Internet Draft

A. Beck
M. Hofmann
Lucent Technologies
M. Condry
Intel Corporation

Expires: May 2001
Document: draft-beck-opes-esfnep-01.txt
Category: Informational

November 21, 2000

Example Services for Network Edge Proxies

Status of this Memo

Table of Contents

1 Introduction

The rapid growth of the Internet and the increasing number of Internet users have resulted in many scaling and growth problems with application designs focused on operations at the ends (i.e. the client or the server). This has led to a wide deployment of network

edge caching proxies as a key strategy to address these problems. These systems have been very successful in accelerating Web content delivery and reducing the load on origin Web servers.

However, the specific role of these network edge caching proxies as a gateway between Web users and content providers suggests utilizing them for intelligent services beyond simple caching.

There are already a variety of existing or proposed approaches that implement particular services on top of a proxy platform. ICAP [5]

extends the basic idea of implementing value-added services on proxies by handling transport of web objects between proxies and content modification servers, thus, enabling remote call out mechanisms. EPSFW [2, 7] describes an extended framework to provide general services on top of an open proxy platform.

This document discusses several service examples possibly being implemented on top of an open proxy platform as described in [2, 7]. Each of the following service description consists of three subsections: a short abstract that describes the service idea, a description of the underlying business model, and finally a section that mentions technical challenges to be addressed when implementing these services.

Section 2 describes virus scanning as an example service, which currently is one of the most frequently cited service ideas. Section 3, 4, and 5 describe services that dynamically assemble personalized content. These services exhibit the use of the proxy device managing information about the client. Sections 6 and 7 present services that adapt content to the capability of client devices and client access bandwidth. Some of the previous service ideas can also be applied to streaming media, which is discussed in Section 8. The services given in Section 9, 10, and 11 operate on client requests rather than on the content itself. More service examples are given in Sections 12 and 13.

2 Virus Scanning

2.1 Abstract

Viruses, Trojan Horses, and worms have always posed a threat to Internet users. Just recently we have seen a number of e-mail based

worms that have hit millions of Internet users worldwide within a few hours.

With the help of a content scanning and filtering system at the caching proxy level, Web pages and also file transfers could be scanned for malicious content prior to sending them to the user. In Web pages active content like ActiveX, Java and JavaScript could be scanned for harmful code (e.g. code exploiting security holes). File transfers could be scanned for known viruses. If a virus is found, the adaptation server could try to remove it or deny the delivery of the infected content. A general rule could be that the caching proxy may store and/or deliver content only, if it has been scanned by the content adaptation server and no viruses are found.

2.2 Business model

This service could be offered as an additional feature to ISP customers who are concerned about security issues. Likewise enterprises could be interested in this solution to prevent any malicious content from entering the company network.

2.3 Technical Challenges

Web pages/files should be scanned for viruses by sending them to a separate server where virus-scanning software would analyze them. ICAP [5] is an example protocol for this purpose. The virus scanning operations should not be performed on the caching proxy as they will probably affect the performance of the caching proxy.

If HTTP file transfers are to be scanned for viruses and the requested file cannot be found in the cache, we have to use a different approach than for Web pages. It would not be feasible if the proxy waited for the requested file to be received completely before sending it over to the content adaptation server for the virus scan. This approach would lead to a long delay at the userÆs end, which is not acceptable. Instead, we would have to scan the file transfer continuously, as it is being sent to the user (similar to streaming media).

3 Insertion of Ad Banners

3.1 Abstract

Many Internet companies rely heavily on revenue made by selling advertisement space on their Web pages. Whenever advertisement banners are inserted dynamically depending on who requests the page, they cannot be cached, even when the content of the page itself is static. This behavior prevents Web pages from being cached, although their static content would allow for it.

Therefore it seems reasonable to cache the static part of those Web pages at a caching proxy near the client and to insert ad banners into the cached Web pages before serving them to the client.

3.2 Business model

This service is a sales item to Internet advertising networks. They obtain a market from customers wishing a low cost network access in return for advertising. This is the free ISP market. Also, content providers who do not want to outsource their ad space management and sales might be interested in providing banner images and insertion rules to proxies/content adaptation servers to accelerate the delivery of their Web pages.

