

WAP: An Introduction

The Wireless Application Protocol (WAP) is a new advanced intelligent messaging service for digital mobile phones and other mobile terminals that will allow you to see Internet content in special text format on special WAP-enabled mobile phones. Enabling information access from handheld devices requires a deep understanding of both technical and market issues that are unique to the wireless environment. The WAP specification was developed by the industry's best minds to address these issues. Wireless devices represent the ultimate constrained computing device with limited CPU, memory and battery life and a simple user interface. Wireless networks are constrained by low bandwidth, high latency and unpredictable availability and stability. The WAP specification addresses these issues by using the best of existing standards and developing new extensions when needed. The WAP solution leverages the tremendous investment in web servers, web development tools, web programmers and web applications while solving the unique problems associated with the wireless domain. The specification ensures that this solution is fast, reliable and secure. The WAP specification is developed and supported by the wireless telecommunication community so that the entire industry and its subscribers can benefit from a single, open specification.

The WAP forum

The WAP specification was developed by the WAP forum, a consortium founded by the telecommunication giants Nokia, Ericsson, Phone.com and Motorola. The WAP forum's membership roster now includes computer industry heavyweights such as Microsoft, Oracle, IBM and Intel along with several hundred other companies. The WAP forum is an industry group dedicated to the goal of enabling sophisticated telephony and information services on handheld wireless devices. The WAP forum has drafted a global wireless protocol specification for all wireless networks and is contributing it to various industry groups and standard bodies. This WAP specification by the WAP forum enables manufacturers, network operators, content providers and application developers to offer

compatible products and secure services on all devices and networks, resulting in greater economies of scale and universal access to information.

Goals of the WAP forum

The goals of the WAP forum are listed as follows.

- To bring Internet content and advanced data services to wireless phones and other wireless terminals.
- To develop a global wireless protocol specification that works across all wireless network technologies.
- To enable the creation of content and applications that scale across a wide range of wireless bearer networks and device types, i.e. to maintain device and bearer independence
- To embrace and extend existing standards and technology whenever possible and appropriate.

Why WAP is necessary

Ensure interoperability

Service providers must feel secure that their investments will yield benefits in the future. They will not be able to do so until equipment and software offered by different suppliers can be made to work together. The WAP specification has been designed to encourage easy, open interoperability between its key components. Any solution component built to be compliant with the WAP specification can interoperate with any other WAP-compliant component. Service providers can choose equipment and software from multiple WAP-compliant vendors, selecting each piece of the solution that is appropriate for the service provider's particular needs.

Bearer and device independence both help foster interoperability. But interoperability goes beyond these two principles to require that each WAP-compatible component will communicate with all other components in the solution network by using the standard methods and protocols defined in the specification.

Encourage and Foster Market Development

The WAP specification is designed to bring Internet access to the wireless mass market. By building open specifications, and encouraging communication and technical exchanges among the industry players, the **WAP Forum** has already begun to open the wireless data market in new ways. Just over a year ago, the idea of a single wireless data standard was unheard of, yet today the WAP specification is available to the public, and dozens of companies are promoting this vision of the future. The revolution is under way to bring information access to any handset, at a reasonable price and in an easy to use form factor.

The Market Is Different

Bringing computing power to a wireless handset opens an extensive new market for information access. This market is very different from the traditional desktop or even the laptop market because the subscriber has a different set of needs and expectations. Some of these differences include:

- Ease of use
- Market size
- Price sensitivity
- Usage patterns
- Essential tasks

The Network Is Different

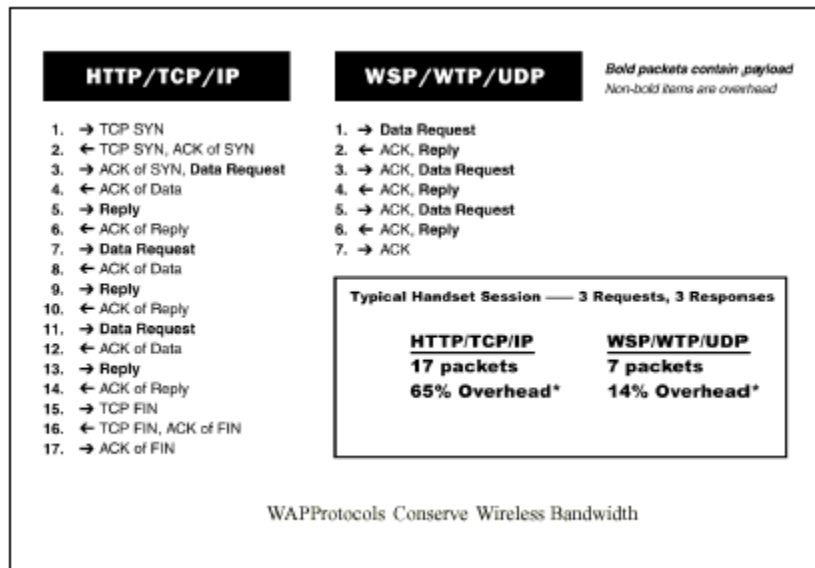
Wireless data networks present a more constrained communication environment compared to wired networks. Because of fundamental limitations of power, available spectrum and mobility, wireless data networks tend to have:

- Less bandwidth:

WAP addresses this issue by minimizing the traffic over the air-interface. WML and WMLScript are binary encoded into a compact form when sent over the air in order to minimise the number of bits and bytes.

