node are detunneled either by the foreign agent or by the mobile node itself. In the latter case, the mobile node obtains a temporary IP address on the foreign agent network to be used for forwarding. This can be done using the IETF Dynamic Host Configuration Protocol (DHCP).

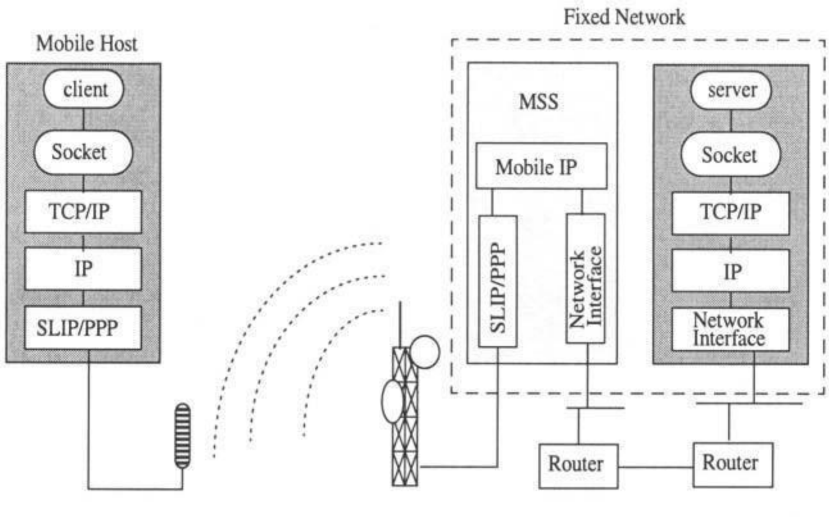


Figure 7.1 Mobile Internet architecture using Mobile-IP

7. The datagrams originating from the mobile node are routed in the normal fashion. The foreign agent may act as a default router in this case.

The routing path of a datagram sent from a fixed host to a mobile node is as follows: (1) the datagram is sent from the fixed host to the home agent using standard IP routing; (2) the home agent encapsulates the received datagram inside another datagram and sends it to the foreign agent (IP-in-IP tunneling); (3) the encapsulated IP packet is received by the foreign agent, DE capsulated, and forwarded to the mobile node; (4) the mobile node replies by sending a datagram to the fixed host through the foreign agent. The Mobile IP protocol stack on the fixed network and on the mobile unit is depicted in Figure 7.1.

The Mobile Host Protocol, known as Mobile-IP, is an evolving standard being developed by the IETF Working Group on IP Routing for Wireless/Mobile Hosts. Standards for both IPv4 and IPv6 have been proposed and are being reviewed for enhancements in scalability and performance. In particular, the triangular routing between the mobile node, the home agent, and the foreign agent (that must be performed every time the mobile node switches over to another communication cell) is a bottleneck that is being removed in IPv6 [81]. Packets addressed to the mobile node's home address are transparently routed to its care-of address. The optimized protocol enables IPv6 nodes to cache the binding of a mobile node's home address with its care-of address, and to then send any packets destined for the mobile node directly to it at this care-of address.

The Mosquito Net project at Stanford aimed at relaxing the requirement of foreign agent availability. Mosquito Net follows the IETF specification of Mobile-IP to support host mobility, but does not require FA support in foreign networks visited by the mobile node.

More details on achieved and ongoing efforts in Mobile IP and its routing optimization can be found in [62, 60, 14, 11, 88, 80].

1.1.3.1 Support for Ad-Hoc Mobility

An ad-hoc mobile network is a collection of wireless mobile nodes forming a temporary network without the aid of any established infrastructure or centralized administration. Examples of ad-hoc networks include wireless portable devices of a group of collaborator, such as an emergency team in a disaster area. No routing is needed between ad-hoc nodes which are within transmission range of each other's. Otherwise, additional nodes must be used to form a sequence of hops from the source to the destination. Routing algorithms in the ad-hoc environment are therefore a necessary support for this mode of mobile connection.

Traditional routing algorithms used in wire line networks use distance vector or link state routing algorithms, which rely on periodically broadcasting routing advertisements by each router node. The distance vector algorithm [55] broad- casts its view of the distance from a router node to each host. The link state routing algorithm broadcasts its view of the adjacent network links. Neither algorithms is suitable for the ad-hoc environment because periodic broadcasts will drain battery power quickly.

Research in ad-hoc routing is dedicated to finding algorithms that avoid the needless battery consumption and the inefficient use of the wireless bandwidth. Dynamic source routing is one such algorithms due to Johnson and Maltz. It allows for route discovery, route maintenance, and the use of route caches. To discover an available route, a source node sends out a route request packet indicating the source, the target nodes, and a request identifier. When a mobile node receives a route request packet, it checks a list of recently processed requests. If a request is found for the same source and request id, the request is dropped and no further action is taken. Otherwise, the address of the node servicing the request is added to the route request packet before the packet is re-broadcasted. However, if the address of the node servicing the request is identical to the target node address, the requested route is discovered, and a reply is sent to the source node.

Due to unpredictable node mobility, cached routes may become incorrect. Route maintenance is therefore necessary in this environment. This is achieved by requiring nodes routing packets to acknowledge successful forwarding and to send error messages to the source node if a route ceases to exist. Active monitoring such as MAC -level acknowledgements, as well as passive monitoring (listening to nearby broadcast, in a promiscuous mode), can be used in route maintenance.

Other recent ad-hoc routing protocols that can be found in the literature include the on-demand distance vector routing, the Location-Aided Routing (LAR) algorithm, and the Distance Routing Effect Algorithm.

1.1.4 Cellular Digital Packet Data (CDPD)

CDPD is a connectionless multi-network protocol, proposed originally by the CDPD Forum (now called the WDF Forum). It is based on the early versions of Mobile-IP . The idea behind CDPD is to share unused channels in existing Advanced Mobile Phone Systems (AMPS) to provide up to 19.2 kbps data channel.

Even though CDPD and Mobile-IP are similar, their terminologies are different. CDPD follows the OSI model terminology. For example, the mobile node is called a Mobile End-System (M-ES); the home and foreign agents are called Mobile Home and Mobile Serving Functions (MHF and SF respectively) and re- side in a mobile data intermediate system (MD-IS). A Mobile Database Station (MDBS) is also defined which deals with the air link communications and acts as a data link layer relay between the M-ES and the serving MD-IS . Two pro-tools, the Mobile Node Registration Protocol (MNRP) and the Mobile Node Location Protocol (MNLP), are responsible for registration of the M-ES with its home MD-IS and the proper routing of packets destined for the M-ES.

The main resemblance between CDPD and Mobile-IP is in the triangular rout- ing approach between the mobile node and the home and foreign agents. The main differences can be summarized as follows:

- The user's IP address must be assigned by the CDPD service provider. Mobile IP makes no such assumptions.
- Mobile IP allows the mobile node to also be a foreign agent. Combining the M-ES and the Serving MD-IS was not considered and is not practical in CDPD.
- CDPD's mobility tunneling is based on CLNP. Mobile IP's mobility tunneling is based on the IP-in-IP protocol, which is IP-based.
- Mobile IP operates completely above the data link layer. CDPD mobility, on the other hand, is mostly above the data link layer.
- Since the infrastructure of the CDPD network is closed there are less secrudity considerations for CDPD.

While the standardization process of Mobile IP has been progressing rather slowly, CDPD has been deployed for a few years now, and is receiving the support of major AMPS carriers. However, due to its lack of openness, the future of CDPD deployment and/or acceptance can only be guessed.

What is CDPD?

Cellular Digital Packet Data (CDPD) is a technique used for transmitting small chunks of data, commonly referred to as packets, over the cellular network in a reliable manner. It allows users to send and receive data from anywhere in the cellular coverage area at any time, quickly and efficiently. CDPD technology provides extensive, high speed (data can be sent over the Airlink at a rate of 19.2 kilobits per second), high capacity, cost effective data services to mobile users. With this technology, both voice and data can be transmitted over existing cellular channels.

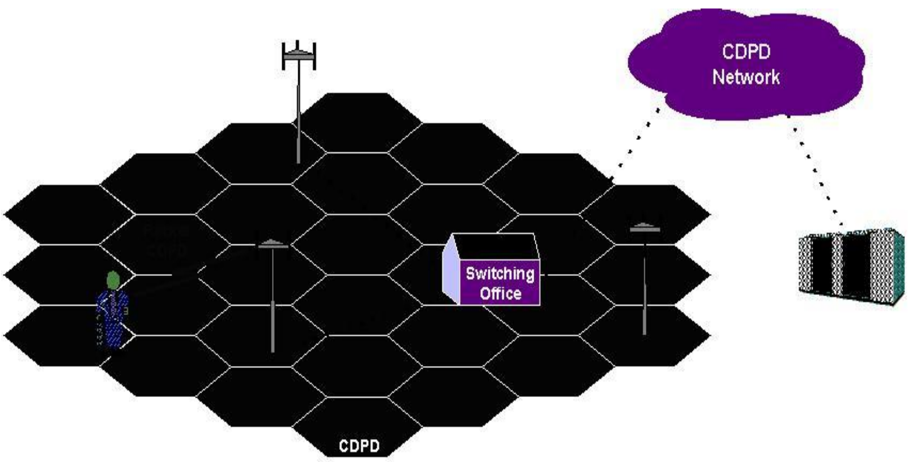
What defines the CDPD Network

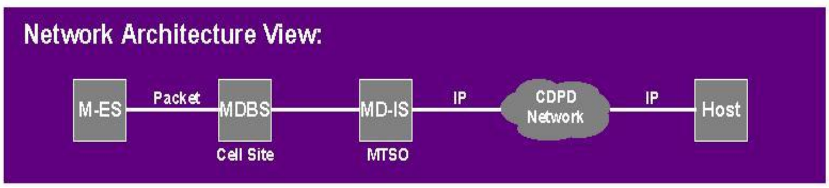
By building CDPD as an overlay to the existing cellular infrastructure, and using the same frequencies as cellular voice, carriers are able to minimize the capital expenditures required to offer the service while offering the same coverage area (footprint) their customer base has grown accustomed to. In comparison, it costs approximately $1 million to build out a new cellular cellsite and only about $50,000 to build the CDPD overlay to an existing site.

The CDPD overlay network is made up of a combination of key components that operate together to provision the overall service. These components are described below:

- The Mobile End System (M-ES) which is defined as any mobile computing device which is equipped with a CDPD modem (e.g. a PC). Unlike voice cellular phones, the decision to initiate a transfer, or hand-off from one cell to another cell is under the control of the CDPD M-ES itself, as it is the M-ES which is responsible for monitoring the received signal strength of the cellular channels being used.
- The Fixed End System (F-ES) which is defined as a stationary computing device, such as a host computer or an on-line information service.
- The Mobile Data Intermediate System (MD-IS) which is a stationary network component with similar responsibilities to the cellular voice switch. It is responsible for keeping track of the M-ES's location and routing data packets to and from the CDPD network and M-ES appropriately. It has been referred to as the "brain" of the network, because of its functionality. Not only is it responsible for ensuring that an M-ES is valid to log on to the network, but it also stores information on the M-ES's last known location, traffic statistics and billing information.
- The Mobile Data Base Station (MDBS) is primarily responsible for RF channel management. It is located at the voice cell sites and is responsible to instruct the M-ES to "hop" to new channels for continued communication in the event voice communication (which is the priority traffic) is detected. It also handles the leg work for the M-ES in locating new channels when a hand off is required between cell sites.
- The Intermediate System (IS) is made up of (off the shelf) routers which are CDPD compatible with the primary responsibility for relaying the data packets.

The way these componenets interact with each other can be seen from the graphic depiction below:





How does CDPD work?

To effectively integrate voice and data traffic on the cellular system without degrading the level of service provided to the voice customer, the CDPD network implements a technique called channel hopping. The way this works is that when a CDPD mobile data unit desires to initiate data transmission, it will check for availability of a cellular channel. Once an available channel is located, the data link is established.

As long as the assigned cellular channel is not needed for voice communications, the mobile data unit can continue to transmit data packet bursts on it. However, if a cellular voice customer initiates voice communication, it will take priority over the data transmission. At such time, the mobile data unit will be advised by the Mobile Data Base Station (which is the CDPD serving entity in the cell and constantly checks for potential voice communication on the channel) to "hop" to another available channel. In the event that there are no other available channels, then data transmission will be temporarily discontinued. It is important to note that these channel hops are completely transparent to the mobile data user. As far as the user can see, there is only one data stream being used to complete the entire transmission.

1.1.5 The GSM General Packet Radio Service (GPRS)

GPRS is a GSM packet data service developed by the European Telecommunication Standards Institute (ETSI) as part of GSM phase 2+ developments. The goal of GPRS was to support data transfer rates higher than the 9.6 kbps achieved through GSM's circuit switching technology. Unlike Mobile-IP, GPRS is not restricted to IP packet data protocols, and offers connection to standard protocols (such as TCP/IP, X.25, and CLNP) as well as specialized data packet protocols. Mobile-IP, however, influenced the design of Mobility management in GPRS.

Figure 7.2 shows the architecture of a GSM system that uses GPRS. In addition to the Base Transceiver Station (BTS), Base Station Controller (BSC), and the Mobile Switching Center (MSC), a new logical network node called the GPRS support node (GSN) was introduced in order to create an end-to-end packet transfer mode. Physically, the GSN can be integrated with the mobile switching center (MSC), or it can be a separate network element based on the architecture of data network routers. GSN is a mobility router that provides connection and interoperability with various data networks, mobility management with the GPRS registers, and delivery of data packets to MSs , independently of their locations.

