Scenarios for the Delivery of Negotiated Content

1. Status of this Memo

This document is an Internet-Draft. Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

To learn the current status of any Internet-Draft, please check the "id-abstracts.txt" listing contained in the Internet-Drafts Shadow Directories on ds.internic.net (US East Coast), nic.nordu.net (Europe), ftp.isi.edu (US West Coast), or munnari.oz.au (Pacific Rim).

Distribution of this memo is unlimited. Please send comments to the author.

2. Abstract

A number of Internet application protocols have a need to provide content negotiation for the resources with which they interact. While MIME media types [1] provide a standard method for handling one major axis of variation, resources also vary in ways which cannot be expressed using currently available MIME headers. This paper asserts that a cross-protocol negotiation framework is needed, both to leverage work already done for specific protocols and to allow for content negotiation when resources will pass through several protocols on their way from provider to recipient. To justify the need, this paper puts forward several content negotiation scenarios and discusses how they affect the requirements for such a framework.

3. Introduction

When a content provider makes a particular resource available in a number of different representations, content negotiation may be used to determine which of the available variants should be provided to a particular recipient. This negotiation entails the exchange of information between the resource provider and the recipient about their respective capabilities and preferences. Among the key pieces of information needed in this exchange are:

1) How the available representations of the resource vary.

2) The preferences of the provider among the available representations.

3) The capabilities and preferences of the recipient.

3.1 Representing variation
The MIME Content-type header can be used to differentiate representations which differ by media type, and it is already used by many protocols for that purpose. Many representations differ, however, by features which are not reflected in media types; two representations using the image/gif type may, for example, differ in color depth. There are several possible methods for referring to these features:

a) using a feature tag registry to record specific feature tags which can be used to describe variants.

b) using URIs to point to descriptions of specific features.

c) using standardized reference documents as pointers to descriptions of the features.

In each case, the association of a particular feature with a registered tag, URI, or standardized reference must take place prior to any content negotiation using that feature. The selection of which method to use to represent variation will depend primarily on how the need for extensibility contrasts with the need for a stable reference. A single feature tag registry provides a highly stable core set of features, but is likely to be relatively slow in deployment of new features. A URI-based system is likely to be highly adaptable to new features, but susceptible to duplication and loss of information. The use of pointers to standardized reference documents allows more than one organization to participate in the vetting of new features, but this method may engender confusion about which organization should review a feature.

3.2 Provider preferences

Provider preferences are commonly expressed as opinions on the quality of various representations of the resource. The determination of quality may be made by reference to a particular standard; the determination of the lossiness of a JPEG image may, for example, be made with reference to the JPEG standards. Quality may also be assigned on a subjective basis, however; an author might, for example, decide that the loss of figures in an ASCII representation of resource implied a loss in quality, but she would have to invent a figure to describe that loss.

The numerical expression of quality indicators can be ascending or descending. In HTTP 1.1 [2], lower qvalues represent increasing degradation of quality; in NTP [3], in contrast, quality indicators like peer.dispersion use increasing values to represent increasing degradation. A standard method for indicating provider preferences for content-negotiating protocols is clearly needed for any system based on numerical expressions of quality to work. Any system based on non-numerical expressions of quality will similarly need standardization of the terms by which quality is expressed.

3.3 Recipient capabilities and preferences

The expression of recipient capabilities and preferences presents all of the problems of feature description and provider preferences, as well as some unique issues. If a standardized method of expressing features is available (cf 3.1, above), recipient capabilities may be expressed with reference to Content-type and one or more features. Similarly, recipient preferences can be expressed using a numerical system like those described above in 3.2, with many of the same
problems of how to apply subjective criteria being present.

When to send indications of recipient preferences and capabilities and how to indicate dependencies among them present problems unique to the recipient side of the equation. A content provider knows when multiple variants are present and on what basis the representations vary. The recipient, in contrast, does not know a particular resource has variants unless the content provider chooses to present that information. Depending on the type of exchange, a recipient may be able to choose when or if to present its capabilities and preferences. If it can and chooses to do so prior to every exchange, it will be wasting network and processor resources whenever the resource has no variants. If it never presents its capabilities it may get a less than optimal variant. If it waits for an indication that variants exist, a lag in the exchange is introduced which may be quite significant for store-and-forward protocols.

Expressing the dependencies among various capabilities also presents problems. A printer may, for example, be able to print either in color or at a high density; if the relationship between the two features is not fixed, however, expressing it appropriately may be very difficult.

4.0 Scenarios

Content negotiation may be provider-based, recipient-based, or based on active content.

