A More Attractive and Interactive TV*

Pablo Rodriguez   Jamel Gafsi   Jörg Nonnenmacher
Institut EURECOM
2229, route des Crêtes. BP 193
06904, Sophia Antipolis Cedex, FRANCE
{rodrigue, gafsi, nonnen}@eurecom.fr
Tel: (33) 04 93 00 2673
Fax: (33) 04 93 00 2627

September 25, 1998

Abstract

The television is currently the most important mass media. On the other hand, the World Wide Web is showing as another important mass media which can easily provide interactivity. The combination of the Web and the television can make the classical TV more attractive as well as more interactive. First, we propose a scheme which offers Web pages related to a TV program as an added-value service. The related Web pages are pushed in advance to a local cache in the receivers' TV boxes. Additionally, we also consider the situation where a Web caching hierarchy is in place and related Web documents are also pushed to the caching hierarchy. We consider the impact of these approaches in terms of latency and bandwidth. Then we discuss some feedback mechanisms that make the classical TV more interactive.

Keywords: TV, WWW, Caching, Push, Satellites, Interactivity

1 Introduction

The Television is the most important mass media communication during the last years. The Web is growing exponentially and becoming another important mass media. The arrival of the digital television will trigger the development of television and web integration. Both, the TV and the Web can benefit from each other. The Web will benefit from the transmission infrastructure of the TV distribution (cable, satellites...) and from the large amount of potential receivers. The TV will become more *attractive* and *interactive*.

In this paper we assume a scenario where there is a TV distribution center transmitting television programs via a satellite distribution. The TV boxes at the receiver's side are provided with a local cache and a connection to the Internet. The Internet is used by the receivers to retrieve Web documents and also could be used as a feedback channel for interactive applications.

First, we analyze a scheme that makes the TV more attractive. We propose to push Web documents related to a certain TV program to the Web caches close to the receivers. We present two different

*Presented at the W3C Workshop on Television and the Web, Sophia Antipolis, France.

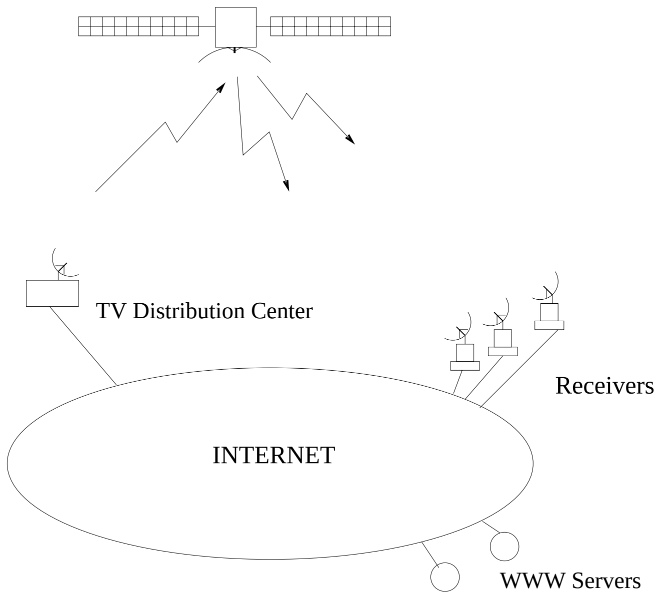

Figure 1: Network Topology.

situations i) TV boxes have a local cache with a finite capacity and ii) in addition to the TV box caches a Web caching hierarchy is in place which is able to satisfy requests not hit at the TV box caches. We discuss why a push operation of documents to caches close to the receivers when the Web and the TV cooperate is more effective than in the Web alone.

Additionally, we discuss some interactive-TV issues. In this situation the Internet is used as a feedback channel. Receivers send feedback information to the TV distribution centers. Then the TV distribution centers can obtain valuable information and implement many attractive applications, i.e. On-line voting.

2 Pushing Web pages related to TV programs into Web Caches

Pushing information to caches close to the receivers decreases the retrieval times because documents can be delivered locally at high transmission rates. If a Web document is placed at the client's cache the network is only used to check the document's consistency but not to transmit the document, which eliminates frustrating transmission times. Additionally, pushing the same information to a high number of caches can be performed very efficiently via a multicast or a broadcast transmission, which save a lot of bandwidth in the network and reduces the burden at the server side.