An ad insertion module at the caching proxy of the Free ISP could insert ad banners (in addition to any ad banners from the content provider) into every Web page requested by a customer. That way the customers of the Free ISP will not have to install any special software in order to use its service.

3.3 Technical Challenges

The caching proxy would have to recognize when and where to insert ad banners into a Web page before serving it to the client. The proxy could for instance scan the Web page for a specific marking (e.g. a special tag). In the case of a Free ISP ad banners would probably always be inserted at the same position (e.g. in a frame at the top of each page) or in a separate pop-up window.

If we wanted to insert advertisements based on the user and his interests, we would have to identify the user (by using cookies for example) and create user profiles. The user profiles could also be provided by the content provider.

A standard model for identifying space where the content providers allow for advertising insertion is critical. This will have to be coordinated with groups defining content structure, such as XML with W3C [4].

4 Insertion of Regional Data

4.1 Abstract

If a content provider wants to add user-specific regional information (weather forecasts for certain areas for example) to his Web pages, he has little choice but to have the user select his location from a list of regions. Usually it is not possible for origin servers to reliably detect from where Web users connect to Web sites because user requests can get routed through a number of proxy servers on their way from the client to the origin server.

In a network edge caching proxy environment user requests are usually redirected to the nearest proxy that is available to respond to the request. Regional information that is relevant to all users who are likely to connect to a certain proxy could be stored at the corresponding caching proxy. Whenever the proxy receives a user request, a module on the caching proxy could insert the regional information into the requested Web page. If the Web page does not contain any user-specific non-cacheable content other than the inserted regional information, the Web page content can now be cached for future requests.

4.2 Business model

This service could be sold to content providers who want to offer regional information on their Web sites and want to accelerate the delivery of their Web content. There are many cases in which a content provider could profit from knowing the location of the user. Users could be targeted with regional advertisement banners (see also ad insertion scenario). Regional distinctions (e.g. sales taxes, differing laws etc.) could be taken into consideration when the Web pages are prepared for the client. It would not be necessary any more to ask the user for his location prior to presenting him relevant information.

4.3 Technical Challenges

The regional content that is to be inserted into the Web pages would have to be distributed to the corresponding caching proxies. Since the regional content represents only a component of a whole Web page, it cannot be cached in the same way a complete Web page can be cached (unless it is an image). We have to find a mechanism to determine when a regional text component needs to be updated (or if the content provider should be responsible for this).

5 Caching of Personalized/Customized Web Pages

5.1 Abstract

Many Web sites (e.g. Yahoo) offer a service where users can create their own personalized version of the Web site (e.g. MyYahoo). It basically means that a user can choose from a number of components (e.g. stock information, weather forecasts, news etc.) and create a personalized Web page with them. This leads to dynamic Web pages that usually cannot be cached. However, the components of the personalized Web page can be cached. Therefore, it is possible to have a service module on the server create the user-specific Web pages by assembling the cached Web site components. In that case the origin server would not have to be contacted again and the page could be served to the client directly from the network edge caching proxy.

5.2 Business Model

This service would be another method of accelerating the delivery of Web content to the user, particularly the delivery of personalized/customized Web pages that would not be cacheable otherwise. It also saves bandwidth between the origin server and the proxy cache.

Content providers who offer their customers the possibility of personalizing their Web pages are likely to be willing to pay for

this kind of service.

5.3 Technical Challenges

We would have to find a caching mechanism for the separate components of the personalized Web pages (unless a component consists of an image only). These components could be stored at the caching proxy.

The page components would have to be refreshed just like complete Web page whenever they become stale.

6 Content Adaptation for Alternate Web Access Devices

6.1 Abstract

There is a growing diversity and heterogeneity in types and capabilities of client devices as well as the forms of network connections that people use to access the Web. Clients include cell phones and PDAs as well as PCs, TVs (with SetTop boxes), etc. However, these appliances have quite diverse display capabilities, storage, processing power, as well as slow network access. As a result, Internet access is still constrained on these devices and users are limited to only a small fraction of the total number of Web pages available in the Internet today. Organizations such as the WAP forum [4] have suggested custom Web page design but this results in special code frequently required on the content server.

Since the number of different access devices is growing constantly content providers cannot be expected to provide different versions of their Web pages for each and every Web access device that is available in the market.