- High latency

Wireless networks have high latency compared to wired networks. This is addressed in WAP by minimizing the roundtrips between the wireless device and the wireless network.



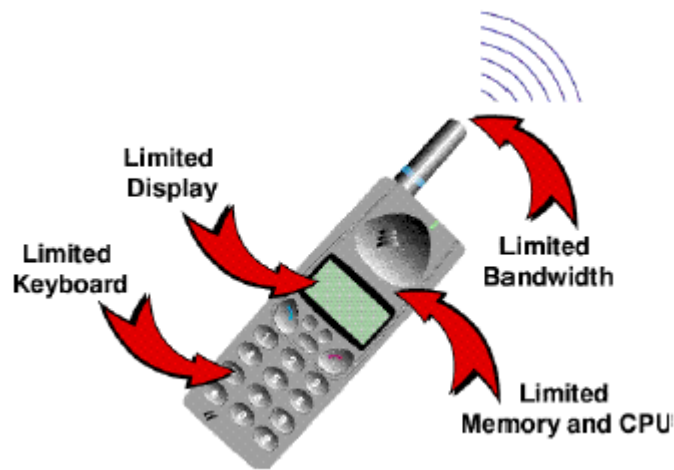
- Less predictable availability

Wired network access provides a more or less reliable connection to the network. That is not the case in wireless networks where bearers might be inaccessible for shorter or longer periods of time due to fading, lost radio coverage or deficient capacity. The problem mentioned above is addressed by allowing lost sessions to be resumed. Selective retransmission is also employed to retrieve small segment of a message that are lost.

Furthermore, as bandwidth increases, the handset's power consumption also increases which further taxes the already limited battery life of a mobile device. A wireless data solution must be able to overcome these network limitations and still deliver a satisfactory user experience.

The Device Is Different

Handheld wireless devices present a more constrained computing environment compared to desktop computers. Because of fundamental limitations of battery life and form factor, mass-market handheld devices tend to have:



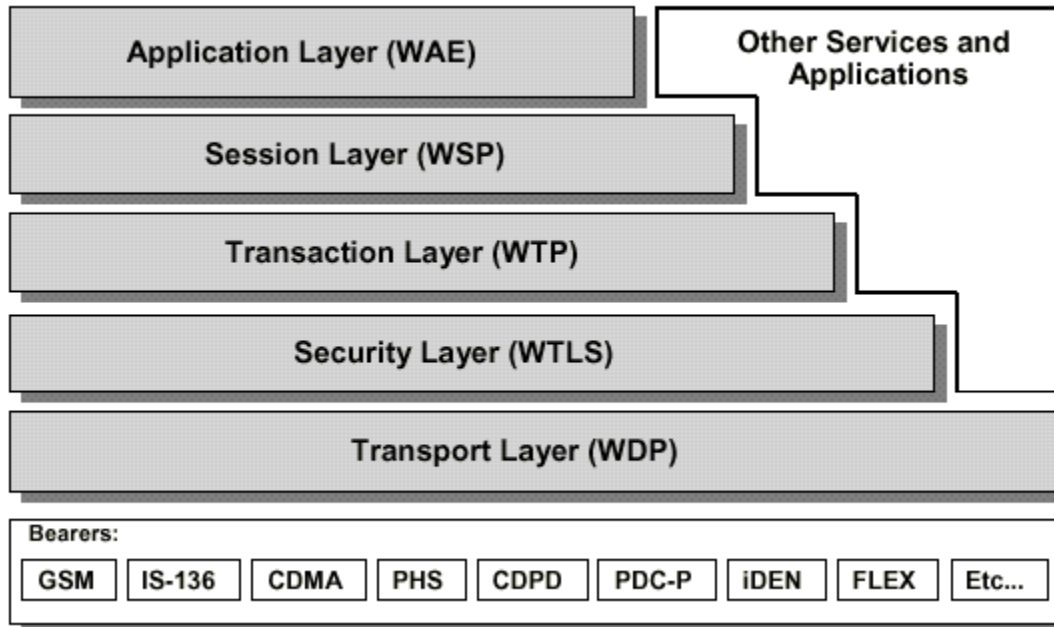
- Less powerful CPUs
- Less memory (ROM and RAM)
- Restricted power consumption
- Smaller displays
- Different input devices (e.g., a phone keypad, voice input, etc.)

