

The Web Proxy Description Format
draft-nottingham-web-proxy-desc-00

Abstract

This specification defines a simple means of configuring Web proxies, and places additional requirements upon them in order to promote improved interoperability, security, and error handling.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on March 5, 2015.

Copyright Notice

Copyright (c) 2014 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	2
1.1.	Notational Conventions	3
2.	WPD Proxies	4
3.	The Web Proxy Description (WPD) Format	5
3.1.	name	5
3.2.	desc	5
3.3.	moreInfo	6
3.4.	proxies	6
3.4.1.	host	6
3.4.2.	port	6
3.4.3.	clientNetworks	6
3.5.	forReferers	7
3.6.	alwaysDirect	8
3.7.	failDirect	8
4.	Discovering WPD Files	9
4.1.	The web-proxy-desc well-known URI	9
5.	IANA Considerations	10
6.	Security Considerations	10
7.	Acknowledgements	10
8.	References	10
8.1.	Normative References	10
8.2.	Informative References	11
Appendix A.	User Experience for WPDs	11
	Author's Address	12

1. Introduction

Web proxies are configured in a variety of ways today, but existing approaches suffer from security, usability and interoperability issues.

This specification defines:

- o A simple format for describing a Web proxy ("WPD"; see [Section 3](#)) to facilitate configuration, and so that it can be represented to users in a consistent way, and
- o A way to discover the proxy description using a well-known URL ([Section 4](#)), so that direct configuration of a proxy is as simple as entering a hostname, and
- o A set of additional requirements for proxies described in this fashion, as well as requirements for User Agents connecting to them, designed to improve security, usability and interoperability. See [Section 2](#).

Nottingham

Expires March 5, 2015

[Page 2]

It can be used in a variety of ways, but is designed to meet the following goals:

- o Users should always be aware of a configured proxy and be able to confidently identify it, and
- o Configuring a proxy should be a deliberate act, but simple to do for non-technical users, and
- o Proxies should always respect the wishes of the end user and Web site, and
- o Proxies should never reduce or compromise the security of connections, and improve it where possible, and
- o The proxy should be able to reliably communicate with the end user regarding its policies and problems that are encountered.

Furthermore, it is designed to be useful in the following cases:

- o An end user wants to use a proxy network that provides improved performance, by re-compressing responses to http:// resources.
- o An end user wants to use a proxy network that provides improved privacy, by routing requests through any number of intermediaries.
- o An end user is required to use a proxy to access Internet resources by their network (e.g., a school, workplace or prison).
- o A network wants to offer enhanced access to selected Web sites, through interposition of a proxy.

Importantly, this specification does not address the automatic discovery of proxy configuration for a given network, because proxy configuration is a security-sensitive action, and ought never be performed without explicit user or administrator action.

It is expected that the mechanisms described could be implemented by a single program (e.g., a Web browser), or through an Operating System facility.

1.1. Notational Conventions

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](#)].

Nottingham

Expires March 5, 2015

[Page 3]

2. WPD Proxies

This specification defines a particular kind of HTTP proxy (as per [\[RFC7230\] Section 2.3](#)) known as a "WPD proxy" that has additional requirements placed upon it, as well as upon those using it.

WPD Proxies MUST support HTTP/2 [[I-D.ietf-httpbis-http2](#)] over TLS for connections from clients. Clients that cannot establish a HTTP/2 connection to a WPD proxy MUST consider that proxy "failed."

WPD Proxies MUST support forwarding requests with the "http" scheme [[RFC7230](#)], and SHOULD support the CONNECT method, as specified in [[I-D.ietf-httpbis-http2](#)] [Section 8.3](#).

[RFC7230] [Section 5.7.2](#) requires proxies to honour the semantic of the "no-transform" cache-control directive, and append the 214 (Transformation Applied) warn-code to other messages that have been transformed; WPD proxies MUST honour these requirements.

When connecting to a WPD proxy, clients MUST use TLS and MUST validate the proxy hostname as per [[RFC2818](#)] [Section 3.1](#). If the proxy presents an invalid certificate, that proxy MUST be considered "failed" and not used (until a valid certificate is presented).

User agents MUST use a CONNECT tunnel when retrieving URLs with the "https" scheme through WPD proxies.