One GSN is designated the Gateway GSN (GGSN) and acts as a logical inter- face to external packet data networks. The GGSN is similar to the home agent in Mobile-IP . It updates the location directory of the mobile station (MS) using routing information supplied by the Serving GSN node (SGSN). The latter is similar to the foreign agent in Mobile-IP. GGSN also routes the external data network protocol packet encapsulated over the GPRS backbone to the SGSN currently serving the MS. It also de-capsulate and forwards external data net- work packets to the appropriate data network and handles the billing of data traffic.

The SGSN is responsible for the delivery of packets to the mobile stations within its service area. The main functions of the SGSN are to detect new GPRS MSs in its service area, handle the process of registering the new MSs along with the GPRS registers, send/receive data packets to/from the GPRS MS, and keep a record of the location of MSs inside of its service area. The GPRS register acts as a database from which the SGSNs can ask whether a new MS in its area is allowed to join the GPRS network. For the coordination of circuit and packet switched services, an association between the GSM MSC and the GSN is created. This association is used to keep routing and location area information up-to-date in both entities.

1.1.6 Security and Authentication Issues in Mobile Networks

In a mobile computing environment, it is desirable to protect information about the movements and activities of mobile users from onlookers. In addition to the basic security concerns in wire line systems (authentication, confidentiality, and key distribution), a new issue is the privacy and anonymity of the user's movement and identity. In fact, a typical situation arises when a mobile user registers in one domain (home domain) and appears in a different foreign do- main; the user must be authenticated and his solvency must be confirmed. Usually during this process the user has to provide a non-ambiguous identity to his home domain and has to verify it. If no care is taken, this identity can be tapped on the air interface in a cellular environment or through the signaling protocols exchanged on the registered wired network.

In CDPD , all the mobility management, as well as security-related activity, are concentrated in the Massage-Data Intermediate System (MD-IS) . Each MD-

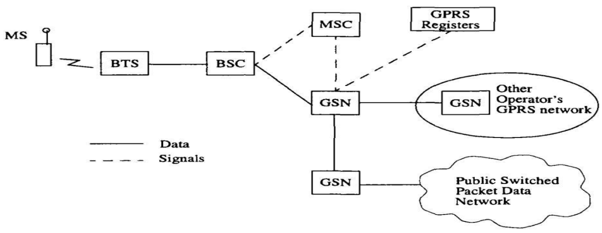


Figure 7.2 General Packet Radio Service

IS controls an area covered by a number of base stations. Upon arrival to a new area, the mobile unit engages in a *Diffie-Hellamn* key exchange protocol with the local MD-IS. As a result, both parties obtain a shared secret key. Subsequently, the mobile unit encrypts its real identity (Network Equipment Identifier) and transmits it to the local MD-IS. This approach allows the local MD-IS to discover the real identity of the mobile unit. Unfortunately, the key exchanging protocol itself is not secure. This means that an active attacker masquerading as the local domain authority can engage in the key exchange protocol with the mobile unit and obtain a shared key.

2. EXPLAIN QUALITY OF SERVICE IN MOBILE NETWORKS

Mobile network protocols such as Mobile-IP and GPRS provide mobility transparency at the network layer level. This allows the higher layers of the protocol stack to be used unchanged. Unfortunately, there are ill consequences to this transparency that are mostly attributed to the constraints of the wireless and mobile environment. For example, transport layer protocols that rely heavily on timeout mechanisms for re-transmission, if used unchanged, will perform poorly under variable delays and limited bandwidth.

This is especially true for applications that require continuous-media streams. Existing session protocols are not of much use under frequent disconnections and reconnections of the same mobile computation. Similarly, existing presentation layer protocols are inappropriate to use unchanged, in the wireless and mobile environment. For example, a user with a limited display and limited battery PDA will not be able to browse the Web unless the presentation of the downloaded data is changed to suite her PDA's capabilities. Regardless of which particular upper layer in the protocol stack suffers the consequences of transparency, the effect on the end-user will always be felt as unacceptable fluctuations in the perceived QoS.

In this section, we describe the following three research efforts that address QoS concerns in the wireless and mobile environment.

- *Optimizing TCP/IP for Mobile Networks.* Transport or network layer solutions to get TCP/IP to work despite the fluctuations in the underlying network QoS. Solutions are not application-sensitive and do not address an overlay of heterogeneous networks.
- *QoS driven, high-level communication protocols.* Session and/or application layer protocols directly addressing QoS parameters. Solutions are sensitive to applications, but do not address network heterogeneity issues.
- *QoS driven, full protocol stacks.* All layers are aware of either QoS or the limitations introduced by mobility and by the wireless networks. Two research efforts will be discussed including, the BARWAN project and the Wireless Application Protocol (WAP) standard.

2.2.1 Optimizing TCP/IP for Mobile Networks

Since mobile users will need connection-oriented communication to obtain re- mote services, they will have to use transport protocols developed for the fixed network. Unfortunately, such protocols like TCP perform poorly when used un- modified in the mobile network. For example, TCP acknowledgment timeout is in the range of tens of milliseconds. A mobile unit crossing cell boundaries blanks out during a hand-off procedure that could last up to 1,000 milliseconds. This leads to sender timeouts and repeated re-transmissions.

Another source of re-transmission is the high error rate inherent in the wireless transmission characteristics. Another problem that can lead to performance degradation un- der standard TCP is bandwidth allocation under unpredictable mobility. An unpredicted number of mobile users can move into the same cell, thus competing on sharing the limited wireless link. Under this scenario, it is difficult to build applications or services that provide performance guarantees or quality of service. A few approaches have been proposed to optimize and extend the standard TCP protocol so that it can be used efficiently under a mobile network protocol such as Mobile IP.

2.2.1.1 Yavatkar et al

Yavatkar et al proposed an approach whereby the communication path be- tween the mobile end and the fixed end is split into two separate connections: one over the wireless link and another over the wired links. The connection over the wireless link may either use regular TCP or a specialized transport protocol optimized for better performance. The splitting of a connection is transparent to an application and no changes are necessary to protocol software on the stationary hosts. A new session layer protocol called Mobile Host Protocol (MHP) is introduced atop standard TCP. MHP compensates for wireless link characteristics and for host migration.

It is located at both the base station and the mobile host. An advantage of this approach is that performance degradation in TCP is limited to a "short" connection over the wireless hop, while traffic over the "long" connection over the wired network can be protected from the impact of erratic behavior over the wireless link. A second alternative is proposed in the same work which is similar to the MHP alternative except that MHP uses a specialized protocol instead of TCP over the wireless hop. The specialized protocol differs from standard TCP in that the former uses selective acknowledgement by the receiver, in which a bitmask is used to indicate all missing segments of the connection stream. This way, the recovery of all losses can be performed via a single round trip message, resulting in a better throughput performance.

Another approach similar to Yavatkar's is the I-TCP protocol (Indirect Trans- port Layer Protocol), which also splits the communication path between the mobile host and the fixed network host into two connections; the first be- tween the mobile host and the base station, over the wireless link, using the I-TCP protocol; and the second between the base station and the fixed network host using standard TCP.

3.2.1.2 Balakrishnan et al

Balakrishnan et al took a slightly different approach to improve the performance of TCP in the mobile network. They focused on the re-transmission behavior of TCP due to hand-off. They redesigned the network layer so that it caches packets at the base stations. Retransmission can therefore be performed locally between the base station and the mobile unit. The gain is that the erratic transmission characteristics of the wireless link are dealt with in isolation of the rest of the fixed network. Experimental evaluation showed a throughput increase of up to 20 times over standard TCP. Their results are based on the Lucent/NCR Waveland network.

2.2.1.3 Caceres et al

Similar research by Caceres and Iftode addressed the problem of communication pauses due to hand-off. They observed that such pauses are interpreted by standard TCP (Tahoe in their experiment) as packet losses due to congestion, which consequently causes retransmissions that get further timed out during the hand-off. They proposed using the fast re-transmission option available in TCP-Tahoe immediately after hand-off is completed. Their experimental verification shows clear smoothening of TCP performance during hand-off.

2.2.2 QoS Driven, High-Level Communication Protocols

Optimizing the behavior and performance of transport protocols is not sufficient to maintain the QoS required by applications. For example, most Web browsers use multiple TCP connections to access a multimedia page. While this parallelism achieves speedup in the fixed network, it is slow and inappropriate in the wireless and mobile environment. In addition to transport optimizations, what was found needed are application-aware (or application-specific) mechanisms to monitor, request, and maintain QoS from the application or user point of view. This section describes high-level, above-transport protocols that understands application QoS requirements and resource limitations.

2.2.2.1 The Loss Profile Approach

Seal and Singh considered the problem of unpredictable mobility and its effect on the degradation of the wireless communication performance. They addressed the case where the aggregate bandwidth required by all mobile units in an overloaded cell exceeds the cell's available bandwidth. Their mechanism is simple and relies on policies and measures for discarding parts of the data of the mobile users. Instead of discarding data in an arbitrary manner, guidelines are proposed to avoid discarding critical portions of the data. A Loss Profile is proposed and is defined to be a description, provided by the application, of an “acceptable” manner in which data for its connection may be discarded. The loss profile is used in the event of bandwidth reduction at the wireless end of the connection. An elaborate example of a loss profile is given on viewer perception of a video clip under data loss. The loss profile is used by a specialized session layer which is transparent to the application.

2.2.2.2 QEX: The QoS Driven Remote Execution Protocol

In the problem of fluctuations in the quality of service (QoS) in a federation of heterogeneous networks is addressed. The work describes a design of a distributed system platform that supports the development of adaptable services. The design allows services to tolerate the heterogeneity of the environment by dynamically adapting to changes in the available communication QoS.

The implementation of the distributed system is based on APM Ltd.’s ANSAware software suite, which is based on the ANSA architecture that has had some influence on the ISO Reference Model for Open Distributed Processing (RM-ODP). The purpose of this effort is to propose extensions to emerging distributed systems standards in order to support mobile services. The basic ANSAware platform is extended to support operation in the mobile environ- ment by introducing the notion of explicit bindings, which is a QoS-aware RPC protocol for objects called QEX.

Explicit bindings allow application programmers to specify QoS constraints on bindings between objects, and to detect violations of these constraints at run time. To support explicit bindings, a new remote procedure call protocol has been developed for ANSAware. The new RPC is able to maintain QoS information on the underlying communications infrastructure and to adapt to changes in the perceived QoS. Moreover, it is able, via explicit bindings, to pass on relevant QoS information to interested applications. This allows the applications themselves to adapt to changes in the QoS. Binding parameters include specification of parameters such as the desired throughput, latency, and jitter associated with the binding. Clients are returned a binding control interface as a result of an explicit bind operation. To control the QoS of the flow once the binding has been established, the control interface includes a pair of operations *setQoS()* and *getQoS()*. These operations take as arguments a set of QoS parameters which can then be passed by the stream binding to the underlying transport protocol. A call-back mechanism is also provided to inform client objects of QoS degradations reported by the underlying transport service.

The work is being put to test using an adaptive collaborative mobile application designed to support field engineers in the U.K. power distribution industry.

2.2.3 QoS Driven, Full Protocol Stacks

Future mobile services will be built upon federations of heterogeneous networks maintained and administered by different providers. The mobility of users will force an application to migrate along overlays of networks that vary in their bandwidth, latency, range, and transmission characteristics. Unless the application adapts to variations in the network overlay, the application performance is bound to suffer. A network overlay can include a cellular network, a personal communication system (PCS), a wireless LAN, an Internet connection, and/or a satellite communication loop, among other networks. In addition to the het- erogeneity of networks, the heterogeneity of the mobile platforms imposes a great impediment to mobile application portability. Unless applications adapt to the capabilities and limitations of the mobile computer with respect to the type and media of communicated data, applications will remain proprietary to the specific mobile computer platforms they were originally designed for. This section describes a research project that proposes a full stack solution as an overlay network stack atop a heterogeneous collection of wireless subnets. This section also describes an ongoing standardization effort called WAP that aims at proposing a specification of a full ISO/OSI-like network stack that is wireless and mobile aware.

2.2.3.1 BARWAN: The Wireless Overlay Network Architecture

The BARWAN project at the University of California at Berkeley developed an architecture that supports applications' graceful adaptation to the available bandwidth and latency of the wireless network. The architecture assumes an overlay of various wireless networks ranging from regional-area, wide-area, metropolitan-area, campus-area, in-building, and in-room wireless networks. A testbed of wireless overlay network management that supports media-intensive applications has been used to demonstrate the adaptability features of BARWAN. The testbed that has been developed in the San Francisco Bay Area includes the participation of over six local carriers including Nextel and Metricom. The testbed integrates the participants' networks and allows full coverage of the greater Bay Area. The BARWAN architecture is gateway- centric, meaning it provides gateway connections from the mobile host to each participating wireless networks. Medical imaging applications have been developed to drive the testbed.

The layered architecture of BARWAN is shown in Figure 7.3. It shows all layers designed for wireless overlay network integration and for providing application support. The lowest layer is the wireless overlay subnets, which are the carrier networks including data link interface, and possibly carrier network rout- ing. The details of this layer depends on the specific subnets being integrated. Next is a layer called the Overlay Network Management Layer which includes network and transport functionalities including location tracking, QoS-based hand-off management, other QoS services, and connection-oriented transport mechanisms. The next higher up layer is the Session Management Layer which provides a "transactional" transport (called message-oriented interface).