Because of these limitations, the user interface of a wireless handset is fundamentally different than that of a desktop computer. The limited screen size and lack of a mouse requires a different user interface metaphor than the traditional desktop GUI.

The WAP protocol stack

WAP is designed in a layered fashion in order to be extensible, flexible, and scalable. With the Open System Interconnection model (OSI model) in mind, the WAP-stack basically is divided into five layers. They are:

- **Application Layer**
Wireless Application Environment (WAE)
- **Session Layer**
Wireless Session Protocol (WSP)
- **Transaction Layer**
Wireless transaction protocol (WTP)
- **Security Layer**
Wireless Transport Layer Security (WTLS)
- **Transport Layer**
Wireless Datagram Protocol (WDP)



Each layer of the WAP protocol stack specifies a well-defined interface to the layer above, meaning that a certain layer makes lower layers invisible to the layer above.

The Application layer: WAE



WML

The Wireless Markup Language is WAP's analogy to HTML used on the WWW. WML is based on the Extensible Markup Language (XML). WML uses a deck/card metaphor to specify a service. A card is typically a unit of interaction with the user, that is, either presentation of information or request for information from the user. A collection of cards is called a deck, which usually constitutes a service. This approach ensures that a suitable amount of information is displayed to the user simultaneously since inter-page navigation can be avoided to the fullest possible extent.

Key features of WML include:

- Variables
- Text formatting features
- Support for images
- Support for soft-buttons
- Navigation control
- Control of browser history
- Support for event handling (for e.g. telephony services)
- Different types of user interactions, e.g. selection lists and input fields

WML can be binary encoded by the WAP Gateway/Proxy in order to save bandwidth in the wireless domain.

WMLScript

Just as JavaScript allows client side processing for use with HTML, WMLScript (Wireless Markup Language Script) is a client side script language for use with WML. JavaScript can be used to make the pages more dynamic, and to perform advanced mathematical calculations in HTML pages. WMLScript is very similar to JavaScript and therefore easy to learn for those who have used that before. WMLScript is actually based on ECMAScript (which is based on Netscape's JavaScript language), however it has been modified in places to support low bandwidth communications and thin clients. WMLScript makes minimal demands on memory and CPU usage, while omitting a number of functions that are not required from other scripting languages. WMLScript is integrated with WML in a particularly flexible way for developers.

Server based computation means that several round trips to and from mobile to server have to make in case of any interaction or computation. This is not very desirable in case of low bandwidth systems. WMLScript allows code to be built into files transferred to mobile client so that many of these round-trips can be eliminated. WML script also allows developer to provide interactivity in WAP pages without taxing the very valuable air interface. WMLScript functions can be called from within WML decks

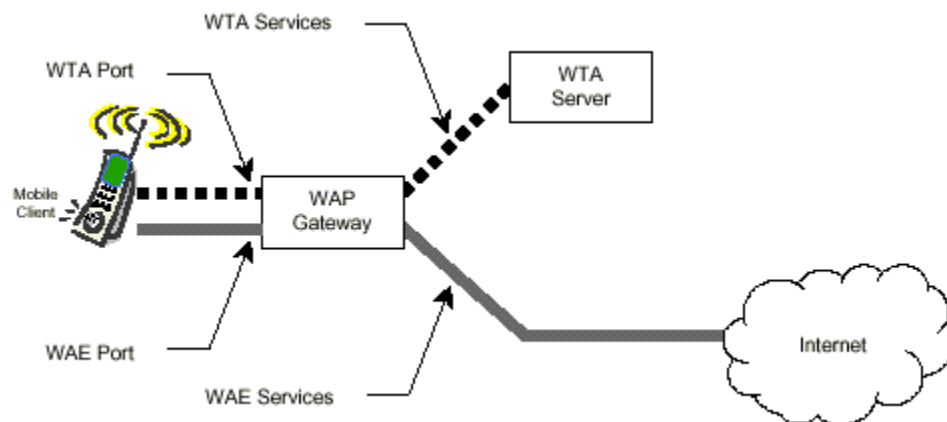
or cards, although the function bodies themselves must be defined in separate WMLScript files. Several predefined core libraries are available in WAP clients that support WMLScript. These libraries provide basic string manipulation and mathematical transforms, as well as functions specific to mobile phones. Just as with WML, WMLScript can be binary encoded by the WAP Gateway/Proxy in order to minimise the amount of data sent over the air.

WBMP

WBMP stands for Wireless BitMaP. It is the default picture format for WAP. The current version of WBMP is called type 0. WBMPs are uncompressed, monochrome black/white bitmaps intended for use in devices with small screens and narrow bandwidth connection. The constraints when using WBMP are the small screen size, limited graphics capabilities and the limited bandwidths available. As a thumb rule, a WBMP should not be wider than 96 pixels and higher than 48 pixels at 72 dots per inch.