Therefore, if it is possible to transcode the general full-fledged Web pages at some point on their way from the origin server to the user so that they are optimized for (or at least adapted to) the end users' specific requirements, it would provide a valuable service for the end customer, the service provider, and the content provider.

6.2 Business model

With the above-mentioned service in place, Web content providers could reach a much wider audience and the manufactures of diverse Web access devices could offer potential customers access to a bigger part of the Internet content, which should make a very good selling point. It would encourage more people to buy non-desktop Web access devices like cell phones and PDAs expanding the market.

We would expect this service would be offered as an additional feature to ISP customers who want to access the Web through different Web-enabled devices. Also, the service might be paid by content providers because they could serve their existing content to more users; likewise, the non-desktop device makers may contribute to this service cost making their client devices more effective at the Web.

6.3 Technical Challenges

Possible adaptations to meet the special requirements of different Web access devices are:

- Conversion of HTML pages to WML (Wireless Markup Language) pages
- Conversion of JPEG images to black and white GIF images
- Conversion of HTML tables to plain text
- Reduction of image quality
- Removal of redundant information

- Stripping of Java applets / JavaScript
- Audio to text conversion
- Video to key frame or video to text conversion
- Content extraction

We have to ensure that the automatic adaptation process will not make changes to a Web page that are unwanted by either the content provider or the recipient. Our suggested strategy to achieve this would be to allow the content provider as well as the client to define their preferences as to how they want Web pages to be adapted. The actual adaptation decisions would then be made based on the given preferences and a set of transformation rules. There would have to be a mechanism of resolving potential conflicts between the content provider's and the user's adaptation preferences. If neither

the content provider nor the client has expressed his preferences, a default adaptation of the requested Web page may be possible but investigation is needed.

A way for preferences to be specified representing the content provider and client customer must be provided. For example, client customers could set their preferences through a Web interface on the ISP Web site. Content providers could express their preferences by adding meta tags to their Web pages. This meta data offers the content provider the ability to specify a number of alternatives and the content adaptation server could then pick the most appropriate one. This meta data should be independent of specific Web content but is likely to depend on the types of content in the pages. Another possibility in the ESPWF [2, 7] framework would be for the content provider would be to provide an adaptation policy to all ISPs that want to adapt Web pages for alternate Web access devices. This policy could consist of general transformation rules or actual code modules that perform the adaptation.

7 Limited Client Bandwidth Adaptation

7.1 Abstract

Different Internet clients can handle different Internet connection speeds. Therefore it seems desirable to adapt the requested Web content to the userÆs bandwidth.

7.2 Business model

One of the main benefits is to decrease the Web access time for users. If a Web site loads too slowly, users tend to leave the site even before it has completed loading the home page. The improved perceived quality of service by adaptive content delivery means that users are more likely to stay and return, thus resulting in a greater profit for e-commerce sites. This can also result in higher hit rates and return rates, which can lead to higher sales for e-commerce sites and higher advertising revenues.

7.3 Technical Challenges

Possible adaptations to reduce the size of Web objects are:

- Reduction of image quality
- Replacement of images by their ALT text
- Removal of redundant information
- Removal of HTML comments
- Stripping of Java applets / JavaScript
- Audio to text conversion
- Video to key frame or video to text conversion
- Text summarizing
- Content extraction

We would have to find a reliable way of determining the bandwidth between the client and the proxy cache. One way of measuring this would be to measure the round trip time (RTT) to determine the connection speed. It is crucial that this bandwidth detection method works more or less exact or otherwise the client will either experience very slow Web browsing or be cut off of some (or all) of the rich Web content. This service requires authorization by the user like any other adaptation service that changes the content and or format of Web pages.

The mapping of a userÆs connection speed to appropriate page adaptations requires defining a set of adaptation rules.

8 Adaptation of Streaming Media

8.1 Abstract

Some of the above-mentioned services could not only be applied to Web pages but also to streaming media like audio and video streams. In particular, media streams could be adapted to meet the bandwidth of the userÆs connection. It would also be possible to insert pre-recorded advertisements into audio or video streams. Even content analysis and content filtering could be applied to streaming media.

