Quality of Service and Electronic Newspaper: The Etel Solution

Valérie Issarny, Michel Banâtre, Boris Charpiot, Jean-Marc Menaud

HAL Id: inria-00072864
https://inria.hal.science/inria-00072864

Submitted on 24 May 2006

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Quality of Service and Electronic Newspaper: The Etel Solution

Valérie Issarny, Michel Banâtre, Boris Charpiot, Jean-Marc Menaud

N˙3795

Octobre 1999

———— THÈME 1 ————

Quality of Service and Electronic Newspaper:
The Etel Solution

Valérie Issarny, Michel Banâtre, Boris Charpiot, Jean-Marc Menaud

Thème 1 — Réseaux et systèmes
Projet Solidor

Abstract: Making electronic newspapers available to users is an attractive business for newspaper editors. However, the success of an electronic newspaper in a commercial setting depends upon the provided quality of service. The newspaper must guarantee to its users, both content quality in terms of reading comfort, and access quality in terms of response time.

This paper presents the Etel electronic newspaper whose main design objective is to guarantee high quality of service to users. Content quality is achieved through the design and production of electronic editions in close collaboration with a newspaper editor. Access quality is realized through the design and implementation of a dedicated client-server distributed system. The proposed distributed system exploits the specifics of electronic newspapers. Precisely, most newspaper readers systematically consult the various editions of their usual newspaper in the same way. In the context of an electronic newspaper, this means that we can safely assume the existence of user profiles in terms of access patterns to newspaper editions. This feature allows us to propose profile-based automatic prefetching and load balancing strategies, which together enable to offer a highly responsive and scalable distributed system.

Key-words: Automatic prefetching, client-server distributed system, data mining, electronic newspaper, load balancing, quality of service, scalability, user profile.

(Résumé : tsvp)

Qualité de service et presse écrite électronique :
La solution Etel

Résumé :   Produire la version électronique d'une édition papier d'un journal permet aux éditeurs d'obtenir des gains financiers importants. En effet, l'édition électronique réutilise les informations de l'édition papier, n'entraînant donc que de faibles coûts de production, tout en étant consultable par un nombre très important d'utilisateurs. Toutefois, le succès commercial d'une édition électronique dépend essentiellement de la qualité de service offerte à ses utilisateurs. Nous entendons par la qualité de service d'une édition électronique, la manière de quantifier la satisfaction d'un utilisateur ordinaire face au service qui lui est offert. De ce fait, la qualité de service recouvre la qualité de contenu, en terme de confort de lecture, et la qualité d'accès, en terme de temps de réponse.

Cet article présente Etel, un système dédié à la diffusion d'un journal électronique, dont le principal objectif est d'offrir aux utilisateurs une qualité de service accrue. Pour ce faire, nous augmentons la qualité de contenu grâce à une étroite collaboration avec l'éditeur, nous permettant ainsi de bénéficier de son expérience dans la mise en valeur et la consultation des informations diffusées. Nous augmentons également la qualité d'accès en proposant un système d'information conçu pour les journaux électroniques et de ce fait, s'appuyant sur les spécificités de ces derniers. En effet, nous avons constaté qu'un grand nombre d'utilisateurs possède des habitudes de lectures i.e., consultent tous les jours les mêmes rubriques du journal et dans le même ordre. En s'appuyant sur cette caractéristique, nous proposons dans cet article, une nouvelle méthode d'anticipation des accès et de répartition de charge, permettant d'améliorer le temps de réponse et l'extensibilité du système donc, sa qualité de service.

Mots-clé :   Anticipation des accès, Système distribué, analyse de donées, journal électronique, répartition de charge, qualité de service, extensibilité, profile utilisateur.