WTA

The Wireless Telephony Application (WTA) framework supports Wireless Telephony Applications that interface with the in-device telephony related functions and the network telephony infrastructure. The WTA framework extends the WAE framework by adding:



- An interface from WTA-WML and WMLScript to a specific set of local, telephony related, functions in the client. This interface is called the “Wireless Telephony Application Interface” [WTAI].
- Network event handling. This means that events originating from the mobile network could be detected by the WTA user-agent and actions in response to the events could be defined.
- A repository, which is a storage container, used by the WTA user-agent, that persistently stores content that executes WTA services.
- A model for WTA user-agent state and WTA context management.

The Binary Presentation Formats: WBXML and WMLScriptc



The binary WML format, WBXML, is an encoded form of the WML language. Binary WML is based on a generic binary XML coding format that the WAP Forum has designed to allow compact transfer of XML markup languages. The format defines a basic header, a string table, and a series of binary encoding for the XML tags, attributes, and end tags. WML content is normally parsed and encoded into this binary format before transmission from a content server to a WAP client.

WMLScriptc, or WMLScript "compiled", is the binary encoded form of the WMLScript language. Unlike ECMA/Javascript, which is interpreted directly in the browser, WMLScript is compiled into a virtual byte code on the WAP gateway or WAP server. This byte code is run in a virtual machine on the WAP client device. In this respect, WMLScript is similar to Java running on a JVM.

The Session Layer: WSP



The WAP Session Protocol/B, WSP/B, is a stateless, binary protocol patterned after the HTTP World Wide Web protocol. It consists of a simple request-response pairing. WSP/B contains fields that describe the contents, origin, and types of the request or response contents. There is a one to one correspondence with a subset of the HTTP 1.1 fields. No state information is maintained between requests. WSP uses the WAP Datagram protocol directly for communication with WAP clients. WSP/B is sometimes called WAP connectionless mode.

Some additions in WSP include

- Capability negotiation.
- Header caching
- Long-lived sessions
- Push

The WAP Session Protocol, WSP, is a session oriented, stateful binary protocol used in conjunction with WTP. WSP is a superset of WSP/B and uses the same fields of information. WSP also defines additional protocol formats to support sessions initiation, suspension, and resumption and to maintain session state information. A session is initiated by a WAP client and is maintained until it is explicitly disconnected. WSP sessions can be suspended and resumed and can even switch WDP bearers mid-stream. All WSP information is exchanged using the WAP Transaction Protocol, WTP, described below. WSP is often referred to as WAP connected mode.

The Transaction Layer: WTP

The logo for WTP (WAP Transaction Protocol) is a rectangular box with a grey background and a black border. The letters "WTP" are centered inside the box in a bold, black, sans-serif font.

The WAP Transaction Protocol, WTP, is a confirmed transaction protocol used in conjunction with WSP. WTP is loosely based on a relatively unknown Internet protocol called TTCP/IP. Three different WTP transaction classes are defined in the protocol.

- Class 0

The simplest transaction class, class 0, is basically not a transaction at all. It is a non-confirmed simple push of information in one direction. This transaction class is used for basic information exchange.

- Class 1

Transaction class 1 is used for WAP 1.1 push transactions and is a simple send-acknowledge exchange.

- Class 2

Transaction class 2 is a three-way handshake used for most WSP/WTP information exchange. This handshake is a send-acknowledge-response trio sent from the initiator to the responder and back again.

WTP also has an optional capability to segment and reassemble data. This is similar to the part of the capability built into TCP/IP. WTP class 1 or class 2 transactions use a timeout and resend mechanism when packets are unacknowledged. The recommended timeouts and maximum resend attempts vary depending on the WDP bearer used. For example, the WAP Forum recommends that fairly low timeout and resend values be used with UDP, while much higher values are recommended with SMS.

The Security Layer: WTLS

(WTLS)

The WAP Transaction Layer Security, WTLS, is a session oriented, secure protocol layer patterned after the web's Secure Session Layer (SSL) and Transaction Layer Security (TLS) protocols. The WTLS layer is optional and is independent of the layers above and below it. One unique feature of WTLS is the ability of both client and server to independently recalculate encryption key information based on an embedded sequence number. WTLS is thus optimized to minimize information exchange between client and server.

WTLS is based on Transport Layer Security (TLS) 1.0 but optimized for narrowband communication channels. Key features include:

- Integrity through the use of Message Authentication Codes (MAC)
- Confidentiality through the use of encryption
- Authentication and nonrepudiation of server and client, using digital certificate
- Denial-of-service protection – WTLS contains facilities for detecting and rejecting data that is replayed or not successfully verified. WTLS makes many typical denial-of-service attacks harder to accomplish and protects the upper protocol layers.