8.2 Business model

The business models for streaming media adaptation are similar to those for Web page adaptation services.

8.3 Technical Challenges

The adaptation of streaming media will add more complexity to the caching proxy platform and the technical challenges of these kind of services have yet to be explored.

9 Request Filtering

9.1 Abstract

The success of Web filtering/blocking systems like NetNanny (http://www.netnanny.com) and WebSense (http://www.websense.com) shows that there is a great need for solutions that let the owner of a Web access device control what kind of Web content can be accessed with his device. Parents, for instance, often demand a means of blocking off offending material when their children browse the Web. Also, companies might want to have control over what kind of Web pages their employees can have access to. Companies might also want to prevent their employees from using the available bandwidth excessively for non-work related activities.

A request filtering service could provide a solution for all of the above. If all Web page requests of a specific user are routed through a caching proxy server, the content adaptation server could analyze the requests prior to fulfilling them. The service module would have to identify the user and determine the userÆs access level. The next step would be to look up the classification of the requested Web page in a database.

9.2 Business model

This service could be offered to enterprises and to ISPs. A database of Web pages that contain offending material could be obtained from companies that have specialized in Web blocking systems.

9.3 Technical Challenges

The database on the proxy caching platform that contains the Web page classifications needs to be updated on a regular basis. If the database is provided by third parties, we have to provide them with a secure way of updating the database.

If a Web access device is shared among different users who have different access levels, it is not sufficient to identify the Web access device. Therefore it will probably be necessary that different users of a Web access device use different user accounts.

The owner of a Web access device must be able to define and change the access rights of the user(s) of his device. This could be done through a Web interface provided by the ISP/company.

10 Request Filtering through Content Analysis

10.1 Abstract

While this service is very similar to the one previously described, it works more dynamically in that the content adaptation server analyzes the Web content once it has been retrieved from either the proxy cache or the origin server prior to sending it to the client.

Through the use of sophisticated content analysis algorithms it should be possible to classify the analyzed Web content. If the classification of the Web page matches the userÆs access level, the page will be delivered to the client. Otherwise, the client will be denied the page. The analyzed page along with its classification should be stored in the proxy cache so that future requests for the same page do not require the cached Web to be analyzed again. This will result in a better Web page delivery performance for popular Web pages. The main benefit of this approach is that there is no need to provide or maintain lists of forbidden Web sites, a process that per definition must always lag behind the creation of new Web sites. If common characteristics of a category of unwanted Web pages can be defined, it should be possible to automatically detect whether a requested Web page falls in a forbidden category.

10.2 Business model

This service could be offered to enterprises and ISPs. The content analysis software could be obtained from software companies that have specialized in this field.

10.3 Technical Challenges

In addition to the technical challenges described in the previous service scenario, we would have to find a way of storing the classification information of Web pages once they have been analyzed. One way to do this would be to add a meta tag (possibly using the Resource Description Framework [6] specification) with content rating information to a Web page before it is cached. Subsequent requests of the same Web page would then require the request filtering service module to scan the cached Web page for this metadata in order to determine the content rating of the requested page.

11 Creation of User Profiles

11.1 Abstract

If all Web requests of a certain Web user were routed through a certain caching proxy platform, it would be easy to log them in order to create a profile of the userÆs Web browsing behavior. These user profiles could be created anonymously with no personal data (e.g. name or e-mail address) stored in the access log files.

Once a sufficient number of requests has been logged by the content adaptation server, the marketing group could start analyzing the log files. In most cases it should be possible to derive the userÆs interests by analyzing what kind of Web sites the user visits and how often he goes there.

11.2 Business model

Companies that want to advertise on Web pages are very interested in knowing more about the recipients of their advertisement campaigns

so that they can target their advertisements at people who are interested in the kind of products/services that the company wants to sell. These companies will pay for information that helps them to target their campaigns at interested users. This money could be offered to users (e.g. in the form of reduced Internet access fees) to give them an incentive to agree to the profiling.

As explained above, we could derive the userÆs interests from his Web browsing behavior and use this information to send the user only those advertisements that match his interests/needs. This will most likely result in a higher ad banner click-rate per user.

This service could be sold separately or in combination with the ad insertion service.