When user agents encounter 5xx responses to a CONNECT request from a WPD proxy, they MUST present the response to the end user, but MUST NOT present or process it as a response to the eventual request to be made through the tunnel (i.e., it has an unidentified payload, as per [[RFC7231](#)] [Section 3.1.4.1](#)).

NOTE: Many user agents refuse to show an error response to a CONNECT to the user, in order to deal with the issues brought to light by [[bad-proxy](#)]. While effective in dealing with those attacks, doing so effectively disallows communication between the proxy and the end user; this requirement is designed to re-open that channel.

If a WPD proxy becomes unresponsive, clients SHOULD consider it failed and attempt to use another proxy (if available) or inform the end user (if not available). Clients SHOULD regularly attempt to re-establish contact with failed WPD proxies (e.g., every minute).

Nottingham

Expires March 5, 2015

[Page 4]

3. The Web Proxy Description (WPD) Format

WPD is a JSON [[RFC7159](#)] format that describes a Web proxy to potential clients. Its root is an object containing WPD-specific object members. For example:

```
{
  "name": "ExampleCorp Web Proxy",
  "desc": "ExampleCorp's Proxy Gateway for Web access. Note that
          all traffic through this proxy is logged, and may be
          filtered for content.",
  "moreInfo": "https://inside.example.com/proxy/",
  "proxies": [
    {
      "host": "proxy.example.com",
      "port": 8080,
      "clientNetworks": ["192.0.2.0/24"]
    },
    {
      "host": "proxy1.example.com",
      "port": 8080,
      "clientNetworks": ["192.0.2.0/24"]
    }
  ],
  "alwaysDirect": ["example.com", "192.0.2.0/24"],
  "failDirect": False
}
```

When configuring a proxy through WPD, a user agent SHOULD present the relevant information contained within (i.e., the 'name', 'desc' and 'moreInfo' members, the latter as a link) to the end user. User agents SHOULD also make this information available to the end user whenever the WPD is in use.

The remainder of this section defines the content of these members. Unrecognized members SHOULD be ignored.

3.1. name

A string containing a short, memorable name for the proxy; typically 64 characters or less. This member MUST be present for the WPD to be considered valid.

3.2. desc

A string containing a textual description of the proxy's function(s); typically 256 characters or less. This member MUST be present for the WPD to be considered valid.

Nottingham

Expires March 5, 2015

[Page 5]

3.3. moreInfo

A string containing a URL [[RFC3986](#)] that leads to more information about the proxy, its operation, who operates it, etc. The URL MUST have a scheme of "https" [[RFC7230](#)], and MUST be able to respond with an HTML [[W3C.CR-htm15-20140731](#)] representation. This member MUST be present for the WPD to be considered valid.

3.4. proxies

An array containing one or more proxy objects; each proxy object represents a HTTP proxy endpoint that can be used when this WPD is configured. See [Section 2](#) for requirements specific to these proxies, as well as those clients connecting to them.

Proxy objects' members are defined by the following subsections; unrecognized members SHOULD be ignored.

The ordering of proxy objects within the proxies array is not significant; clients MAY choose any proxy they wish (keeping in mind the requirements of clientNetworks), and MAY use more than one at a time.

NOTE: the array of proxy objects is functionally similar to, but not as expressive as, the commonly-used "proxy.pac" format. While it would be expedient for WPD to just reference a proxy.pac, feedback so far is that proxy.pac has a number of deficiencies, and interoperability over it is poor. Therefore, this document specifies the proxy object instead, in order to gather feedback on an alternative approach.

3.4.1. host

A string containing the host (as per [[RFC3986](#)], [section 3.2.2](#)) of the proxy. This member MUST be present.

3.4.2. port

A number representing the port that the proxy is listening on. This member MUST be present, and MUST be an integer.

3.4.3. clientNetworks

An array containing strings; each string contains a classless prefix (see [[RFC4632](#)]) which the proxy can be used within. Clients MUST NOT attempt to use the proxy if their IP address is not within one of the stated ranges.

Nottingham

Expires March 5, 2015

[Page 6]

This member is optional.

For example, if the value of `clientNetworks` is

```
[ "192.168.1.0/32", "192.168.2.0/24" ]
```

then the only clients that could use the proxy would have IP addresses in the ranges 192.168.1.0 to 192.168.1.3 and 192.168.2.0 to 192.168.2.255.

