WAP Caching Model
Version 11-Feb-1999

**Wireless Application Protocol
Cache Model Specification**

Disclaimer:

This document is subject to change without notice.

Contents

1. Scope

Wireless Application Protocol (WAP) is a result of continuous work by the WAP Forum to define an industry-wide specification for developing applications that operate over wireless communication networks. The scope of the WAP Forum is to define a set of specifications to be used by wireless communication device service applications. The wireless market is growing very quickly and reaching new customers and services. To enable operators and manufacturers to meet the challenges in advanced services, differentiation and fast/flexible service creation, WAP defines a set of protocols in transport, session and application layers. For additional information on the WAP architecture, refer to "*Wireless Application Protocol Architecture Specification"* [WAP].

This specification defines the caching model for the WAP user agent. The WAP user agent caching model is an implementation of the HTTP/1.1 caching model used on the World Wide Web. The design is intended to allow the adoption of the HTTP/1.1 caching model with no loss of semantics or function while providing support for low-end WAP phones. See [RFC2068] for more information on HTTP/1.1 caching.

This specification addresses the following issues:

- User agent caching of resources fetched from the network. This may include WML decks, WMLScript compilation units, or any other resources named with a URL and fetched over WSP.

- WAP Gateway behavior – the role and responsibility of the WAP Gateway in the implementation of reliable caching.

- Support infrastructure needed to implement HTTP/1.1 caching on a WAP user agent.

2. Document Status

This document is available online in the following formats:

- PDF format at http://www.wapforum.org/.

2.1 Copyright Notice

© Copyright Wireless Application Forum Ltd, 1999 all rights reserved.

Terms and conditions of use are available from the Wireless Application Protocol Forum Ltd. Web site at http://www.wapforum.org/docs/copyright.htm.

2.2 Errata

Known problems associated with this document are published at http://www.wapforum.org/.

2.3 Comments

Comments regarding this document can be submitted to the WAP Forum in the manner published at http://www.wapforum.org/.

3. References

3.1 Normative References

[NTP]	"Network Time Protocol, Version 3, Specification Implementation and Analysis", D. Mills, University of Delaware, March 1992. URL: ftp://ftp.isi.edu/in-notes/rfc1305.txt
[RFC2068]	"Hypertext Transfer Protocol - HTTP/1.1", R. Fielding, et al., January 1997. URL: ftp://ftp.isi.edu/in-notes/rfc2068.txt
[RFC2119]	"Key words for use in RFCs to Indicate Requirement Levels", S. Bradner, March 1997. URL: ftp://ftp.isi.edu/in-notes/rfc2119.txt
[WAE]	"Wireless Application Environment Specification", WAP Forum, 30-April-1998. URL: http://www.wapforum.org/
[WAP]	"Wireless Application Protocol Architecture Specification", WAP Forum, 30-April-1998. URL: http://www.wapforum.org/
[WSP]	"Wireless Session Protocol", WAP Forum, 30-April-1998. URL: http://www.wapforum.org/

3.2 Informative References

[WML]	"Wireless Markup Language", WAP Forum, 30-April-1998. URL: http://www.wapforum.org/

4. Definitions and Abbreviations

4.1 Definitions

The following are terms and conventions used throughout this specification.

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY" and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

Cache - a device's local store of response messages and the subsystem that controls its message storage, retrieval and deletion. A cache stores cachable responses in order to reduce the response time and network bandwidth consumption on future, equivalent requests.

Content - subject matter (data) stored or generated at an origin server. Content is typically displayed or interpreted by a user agent in response to a user request.

Origin Server - the server on which a given resource resides or is to be created. Often referred to as a web server or an HTTP server.

Request - a WSP request (see [WSP]). Also known as a "WSP method request".

Response - a WSP response (see [WSP]).

Resource - a network data object or service that can be identified by a URL. Resources may be available in multiple representations (eg, multiple languages, data formats, size and resolutions) or vary in other ways.

User - a user is a person who interacts with a user agent to view, hear or otherwise use a resource.

User Agent - a user agent is any software or device that interprets resources. This may include textual browsers, voice browsers and search engines.

4.2 Abbreviations

For the purposes of this specification, the following abbreviations apply.