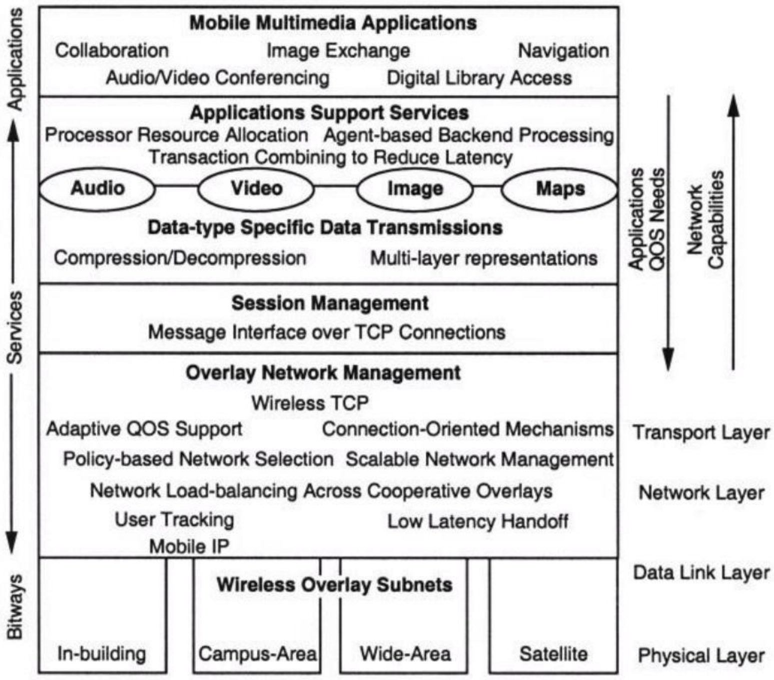


Figure 7.3 The BARWAN conceptual layered architecture

The layer attempts to optimize transport connections related to the same application by session sharing whenever possible. On top of the session layer is the Application Support Services including support for various data types and continuous media such as audio and video. Finally, the mobile multimedia application is on top of the stack. Figure 7.3 also shows how the quality of services needs pass down the layers from applications towards the network management layers, while information about network capabilities propagates up the layers.

2.2.3.2 The Wireless Application Protocol (WAP)

In June 1997, Ericsson, Nokia, Motorola, and Phone.Com (previously Unwired Planet) formed a consortium for the standardization of an open middleware architecture for wireless application. The objective was to create the specification of a wireless application environment and a wireless ISO/OSI-like protocol stack. The goal was to provide the needed interoperability to connect different portable devices, via heterogeneous wireless networks, into the internet and corporate intranets. The focus was to bring the internet content and advanced services to digital cellular phones and other hand-held devices such as smart communicators and PDAs.

In January 1998, the consortium created a nonprofit company named the WAP Forum with the mission of enabling: (1) interoperability across heterogeneous portable devices, wireless networks, and internet contents, and (2) portability of third party wireless software and applications across different portable devices that are WAP-compliant. Currently, the WAP Forum is creating a set of specifications for the Wireless Application Environment and for each layer in the WAP protocol stack.

The architectural infrastructure of WAP is depicted in Figure 3.4 and consists of: (1) hand-held devices ranging from digital cellular phones, to smart communicators such as the Nokia 9000, to palmtop computers. Only devices that will be WAP-compliant (implement the WAP stack and wireless application environment) are part of the WAP infrastructure, (2) Wap-compliant wireless networks, which are carrier networks augmented with the WAP stack on top of the air link interfaces, (3) WAP-compliant internet information providers such as Web servers, that must conform to levels of presentations of information suitable to the capabilities of the hand-held device requesting the information, and (4) WAP-compliant TeleVAS providers.

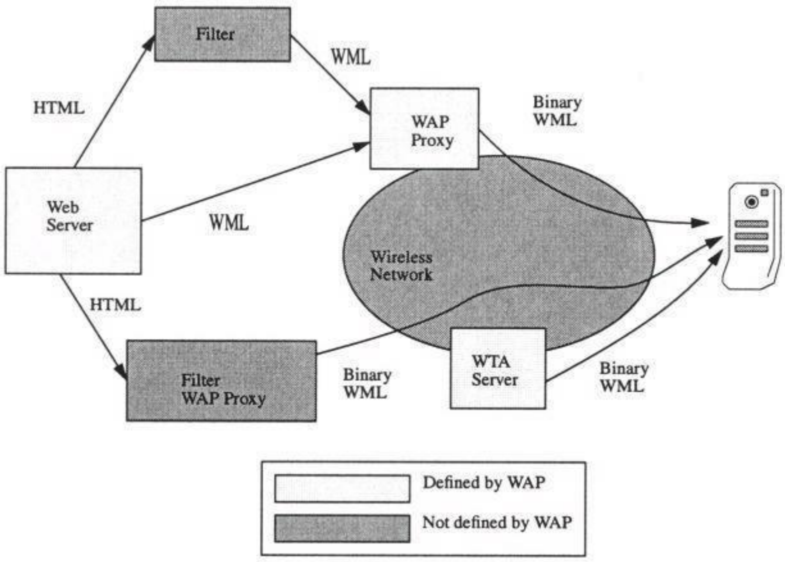


Figure 7.4 The WAP architectural infrastructure

In the heart of the WAP standard is the WAP protocol stack shown in Fig- ure 7.5. The stack is similar to the ISO/OSI stack and consists of a lowest layer containing air link interfaces such as GSM's GPRS, CDPD, D-AMPS, among others. This lowest layer corresponds to both the physical and the data link layers combined in the OSI stack. On top of the air link is the transport layer, in which datagram and connection-oriented streams are supported. Also, *transactional* connections are supported to enable electronic commerce applications. This layer corresponds to both the network and the transport layers of the OSI stack combined. On top of the transport, WAP dedicates a layer for security. This includes encryption, authentication, and capabilities. On top of security is the session layer which is responsible for enabling multi-tasking on the hand-held device. This is because multiple connections can be maintained as multiple sessions managed by the session layer. The session layer, which is the most elaborate layer, also contains critical QoS features including:

- exception mechanisms to allow applications to register interest in QoS related network events and parameter thresholds. This allows the application to be mobility-aware, by using QoS API to program how to adapt to changes in the environment.
- mechanisms for capability and content negotiation. This will enable the WAP stack itself to partner through its pieces (on the fixed network, on the wireless network gateways, and on the hand-held device) to perceive and adapt to the mobility and the changes in network characteristics. When certain information content is being delivered, the WAP stack negotiates with the device the capability to receive and display the contents. The negotiation decides for the feasibility of the transfer and for the level of filtering that might be needed to deliver the the information while maintaining QoS.

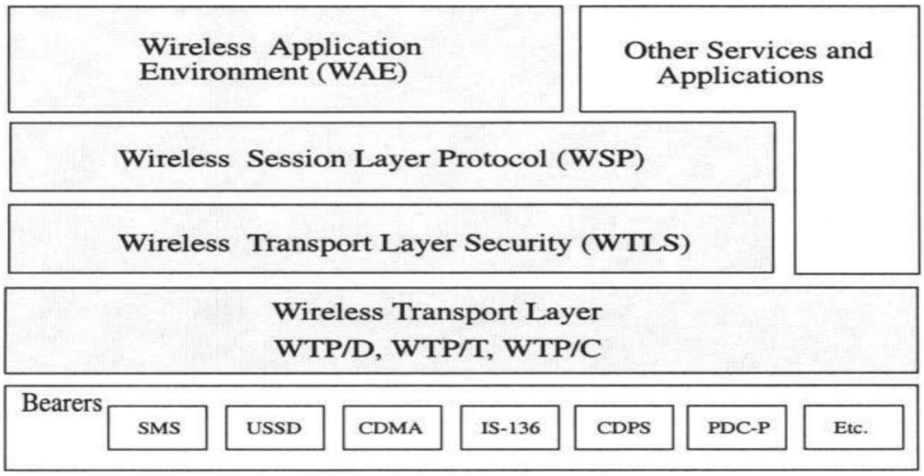


Figure 7.5 The WAP protocol stack

The first capability provides applications with the environment awareness needed to initiate QoS adaptations. The second capability, on the other hand, provides the system with automated awareness mechanisms not only of the environment, but also of the device capabilities and the characteristics of the information content.

Another standardization effort similar to WAP is the Mobile Network Computer Reference Profile (MNCRF), which is based on the NCRF standard developed by the Open Group. The first draft of the standard has been released addressing the unique requirements of mobile network computing. Details of this initiative are available in a white paper and a reference specification document.

3. EXPLAIN MOBILE ACCESS TO THE WORLD WIDE WEB

More and more users are becoming increasingly dependent on information they obtain from the World Wide Web. Users are also demanding ubiquitous access, anytime, anywhere, to the information they rely on. Several research efforts explored the problems associated with wireless access to the Web. Most solutions used a Web proxy that enabled Web browsing applications to function over wireless links without imposing changes on browsers and servers. Web proxies are also used to prefect and cache Web pages to the mobile client's machine, to compress and transform image pages for transmission over low-bandwidth links, and to support disconnected and asynchronous browsing operations.

3.3.1 The Wireless WWW (W4)

In a prototype consisting of commercially available PDAs and a wireless LAN has been used to provide a "proof of concept" for the Wireless World Wide Web (W4). A simplified version of Mosaic was ported to the PDA for the purpose of experimenting with response time performance and to sort out design choices. A PDA cache was used to improve the performance.

3.3.2 Dynamic Documents

The concept of dynamic documents was introduced in as an approach to extending and customizing the WWW for mobile computing platforms. Dynamic documents are programs executed on a mobile platform to generate a document; they are implemented as Tcl scripts as part of the browser client.

A modified version of the NCSA Mosaic browser was used to run the dynamic documents it retrieves through a modified Tcl interpreter. The interpreter is designed to execute only commands that do not violate safety. By using dynamic documents, an adaptive e-mail browser that employs application-specific caching and prefetching is built. Both the browser and the displayed e-mail messages are dynamically customized to the mobile computing environment in which they run.

Dynamic documents can solve the problem of limited resources in the mobile host. For example, the Tcl script could be a filter that reduces an incoming image so that it fits the screen size or resolution. Unfortunately, dynamic documents being placed at the client side are not wireless-media sensitive. This is because filtering occurs after all transmitted information is received by the client. Although caching and prefetching can alleviate some of the communication overhead, excess data (that would be reduced by the dynamic document) is, however, communicated, leading to inefficient utilization of the wireless bandwidth.

3.3.3 Dynamic URLs

The Mobisaic project at the University of Washington extends standard client browsers to support dynamic URLs and active documents. The Mosaic Web client and the URL syntax are modified so that when the user traverses a dynamic URL, the client resolves any references to dynamic information it may contain and sends the result back to the server. This is helpful in defining location-sensitive resources. Active documents are Web pages that notify the client browser when dynamic information changes. This feature also supports location-sensitive information by keeping the mobile client aware of service relocation or of services offered by a mobile server.

3.3.4 Mobile Browser (MOWSER)

In [65], a design of a mobile-aware Web browser is discussed. The design is based on a mediator server that filters retrieved information according to the limitations of the mobile unit. Color, resolution, display mode, sound capability, and maximum file size are among the factors considered. The browser, called MOWSER, connects to two servers in the fixed network. The first is the preference server that maintains the user profile; the second is a proxy server that implements all the filtering indicated by the preference server.

MOWSER assumes that the user is aware of the mobile unit limitations, which in a way sacrifices transparency. Similar to the dynamic document approach, MOWSER does not directly consider the limitations of the wireless media (although the maximum file size indirectly preserves the limited bandwidth).

3.3.5 Web Express

Web Express uses the proxy approach to intercept and control communications over the wireless link for the purposes of reducing traffic volume and optimizing the communication protocol to reduce latency. Two components are inserted into the data path between the Web client and the Web server: (1) the Client Side Intercept (CSI) process that runs in the client mobile device and (2) the Server Side Intercept (SSI) process that runs within the wired and fixed network (see Figure 7.6). The CSI intercepts HTTP requests and, together with the SSI, performs optimizations to reduce bandwidth consumption and transmission latency over the wireless link. From the viewpoint of the browser, the CSI appears as a local Web proxy that is co-resident with the Web browser.

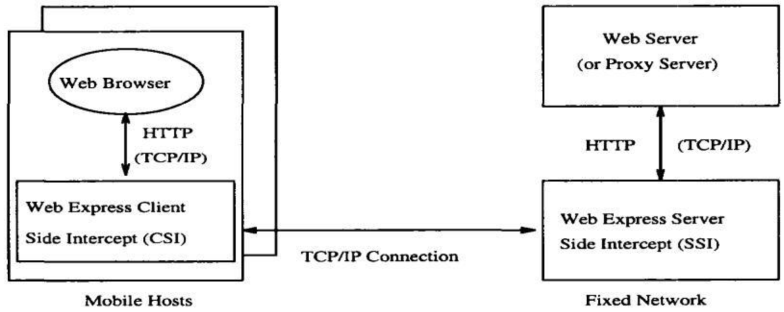


Figure 7.6 WebExpress proxy intercept model

On the mobile host, the CSI communicates with the Web browser over a local TCP connection (using the TCP/IP "loopback" feature) via the HTTP protocol. Therefore, no external communication occurs over the TCP/IP connection between the browser and the CSI. No changes to the browser are required other than specifying the (local) IP address of the CSI as the browser's proxy address. The CSI communicates with an SSI process over a TCP connection using a reduced version of the HTTP protocol. The SSI reconstitutes the HTML data stream and for- wards it to the designated CSI Web server (or proxy server). Likewise, for responses returned by a Web server (or a proxy server), the CSI reconstitutes an HTML data stream received from the SSI and sends it to the Web browser over the local TCP connection as though it came directly from the Web server.