Client point of view

Pushing a document in advance is based on the assumption that the document will be requested soon. If receivers do not show any interest in the document, pushing the document in advance has little benefit and may result in bandwidth waste. In the Web it is not an easy task to choose which documents to push into the Web caches. Usually receivers surf between many different documents, which makes very difficult to predict their future request pattern.

In the case that the Web and the TV cooperate, it is easier to predict and suggest some Web documents that may be of interest for the receivers making the push operation more effective. Imagine the situation where a certain program (film, news, sports...) is being played in TV. At the TV distribution center someone previously fetched the most interesting Web pages related with that TV program (this could be also done automatically). The related Web pages are pushed to the receivers box cache. As the pushed pages are very related with the actual TV program (i.e. additional weather maps on the weather report, information about the players in a football match..), the probability that receivers request the proposed pushed-documents is higher. Additionally receivers know that the pushed pages can be browsed locally with minimum delay and without using the Internet, which makes these pages even more attractive. Receivers do not need to use any network resources avoiding any monetary expenses. The pushed pages come as a value-added service to the TV information. In the case that a receiver is interested in additional information to the one already pushed on its TV box, the receiver sends a request on the Internet to fetch the document from the network.

Network Point of view

If the same information needs to be pushed to a high number of receivers a unicast distribution from the source to the receivers wastes a lot of bandwidth in the network. A more efficient distribution scheme is a multicast distribution [5]: in order to deliver one document to M receivers, only one copy is sent from the source and the document is duplicated at the points where the network path forks off. We consider two different possibilities to push a Web document to many receivers: IP Multicast, IP Satellite distribution.

IP Multicast

IP multicast requires that the network routers connecting a server with its clients are multicast capable. While today the multicast routing software is part of any new router that is installed, not all the existing routers have been enabled to be multicast capable. Therefore, multicast routing on the Internet is not yet everywhere available.

Additionally, a significant obstacle to the deployment of an IP multicast distribution on the Internet is the lack of proper congestion control mechanisms such as those present in TCP. Without such mechanisms, applications cannot adapt to variable network conditions, and tend to use an uneven share of the network's capacity. As a consequence, Internet providers and network managers might choose not to support multicast services in order to eliminate a potential source of congestion.

IP Satellite distribution

Pushing documents via a satellite distribution has significant advantages to pushing documents via the current Internet [6] [1]. The satellite distribution has experienced a great deployment during last years mainly due to the TV market. A multicast distribution on the Internet is still in its infancy as an available service.

When a document is sent through a satellite distribution the document is broadcasted to all receivers instead of being multicasted. The cost of a satellite distribution has decreased considerably and currently it is cheaper than using the transoceanic optical fibers. Additionally a satellite distribution has almost zero losses compared to the high loss rate (5-20 %) on the Internet. Reliability can be obtained very efficiently with a forward error correction code (FEC) [3]. With such small loss rates the number of retransmissions needed to achieve reliability is very low. Therefore pushing Web documents to a high number of receivers via a satellite distribution instead of an IP multicast distribution can be done with fewer losses, higher transmission bandwidths and lower costs.

Server Point of view

A multicast/broadcast distribution greatly alleviates the load on the origin servers. The origin servers do not deal directly with the receivers because a multicast/broadcast transmission is open-loop and no connection set-up is established between the receivers and the source. This reduces the server complexity and scales very well.

Once we have analyzed how pushing Web documents via a satellite distribution is beneficial for the clients, the network operators, and the servers we will focus in the Web caches that store the pushed Web documents.