1 Introduction

The ever-increasing use of computing facilities by the general public makes attractive the creation of online newspaper services. Among benefits, this allows a newspaper editor to attain a much larger audience, and to offer additional services to its customers by exploiting together its information base and processing capacities. This has led to make available various electronic newspapers over the

INTERNET, either *via* proprietary online services (*e.g.* AMERICA ON LINE) or the World Wide Web (*e.g.* USA TODAY). The success of an online newspaper service depends upon the *quality of service* it provides to its users. The service must guarantee a quality that makes the online newspaper version at least as attractive as a paper version. In other words, reading a newspaper from an online service should be as convenient as reading a conventional newspaper from the customers' standpoint. Meeting such a requirement implies to offer an adequate layout for pages of the online newspaper as well as to display requested pages in a timely manner. While the provision of a newspaper having a satisfactory layout may be considered as being achieved by existing online newspaper services, the delivery of an online newspaper in a timely manner is yet to be solved. Newspapers available *via* the Web cannot guarantee satisfying response times due to the current limitation of the Web from this standpoint. Newspapers available *via* proprietary online services in general exhibit better response time through adequate server configuration, but still remain dependent upon the INTERNET bandwidth. This article presents the Etel online newspaper service that has been designed so as to offer a high quality of service to users. The main novelty of the Etel service comes from its associated distributed system, which provides support for the timely delivery of newspaper editions in the presence of a large number of users.

1.1 The Etel Quality of Service

Let us first consider the quality of the Etel newspaper editions with respect to their content relevance and presentation. Although this is a user-specific matter, we may foresee the adhesion of a population of users by producing the editions of an online newspaper service in collaboration with a newspaper editor. The provided newspaper is expected to gain acceptance at least with the newspaper's usual readers. The Etel online newspaper service has been designed in close collaboration with the *Ouest-France* editor, that is the major French local newspaper editor (*i.e.*, *Ouest-France* sells about 1 million copies per day). Thus, Etel editions are produced using the *Ouest-France* information base. Furthermore, in order to best match users interests,

every customer subscribing to the Etel service may specify the topics in which he is interested so as to give him access to individualized editions. This user-specific information is known as the *user's profile* [14]. An individualized online edition is then organized as a set of pages where each page corresponds to either a summary, an article, or a photograph (see Figure 1). A summary page roughly corresponds to a page of the paper version. In addition, except for the newspaper's first page, each summary page relates to a specific topic (*e.g.* the summary page of Figure 1.*a* relates to the *Pays d'Alençon* local topic). The content of a summary page differs from the one of its corresponding paper page by summarizing every article into the article's title and photograph (if any), and hyperlinks to the corresponding article and photograph (*e.g.* Figures 1.*b* and 1.*c* give the article and photograph pages for the article entitled *Belote : les qualifiés pour la finale* relating to the local sub-topic *Le Méle-sur-Sarthe* of the summary page). Every page further integrates a frame on the right-hand side of the screen, listing customized hyperlinks to a set of topics according to the user's profile. Each article of a page decomposes into a number of PDF files, which are produced using the XPress[1] and Acrobat[2] software tools. These files are: (*i*) a file containing the text of the article for both the summary and article pages, and (*ii*) two files per photograph associated to the article, embedding respectively the full-page-size and small-size photographs. A summary page is then made of the sets of PDF files associated to its articles but files embedding full-page size photographs.

Ensuring quality of the newspaper editions from the standpoint of their presentation and of their content relevance is not sufficient for an online newspaper service to be successful and competitive, editions should be delivered to users in a timely manner. Ideally, this means that a requested page of an online edition is made available to a user in a time interval that does not exceed the one taken to select a paper page. The only way to meet this ideal timeliness requirement is to have the newspaper edition locally available on the user's machine before the user accesses the newspaper. However, such a solution is not practical due to the newspaper's storage needs. Precisely, a *Ouest-France* daily edition actually decomposes into 40 regional editions whose topics may either be shared by a subset of regional editions or be specific to one edition in the case of village-specific topics. The ensemble of *Ouest-France* daily editions leads to the production of about 400 pages, each having an average size of 80 Kbytes. The Etel newspaper further enables a user profile to embed an evolving set of topics relating to distinct *Ouest-France* regional editions.

---

[1]http://www.quark.com
[2]http://www.adobe.com/Acrobat/

a) A summary page

b) An article page

c) A photograph page

Figure 1: A sample of Etel pages

In addition, we want the service to be accessible from a large variety of terminals including those having a low storage capacity (*e.g.*, Web phone, set-top-box). Thus, we had to design and implement a client-server distributed system, dedicated to the delivery of newspaper editions. In particular, the proposed system had to guarantee timely delivery of requested pages to clients from the server, in a way that approaches the ideal case without requiring large storage capacity from user machines.