There are three levels of WTLS secure sessions.

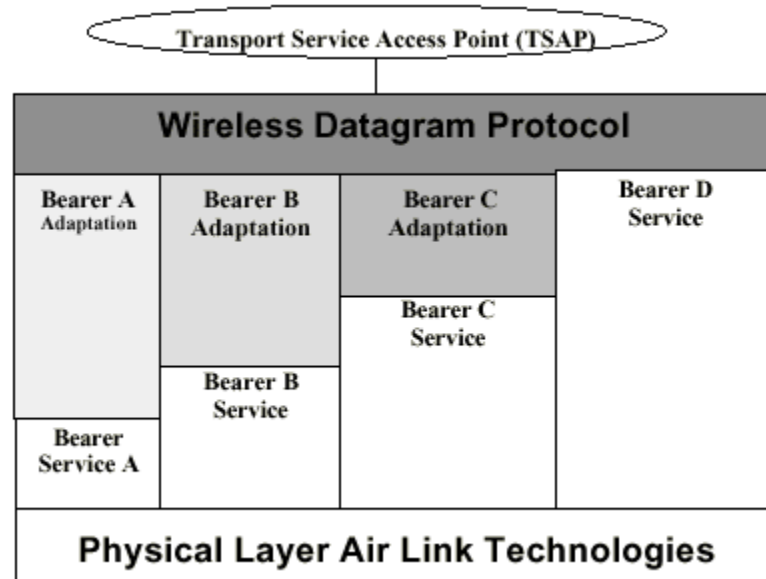
- Level one is anonymous encryption where neither client nor server is authenticated.
- Level two supports server certificates where clients authenticate the server.
- Level three supports client certificates where the server can authenticate the client.

WTLS supports three certificate types: x.509, WTLS, and x.968. The WTLS certificate format is unique to WAP and is designed to minimize information transfer. The x.509 certificate is the same format as that used on the web in SSL and TLS transactions. And the x.968 format is currently not fully specified, but will be supported in the future. WTLS is compatible with both WSP/B and WSP with WTP and can be activated as an additional protocol layer between either of these higher layers and the WDP protocol.

The Data Transport Layer: WDP



The WAP Datagram Protocol, WDP, is a datagram oriented, network layer protocol modeled after the User Datagram Protocol (UDP) used on the Internet. UDP is a member of the TCP/IP protocol suite and is a simple, "best effort" data delivery protocol. On those networks where Internet protocols are present, WDP and UDP are identical. On networks where UDP is not available, WAP defines a UDP equivalent. These UDP equivalents are known as "mappings". The currently defined mappings create the equivalent of UDP over SMS, USSD, and other mobile data transports. WDP makes no attempt to confirm delivery, resend lost packets, or correct errors in transmission. This is left to the higher layer protocols.



The WDP layer operates above the data capable bearer services supported by the various network types. As a general datagram service, WDP offers a consistent service to the upper layer protocol (Security, Transaction and Session) of WAP and communicate transparently over one of the available bearer services. Since the WDP protocols provide a common interface to the upper layer protocols, they are able to function independently of the underlying wireless network. This is accomplished by adapting the transport layer to specific features of the underlying bearer.

Bearers

In order to allow the WAP technology to be exploited as flexibly as possible, use of the protocol can be based on several different bearers. One of these is part of the GSM (Global Systems for Mobile) standard. Other mobile communication systems can also be used to make WAP sessions possible. It is also possible to use the Short Messaging Service (SMS) or a packet oriented data link. If SMS is used, communications between the client and server are maintained with the aid of SMS messages. Up to 140 bytes or 160 characters can be sent in each message. When SMS is used it does not matter how much time elapses between two successive actions in the terminal. Only

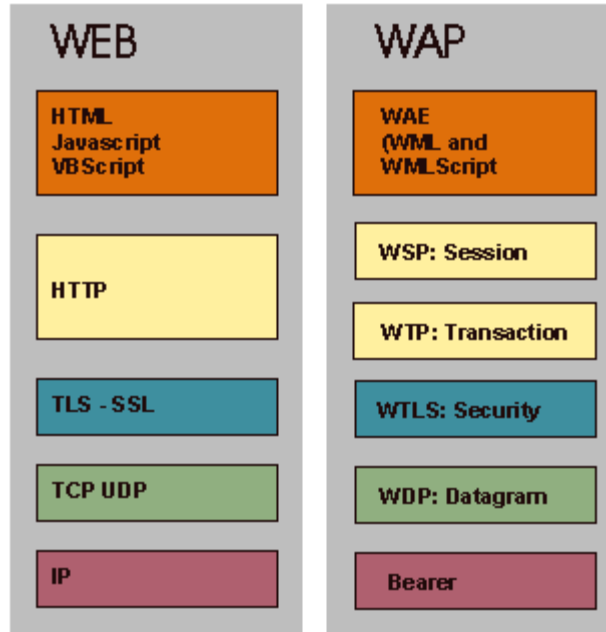
when information is requested or supplied is there any activity on the mobile network and thus charged.