Web caches are being extensively deployed by ISP and backbone operators to reduce their bandwidth expenses. Web caches are usually placed at the access points between different networks. Web caches need to cooperate to increase their hit rate. One of the most popular schemes to join several caches is placing caches on a hierarchical structure [2]. At the bottom the client's cache (on the TV box). Requests not satisfied by this cache are forwarded to a metropolitan cache. If the document is not hit at the metropolitan cache the request is forwarded to the regional cache which in turn does the same with not satisfied requests towards a national cache. At the top is the national cache harvests all requests from a very large community. If the document is not hit at the national cache, the request is forwarded to the origin server. Thus, when a client desires a document, a series of requests is sent up the caching hierarchy until the document is found. When the document is found, either at a cache or at the origin server, it travels down the hierarchy, leaving a copy at each of the intermediate caches.

In next sections we analyze how TV-related Web documents can be pushed from the source to the local caches in the TV box. We consider two different scenarios: i) the receiver has a finite local cache, ii) the receivers have a finite local cache and a Web caching hierarchy is in place.

2.1 Finite Client-Caches

The price of the disks is decreasing faster than the price of the bandwidth. For example, until recently the inbound bandwidth for an ISP in Australia or Europe costs aprox. us $.11/megabyte. But disk space only costs aprox. us $0.04/megabyte. Therefore, it is cheaper to store in a disk all documents ever retreived than fetching them again from the network if someone is going to access them just once more.

Due to the low price of disks, receivers' boxes can have very large caches, practically of infinite size. However, Web pages tend to include more and more multimedia applications which may consume the available disk space again.

If receivers have a finite local cache may be they can not keep all pushed documents. Imagine the situation where there are 40 TV channels each pushing the top-ten related Web pages with the current program. Assuming an average page size of 20 KB, the space needed on every local cache to keep all the pages suggested for every channel is around 8 MB. Even if this number is not very high it can happen that the client cache does not have such a free space. In this situation the local cache needs to take some decision on which documents to keep.

If 40 channels are being sent at the same time it is clear that a receiver will not show interest in all the 40 channels at the same time. Depending on the time of the day and the day of the week a receiver shows different preferences for one TV channel or other. The local cache can keep some statistics showing the TV preferences of a certain receiver. Based on this statistics, when the client cache is limited the cache only keeps the Web pages related with those channels that are more frequently seen by the client. Additionally, these documents should be removed from the cache after some time to free space for others. After the TV program ends, the TV-related documents are given a time-to-live and then they are removed from the local cache. In general all the caching replacement policies that have already been suggested for Web caches can be applied.

If the cache space is limited to keep all related documents, we propose that every Web related document is pushed with some additional *meta-information* including: The context of the document, the URL, etc. Therefore, the TV distribution center pushes all TV-related Web documents and additionally their meta-information. A local cache may not be able to store the pushed documents but it may still keep the documents' meta-information. This information is useful for the receiver to easilly know which are the most interesting Web pages related with a certain TV program.

Due to the limited space on the local cache, a receiver may request a document that is not kept locally. When the document is not met at the local cache, a request is sent through the network to fetch the document from the origin Web server. In the case that a caching hierarchy is in place the document can be hit at close caching levels saving bandwidth on the more congested up links and reducing the latency to the receivers.

2.2 Finite Client-Caches and a Caching Hierarchy

Assume the situation that the client-cache is finite and a caching hierarchy is in place. The TV distribution center pushes the TV-related documents with the meta-information to the receivers' cache and additionally places a copy of the Web documents at the different caches in the caching hierarchy. The caches at the different levels of the caching hierarchy are usually very large. Additionally the ISPs at the different hierarchy levels are very interested in keeping the TV related Web documents because they have a big number of potential receivers. As we discussed before it is more beneficial for an ISP to store a document in its cache if it is going to be requested at least once more than to retreive it again from the network.