1.2   The Etel Client-Server Distributed System

Our primary objective in the design of the Etel distributed system was to guarantee low client latency in the presence of a large number of clients. In particular, the system's scale in terms of both the number of supported users and the users' geographical dispersion hardly enables to offer convenient client latency by simply issuing client requests in response to the users' ones. A well known technique for meeting the low latency requirement in a large scale setting is then predictive prefetching, *i.e.*, issuing the client requests before the user actually requests for data [3, 22, 8]. However, predictive prefetching is also recognized as a consuming technique in terms of bandwidth and processing usage. Since, access predictions cannot be 100% accurate, prefetching requests lead to waste network and server resources each time a predicted request does not match the user's one. Hence, for predictive prefetching to be beneficial, it is necessary to have a highly accurate prediction algorithm. It is further desirable to have network and server resource management enabling to alleviate the penalty of the increased resource usage –even if low– that is incurred by predictive prefetching. Regarding network management, the most reliable solution to date from the standpoint of reducing the client latency is to ensure availability of a fixed network bandwidth between the client and the server. Such a requirement can be met by using either a proprietary network (*e.g.*, ISDN, ASL/RTC) or a resource reservation protocol. The latter solution may possibly lead to reject user connections and is further not currently supported over the INTERNET. Hence, the former solution is currently the most appropriate for an online newspaper service. With respect to the ability of the server to handle a large number of client requests (including extra prefetching requests), this may be achieved by using a powerful hardware infrastructure such as the Web cache products that are entering the market place. However, such products are very costly and not easily extensible. On the other hand, an infrastructure made of clusters of PCs interconnected by a high speed network such as ATM provides a low-cost, extensible solution for implementing powerful servers. In addition, such an architecture matches the organization of the *Ouest-France* company that has a number of regional sites distributed over the West
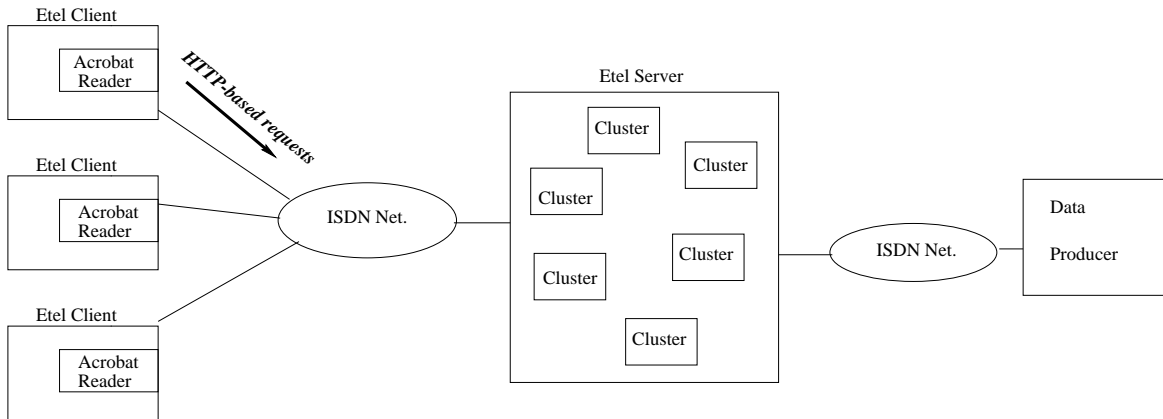
Figure 2: Architecture of the Etel distributed system

part of France, which may host the server's clusters. It is then necessary to implement adequate software solutions so as to actually enforce server scalability and fully benefit from the server's processing power.

The overall architecture of the Etel distributed system is depicted in Figure 2. The specifics of the Etel distributed system compared to existing client-server systems come from exploiting the habits of newspaper readers for the design of the Etel clients and server. Precisely, according to information provided to us by *Ouest-France*, most newspaper readers systematically consult the various editions of their usual newspaper in the same way (*e.g.* a reader may consult his newspaper by first reading local and then sport pages). In the context of an online newspaper service, this means that we can safely assume the existence of user profiles in terms of recurring patterns of accesses to newspaper editions. In the remainder, such profiles are qualified as *access profiles*, while the term *profile* is used to refer to the set of topics usually consulted by a user. For the design of the Etel distributed system, we exploited the existence of both user profiles and user access profiles to conveniently drive data communication and distribution. Within an Etel client, the user's access profile serves to anticipate the user's accesses and hence to realize efficient predictive

prefetching [18]. Within the Etel server, the set of user profiles may be processed so as to group together users having close profiles, which further serves to define corresponding data groups that are distributed over the server's clusters [5].