The proxy approach implemented in Web Express offers the transparency ad- vantage to both Web browsers and Web servers (or proxy servers) and, there- fore, can be employed with any Web browser. The CSI/SSI protocols facilitate highly effective data reduction and protocol optimization without limiting any of the Web browser functionality or interoperability. Web Express optimization methods are summarized below:

- *Caching:* Both the CSI and SSI cache graphics and HTML objects. If the URL specifies an object that has been stored in the CSI's cache, it is returned immediately as the response. The caching functions guarantee cache integrity within a client-specified time interval. The SSI cache is populated by responses from the requested Web servers. If a requested URL received from a CSI is cached in the SSI, it is returned as the response to the request.

- *Differencing:* CSI requests might result in responses that normally vary for multiple requests to the same URL (e.g., a stock quote server). The concept of differencing is to cache a common base object on both the CSI and SSI. When a response is received, the SSI computes the difference between the base object and the response and then sends the difference to the CSI. The CSI then merges the difference with its base form to create the browser response. This same technique is used to determine the difference between HTML documents.

- *Protocol reduction:* Each CSI connects to its SSI with a single TCP/IP connection. All requests are routed over this connection to avoid the costly connection establishment overhead. Requests and responses are multiplexed over the connection.

- *Header reduction:* The HTTP protocol is stateless, requiring that each request contain the browser's capabilities. For a given browser, this information is the same for all requests. When the CSI establishes a connection with its SSI, it sends its capabilities only on the first request. This information is maintained by the SSI for the duration of the connection. The SSI includes the capabilities as part of the HTTP request that it forwards to the target server (in the wire line network).

SRI VENKATESHWARAA COLLEGE OF ENGINEERING & TECHNOLOGY
(Approved by AICTE, New Delhi & Affiliated to Pondicherry University, Puducherry.)
13-A, Villupuram – Pondy Main road, Ariyur, Puducherry – 605 102.
Phone: 0413-2644426, Fax: 2644424 / Website: www.svcetpondy.com

Department of Computer Science and Engineering

Subject Name: **MOBILE COMPUTING** Subject Code: **CS E84**

UNIT IV

Mobile Data Management: Mobile Transactions - Reporting and Co Transactions –Kangaroo Transaction Model - Clustering Model –Isolation only transaction – 2 Tier Transaction Model – Semantic based nomadic transaction processing.

2 Marks

1. Classify the mobile data management.

Mobile data access can be broadly classified into two categories:

(1) Data access in mobile client/server.

(2) Data access in ad-hoc networks.

2. What is meant by mobile transaction?

A mobile transaction is a long-live transaction whose locus of control moves along with the mobile user. Mobile transactions may access remote data wirelessly, through a weak connection, or may access local replicas of data in disconnected mode. The differences between mobile and distributed transaction management are significant because their goals are different.

3. Write about ACID properties.

ACID (Atomicity, Consistency, Isolation, Durability)

ACID (Atomicity Consistency Isolation Durability) is a concept that Database Professionals generally look for when evaluating databases and application architectures. For a reliable database all this four attributes should be achieved. **Atomicity** is an all-or-none proposition.

Consistency guarantees that a transaction never leaves your database in a half-finished state.

Isolation keeps transactions separated from each other until they're finished.

Durability guarantees that the database will keep track of pending changes in such a way that the server can recover from an abnormal termination.

4. Differentiate between distributed transaction and mobile transaction.

In distributed transactions, the main goal is maximizing availability while achieving ACID properties. In mobile transactions, maximizing reliability while achieving some sort of consistency is the main goal.

5. Write short notes on reporting and co-transaction.

This model is based on the Open Nested transaction model. A computation in the mobile environment is considered to consist of a set of transactions, some of which may execute on the mobile node and some of which may execute on the fixed host. The model addresses sharing of partial results while in execution, and maintaining computation state in a fixed node so that the communication cost is minimized while the mobile host relocates. The model proposes to modify Reporting and Co-Transactions to suit mobile environments. The model defines a mobile transaction to be a set of relatively independent transactions which interleave with other mobile transactions.

6. What are the types of transactions? Atomic transactions

Compensable transactions

Reporting transactions Co-transactions

7. What is mean by Kangaroo transaction model? APRIL/MAY2014

This model introduced in is based on global transactions and the split transaction models, where transaction relocation is achieved by splitting the transaction at the point of hand-off. A mobile transaction (called Kangaroo transaction) is considered a global transaction in a multi database environment. A Kangaroo transaction (KT) is a global transaction that consists of a set of Joey Transactions (JT). A JT is associated with the base station or the cell in which it executes.

8. Define clustering model.

This model described in assumes a fully distributed system. The database is divided into clusters . A cluster defines a set of mutually consistent data. Bounded inconsistencies are allowed to exist between clusters. These inconsistencies are finally reconciled by merging the clusters. The model is based on the open nested transaction model, extended for mobile computing.

9. What is meant by m-degree consistent?

Consistency between clusters can be defined by an m-degree relation, and the clusters are said to be m-degree consistent. The m-degree relation can be used to define the amount of deviation allowed between clusters. In this model, a mobile transaction is decomposed into a set of weak and strict transactions. The decomposition is done based on the consistency requirement.

10. What is mean by isolation only transaction.

The Coda file system at CMU provides an application-transparent file system for mobile clients by using file hoarding and optimistic concurrency control. A proxy logs all updates to the file system during disconnection and replays the log on reconnection. Automatic mechanisms for conflict resolution are provided for directories and files through the proxy and the file server. Hoarding is based on user-provided, prioritized list of files. Periodically, the proxy walks the cache to ensure that the highest priority files are present and consistent with the server. Coda provides Isolation-only Transactions (IOT) to automatically detect read/write conflicts that could occur during disconnection.

11. Define 2 tier transaction model.

A two-tier replication scheme has been proposed in whereby mobile disconnected applications are allowed to propose tentative update transactions. On connection, tentative transactions are applied to (re-processed at) the master data copy in the fixed network. At the re-processing stage, application semantics are used (such as finding commutative operations) to increase concurrency. To reduce re-processing costs that can be high in certain occasions, the work in uses a history-based approach

12. Write short notes on semantic based nomadic transaction process.

The semantics-based mobile transaction processing scheme views mobile transactions as a concurrency and cache coherence problem. It introduces the concepts of fragment able and reorders able objects to maximize concurrency and cache efficiency exploiting semantics of object operations. The model assumes the mobile transaction to be long-lived with unpredictable disconnections. Traditional definitions of concurrency and serializability are too restrictive for most operations. Commutatively of operations is an important property which allows concurrent operations on an object. If certain operations on an object are commutative, then the database server can schedule these operations in an arbitrary manner.

13. Explain briefly about compensable transaction.

Atomic transactions whose effects cannot be undone at all. When ready to commit, the transaction delegates all operations to its parent. The parent has the responsibility to commit or abort the transaction later on.

14. What is atomic transaction?

Atomic transactions are the Normal components and may be compensable with atomic compensating dual steps.

15. What is reporting transaction?

Reporting transactions can make its results available to the parent at any point of its execution. It could be a compensating or a non-compensating transaction.

16. Define co-transaction.

Co-transactions: behave in a manner similar to the co-routine construct in programming languages. Co-transactions retain their current status across executions; hence they cannot be executed concurrently.

17. What is meant by RDO?

An RDO is an object (code and data) with a well-defined interface that can be dynamically loaded into a mobile client from a server computer, or vice versa, to reduce client-server communication requirements, or to allow disconnected operation.

18. What is meant by QRPC?

Queued remote procedure call is a communication system that permits applications to continue to make non-blocking remote procedure calls even when a mobile client is disconnected; requests and responses are exchanged upon network reconnection.

19. Explain Mobile Data Management strategy.

A mobile data management strategy is a structure imposed on a complex data model that is to be navigated by a user on a mobile device. This is a relatively new process born from the popularity of mobile applications that requires a flexible and in-depth navigation structure. There are several methods for such strategy, each describing approaches to a variety of tasks or activities.

11 Marks

1. EXPLAIN MOBILE DATA MANAGEMENT.

Mobile data access can be broadly classified into two categories: (1) data access in mobile client/server, and (2) data access in ad-hoc networks. Several research projects from each category are presented in the following subsections.

Mobile Client/Server Data Access

In the first category, mobile data access enables the delivery of server data and the maintenance of client-server data consistency in a mobile and wireless environment. Efficient and consistent data access in mobile environments is a challenging research area because of the weak connectivity and resource constraints. The data access strategies in a mobile information system can be characterized by delivery modes, data organizations, and consistency requirements, among other factors. The mode for server data delivery can be *server-push, client-pull,* or a hybrid of both. The server-push delivery is initiated by server functions that push data from the server to the clients. The client-pull delivery is initiated by client functions which send requests to a server and "pull" data from the server in order to provide data to locally running applications. The hybrid delivery uses both server-push and client-pull delivery. The data organizations include mobility-specific data organizations like mobile database fragments in the server storage and data multiplexing and indexing in the server-push delivery mode. The consistency requirements range from weak consistency to strong consistency.

Broadcast Disks: A Server PUSH Approach

When a server continuously and repeatedly broadcasts data to a client commu-nity, the broadcast channel becomes a "disk" from which clients can retrieve data as it goes by. The broadcasting data can be organized as *multiple disks* of different sizes and speeds on the broadcast medium. The broadcast is created by multiplexing chunks of data from different disks onto the same broadcast channel. The chunks of each disk are evenly interspersed with each other. The chunks of the fast disks are repeated more often than the chunks of the slow disks (see Figure 7.7). The relative speeds of these disks can be adjusted as a parameter to the configuration of the broadcast. This use of the channel effectively puts the fast disks closer to the client while at the same time pushing the slower disks further away

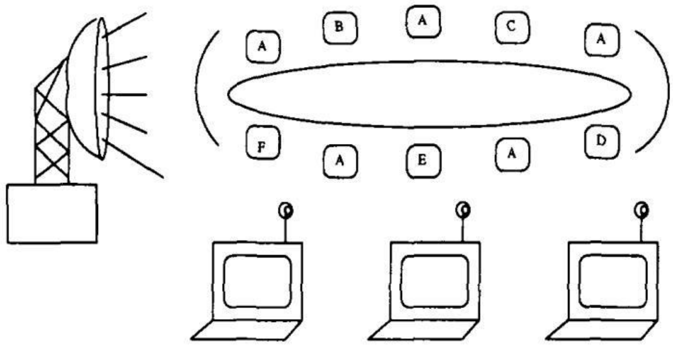


Figure 7.7 A simple broadcast disk

Odyssey: A Client PULL Approach

Odyssey is a CMU research project led by M. Satyanarayanan. It ad- dresses an application-aware adaptation approach to deal with application diversity and concurrency in mobile environments. The application-aware adaptation is implemented with the support of system-coordinated, type-specific operations. It supports concurrent execution of diverse mobile applications that execute on mobile clients, but read or update remote data on servers. The data accessed by an application may be stored in one or more general-purpose repositories such as file servers, SQL servers, or Web servers. It may also be stored in more specialized repositories such as video libraries, query-by-image- content databases, or back– ends of geographical information systems.

Ideally, a data item available on a mobile client should be indistinguishable from that available to the accessing application if it were to be executed on the server storing that item. But this correspondence may be difficult to preserve as mobile resources become scarce; some form of degradation may be inevitable. In Odyssey, *fidelity* is used to describe the degree to which data presented at a client matches the reference copy at the server. Fidelity has many dimensions. One well-known, universal dimension is consistency. For Video applications, data has at least two additional dimensions: frame rate and image quality of individual frames. Odyssey provides a framework within which diverse notions of fidelity can be incorporated.

Rover: A Mobile Objects Approach

The Rover project at MIT provides mobility support to client server applications based on two ideas: reloadable dynamic object (RDO) and queued remote procedure calls (QRPC). An RDO is an object (code and data) with a well-defined interface that can be dynamically loaded into a mobile client from a server computer, or vice versa, to reduce client-server communication requirements, or to allow disconnected operation. Queued remote procedure call is a communication system that permits applications to continue to make non-blocking remote procedure calls even when a mobile client is disconnected; requests and responses are exchanged upon network reconnection. Rover gives applications control over the location where the computation is performed. By moving RDOs across the network, applications can automate the movement of data and/or computation from the client to the server and vice versa.

Mobile Data Access in Ad-hoc Networks

The Bayou project at Xerox PARC developed a system to support data sharing among mobile users. The system is intended to support ad-hoc mobility, where no network infrastructure is assumed to be available. In particular, a user's mobile computer may experience extended disconnection from other computing devices. Bayou allows mobile users to share their appointment calendars, bibliographic databases, meeting notes, evolving design documents, news bulletin boards, and other types of data in spite of their intermittent network connectivity. The Bayou architecture supports shared databases that can be read and updated by users who may be disconnected from other users, either individually or as a group. Bayou supports application-specific mechanisms that detect and resolve the update conflicts, ensures that replicas move towards eventual consistency.