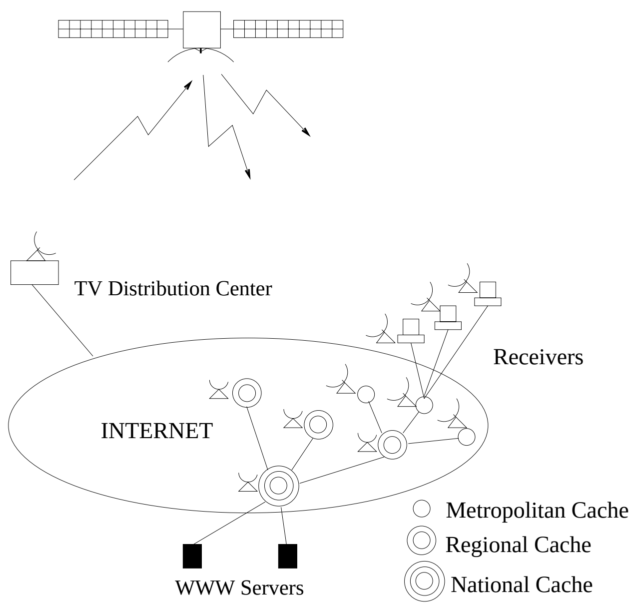

Figure 2: Network Topology and Caching hierarchy.

Receivers that do not have enough cache space to keep all pushed Web documents still keep a list with meta-information about the rest of pushed documents which were not cache locally. The meta-information is useful for the receivers because they can easilly know which are the most interesting Web documents related to a TV program.

Additionally, in the case that a caching hierarchy is in place the meta-information list gives information about which documents have also been pushed to the caching hierarchy and therefore may also have low retreival latencies. If a client wants to request a Web document that has been pushed from the TV center but could not be cached locally, the request is sent through the network. This request has a high probability to hit the document at the caching hierarchy and obtain low latencies [5]. Therefore, receivers have a high incentive to request documents that have been suggested by the TV distribution center, even if they were not cached locally, because it is very likely that these documents are hit at close caches in the caching hierarchy.

3 Interactive TV

Interactivity gives the TV distribution center the possibility to receive valuable feedback information from the receivers. We have a method based on exponentially distributed timers that estimates the size of a receiver population that changes in size dynamically by orders of magnitude. A low bandwidth feedback channel is shown to be sufficient to provide for a fast adapting receiver estimate [4]. The feedback bandwidth is unaffected for changes by orders of magnitude in the number of receivers.

Using this method, the TV distribution center has the possibility to continuously estimate the number of receivers. Required is only a unicast feedback channel (e.g. dial-up line) from the receivers back to the sender.

The given estimation method can also be applied to determine the size of a specific subgroup of all receivers. The request for feedback must contain a discriminator that allows a receiver to determine if it belongs to the subgroup requested. Estimating the size of dynamic subgroups enables applications such as On-line voting.

4 Conclusions

Web and TV are two important mass medias. The arrival of the digital television will foster the integration of the Web and the TV. Both, the Web and the TV will benefit from this marriage. The Web will benefit from the transmission infraestructure of the TV and from its large audience. The TV will become more attractive and interactive.

In this paper we have presented one scheme that pushes Web pages related to a certain TV program as additional information to the program itself. The distribution of the Web pages is done via a satellite distribution to caches close to the receivers. We also consider the case that a caching hierarchy is in place. We discuss how to deal with finite client caches which can only cache a limited number of Web pages. We also discuss the impact of this model in terms of receivers' latency and bandwidth used in the network.

Additionally we how to make the classical TV more interactive. The Web and the Internet infraestruture is used to send valuable feedback information to the TV distribution centers. Based on this information the TV distribution centers can better addapt their emmissions as well as implement interesting applications (i.e. On-line voting).

References

[1] Broadcast Satellite Services, "http://www.isp-sat.com".

[2] A. Chankhunthod et al., "A Hierarchical Internet Object Cache", In *Proc. 1996 USENIX Technical Conference*, San Diego, CA, January 1996.

[3] J. Nonnenmacher, E. W. Biersack, and D. Towsley, "Parity-Based Loss Recovery for Reliable Multicast Transmission", In *SIGCOMM '97*, pp. 289–300, Cannes, France, September 1997.

[4] J. Nonnenmacher and E. Biersack, "Scalable Optimal Multicast Feedback", , Institute EURECOM, BP 193, 06904 Sophia Antipolis, FRANCE, July 1997.

[5] P. Rodriguez, E. W. Biersack, and K. W. Ross, "Improving the Latency in the Web: Caching or Multicast?", In *3rd International WWW Caching Workshop*, June 1998.

[6] SkyCache, "http://www.skycache.com".