The best option for working with the WAP is to use packet switched data links. This communication protocol can be used once the mobile telephone network operator has implemented General Packet Radio Services. (GPRS) With a GPRS connection, the mobile phone is connected to the network via a packet oriented data link (IP). This effectively means that the GSM network is a subset of the Internet, and all of its mobile phones are IP terminals. This approach makes very efficient use of the available network bandwidth. The fees that are charged in this case are based on the amount of data received, rather than on the amount of time used. Other technologies used for bearer services are Code Division Multiple Access (CDMA), Cellular Digital Packet Data (CDPD), Interim Standard (IS)-136 and Time Division Multiple Access (TDMA).

WAP- An Extension of Internet Model

WAP uses some new technologies and terminologies, which may be foreign to the software developer, however the overall concepts should be very familiar. WAP client applications make requests very similar in concept to the URL concept in use on the Web. The WAP model closely resembles the Internet model of working. In Internet a WWW client requests a resource stored on a web server by identifying it using a unique URL, that is, a text string constituting an address to that resource. Standard communication protocols, like HTTP and Transmission Control Protocol/Internet Protocol (TCP/IP) manage these requests and transfer of data between the two ends. The content that is transferred can either be static like html pages or dynamic like Active Server Pages (ASP), Common Gateway Interface (CGI), and Servlets.

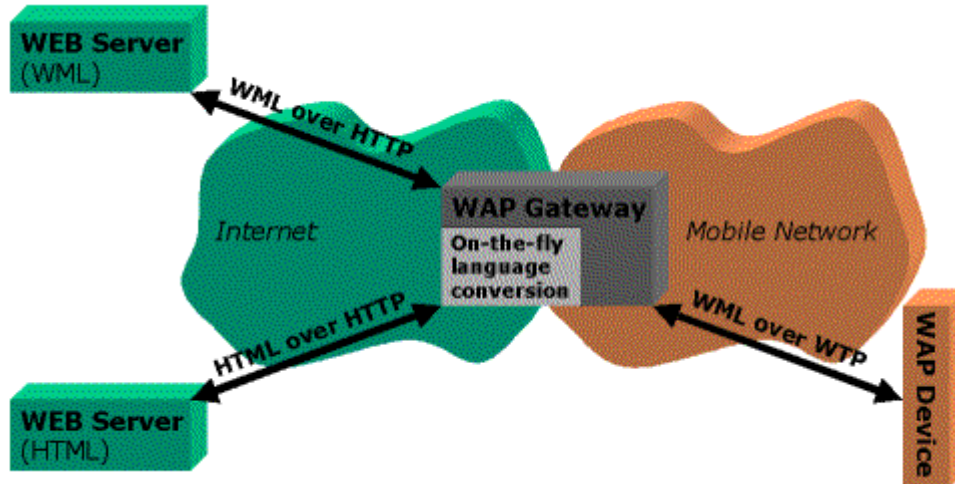
The following figure helps draw a parallel to the Internet protocols. You can see how WAP extends or reuses Internet protocols to achieve mobile Internet access.



WAP Gateway

WAP Gateway is a piece of software that sits between the mobile device and the external network like the Internet. A WAP Gateway typically includes the following functionality:

- Protocol Gateway – the protocol gateway translates requests from the WAP protocol stack to the WWW protocol stack (HTTP and TCP/IP).
- Content Encoders and Decoders – the content encoders translate Web content into compact encoded formats to reduce the size and number of packets traveling over the wireless data network.