Bayou includes consistency management methods for conflict detection called dependency checks and per-write conflict resolution based on client-provided merge procedures. To guarantee eventual consistency, Bayou servers are able to rollback the effects of previously executed writes and redo them according to a global serialization order. Furthermore, Bayou permits clients to observe the results of all writes received by a server, including tentative writes whose conflicts have not been ultimately resolved. In the Bayou system, each data collection is replicated in full at a number of servers. Applications running as clients interact with the servers through the Bayou API, which is implemented as a client stub bound with the application. This API, as well as the underlying client-server RPC protocol, supports two basic operations: Read and Write. Read operations permit queries over a data collection, while Write operations can insert, modify, and delete a number of data items in a collection.

2. EXPLAIN MOBILE TRANSACTIONS.

A mobile transaction is a long-live transaction whose locus of control moves along with the mobile user. Mobile transactions may access remote data wirelessly, through a weak connection, or may access local replicas of data in disconnected mode. The differences between mobile and distributed transaction management are significant because their goals are different. In distributed transactions, the main goal is maximizing availability while achieving ACID properties. In mobile transactions, maximizing reliability while achieving some sort of consistency is the main goal.

In this section we describe some of the existing approaches to mobile transaction management. The transaction models considered here have been proposed by Chrysanthis, Dunham, Pitoura, Satyanarayanan, Gray, Walborn, and Nielsen. All the models described assume the mobile computing reference model.

Reporting and Co-Transactions

This model is based on the Open Nested transaction model. A computation in the mobile environment is considered to consist of a set of transactions, some of which may execute on the mobile node and some of which may execute on the fixed host. The model addresses sharing of partial results while in execution, and maintaining computation state in a fixed node so that the communication cost is minimized while the mobile host relocates.

The model proposes to modify Reporting and Co-Transactions to suit mobile environments. The model defines a mobile transaction to be a set of relatively independent transactions which interleave with other mobile transactions. A component transaction can be further decomposed into other component transactions allowing arbitrary levels of nesting. Component transactions are allowed to commit or abort independently. If a transaction aborts, all components which have not yet committed may abort. Some of the transactions may have compensating duals and may be compensated. The model classifies mobile transactions into the following four types:

- Atomic transactions: normal components and may be compensable with atomic compensating dual steps.
- Compensable transactions: atomic transactions whose effects cannot be undone at all. When ready to commit, the transaction delegates all operations to its parent. The parent has the responsibility to commit or abort the transaction later on.
- Reporting transactions: can make its results available to the parent at any point of its execution. It could be a compensating or a non-compensating transaction.
- Co-transactions: behave in a manner similar to the co-routine construct in programming languages. Co-transactions retain their current status across executions; hence they cannot be executed concurrently.

In an atomic transaction, a series of database operations either all occur, or nothing occurs. A guarantee of atomicity prevents updates to the database occurring only partially, which can cause greater problems than rejecting the whole series outright. In other words, atomicity means indivisibility and irreducibility. A compensable transaction can be formed from a pair of programs: one that performs an action and another that performs a compensation for that action if and when required. The forward action is a conventional atomic transaction: it may fail before completion, but before failure it guarantees to restore (an acceptable approximation of) the initial state of the machine, and of the relevant parts of the real world.

A reporting transaction reports its results to other transactions by delegating the results. A reporting transaction can have only one recipient at any given point of time. The changes made by a reporting transaction are made permanent only when the receiving transaction commits. If the receiving transaction aborts, the reporting transaction aborts as well. A co-transaction, on the other hand, reports its results in a way similar to reporting transactions. But upon delegation, the transaction stops execution and is resumed from the point it left off. For any pair of co-transactions, either both commit or both abort.

Co transactions are one of the conventional methods to value a company for sale. The main approach of the method is to look at similar or comparable transactions where the acquisition target has a similar business model and similar client base to the company being evaluated. This approach is fundamentally different from that of DCF valuation method, which calculates intrinsic value.

When executing transactions in mobile environment, it is necessary to maintain logs to enable recovery after a system crash. A MH is highly vulnerable to failures due to the loss or theft of equipment, memory loss, etc. In order to recover from such failures, it is necessary to store data objects and their logs at the mobile service stations (MSS) rather than on mobile host (MH). A MH transfers a transaction's execution to the MSS by moving all the prewrite values and the log records. A separate pre-commit log is maintained for each transaction.

3. THE KANGAROO TRANSACTION MODEL (APRIL/MAY2014)

This model is based on the global transactions and the split transaction models, where transaction relocation is achieved by splitting the transaction at the point of hand-off. A mobile transaction (called Kangaroo transaction) is considered a global transaction in a multi database environment. A Kangaroo T ransaction (KT) is a global transaction that consists of a set of Joey Transactions (JT). A JT is associated with the base station or the cell in which it executes. When the mobile unit moves to a new cell, the JT in the previous cell is split, and one of the JTs is moved to the current cell of the mobile unit. Each JT may consist of a set of local and global transactions.

The model is built upon the existing databases. The transactions are micro- managed by the individual database transaction managers. A Joey Transaction should terminate in an abort, commit, or a split. For a KT to be successful, the last JT in the order of execution should end in a commit or abort, whereas all other JTs should be split. Based on the ability to compensate the split transaction component, a KT can be executed as a whole atomic transaction or in a relaxed mode where only component transactions are executed atomically. Uses a Data Access Agent (DAA) at each BS to manage mobile transactions and the movement of the MH. Mobile transaction's execution is coordinated by the BS with the MH is currently assigned. When MH hops from one cell to another, coordination of the mobile transaction moves to the new BS. Original transaction is split into Joey transactions (JT). One JT at each base station. Each BS coordinates the operations that are executed while the MH was in its cell.

Three Layers:

- source system,
- data access agent, and the
- mobile transaction.

Two Processing Modes: compensating and split.

DAA (Data Access Agent) acts as transaction manager at base station. For each transaction request DAA generates:

- A Kangaroo Transaction (KT) at MH
- A set of Local Transactions (LTs) & Global Transactions (GTs) at local base station called as Joey Transaction (JT).

For each hop a new Joey is created. When a Joey fails, all previous Joeys and KT will abort.

Kangaroo Transactions:

- Data Access Agent (DAA) tracks MH movement by maintaining a linked list of all BS that have been coordinators of the KT.
- List will be used in case of cascading aborts.

THE CLUSTERING MODEL

This model assumes a fully distributed system. The database is divided into clusters. A cluster defines a set of mutually consistent data. Bounded inconsistencies are allowed to exist between clusters. These inconsistencies are finally reconciled by merging the clusters. The model is based on the open nested transaction model, extended for mobile computing. A transaction submitted from a mobile host is composed of a set of weak and strict transactions. Transaction proxies are used to mirror the transactions on individual machines as they are relocated from one machine to another. A cluster is defined as a unit of consistency in that all data items inside a cluster are required to be fully consistent, while data items residing in different clusters may exhibit bounded inconsistency.

Clusters can be defined either statically or dynamically. A wide set of parameters can be used for defining clusters. This could include the physical location of data, data semantics, and user definitions. Consistency between clusters can be defined by an m-degree relation, and the clusters are said to be m-degree consistent. The m-degree relation can be used to define the amount of deviation allowed between clusters. In this model, a mobile transaction is decomposed into a set of weak and strict transactions. The decomposition is done based on the consistency requirement. The read and write operations are also classified as weak and strict. The weak operations are allowed to access only data elements belonging to the same cluster, whereas strict operations are allowed database-wide access. For every data item, two copies can be maintained–one of them strict and the other weak. As mentioned above, a weak operation can access only the local copies of a data item. Weak operations are initially committed in their local clusters. When the clusters are finally merged, they are once again committed across the clusters.

ISOLATION-ONLY TRANSACTIONS (APRIL/MAY2014) (APRIL 2013)

The Coda file system at CMU provides an application-transparent file system for mobile clients by using file hoarding and optimistic concurrency control. A proxy logs all updates to the file system during disconnection and replays the log on reconnection. Automatic mechanisms for conflict resolution are provided for directories and files through the proxy and the file server. Hoarding is based on user-provided, prioritized list of files. Periodically, the proxy walks the cache to ensure that the highest priority files are present and consistent with the server. Coda provides Isolation-only Transactions (IOT) to automatically detect read/write conflicts that could occur during disconnection. Unlike traditional transactions, it does not guarantee failure atomicity and only conditionally guarantees permanence.

The SEER hoarding system developed at UCLA is based on the Coda file system. It operates without user intervention by observing user activities and predicting future needs. It defines and uses a measure called "semantic distance" between files to determine how best to cluster files together in preparation for hoarding. The semantic difference between two files is based on the time elapsed between the events of opening the files, and on how many reference to other files occurs in between. SEER does not actually hoard files, but instead interfaces with Coda (and other replicated systems) to do the hoarding. SEER also detects hoard misses during disconnection.

THE TWO-TIER TRANSACTION MODEL (APRIL 2013)

A two-tier replication scheme has been proposed in whereby mobile disconnected applications are allowed to propose tentative update transactions. On connection, tentative transactions are applied to (re-processed at) the master data copy in the fixed network. At the re-processing stage, application semantics are used (such as finding commutative operations) to increase concurrency. To reduce re-processing costs that can be high in certain occasions, the work in uses a history-based approach. On reconnection, tentative transactions, which are represented as histories, are merged with base transactions' histories. The merging process quickly identifies the set of tentative transactions that need to be backed out to resolve conflicts.

SEMANTIC-BASED NOMADIC TRANSACTION PROCESSING (APRIL 2013)

The semantics-based mobile transaction processing scheme views mobile transactions as a concurrency and cache coherence problem. It introduces the concepts of fragment able and reorders able objects to maximize concurrency and cache efficiency exploiting semantics of object operations. The model assumes the mobile transaction to be long-lived with unpredictable disconnections. Traditional definitions of concurrency and serializability are too restrictive for most operations. Commutatively of operations is an important property which allows concurrent operations on an object. If certain operations on an object are commutative, then the database server can schedule these operations in an arbitrary manner. Recovery also becomes quite simplified. Operations may be commutative either for all states or part of the states of the objects. The I/O values of the operations can be used to redefine serial dependencies of the operations. Though this may improve concurrency, it may require complex recovery mechanisms than normal schemes. Organization of the object can be used for selective caching of the object fragments, necessary for continuing the operation during the disconnected state. This approach reduces the demand on the limited wireless bandwidth and provides better utilization of the cache space available on the mobile host. Application semantics can also be utilized to define the "degree of inconsistency," "degree of isolation," and the "degree of transaction autonomy". Techniques like epsilon serializability and quasi copies can be used to specify allowable inconsistencies in the system.

This approach utilizes the object organization to split large and complex objects into smaller easily manageable pieces. The semantic information is utilized to obtain better granularity in caching and concurrency. These fragments are cached and/or operated upon by the mobile hosts and later merged back to form a whole object. A stationary server sends out the fragments of an object when requested by mobile units. The objects are fragmented by a split operation. The split is done using a selection criteria and a set of consistency conditions. The consistency conditions include the set of allowable operations on the object and the conditions of the possible object states. On completion of the transaction, the mobile hosts return the fragments to the server. These fragments are put together again by the merge operation at the server. If the fragments can be recombined in any order, then the objects are termed "re- orderable" objects. Aggregate items, sets, and data structures like stacks and queues are examples of fragment able objects.

SRI VENKATESHWARAA COLLEGE OF ENGINEERING
& TECHNOLOGY
(Approved by AICTE, New Delhi & Affiliated to Pondicherry University, Puducherry.)
13-A, Villupuram – Pondy Main road, Ariyur, Puducherry – 605 102.
Phone: 0413-2644426, Fax: 2644424 / Website: www.svcetpondy.com

Department of Computer Science and Engineering

Subject Name: **MOBILE COMPUTING** Subject Code: **CS E84**

UNIT V

Mobile Computing Models: Client Server model – Client/Proxy/Server Model – Disconnected Operation Model – Mobile Agent Model – Thin Client Model – Tools: Java, Brew, Windows CE, WAP, Sybian, and EPOC.

2 Marks

1. What are the Different types of mobile computing models? (April'13)

The following models of computing in the mobile environment are currently being researched and investigated:

Client/Server

Client/Proxy/Server

Disconnected Operation

Mobile Agents

The Thin Client model

2. Write about client-server model.

In this model, neither the client nor the servers a r e aware of the client (or server) mobility. The conventional client/server model is used without any modifications made in the application or the transport layer. The wireless media is transparent since it is handled in the data link layer. The mobility, on the other hand, is made transparent by handling the variable client/server location through location-based routing in the network layer. Mobile IP is an example of a network protocol that hides the C/S mobility.

3. What is the advantage in client server model? (UQ April'13)

The advantage of this model is its portability. The client or the server need not be changed in any way. Simply, the client (server) is ported to the mobile host, with a fixed address specified for the server (on the fixed network or on a mobile host).

4. What are the disadvantages of client server model? [April 2013][April 2014]

The mobile client may suffer from a slow and unpredictable response time, especially when large server replies such as query results are transmitted without any considerations to the limited wireless bandwidth.

The server caching strategy may not work properly in this model because the majority of the fixed network caching algorithms use call backs to invalidate the client cache. Most invalidation algorithms rely on the continuous availability of the client. Since the client could be disconnected or temporarily inaccessible (during hand-offs for example), the cache invalidation process could fail.

5. What is meant by client/proxy/server model?

To overcome the shortcomings of the conventional client/server model, the client/proxy/server (C/P/S) model introduces a mobility-aware middle layer to mediate the interactions between the client and the server. The basic idea behind this model was introduced in the Mowgli architecture [72], which is to split the communication path between the client and the server into two parts by using a store-and-forward interceptor.