Note that by their nature private networks (as specified in [\[RFC1918\]](#)) are not unique, and therefore there may be false positives. As such, clients SHOULD NOT automatically configure a WPD based upon `clientNetworks` when the IP address is in these ranges, although they MAY notify the user of a WPD's possible applicability, and MAY use additional information to correlate a WPD to its proper network.

[3.5.](#) forReferers

An array containing strings; each string is a host (as per [\[RFC3986\]](#) [Section 3.2.2](#)).

When `forReferers` is present, Clients MUST use the WPD's proxies to access these hosts, hostnames that have the host as a root, and for traffic generated by that content. They MUST NOT be used for other traffic.

This member is optional.

For example, if the value of `forReferers` is

```
[ "friendface.example.com" ]
```

then requests to "friendface.example.com", "www.friendface.example.com", "app.friendface.example.com" etc. would use the associated proxies; likewise, if processing a response from one of these hosts generated further requests to "images.example.net" and "scripts.example.org", they would also use the proxies.

Note that `alwaysDirect` takes precedence over `forReferers`.

TODO: tighten up what "processing" means here; the intent is to omit a href

3.6. alwaysDirect

An array containing strings; each string is one of:

- o a host (as per [\[RFC3986\] Section 3.2.2](#)),
- o a classless prefix [\[RFC4632\]](#).
- o the string "CONNECT".

Clients MUST NOT use the WPD's proxies to access nominated hosts and hostnames that have the a nominated host as its root. Likewise, clients MUST NOT use the WPD's proxies to access bare IP addresses that fall within the classless prefix.

If the string "CONNECT" appears in alwaysDirect, it indicates that requests that require establishment of a tunnel (e.g., for "https" URLs) MUST NOT use the WPD's proxies, but instead ought to be made directly to the origin (i.e., without a tunnel).

Note that when a "bare" IP address or classless prefix is used in alwaysDirect, clients are not required to perform a reverse lookup on hostnames; these forms are only intended for accessing URLs that use the IP-literal or IPv4address forms.

This member is optional.

For example, if the value of alwaysDirect is:

```
[ "example.com", "192.168.5/24" ]
```

then requests to "example.com", "www.example.com", "foo.example.com" etc would not use any proxy. Likewise, requests whose URL authority were bare IP addresses in the range 192.168.5.0 to 192.168.5.255 would not use any proxy.

3.7. failDirect

A boolean indicating whether the client should attempt to directly access the origin server if all applicable proxies are unavailable.

When False, clients MUST NOT attempt to directly access the origin server when no proxy is available, but instead SHOULD inform the user that the proxy is unavailable.

When True, clients MAY do so. If failDirect is not present, clients MAY default to this behavior.

Nottingham

Expires March 5, 2015

[Page 8]

4. Discovering WPD Files

To facilitate easy configuration of WPD proxies, this specification defines a well-known URI [[RFC5785](#)]. Doing so allows a proxy's description to be found with a simple hostname; e.g., "proxy.example.net" or even just "example.net".

Clients MUST NOT use the DHCP "WPAD" mechanism to discover WPDs.

4.1. The web-proxy-desc well-known URI

The "web-proxy-desc" well-known URI allows discovery of a Web Proxy Description ([Section 3](#)).

This well-known URI is only valid when used with the "https" URI Scheme [[RFC7230](#)]; it MUST NOT be used with "http" URIs. In other words, WPD discovery is always protected by TLS [[RFC5246](#)].

The description found at this location is considered valid for its freshness lifetime, as defined in [[RFC7234](#) [Section 4.2](#)]. Once stale, clients SHOULD refresh it and apply any changes.

If the WPD is not retrievable (e.g., a 404 response status), invalid (as per JSON [[RFC7159](#)] or the requirements in [Section 3](#)), or its certificate is not valid for the host (as per [[RFC2818](#) [Section 3.1](#)]), the client MUST NOT use the WPD, and if a user agent, SHOULD inform the end user.

The well-known URI MAY use proactive content negotiation ([\[RFC7231](#) [Section 3.4.1](#)]) to select an appropriate language for the response representation. Therefore, clients SHOULD send an Accept-Language request header field ([\[RFC7231](#) [Section 5.3.5](#)]) when they wish to advertise their configured language.

The registration template is:

- o URI suffix: web-proxy-desc
- o Change controller: IETF
- o Specification document(s): [this document]
- o Related information: only to be used with 'https' scheme

5. IANA Considerations

This specification registers a new well-known URI, as per [\[RFC5785\]](#). See [Section 4.1](#) for the template.