HTTP	Hypertext Transfer Protocol [RFC2068]
RFC	Request For Comments
TOD	Time Of Day
URI	Universal Resource Identifier
URL	Universal Resource Locator
W3C	World Wide Web Consortium
WAE	Wireless Application Environment [WAE]
WAP	Wireless Application Protocol [WAP]
WSP	Wireless Session Protocol [WSP]
WML	Wireless Markup Language [WML]

5. Caching Model

The WAP caching model is based on HTTP/1.1 caching as defined in [RFC2068]. In addition, a number of extensions and clarifications have been specified to facilitate the operation of HTTP/1.1 caching on limited function devices.

This document applies to all WAP user agents, including those with no cache (ie, a zero byte cache).

5.1 WAP User Agent Responsibilities

The user agent MUST faithfully implement resource caching as described in [RFC2068].

5.1.1 Interaction with the User Agent History Mechanism

The interaction between the user agent cache and the WML history is controlled by the `Cache-control: must-revalidate` header. When going "back" to a URI (e.g., a WML `PREV` task), the user agent should attempt to use the cached resource. The browser must support the following interaction between cached resources and the WML history:

- If the cached resource has the `must-revalidate` cache control attribute set, and the cached resource is stale, then the user agent must revalidate the resource when going "back" in the history. The original request must be made in precisely the same way (eg, the same method, the same request entity if a POST, etc.). This re-fetch must occur without any user interaction.

- If the cached resource does not have the `must-revalidate` cache control attribute, and the cached resource is stale, the user agent must not revalidate the entity when doing "back" in the history.

In other words, the history normally shows the way things were at the time the user retrieved the resource. If the `must-revalidate` attribute is set on a resource, the history will show the "up-to-date" version of the resource.

5.1.2 Intra-Resource Navigation

Navigation and processing within a single cached resource does not require revalidation unless the content type specifies some other cache validation model. For example:

- Function calls within a single WMLScript compilation unit may occur without revalidation. Only the initial fetch of the compilation unit requires a check of the resource validity.

- Intra-deck navigation within a single WML deck may occur without revalidation. Only the initial fetch of the deck requires a check of the resource validity.

5.2 WAP Gateway Responsibilities

The WAP Gateway must faithfully implement the role of an HTTP/1.1 proxy with respect to caching and cache header transmission. See [RFC2068] for more information.

5.2.1 Time-of-Day

The HTTP/1.1 caching model is somewhat sensitive to a loss of time synchronisation. The WAP Gateway should maintain a reliable time-of-day clock, and it is recommended that NTP (or an equivalent) be used by the WAP Gateway to maintain a reliable real-time base. See [NTP] for more information.

5.2.2 Caching Proxy

If the WAP Gateway caches its WSP responses, it will fulfil the role of an HTTP/1.1 caching proxy.

6. Time of Day Clock

The HTTP/1.1 caching model has built-in assumptions about the existence of a time-of-day (TOD) clock in the user agent, and even suggests in section 13.2.3 of [RFC2068] that the client should use NTP to synchronise its clock with the origin server. It is clear that many WAP user agents will not have access to a time-of-day clock, and even fewer will have need of a high-quality time synchronisation protocol such as NTP. However, nearly all wireless communication devices have a real-time clock, and are capable of maintaining a reasonable sense of the TOD when given a means to synchronise their clocks to an outside source.

6.1 TOD Requests

The WAP Gateway must support requests for the current time of day. When the gateway receives a WSP method request that includes a header named `X-WAP.TOD`, it will include that header in the response, with the header value set to the Gateway's current time of day.

The following rule is used to encode the `X-WAP.TOD` values:

```
x-wap-tod = Date-value
```

See [WSP] for the definition of `Date-value`.

The user agent should encode a date value of zero (0) when using `X-WAP.TOD` in a request. The WAP Gateway must ignore the value of the `X-WAP.TOD` header in a request.

6.2 User Agent TOD Clock

If the user agent lacks an accurate TOD clock, it should use the response from a TOD request as the basis for synchronising its real-time clock with the WAP Gateway. For further accuracy, the user agent should estimate the round-trip time of the request, and add one-half of this value from the TOD response value.

It is recommended that user agents containing a TOD clock use this mechanism as a means of determining the skew between the device clock and the network time base.

7. Security Considerations

The storage of sensitive information in non-volatile storage raises a variety of security considerations. Implementers must ensure that private information in the user agent cache is protected from unintended or malicious access. WAP Gateways implementing a caching function must obey all security-related considerations defined in [RFC2068].