11.3 Technical Challenges

The creation of user profiles requires a mechanism to identify Web users. The ISP could provide a mapping from the userÆs (possibly dynamic) IP number to some unique user ID. Another alternative would be to use cookies, provided that the user has not disabled them in his Web browser.

12 Search Engine Index on Cached Web Pages

12.1 Abstract

A proxy usually contains the most frequently requested Web pages of the Web users whose Web requests are routed through it. If we indexed the content of all Web pages currently contained in one or more proxies, we would have an index of Web pages that Web users are very likely to request (since they have been the most popular in the past). A search engine based on this index could therefore yield a high hit rate when used by a group of users who have similar interests and usually connect to the same caching proxies. The benefit of this approach would be that the index could be created very fast (there is no Web crawling to do) and that the search results could be returned to the user directly from the network edge caching proxy. The drawback, however, is that this search engine would index only a small fraction of the existing Web pages. Web users have to be aware of this fact when they use the cache-based search index service. Another approach would be to display the proxy search results first while a global search engine prepares the results of a global search in the meantime. As soon as the global search results become available, they will be sent to the user.

12.2 Business model

The search engine service described above could be sold to big companies who have users with similar interests and want to provide

a fast search engine. Companies offering traditional search engines could be interested in combining their services with a cache-based search engine service to accelerate the delivery of their search results.

12.3 Technical Challenges

If the cached Web pages of more than one caching proxy were to be indexed, we would have to find a way of replicating the search index to all affected caching proxy servers.

13 Language Translation

13.1 Abstract

Soon the majority of all Internet users will be non-English speaking. As most of the current Web content is written in English, it becomes desirable to be able to translate the English content to the Web userÆs local language, even if the content provider does not offer translations of his Web content. An automatic translation service for all Web pages could be implemented with a content adaptation server.

The proxy server will determine the Web user's preferred language(s) and ask whether the content requested should be translated to the user's preferred language. If the content is to be translated, the proxy cache will forward the Web content to a translation server where the page then is automatically translated. The proxy could also locally store translated content eliminating the need to repeat translations for different users.

13.2 Business model

The automatic language translation service will help break language barriers and open new markets for e-commerce. The average non-English speaking Web user will have access to more Web content. ISPs, especially those with customers in non-English speaking countries, could offer this service to their customers.

13.3 Technical Challenges

The automatic translation of text found on Web pages is not a trivial task. It will not be possible to translate a Web page automatically without running the risk of rendering parts of it incomprehensible. Worse yet, the original meaning could be changed and it is not said the reader of the translated page will notice the change in meaning. It is questionable whether content providers would even tolerate this kind of translation service.

Therefore it is very important that the client authorizes this translation service and is fully aware of its potentially faulty

behavior. It should also be considered to mark translated pages in a specific way to remind the user of the machine translation.

Other technical challenges include the automatic detection of the language used in the original document and the clientÆs local language.

14 Author's Addresses

Andre Beck
Markus Hofmann
Bell Labs Research
Lucent Technologies
101 Crawfords Corner Rd.
Holmdel, NJ 07733
Phone: (732) 332-5983
Email: {abeck, hofmann}@bell-labs.com

Michael W. Condry
Intel Corporation

2111 NE 25th Avenue
M/S JF3-206
Hillsboro, OR 97124
Phone: 503-264-9019
Email: condry@intel.com

15 References

1 Bradner, S., "The Internet Standards Process -- Revision 3", BCP 9, RFC 2026, October 1996.

2 Tomlinson, G., et al., "Extensible Proxy Services Framework", Work in Progress, Internet Draft draft-tomlinson-epsfw-00.txt, July 2000.

3 World Wide Web Consortium (W3C), http://www.w3.org.

4 The Wireless Application Protocol (WAP) Forum, http://www.wapforum.org/.

5 ICAP Protocol Group, "ICAP - the Internet Content Adaptation Protocol", submitted as Internet Draft draft-elson-opes-icap-00.txt, (previous version available at http://www.i-cap.org/), November 17, 2000.

6 Resource Description Framework (RDF), http://www.w3.org/RDF.

7 Open Proxy Extensible Services (OPES), http://www.extproxy.org.

Full Copyright Statement

Acknowledgement

Funding for the RFC editor function is currently provided by the

Internet Society.