6. Security Considerations

If a user can be convinced to configure a WPD hostname as their proxy, that host can observe all unencrypted traffic by the client. As such, WPD configuration interfaces ought only allow configuration of proxies once their identity is validated (as required), and the user ought to be given access to all relevant information about the WPD proxy (i.e., 'name', 'desc' and 'moreInfo', the latter as a link). Furthermore, WPD proxies ought only be configured as the result of an intentional act, not as a side effect of normal Web browsing.

7. Acknowledgements

Thanks to Patrick McManus for his feedback and suggestions.

8. References

8.1. Normative References

- [I-D.ietf-httpbis-http2] Belshe, M., Peon, R., and M. Thomson, "Hypertext Transfer Protocol version 2", [draft-ietf-httpbis-http2-14](#) (work in progress), July 2014.
- [RFC1918] Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G., and E. Lear, "Address Allocation for Private Internets", [BCP 5](#), [RFC 1918](#), February 1996.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC2818] Rescorla, E., "HTTP Over TLS", [RFC 2818](#), May 2000.
- [RFC3986] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax", STD 66, [RFC 3986](#), January 2005.
- [RFC4632] Fuller, V. and T. Li, "Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan", [BCP 122](#), [RFC 4632](#), August 2006.

Nottingham

Expires March 5, 2015

[Page 10]

- [RFC7159] Bray, T., "The JavaScript Object Notation (JSON) Data Interchange Format", [RFC 7159](#), March 2014.
- [RFC7230] Fielding, R. and J. Reschke, "Hypertext Transfer Protocol (HTTP/1.1): Message Syntax and Routing", [RFC 7230](#), June 2014.
- [RFC7234] Fielding, R., Nottingham, M., and J. Reschke, "Hypertext Transfer Protocol (HTTP/1.1): Caching", [RFC 7234](#), June 2014.
- [W3C.CR-html5-20140731]
Berjon, R., Faulkner, S., Leithead, T., Navara, E.,
O'Connor, E., and S. Pfeiffer, "HTML5", World Wide
Web Consortium CR CR-html5-20140731, July 2014,
<<http://www.w3.org/TR/2014/CR-html5-20140731>>.

[8.2.](#) Informative References

- [RFC5246] Dierks, T. and E. Rescorla, "The Transport Layer Security (TLS) Protocol Version 1.2", [RFC 5246](#), August 2008.
- [RFC5785] Nottingham, M. and E. Hammer-Lahav, "Defining Well-Known Uniform Resource Identifiers (URIs)", [RFC 5785](#), April 2010.
- [RFC7231] Fielding, R. and J. Reschke, "Hypertext Transfer Protocol (HTTP/1.1): Semantics and Content", [RFC 7231](#), June 2014.
- [bad-proxy]
Chen, S., Mao, Z., Wang, Y., and M. Zhang, "Pretty-Bad-Proxy: An Overlooked Adversary in Browsers' HTTPS Deployments", January 2009, <research.microsoft.com/jump/79323>.

[Appendix A.](#) User Experience for WPDs

There are a variety of ways to present proxy configuration to users and administrators, so this specification does not constrain how this is done. That said, guidance for the common case (visual Web browsers) can be helpful in assuring consistent user experience.

One of the core principles of this specification is that WPDs need to be explicitly configured, either by the end user or an administrator on their behalf. This is because using a proxy is a security-sensitive operation; if an attacker can automatically configure a proxy, or convince a user to do so as part of accessing a site, they

Nottingham

Expires March 5, 2015

[Page 11]

can gain access to the user's traffic, even when the user leaves the attacking network.

Therefore, a user agent might allow configuration by entering a hostname (e.g., "example.net"), whereupon it retrieves the WPD, validates its certificate and contents, and present its information to the end user for confirmation.

Once a WPD is confirmed, a user agent might "remember" it for future use; e.g., by allowing quick configuration through a drop-down menu. When a WPD nominates clientNetworks and the client does not have a suitable IP address, the drop-down might make that option unavailable.

It is envisioned that only a single WPD ought be configured at a time; combining WPDs leads to ambiguity regarding precedence and therefore user confusion.

When a WPD is active, its ought be visible to the end user, to remind them of its presence, and to offer more information about the configured proxy.

Author's Address

Mark Nottingham

Email: mnot@mnot.net

URI: <http://www.mnot.net/>