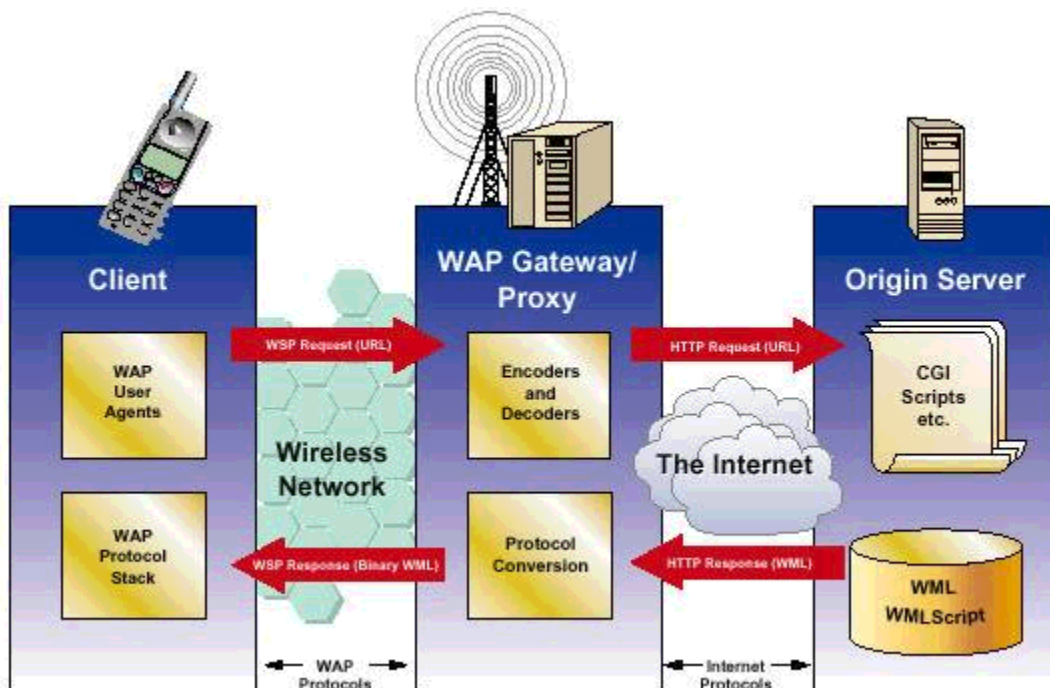
This infrastructure ensures that mobile terminal users can browse a variety of WAP content and applications regardless of the wireless network they use. Application authors are able to build content services and applications that are network and terminal independent, allowing their applications to reach the largest possible audience. Because of the WAP proxy design, content and applications are hosted on standard WWW servers and can be developed using proven Web technologies such as CGI scripting. The WAP Gateway decreases the response time to the handheld device by aggregating data from different servers on the Web, and caching frequently used information. The WAP Gateway can also interface with subscriber databases and use information from the wireless network, such as location information, to dynamically customize WML pages for a certain group of users.

WAP Server

A WAP server is simply a combined web server and WAP gateway. When you are accessing a secure service (using WTLS) the data is sent encrypted to the WAP gateway. The gateway decrypts it and encrypts it using SSL before passing it on to the web server. When the data returns, it will be decrypted and encrypted using WTLS before being sent to your WAP device. By combining the web server and the WAP gateway it is possible to enhance the overall security in these operations involved.

Working of WAP

A WAP request is routed through the WAP gateway, which acts as an intermediary between the “bearer” used by the client (GSM, CDMA, TDMA, etc.) and the computing network that the WAP gateway resides on (TCP/IP in most cases). The gateway then processes the request, retrieves contents or calls CGI scripts, Java servlets, or some other dynamic mechanism, and then formats data for return to the client. This data is formatted as WML (Wireless Markup Language), a markup language based directly on XML. Once the WML has been prepared (known as a deck), the gateway then sends the completed request back (in binary form due to bandwidth restrictions) to the client for display and/or processing. The client retrieves the first card off of the deck and displays it on the screen.



The *deck of cards* metaphor is designed specifically to take advantage of small display areas on handheld devices. Instead of continually requesting and retrieving cards (the WAP equivalent of HTML pages), each client request results in the retrieval of a deck of one or more cards. The client device can employ logic via embedded

WMLScript (the WAP equivalent of client-side JavaScript) for intelligently processing these cards and the resultant user inputs.

To sum up, the client makes a request. This request is received by a WAP gateway that then processes the request and formulates a reply using WML. When ready, the WML is sent back to the client for display.

Benefits

Operators

For wireless network operators, WAP promises to decrease churn, cut costs, and increase the subscriber base both by improving existing services, such as interfaces to voice-mail and prepaid systems, and facilitating an unlimited range of new value-added services and applications, such as account management and billing inquiries. New applications can be introduced quickly and easily without the need for additional infrastructure or modifications to the phone. This will allow operators to differentiate themselves from their competitors with new, customized information services. WAP is an interoperable framework, enabling the provision of end-to-end turnkey solutions that will create a lasting competitive advantage, build consumer loyalty, and increase revenues.

Developers

Application developers can reach the largest possible audience when they write their applications in WML because they are writing to an industry standard. Additional benefits for developers include:

- Access to an entirely new, immense market of information-hungry wireless subscribers, while complementing their existing Internet services.
- Because WML is an XML-based language, it is an easy markup language for existing Web developers to learn.
- WML's basis in XML also positions it well as a future target markup language for automatic content transformation.

Subscribers

The WAP specification delivers significant value to the subscriber. The WAP specification pulls together existing technologies and defines new standards to provide subscribers with:

- Fast, efficient access to essential information from a wireless handset.
- Peaces of mind that all transactions are completely secure.
- An easy to use interface metaphor that meets the needs of the user within the restrictions of a constrained network and device.