6. What are the advantages of proxy model?

The main advantage of this model is that the proxy allows the client and the server to be designed without any built-in mobility assumptions. The proxy assumes that the client is mobile and the server is in the fixed network. The result from the server is sent back to the proxy. The proxy filters the results according to the limitations of the wireless media and/or the client's mobile unit. The proxy may also store the filtered results until the client is connected.

7. What is disconnected operation model.

Mobile clients may face wide variations in network conditions and local resource availability when accessing remote data. This is true in the C/S and the C/P/S models. Disconnected operations are a variation of the C/S model where, instead of working under the extreme case of weak-connectivity, the mobile client effectively switches to use a network of zero bandwidth and infinite latency.

8. Explain briefly about mobile agent's model. [April 2014]

Mobile agents are mobile scripts with associated execution state information. A mobile agent could either be relocated along with the user, or it could be relocated during the execution of the agent. The relocation of the agent involves saving the state before initiating relocation and later restarting the mobile agent at the new location.

9. Write about thin client model. [NOV 2013]

The thin client computing model attempts to offload most application logic and functionality from mobile clients to stationary servers. In this model, applications in stationary servers are usually mobile-aware and optimized for mobile client devices. This model is especially suitable for dumb terminal or small PDA applications.

10. Write short notes about the architecture of thin client model.

The thin client architecture from CITRIX Corporation allows a variety of re- mote computers, regardless of their platform, to connect to a Windows NT terminal server to remotely access a powerful desktop and its applications. A server called Meta Frame runs under Windows NT in the desktop machine

11. What is a mobile agent?

The mobile agent is an emerging new model that provides an alternative to the C/P/S model. A mobile agent i s an active entity that is knowledgeable of both the limitations of the mobile environment and the mobile user.

12. Classify the mobile agents.

Agents can be classified as static, mobile scripts, or mobile objects. Static agents are those which execute just on a single site, either as a client or as a server. A static agent could be carrying out some activity like mail filtering. Mobile scripts are those that are downloaded from a server and executed on a client. Java applets, perl, or python scripts can be classified as mobile scripts.

13. Expand and explain MDCS.

In the proxy architecture, the mobile host is provided with a specialized transport service, the Mowgli Data Channel Service (MDCS). It provides prioritized data channels with flow control between the mobile host and the base station. Existing TCP/IP protocols are used between the base station and a fixed host so that the protocol software in the fixed network remains unmodified.

14. When does the agent act as a local client?

The mobile agent is an execution context initially loaded with the queries or data access requests. Once the agent moves to the data source (server), it acts as a local client to the server.

15. What is Wireless Application Protocol (WAP)? Wireless-
that does not require wires, enabling radio transmission.

Application- software designed to complete a particular task.

Protocol- means a set of rules.

These three simply means a set of rules governing the transmission and reception of signals through computer applications. Computer may include mobiles, tablets also. WAP has spearheaded today's Internet communication and advanced telephony services. WAP allows allow the devices to view pages from Internet. But, the pages displayed have only a plain text and the images are black and white. WAP is similar to HTML, except that it has been optimized for-

- ✓ Low-display capability.
- ✓ Low-memory.
- ✓ Low band width devices like mobile phones etc.

16. Why do we need WAP?

Long before the first WAP devices came, Internet was limited to your computer only. Now with WAP, you can communicate to your friends using Internet through your mobile phone too. Thus globally expanding massive communication and sharing of data.

17. What is a Micro WAP Browser?

Similar to your own internet browser, there is a browser for WAP too. This browser is called Micro WAP Browser that is used for visiting web-sites through a WAP device. What's special about it is, it makes minimal demands on hardware, memory and CPU and it displays the information in WML (it a restricted mark-up language).

18. What are the different layers of WAP architecture?

The whole WAP architecture has been divided into 5 main layers, namely:-

1. Application Layer
2. Session Layer
3. Transaction Layer
4. Security Layer
5. Transport Layer

19. What are the various protocols in a WAP protocol suite? The WAP protocol suite consists of following protocols-

1. Wireless Application Environment (WAE)
2. Wireless Session Protocol (WSP)
3. Wireless Transaction Protocol (WTP)
4. Wireless Transport Layer Security (WTLS)
5. Wireless Datagram Protocol (WDP)

20. What is WAP 2.0?

WAP 2.0 is simply a blended mixture of XHTML, end to end HTTP, which was released in 2002. It has dropped the gateway and custom protocol suite used to communicate with it.

21. On which Networking Model is WAP based upon?

Open System Interconnection (OSI) model.

22. Describe the Transport Layer of WAP

The Transport Layer consists of Wireless Datagram Protocol (WDP).WDP allows WAP to be bearer-independent by adapting the transport layer of the underlying bearer. The WDP presents a consistent data format to the higher layers of the WAP protocol stack, thereby offering the advantage of bearer independence to application developers.

23. Define in brief, "Wireless Application Environment (WAE)"

The Wireless Application Environment, or WAE, provides architecture for communication between wireless devices and Web servers. It consists of device specifications and the content development programming languages like WML

24. How you define WAP

WAP is stands for **Wireless Application Protocol**. Using WAP protocol we display internet contents on wireless users. Example: mobile phone, i-pod etc.

I have given you some basic information about WAP are given below:

1. WAP is used as an application communication protocol.
2. Using WAP we can access services and information
3. We can inherited WAP from Internet standards.
4. WAP is designed for handheld devices Like: mobile phones, i-pod etc.

25. Give examples of WAP

1. Using WAP we can get information of train time-table.
2. Using WAP can purchase tickets. Like: movie, journey ticket etc.
3. Using WAP we perform task like that Flight check in
4. We can also viewing traffic information.
5. We can get information about current weather condition.
6. Using WAP we can do also trading of shares.
7. We can also display sport results on our small wireless devices

26. Does WAP runover GPRS

Yes, it can do. GPRS is a new over-the-air service that transmits data packets to hand-held devices. It will allow much faster WAP transmission than currently available over SMS or CSD when using GSM.

27. How secure is WAP

One of the layers of the WAP stack, known as WTLS, provides encryption and authentication for server-to-client security. This prevents fraudulent access to WAP transactions and opens the way for e-commerce- and intranet-type applications.

29.What does Windows CE means?

Windows CE is a Windows operating system, also called Windows Embedded Compact. It is mainly used for mobile devices like Smart Phones and PDAs. Windows CE is an operating system created by Microsoft for entrenched systems. It is a different operating system and kernel, rather than a trimmed down version of desktop Windows. It is not to be confused with Windows XP Embedded which is NT-based.

30. What is EPOC? [NOV 2013]

EPOC[Electronic Piece Of Cheese] is an operating system designed for small, portable computer-telephones with wireless access to phone and other information services. operating system from Psion Software, designed specifically for mobile, ROM -based computing devices. *EPOC16* is a 16-bit version of the operating system that has been available for several years and is embedded in many handheld devices. *EPOC32* is a newer, 32-bit operating system that supports preemptive multitasking

11 Marks

1. Explain mobile computing models (11)][April 2014] [April 2013][NOV 2013]

Computing in the mobile environment is different from the conventional fixed- network computing. This is partially due to the movement of the mobile hosts that require remaining connected from different access points while moving. The difference also stems from the nature of the wireless links that are relatively unreliable and offer low communication bandwidth. Furthermore, mobile hosts equipped with rechargeable batteries suffer from limited operation time constraints. As a consequence, new models in the mobile environment are needed to support information access for mobile users.

The following models of computing in the mobile environment are currently being researched and investigated:

1. Client server model
2. The Client/Proxy/Server Model
3. The Thin Client Model
4. The Disconnected Operation Model
5. The Mobile Agent Model

THE CLIENT/SERVER MODEL

In this model, neither the client nor the server are aware of the client (or server) mobility. The conventional client/server model is used without any modifications made in the application or the transport layer. The wireless media is transparent since it is handled in the data link layer. The mobility, on the other hand, is made transparent by handling the variable client/server location through location-based routing in the network layer. Mobile IP is an example of a network protocol that hides the C/S mobility. The advantage of this model is its portability. The client or the server need not be changed in any way. Simply, the client (server) is ported to the mobile host, with a fixed address specified for the server (on the fixed network or on a mobile host). The disadvantages of this model are listed below:

The mobile client may suffer from a slow and unpredictable response time, especially when large server replies such as query results are transmitted without any considerations to the limited wireless bandwidth.

The server caching strategy may not work properly in this model because the majority of the fixed network caching algorithms use call backs to invalidate the client cache. Most invalidation algorithms rely on the continuous availability of the client. Since the client could be disconnected or temporarily inaccessible (during hand-offs for example), the cache invalidation process could fail.

THE CLIENT/PROXY/SERVER MODEL

To overcome the conventional client/server model, the client/proxy/server (C/P/S) model introduces a mobility-aware middle layer to mediate the interactions between the client and the server. The basic idea behind this model was introduced in the Mowgli architecture], which is to split the communication path between the client and the server into two parts by using a store-and-forward interceptor.

The main advantage of this model is that the proxy allows the client and the server to be designed without any built-in mobility assumptions. The proxy assumes that the client is mobile and the server is in the fixed network. The result from the server is sent back to the proxy. The proxy filters the results according to the limitations of the wireless media and/or the client's mobile unit. The proxy may also store the filtered results until the client is connected. Examples of data filtering include: color and resolution reduction, audio file removal, and file size reduction. The programming of the proxy involves knowledge of the mobile host hardware specifications. For example, a mobile host that does not have audio capability will benefit from audio file removal filtering. In addition to the mobile host, the mobile user profile can be useful in providing the proxy with user preferences such as no-images and no-colors.

THE DISCONNECTED OPERATION MODEL

Mobile clients may face wide variations in network conditions and local resource availability when accessing remote data. This is true in the C/S and the C/P/S models. A disconnected operation is a variation of the C/S model where, instead of working under the extreme case of weak-connectivity, the mobile client effectively switches to use a network of zero bandwidth and infinite latency. The operations that enable a client to continue accessing critical data during the disconnection (switch off) period are called disconnected operations. The ability to operate when disconnected can be useful even when connectivity is available. For example, disconnected operation can extend battery life by avoiding wire- less transmission and reception. It allows radio silence to be maintained, a vital capability in military applications. And, it is a viable fallback position when network characteristics degrade beyond usability. Voluntary disconnection can be treated as planned failures which can be anticipated and prepared.

THE MOBILE AGENT MODEL

Agents can be classified as static, mobile scripts, or mobile objects. Static agents are those which execute just on a single site, either as a client or as a server. A static agent could be carrying out some activity like mail filtering. Mobile scripts are those that are downloaded from a server and executed on a client. Java applets, perl, or python scripts can be classified as mobile scripts. Mobile agents are mobile scripts with associated execution state information. A mobile agent could either be relocated along with the user, or it could be relocated during the execution of the agent. The relocation of the agent involves saving the state before initiating relocation and later restarting the mobile agent at the new location. The mobility of agents raises a large number of issues like security, authorization mechanisms, access mechanisms, and relocation mechanisms.

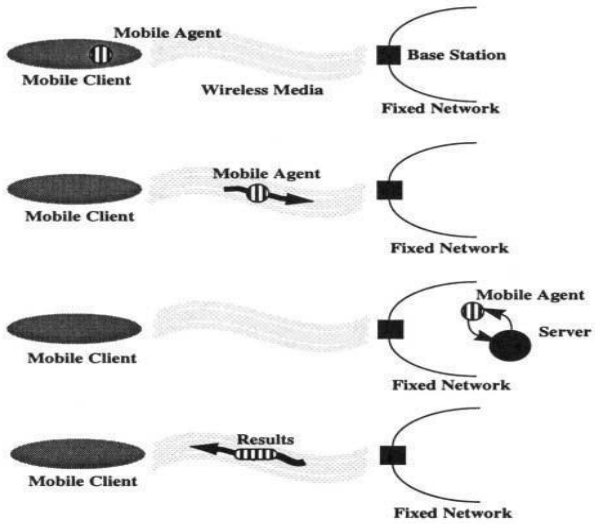


Figure 7.8 Mobile agents for mobile computing

The mobile agent is an emerging new model that provides an alternative to the C/P/S model. A mobile agent is an active entity that is knowledgeable of both the limitations of the mobile environment and the mobile user. To access remote data, the mobile user sends a mobile agent on his behalf to the data source in the fixed network. The mobile agent is an execution context initially loaded with the queries or data access requests.

THE THIN CLIENT MODEL

The thin client computing model attempts to offload most application logic and functionality from mobile clients to stationary servers. In this model, applications in stationary servers are usually mobile-aware and optimized for mobile client devices. This model is especially suitable for dumb terminal or small PDA applications.

The thin client architecture from CITRIX Corporation allows a variety of re- mote computers, regardless of their platform, to connect to a Windows NT terminal server to remotely access a powerful desktop and its applications. A server called MetaFrame runs under Windows NT in the desktop machine and communicates with the thin clients executing at the remote computers using the Independent Computing Architecture protocol (ICA). The ICA client and the MetaFrame server collaborate to display the virtual desktop on the remote computer screen. They also collaborate to process mouse and keyboard events and to execute programs and view data stored at the server. All executions are remote and none take place at the client portable computer.

2. Write in detail about JAVA (11)

Java is a programming language offers the most portable commercial environment for writing software applications. The success of Java has been mostly in providing standard Application Program Interfaces (APIs), a very thoughtfully designed infrastructure for OOP that prohibits many bad design and implementation habits such as multiple inheritances. Standard and open APIs offer a process of evolving a language that is open to many vendors. There are three major categories of Java APIs and virtual machines, namely J2ME, J2SE, and J2EE.