Section 2 details the design and implementation of Etel clients, defining the embedded predictive prefetching strategy together with associated cache management and client-server communication protocol. Section 3 then focuses on the design and implementation of the Etel server, which integrates a load balancing strategy among clusters based on the grouping of both user profiles and newspaper topics. Every cluster further runs an efficient cooperation protocol among its composing PCs so as to minimize both the memory and processing load of each PC. Evaluation of the Etel distributed system with respect to offered client latency and scalability compared to related work is addressed in Section 4. It is shown that our system performs better than alternative solutions when concerned with the timely delivery of newspaper editions in a large scale setting. Finally, we conclude in Section 5, summarizing our contribution and considering one perspective for our work, which is the benefit that could bring the novel features of the Etel distributed system when applied to the Web.

2 Client Design and Implementation

A newspaper edition can be characterized by a hyperdocument, which is structured as a set of *pages* where each page contains a set of *hyperlinks* that determine the pages that are accessible from it. Each page further gives information relating to a particular topic (*e.g.* international politics, national economics, sports, ...). Hence, the content of each page is updated daily but the hyperdocument structure is invariant. A user then usually consults the hyperdocument pages according to a recurrent order that reflects his personal interest for the corresponding topics. Such a recurring access pattern is exploited for predicting the newspaper pages that the user is likely to access next with respect to the previously accessed pages, and hence prefetch the newspaper pages within the client cache before the user actually request for them. Given the execution of a predictive prefetching strategy within the client, it is necessary to devise adequate cache management policy and client-server communication protocol. The former should avoid replacement from the cache of a page that is about to be accessed; the latter should allow an actual user request request to take precedence over a prefetching request when they do not match. The following subsections detail the key elements that are integrated within Etel clients and that

together enable to achieve low client latency, *i.e.*, the predictive prefetching strategy, cache management policy, and client-server communication protocol.

2.1 Predictive Prefetching Strategy

An automatic predictive prefetching strategy relies on the management of a directed *access graph* that captures the correlation between data accesses: a node in the graph corresponds to an accessible data and an arc between two nodes denotes an *access dependency*. Then, given the access to a piece of data (*e.g.* a file) denoted by a node A of the graph, all the pieces of data denoted by nodes that are pointed by arcs leaving A are likely to be accessed and hence are eligible for prefetching. In order to differentiate eligible data pieces, the arcs of the access graph are generally weighed by the number of times (or frequency) that the pointed node is accessed after the origin node. Thus, each arc gives the probability of a particular piece of data being accessed soon after another one, which serves ordering data pieces that are eligible for prefetching. Such an ordered sequence of data pieces constitutes an *access prediction*, and it is taken as input by a *prefetching algorithm* for issuing prefetching requests to the data server.

Patterned access graph: Let us first define the access graph used for achieving predictive prefetching within Etel clients. As shown in Figure 3.*a*, the hyperdocument characterizing newspaper editions can be represented as a directed graph where each node corresponds to a page and an arc to a hyperlink. Notice that the hyperdocument graph is fully connected; this is due to the presence of the frame listing customized hyperlinks within each page (see Subsection 1.1). Given the hyperdocument graph, the user's access profile may be maintained by weighing each arc with the *access frequency* to the pointed node. However, the resulting access graph gives only the frequency with which a page is accessed with respect to the currently accessed page. It does not take into account the order in which pages are accessed. The frequency of accessing a page given the preceding sequence of accesses is obtained by introducing a path within the access graph for each sequence of accesses [15, 17, 8]. For instance, Figure 3.*b* illustrates an access graph encoding the access pattern of a user for 3 sessions where the user accessed $< A, E, D, C, B >$ for 2 sessions and $< A, E, F, G, H >$ for 1 session. In order to minimize the size of the graph, only sequences made of distinct nodes are stored within the access graph. Hence, the re-access to a node that already belongs to the current sequence leads to introduce the corresponding cycle. In the following, the resulting access graph is called *patterned access graph*. For illustration, let us assume the hyperdocument graph given