4.1 Provider-based negotiation

Provider-based content negotiation occurs when the provider of a resource selects among the available variants without (or prior to) notifying the potential recipient that alternatives exist. In HTTP, for example, content providers may configure their servers to deliver certain variants based on the user-agent of the requester or the Accept headers available in the request.

This method is obviously limited to protocols in which recipients initiate requests for resources and provide information about their capabilities in those requests. Beyond that limitation are several problems of scalability. As features multiply, the overhead of providing information on capabilities for each potential feature grows, and since that overhead is incurred for every exchange it can rapidly lead to network resource problems. The use of feature "bundles" would ameliorate this somewhat, but it would increase the problems associated with maintenance for both provider and recipients.

Provider-based selection may also present privacy problems, as the full set of preferences and capabilities must be released to the provider for the negotiation to operate at its best. This creates an effective user profile, which could be combined with other information to reveal more about potential recipients than they intend or expect.

Lastly, it is common for a single provider to serve multiple recipients. As the number of potential recipients grows, the processing power required to handle variant selection also grows, and this can strain providers.

4.2 Recipient-based negotiation

Recipient-based selection occurs whenever the provider notifies
the recipient that alternatives exist and allows the recipient to select among them. The Multipart-alternative MIME type is currently the chief method by which providers notify potential recipients that alternatives exist. This message type allows the provider to send either all of the available variants or, with external message bodies, pointers to each of them. The alternatives are ranked by the provider according to the provider’s preferences, and the recipient selects the first of the alternatives that it can handle.

An Alternates header [4] has been proposed to both extend and simplify the Multipart alternative method for delivering information on available variants. Using an Alternates header, a provider can notify a potential recipient of available variants without the overhead of MIME multipart. The Alternates header also proposes a method by which the provider can indicate on what basis the variants differ, so that recipients can select based on the variation rather than the rank assigned by the provider.

This form of negotiation limits the process of selection among variants to the axes along which they vary. This eliminates the need for recipients to enumerate their capabilities and preferences, which increases the recipients’ privacy and reduces the possibility of bandwidth wastage or provider processor strain. This shift to recipient-based selection increases, however, the processing demanded of the recipient. Depending on how the alternatives are presented, it also either wastes bandwidth in the delivery of all available variants or requires an additional step in the secondary retrieval of the one selected. If no resources are available or if the available methods for selecting among them are inadequate, this method may also require direct user intervention.

4.3 Active content negotiation

Active content negotiation takes place when the resource provided contains content for execution in the recipient’s environment; this active content then determines the capabilities and preferences of the recipient and retrieves or produces the best available variant.

This method has several advantages: it reduces the possible need for active user intervention; it has the potential to use more complex algorithms for selection than are available with Multipart-alternative or Alternates-based systems; and the active content can be cached for later use.

There are, however, several limitations to the usefulness of active content negotiation. The most severe problem is that there is no cross-platform or cross-protocol standard for active content. Without such a standard, providers must engage in other forms of content negotiation before providing the correct form of active content. The use of active content also raises the bar for recipients again, and some recipients will be unable to use it, either because of limited processing power or because of security considerations (cf 6, below).

5. Requirements and Desiderata

The limitations of each of the current methods of content negotiation signal the need for a comprehensive framework for negotiation. For that framework, this author believes that there are three principal requirements and three additional desiderata:
1) A cross-protocol negotiation framework must include a standardized method for reflecting content features.

2) A cross-protocol negotiation framework must include a standardized method for associating variants with features.

3) A cross-protocol negotiation framework must provide a method for indicating provider and recipient preferences for specific features

4) A cross-protocol negotiation framework should have the minimum possible impact on network resource consumption, especially in terms of bandwidth and number of round-trips required.

5) A cross-protocol negotiation framework should protect the privacy of users’ profiles and providers’ inventories of variants.

6) A cross-protocol negotiation framework should allow intelligent gateways, proxies, or caches to participate in the negotiation.

6. Security Consideration

Active content negotiation presumes that recipients will execute code provided and will allow that code to examine user preferences and capabilities in order to select among variants. Any such system would have to be tightly bounded to prevent the active content from executing arbitrary code on the client.

Detailed listings of user preferences and capabilities present possible privacy problems, whether that listing occurs in provider-based negotiation or via active content. Some content providers may also find that exhaustive listings of available representations may compromise their privacy.

7. Acknowledgments

Larry Masinter, Koen Holtman, Graham Klyne, Scott Lawrence, Dan Wing, Andy Mutz, and Neil Joffe provided valuable comments on previous versions of this draft. Discussions within the HTTP working group and the IETF-FAX mailing list also provided insights into the scope of negotiation needed in different contexts.

8. References


9. Author’s Address

Edward Hardie