Java offers three distinct features as a mobile application platform:

1. Java is an object oriented programming language. As any other programming language, it can be used to write applications.

2. Java offers complete code mobility and weak mobile agent ability. Java allows for platform-independent programming.

3. Java is a platform.

First, Java, as with any other programming language, is just that: a programming language. It allows us to program a set of instructions. Perhaps just as importantly,

Java is somewhat of a *vendor-neutral language-based platform*."

Java Database Connectivity (JDBC) APIs present the same interface to the developers regardless of what database is being used.

Java, as a platform and programming language, offers mobile code. But, the standard Java Virtual Machine was designed for desktop computers and requires far too many resources for the typical cell phone, PDA, or mobile device. The standard Java Virtual Machine is packaged, along with accompanying tools and class libraries, into Java 2 Standard Edition (J2SE).

J2ME

J2ME is a specification for a virtual machine and some accompanying tools for resource-limited devices. J2ME specifically addresses those devices that have between 32 kB and 10 MB of memory. J2ME addresses the needs of two categories of devices [Sun Micro J2ME Spec 2000]:

1. *Personal, mobile, connected information devices*. This portion of J2ME is called CLDC for Connected, Limited Device Configuration. These types of devices include cell phones, PDAs, and other small consumer devices. CLDC addresses the needs of devices with 32 to 512 kB of memory. The virtual machine for the CLDC is called KVM for K-Virtual Machine.

2. *Shared, fixed, connected information devices*. Internet-enabled appliances, mobile computers installed in cars, and similar systems that have a total memory of 2 to 16 MB and can have a high bandwidth and continuous connection to the network are in this group. CDC, or Connected Device Configuration, is the part of J2ME that addresses such devices. CDC is a superset of CLDC.

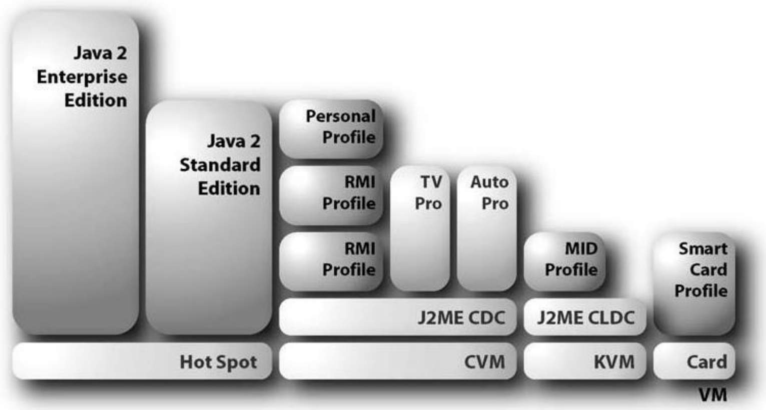


FIGURE 2.2. J2ME Stack (CLDC/CDC and MIDP).

CLDC addresses the following features:

1. *Providing a virtual machine for providing language features*. Perhaps the most important thing to keep in mind for those who have built applications using the Java Virtual Machine on desktops and servers is that the J2ME/CLDC Virtual Machine is not at all like the version that comes with J2SE. Some features not offered on the KVM are the following:

a. *Floating point arithmetic*: Floating point operations are expensive or require the chipset on the device to have specific implementations for them.

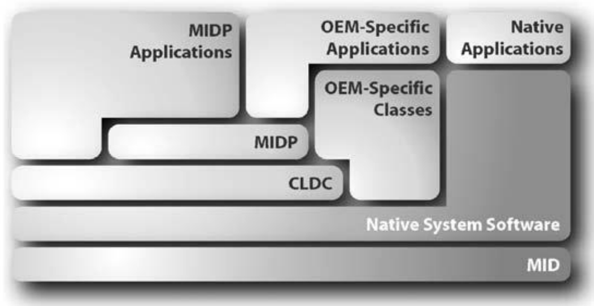


FIGURE 2.3. Layering of Functionality between CLDC and MIDP.

b. *Support for JNI*: Java Native Interfaces (JNI) allow developers to write applications that use C/C++ programming languages along with Java in providing Java APIs to modules or applications not written in Java.

c. *Thread grouping*: Advanced threading features are not offered on the KVM and CLDC. Multithreading requires a baseline amount of resources to be dedicated to creating, maintaining, and destroying threads. Each thread takes up a certain amount of resources by simply existing, even if it never does any actual work.

d. *Full-blown exception handling*: Exception and error handling seems to be one of the first places that platform providers trim when building frameworks and tools for limited devices.

e. *Automatic garbage collection of unused objects*: Though the KVM does offer some of the memory management features of the J2SE Virtual Machine, it does not offer *finalization* of objects. This means that you have to tell the KVM when you are done with an object.

f. *Weak references*: The J2SE Virtual Machine does not allow finalization of an object until all weak and strong references to that object are cleared. The KVM does not provide this functionality for weakly referenced objects.

2. *Providing a security framework for tasks such as downloading MIDlets (J2ME CLDC/MIDP applications)*. Security is one of the most troublesome and complicated features for providers of mobile application frameworks and tools

3. *Providing a reasonable amount of functionality for input and output*. Most pro- grams need a persistence mechanism. CLDC provides a very limited and yet sufficient set of APIs to read and write to the nonvolatile memory provided by devices

4. *Providing some internationalization capabilities*. CLDC's input/output (I/O) pack- age provides input and output stream readers that can handle different character encoding schemes. This allows internationalization in two ways:

a. *Dynamic*: The program can determine the required character set dynamically and use the proper character set at run time.

b. *Static*: There can be multiple versions of the J2ME application ready to be loaded onto the device.

5. *Providing a reasonable amount of networking capabilities*. CLDC provides a con- nection framework to provide basic networking capabilities. The areas addressed by profiles are the following:

1. Download and installation of applications,

2. life-cycle management of applications,

3. User interface feature,

4. Database functionality, and

5. Event handling.

Sun Microsystems has a free tool kit offers the components for development of J2ME applications:

1. *KToolbar*

2. *Preverifier*

3. *Compiler*

4. *Emulators*

5. *Emulation of Performance*

3. Explain BREW (11)

Qualcomm's **BREW** (**Binary Run-time Environment for Wireless**) gives application developers a new and different approach in producing mobile applications. BREW is built directly into the hardware. It is offered as an API to access the CDMA, GSM/GPRS, or UMTS chip sets that provide the support for it. But, it is primarily intended for the variations of CDMA, a technology owned and licensed by Qualcomm. BREW applications can be written on a PC using the BREW Software Development Kit (SDK). Once the application is developed, it must be tested, and then deployed. Deployment of BREW applications is a process done jointly by Qualcomm and telecommunications carriers and not just the developer.

Though the creators of BREW say that they first came up with the acronym BREW and then found the words to fit the acronym, the platform is somewhat biased toward wireless applications that run on phones. And this may be the only weakness of BREW as a mobile development platform. Although today developing mobile applications means targeting cell phones or PDAs, this is changing rapidly with new devices being introduced to the market.

BREW applications, also referred to as BREW applets, are written in C though some support for C++ is provided (although some fundamental things such as extending the base API through inheritance are not possible) and, using code generation or virtual machine technologies, other languages such as Java can be supported. One of the most impressive things about BREW is its near-full treatment of dimensions of mobility in its architecture, feature implementation, and SDK. Let us look at the various components that allow the developer to build a BREW application.

BREW SDK Overview

Once you have installed the BREW SDK, you will have the following set of applications available for development:

1. *BREW MIF Editor*: Every BREW module, defined as the classes that make up one or more BREW applications, has an associated Module Information File (MIF). MIFs are *required*. Every BREW module must have a MIF. The MIF Editor provides a GUI tool for editing the MIF file associated with the classes that make up a module. The MIF Editor that comes with BREW SDK version2.0 can be started as a wizard inside Visual C++ 6.0 or independently as a stand- alone application. We will look at the use of the MIF Editor and building a simple application.

2. *BREW Device Configurator*: This is a stand-alone application that allows developers to make up their own handset by configuring a vanilla mobile phone and specifying the behavior of the keys, the look and feel of the screen, and other specifics of the device. This development tool addresses the large variety of existing devices by allowing developers to create their own device emulator and testing the application. Remember, also, that because BREW is a platform for writing application for the handset,

it is possible to use the application to build some adaptive behavior to adapt to each type of device. Still, the Device Configurator is invaluable in that it allows developers to test the application on their own emulated device environment.

3. *BREW Emulator*: For those who have designed and implemented any mobile application, it is obvious that one of the most difficult steps in the development process is the incremental unit testing. Although most platforms provide some sort of a generic emulator, most do not allow for custom configuration of a device (done by the Device Configurator) or using the custom configuration to simulate running an application **4.** *Register as a BREW developer*: This is not a simple sign up for notifications and other news about BREW. You will have to have the VeriSign Class 3 certificate before you can become an “Authenticated Developer.” Once you are an authenticated developer, you are ready to work.

5. *Obtain a Class ID for your application*: During the development process, you can use a dummy Class ID. But, to get the application out on a real device, you need to get a Class ID for the application. This Class ID uniquely distinguishes your application from all other BREW applications. Every BREW application has to have a Class ID and the Class IDs are issued and provisioned by Qualcomm centrally to avoid ID collision.

6. *Perform a unit test and send it to a testing lab*: If you are done with steps 1–5, you are ready to submit your application for testing. To get the application onto your phone, a Qualcomm-approved testing center needs to test your application. It should be obvious that you want to bullet-proof the application before submitting it to the testing center. This process is put in place to avoid “crashing” the mobile device.

7. *Perform a pricing and carrier evaluation*: Once the testing lab approves an application, it is ready to be provisioned. Deployment of an application is done by Qualcomm and the carriers supporting BREW. Therefore, the application developer must submit the software to Qualcomm and the carriers for actual deployment. Of course, deployment is done after the carrier and Qualcomm approve of the application.

Unfortunately, deploying a BREW application onto a device is not free. Before you get an application up and working on a device, you need to pay various types of fees for the VeriSign certificate and, in practicality, become a Qualcomm developer, requiring a significant membership fee depending on what type of membership you want to sign up for. The positive spin on this, of course, is that the membership fee pays for some marketing and weeds out those developers who do not have a product and are just playing around with the platform. If you want to just learn the platform, the best thing is probably just to download the SDK and the tools after reading this section and then experiment with it. Now, let us go on to the actual code.

Hello BREW

Once you have downloaded and installed the BREW SDK, you can get started. Here is the procedure:

1. Click on File, New, and then Projects. You will see the BREW Application Wizard. Choose it.

2. The wizard will ask you if you want File, Network, Database, TAPI, or Sound functionality. These selections correspond to the organization of the BREW standard libraries:

a. *Files*: BREW provides an API that allows storage of small amounts of information in structures with which application developers are familiar: files and directories of files.

b. *Database*: Information is often better stored in a database instead of files if the data must be searched, sorted, or indexed. BREW provides a set of APIs to store, manipulate, and retrieve data that have a small amount of database- like functionality. What BREW offers, it should be noted, is not nearly as complete as a full-blown database system. However, it offers enough for useful functionality.

c. *TAPI (Telephony API)*: Because BREW is built on a wireless telephony platform (CDMA), it is natural that it provides telephony functionality. At the time of release of BREW SDK 2.0, functionality is limited to sending SMS messages and switching back and forth between incoming and outgoing telephony calls. However, being built on a telephony platform, it is almost certain that BREW will offer functionality that provides control and manipulation of the audio over the telephony channel, integration with voice recognition, and other useful functionality.

d. *Sound functionality*: Sound functionality is provided through a set of multimedia APIs. Sound can be stored on the device in BREW's own format of QCELP (a Qualcomm technology).

3. Once you have selected which libraries you will be using in your application, you will need to create a MIF file for it. If you do not yet have a true valid Class ID, you have to get the Verisign certificate, become a Qualcomm developer, and go through the steps that we mentioned previously. Once again, every BREW application must have an MIF file. Click on the MIF Editor. You will see these different tabs on the wizard:

a. *Applets*: This is where you generate a test Class ID (or if you have already become a Qualcomm developer, get a real Class ID from Qualcomm). The other basic properties of the BREW application (as we mentioned bef or e , interchangeably referred to as BREW applet) are set in this pane. Note that if you click on the *advanced* button on this pane, another window pops up with some features that seem to be programmatic. In BREW, certain behaviors of the application such as its treatment of incoming telephony events (when a call comes in while the application is running) have to be specified. Some behaviors are fundamental to the behavior of the application and, therefore, are required to have a Footprint in the MIF file.

b. *General*: This pane is for entering the security-related information. Every BREW application m a y provide access to other BREW applications a n d modules or require a particular access level for certain functionality on the device. It is important for this information to be on the MIF file because other applications must know whether they are usable by other applications or not and so that the application container (the BREW environment running on the BREW device) knows whether it can load and execute the application.

c. *Extensions and Dependencies*: Because BREW applications can come in several modules or have interdependencies among themselves, the MIF files allow for specifying these dependencies. The Extensions and Dependencies panes provide a graphical way of manipulating these dependencies.