Applications

There are various business applications for mobile computing. This includes both horizontal applications that are used by workers and professionals across all the industries as well as those specific to business processes in a vertical industry.

Horizontal applications

Electronic mail

Wireless-network-based e-mail is becoming a popular application available now. In order to provide a high level of customer service, mobile workers and sales professionals must stay in touch with home offices and customers. This is possible only through wireless network support.

Wireless workgroup applications

These applications allow members of a workgroup to access information on workgroup calendaring (scheduling meetings), status of collaborative projects, research and development, time and expense reporting, customer service and other activities where multiple members of a workgroup participate in approval process. Because many of these people are mobile, they need to access this information wirelessly from the field or from their vehicles while they are moving from customer to customer.

Mobile data collection

These solutions are based on some sort of handheld device scanning information on an item and either storing it locally or transmitting it to a central processor. The device might range from simple portable bar code readers to more sophisticated PDTs (Portable Data Terminals) with RF capability that will read information from various devices and send this information automatically through wireless local area networks or wide area networks.

Vertical Applications

Banking

Many banking industry customers are developing wireless applications to improve bottom-line costs. Even the big banks are realizing that their sales people must leave their offices to sell directly to customers. The features provided include:

- Wireless banking transactions - account balance, funds transfer, bill payment
- Sales Professional Automation in financial industry
- Credit card authorization via POS terminals equipped with wireless adapters

Stock Trading

The New York stock exchange has made a significant change to the classical methods used by traders in the past. This include:

- Hand-held PDAs connected to wireless networks, accessing information from stock exchange servers.
- Wireless mobile computing trading from Palm and Pocket PC by large active investors.

Airline and Railway industries

The application in airline and railway industries include:

- Data access for staff via cellular circuit switched network including: ticketing and schedule information, maintenance - fueling and de-icing information and baggage handling information.
- Mobile scanners to scan bar-coded information from baggage tickets directly into a database.
- Pen based work order application using mobile data network.
- Airport security and monitoring.
- Airline Baggage and Cargo Control.

WAP in India

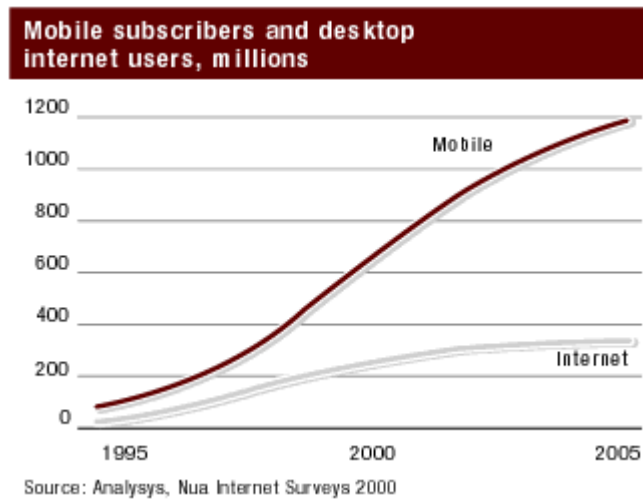
In the past few months, a number of sites such as Rediff.com, Clubgreetings.com and Sharekhan.com have come out with WAP versions, offering a range of services from daily news, stock quotes and weather reports to airline schedules, restaurant listings, e-cards and e-mail to people on the move. All the players hope to be able to take a fair share of the WAP market once it takes off, hopefully by the end of this year. Even so, issues such as low mobile phone penetration, small screen size, low memory of phones and smaller bandwidth continue to dampen this optimism. The major mobile operators providing mobile Internet services in India at present are Orange, Airtel and Tata Cellular. And in a few months, probably all the major mobile operators in this country would have enabled their cellular phones to connect to the WAP-enabled Internet sites.

Future Outlook for WAP

The limitations in mobile Internet access are not just the low bandwidths available. The very nature of mobile devices presents limitations like display and power consumption. Power consumption is a very critical issue and even if high speeds are available power considerations may limit the data speeds. Services like GPRS are bearer services. Internet access via WAP should infact become much more easier with GPRS. Currently, WAP access needs a specific connection via an Internet service provider (ISP) in much the same way as a PC accesses. But the system will come into its own with the

introduction of another enabling technology, general packet radio services (GPRS), a method of sending Internet information to mobile telephones at high speed. By allowing mobile to be in always connected state GPRS (or other services like CDPD) will bring Internet more closer to mobile.

Mobile commerce is one such application that can open up lots of opportunities for WAP. By 2004, there could be more than 700m mobile commerce users. M-commerce is emerging more rapidly in Europe and in Asia, where mobile services are relatively advanced.



WAP is one of the families of technologies that have the potential of bringing about the convergence of mobile communications and the Internet. Technologies like Bluetooth will connect the mobile to the personal computers. All this should make adoption of WAP much more attractive and desirable.