4. Explain WINDOWS CE (11)

Windows CE is a Windows operating system, also called Windows Embedded Compact. It is mainly used for mobile devices like Smart Phones and PDAs. Windows CE is an operating system created by Microsoft for entrenched systems. It is a different operating system and kernel, rather than a trimmed down version of desktop Windows. It is not to be confused with Windows XP Embedded which is NT-based. An operating system is the master control program that enables the hardware by abstracting it to the application via drivers [Development tools for Mobile and Embedded Applications]. Microsoft's various products revolve around different versions of an operating system.

These two operating systems are designed for two different purposes. There are different flavors of the Windows CE operating systems, of course, depending on the hardware platform. Some of these flavors are the Pocket PC, Windows CE .NET, and Pocket PC 2002. These flavors largely depend on the commercial bundling of different feature sets and hardware platforms with which they are shipped (such as Compaq's IPAQ). Embedded Windows XP, in contrast, is a subset of the desktop version of Windows XP components. Development for Embedded Windows XP is a bit more straightforward than developing for Windows CE.

Mobile application frameworks that are based on an operating system treat developing mobile applications in the same way as they treat their stationary counterparts on PCs. As we mentioned previously, the operating system provides basic access to the hardware such as I/O, networking, etc. So, the applications that run on Windows CE and Embedded Windows XP are controlled by them, respectively. Microsoft provides tools to build applications for each environment too. These are as follows:

1. *Embedded Visual C++*: This is a tool set separate from Visual Studio, the typical development environment for PC-based Windows applications. It allows for authoring mobile applications in C++.

Emulators and a debugger are provided. The latest version of this tool provides advanced features such as exception handling and run-time debugging, features you will cherish if you are actually developing an application in C++ for mobile devices.

2. *Embedded Visual Basic*: This tool provides the ability to write applications using Visual Basic. Visual Basic applications can be developed faster but do not offer the developer the ability to tune and optimize the application for resource- starved mobile devices. Therefore, Embedded Visual Basic is really not a suitable tool for developing large commercial applications, but it does well for proof-of- concept and prototype applications.

3. *Smart Device Extensions for .NET*: The .NET application programming platform, the newest set of tools for building Microsoft Windows-based applications, can be complemented with a set of extensions that allow developers to author .NET applications for mobile devices.

4. *Microsoft Mobile Internet Toolkit*: This is really a server-side framework.

As in the other Microsoft Windows platforms, Windows CE allows the use of COM and ActiveX components in addition to the Win32 API. The other significant features are markup language processing (HTML, XML, XSL, etc.), security (e.g., SSL), a subset of the Windows ADO database access framework in ADOCE (ADO for Windows CE), and limited functionality in coupling with the Microsoft messaging queue in MSMQ.

Just like Windows 2000, Windows CE utilizes protected memory architecture. When a Windows CE machine first boots, it creates a single 4 GB virtual address space [Introduction to eVC++]. This, however, does not mean that there is 4 GB of random-access memory (RAM) available! In fact, currently most Windows CE devices are limited to under 64 MB. Moreover, although we expect mobile devices to continue to grow in their processing ability and memory, it is unlikely that we will ever want to run memory-intensive applications on mobile devices because physical size is a limiting factor and battery life is not growing as fast as processing power. This virtual address space is then divided into 33 different "slots," each of which is available for use by a process. The maximum size of each slot is 32 MB. This is simply the model with which the memory is managed; it does not mean that the device is required to have 4 GB of memory or that it offers 32 MB of memory per process. Also, keep in mind that Windows CE does not allow paging (file swaps), so you can exceed the available RAM. File swapping is something that is typically not implemented as a strategy for improving the memory limitations of mobile devices as it is cost prohibitive to the battery life and takes considerable processing power.

When building eVC applications, keep the following in mind:

1. Graphics are expensive. Whether you decided to use GDI or another method of rendering graphics, delivering them takes more memory, more CPU, more time, and more power from the battery. So, try to avoid graphics when possible.

2. Use events instead of polling when possible. Figure 2.8 shows some sample code that involves events in Windows CE environment. Polling is expensive for the same set of reasons as graphics are. Sleeps and event notifications are both features available to the eVC programmer to produce efficient applications.

3. Be very frugal in the use of RAM in your applications. Remember that persistent and RAM memory are typically handled by one set of hardware on mobile devices. Today, most mobile devices, including Windows CE devices, do not have hard drives (though this is changing).

4. As mentioned in item 3, because RAM and persistent memory often share the same hardware, being frugal in persisting data or handling data in memory pays off in reducing power consumption as well.

5. There is some functionality provided to the application developer to get the status of the power consumption. You can use this functionality in two ways. First, you can use it while designing and testing to see the power consumption during the life cycle of the usage of the application. Alternatively, you may want to use the power status (called up by GetPowerStatusEx function) to change the behavior of the application. For example, if the battery is getting low, you might want to persist the data to the network or locally and shut down the application after warning the user.

6. Make sure that you clean up memory resources whenever you get a WM HIBERNATE event (which sends the device into hibernation). Failing to do a good memory cleanup there will lead to memory leaks and application instability.

5. Explain WAP in detail.

Wireless Application Protocol (WAP) is the single framework most used in building mobile applications today. Despite all of its initial high promises, its lack of meeting those promises, and being written off for dead, WAP seems to have survived the critics and continues to improve. WAP, which was initially intended to be as pervasive for wireless and mobile applications as HTTP has been for the Web, never achieved the level of success initially expected. However, to date, WAP has the largest install base of all open application development platforms (second to NTT Docomo's closed and proprietary i-mode system) on mobile phones, meaning that WAP is installed on more mobile phones than any other software.

1. *WAP is intended for thin clients*. Like HTTP, the designers of WAP 1.x were thinking about a thin-client technology: a case where nearly all logic is calculated on the server and very simple display instructions are bundled in some markup language to be displayed by the client.

2. *WAP is built on its own lower level communication protocol*. The HTTP assumes the existence of TCP/IP (which in turn provides persistent connections), WAP is built on its own set of communication protocols that wrap around TCP, UDP, or a variety of other possible protocol implementations.

3. *Typical deployment of WAP includes a proxy or a gateway*. Wireless carriers (also referred to as *bearer networks*) used to control every single incoming and outgoing bit of data that travels on their network. Interoperability or functionality.

4. *WAP is a complete framework for mobile applications*. Most tools created for development of applications treat a part of the mobile application chain; WAP treats, or at least attempts to treat, all parts of the mobile equation. The fact that WAP has various parts, however, does not mean that some of its parts cannot be used individually. For example, J2ME applications can use WAP as a communication protocol while not using the WAP browser (instead using a client-side J2ME application to provide a richer and better user experience).

WAP Architecture

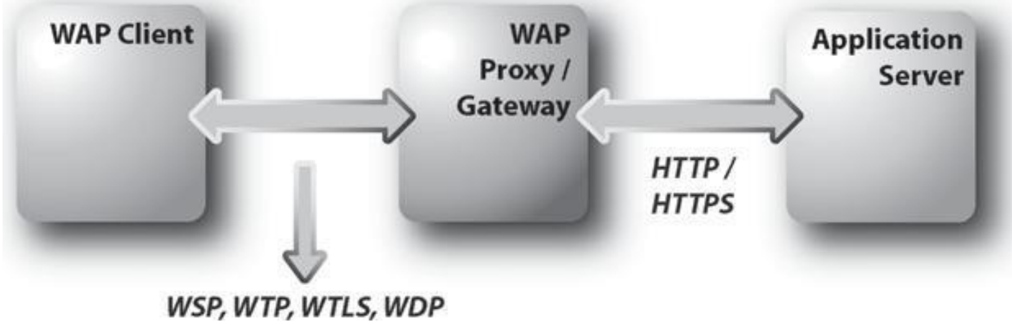


FIGURE 2.11. Basic Communication Architecture in WAP.

WAP architecture shows the functionality to the communication between the client and the server (WAP Gateway or WAP Proxy):

1. *Handling of Telephony on the Device*: Currently, devices and networks treat voice and data differently. Most voice is processed, and will continue to be processed for a very long time, through a telephony system that does not provide much in the way of complex operations.

2. *Push*: Push offers some degree of treatment for one of the dimensions of mobility, namely support for active behavior. It needs to be reiterated that the WAP architecture was designed with a subset of the mobile application arena in mind: wireless applications.

WAP Proxies and Gateways

The difference between a proxy and a gateway is that a client determines when it will use a proxy. Some WAP-enabled devices do have the ability to change their proxy settings; however, this feature is typically disabled from access by the user by the network provider or the device manufacturer as it can circumvent billing mechanisms (especially for the United States compared to Europe). Therefore, whether used as a proxy or a true gateway, WAP intermediaries are typically referred to as WAP Gateways.

WAP gateways provide six important features:

1. *Security*: The WAP gateway provides a secure handoff point between WTLS (Wireless Transport Layer Security) to external security mechanisms such as SSL for HTTP in the form of HTTPS.

2. *Network Access*: The WAP gateway is the access point for the WAP client devices. Network providers are able to restrict access to users and connect their usage to billing systems by using the WAP gateways.

3. *Protocol Conversion*: The gateway is responsible for converting Wireless Session Protocol (WSP) to HTTP. This allows WAP to be based on a non- TCP/IP application layer protocol (WDP) and yet interact with the HTTP-based

4. *Caching*: The WAP gateways provide a caching mechanism equivalent to (and even surpassing) that provided by HTTP.

5. *Preparation of Content and Scripts*: WML is textual, and text is not a very efficient format for transfer of data. Therefore, the gateway encodes WML into “Compiled WML” (WMLC) before shipping it to the WAP-enabled device.

6. *Functionality Offered through WAP 2.x and Higher*: Although WAP 2.x is not yet deployed in the United States and is just beginning to be rolled out in Europe, it offers a variety of changes to the role of the proxy/gateway in WAP deployments.

WAP 2.x offers WAP push, support for UAProf ,additional functionality for WTA (Wireless Telephony Application), an External Functionality Interface (EFI) offering an extensibility mechanism for WAP APIs, data synchronization using SyncML, a persistent storage interface for storing data on the device, a multimedia messaging service, pictograms (small images used in messages), and provisioning.

The Multimedia Messaging Service (MMS) is the more mature child of the ever popular Short Messaging Service (SMS).

6. Explain SYMBIAN EPOC (5)

Symbian, one of the most powerful and popular platforms for mobile development, was created jointly by Ericsson, Nokia, Panasonic, Psion, Samsung Electronics, and Siemens. The effort in creating this new operating system targeted at mobile devices started in 1998 and the first Symbian phones became available in 2001. The majority of the user base of Symbian devices is in Europe with very little user base in the United States; however, the market share in Europe is large and growing, with other markets wide open between the various contenders of mobile operating systems including Symbian. The Symbian OS 7.0 comes with considerable basic functionality for mobile applications: support for MMS, HTTP communication, SyncML synchronization, SMS, support for Mobile IP (through support for IPv6), and short-range wireless networking with IrDA and Bluetooth.

Symbian started as an operating system that supported primarily C++, but it evolved to providing support for Java as well. Like the other tools that we have looked at, you can download the development SDK for free from the Symbian site and there are a variety of commercial IDE (Integrated Development Environment) that support application development for Symbian. The Java Virtual Machine implementation of EPOC is based on the Personal Java standard.

Deploying Java applications to Symbian is much easier than deploying BREW or J2ME and more like deploying them onto a Windows CE device. This is because Symbian is designed more as a PDA operating system than as an ultra-light mobile environment. Symbian’s latest operating system (Symbian OS 7.0) supports multithreading.

7. Briefly explain about The Tools used in Mobile Computing Model.[NOV 2013]

Mobile Computing is "taking a computer and all necessary files and software out into the field". Mobile computing is any type of computing which use Internet or intranet and respective communications links, as WAN, LAN, WLAN etc. Mobile computers may form a wireless personal network or a piconet.

There are at least three different classes of mobile computing items:

- *Portable computers*, *compacted* lightweight units including a full character set keyboard and primarily intended as hosts for software that may be parameterized, as laptops, notebooks, notepads, etc.

- *Mobile phones* including a restricted key set primarily intended but not restricted to for vocal communications, as cell phones, smart phones, phonepads, etc.
- *Wearable computers*, *mostly* limited to functional keys and primarily intended as incorporation of software agents, as watches, wristbands, necklaces, keyless implants, etc.

The existence of these classes is expected to be long lasting, and complementary in personal usage, none replacing one the other in all features of convenience..

Limitations:

- **Range & Bandwidth**: Mobile Internet access is generally slower than direct cable connections, using technologies such as GPRS and EDGE, and more recently HSDPA and HSUPA 3G and 4G networks. These networks are usually available within range of commercial cell phone towers. Higher speed wireless LANs are inexpensive but have very limited range.
- **Security standards**: When working mobile, one is dependent on public networks, requiring careful use of VPN. Security is a major concern while concerning the mobile computing standards on the fleet. One can easily attack the VPN through a huge number of networks interconnected through the line.
- **Power consumption**: When a power outlet or portable generator is not available, mobile computers must rely entirely on battery power. Combined with the compact size of many mobile devices, this often means unusually expensive batteries must be used to obtain the necessary battery life.
- **Transmission interferences**: Weather, terrain, and the range from the nearest signal point can all interfere with signal reception. Reception in tunnels, some buildings, and rural areas is often poor.
- **Potential health hazards**: People who use mobile devices while driving are often distracted from driving and are thus assumed more likely to be involved in traffic accidents (Cell phones may interfere with sensitive medical devices. Questions concerning mobile phone radiation and health have been raised.
- **Human interface with device**: Screens and keyboards tend to be small, which may make them hard to use. Alternate input methods such as speech or handwriting recognition require training.