ARÃ : A User-Centric Mirror Locating Service for the World Wide Web

Neilze Dorta, Pierre Sens, Christian Khoury
Université Paris 6, LIP6/CNRS
firstname.lastname@lip6.fr

Abstract

Information sharing via the Web has risen to dominance in the last few years due to applications such as cooperative engineering and network games. The latters and likewise applications' viability rely on high availability of shared data and low access latency to it. Mirroring is one of the techniques to make such requirements possible; for that to work, users should be able to locate mirrors. In this paper, we present *ARÃ*, a system for managing, sharing and transparently locating mirrors based on users' requirements; it helps make mirroring trendy, not only to large groups of users as it is common today, but to small ones too. Preliminary measurements show encouraging figures in favour of using *ARÃ* in large scale information sharing systems.

KEY WORDS
Algorithms, Protocols, Large Scale, Caching, Mirroring, Replication, Discovery, Distributed Location Service, Lazy Replication, Cooperation

1 Introduction

Large scale systems suffer from (1) high latency due to large distances between users and servers and (2) low availability due to higher probability of network partitions and servers being overloaded by a large number of users. Data replication in nearby mirrors (mirroring) is a common answer to such problems, as it decreases latency and enhances availability; mirroring also helps in reducing servers' load as well as inter-network traffic.

Benefitting fully of data replication entails that users be able to locate and access nearby copies of sought data. Still nowadays, locating mirrors is mostly done manually by users or at the very best, by Content Delivery Networks (CDN), e.g. Akamai. The former case inherently relies on users' patience and willingness in perpetually searching for *interesting* mirrors. However, it also introduces a new kind of latency during the search process, paradoxically to the raison d'être of mirrors. Besides, it is practically proven that most users tend to flee such difficulties by going directly on the source servers, thus generating more inter-network traffic and sacrificing performance. The latter case transparently locates "mirrors" of some of the servers that managed to pay the services of such CDNs; this leaves small organizations out of the picture and "unreachable". Besides, what happens to large organizations that wish to deploy mirrors inside their world-wide intranet ?

In this paper, we present the conception and raison d'être of *ARÃ*, a novel locating system for mirrors in large scale information systems; it is based on cooperation among groups with related interests sharing mirroring information using an epidemic and lazy replication protocol. It opens a path towards making mirroring ubiquitous as much for large organizations and small ones, thus fostering a better use of the Internet and its resources.

The remainder of this paper is organized as follows. The next section gives a brief summary of related work. Section 3 details the architecture of *ARÃ*, its inter-group cooperation protocol for discovering and publicizing mirrors. Section 4 presents the rationale for such work in terms of preliminary measurements and section 5 presents our conclusions and directions for future research.

2 Related Work or Prior Work

In this section, we present some of the most relevant works in the domain of locating documents on Internet.

Work done by [APS98] proposes three solutions for "seamlessly locating replicas"; an HTTP, DNS and IP based schemes. The HTTP [BL95] scheme uses the HTTP "redirect" header to redirect a client request to the "best"replica; this notion of "best", in this case, is relative to the application and not to the requesting client, thus not necessarily having a positive impact on the access time seen by the user. Using such a mechanism, the semantics of the clients' requests could be taken into account. The DNS [Moc87, VtYB97] scheme determines the "closest" replica by taking profit of the fact that a client's DNS server saves the round trip time (RTT) of packets sent to the DNS servers of the replicas and chooses the server with the lowest response time; therefore, putting the IP address of the "closest" replica in the DNS servers (primary and secondary)- using the *DNS Update* feature [VtYB97]- results in the best access time. This scheme has the advantage of using the current Internet infrastructure as it is; however, it can be noted that such a scheme gives the closest server's replica to a DNS server, but not necessarily the closest to the client. Its primary disadvantage is the time necessary to establish which DNS server is the "closest". The IP scheme is based on the fact that routing inside an organization is invisible to outside requests; that gives the administrators of the replicas the possibility of redirecting the requests to the "closest" replica *from their point of view*.

In [KRR99], the authors propose a system called *LDS*. LDS is an extension of DNS. The client's URL is given to LDS which return a list of servers holding the document. As in DNS, LDS is hierarchical with servers holding chunks of a distributed database mapping, not only names to IP addresses, but even URLs of documents to an IP address. This approach is unrealistic when it comes to handling a huge number of documents, e.g. all documents on Internet, but it may be interesting for selecting the "best" mirror [1]. However, it is still a global solution, as the authors suggest that all URLs be mapped in this database.

The solution presented in [Pul98] implements a scheme using the available routing information exchanged by autonomous systems (*Internet Routing Register (IRR)*) in order to determine the shortest path- the number of autonomous systems traversed- for all destination IP addresses. The routing information is used by each server to build a graph linking all ASs; when a client's request is received, its IP address is examined by the server, looks into its graph and finds the shortest path to a replica in some AS using Diskjstra's algorithm. This solution presents an important flaw : the notion of "closest" is measured by the number of ASs traversed, which is a very approximative figure, if not a false one.

Content Distribution Networks, such as Akamai [Aka], propose a global solution for locating the "best" replica. We will discuss the case of Akamai, the first in the area, knowing that other CDNs proceed in almost the same way. The solution is based on a proprietary and paying network of servers. The first stage is to "akamize" the web site or documents to be served which entails changing the links inside local web pages to point to the akamai network holding replicas of theses links' contents. Upon receiving a client's request, the origin server sends the modified web document to the client which then requests the rest of the document from the Akamai network.

All these solutions tackle the issue of locating documents on the web from its server's side; the latters decide when and where to instanciate a mirror and how/when to make it or not known. In the next section, we will present a complementary side which looks at locating from the users' standpoint, giving them the right to decide on the mirrored sites they wish to access and the ones they don't really care about.

3 $AR\tilde{A}$

In this section, we will succinctly present the main features of the $AR\tilde{A}$ protocol. $AR\tilde{A}$ locates mirrors using intra/inter-group cooperation (see figure 1); a group consists of cooperating organizations having common interests, each organization being composed of sites geographically close with common interests, e.g. companies or specialized groups in research centers.

[1] In this case, it is the same as the DNS scheme proposed by [APS98]

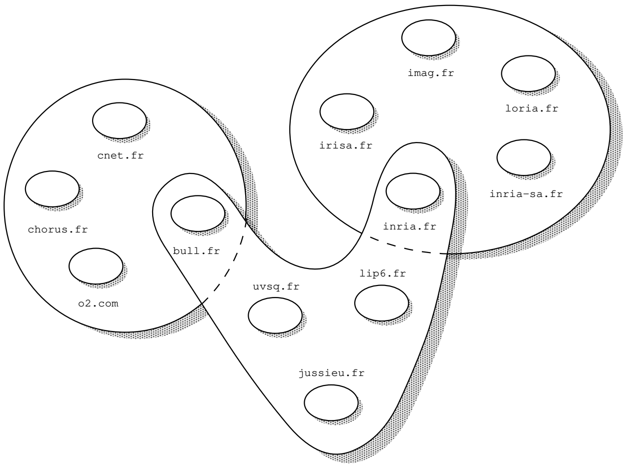

Figure 1: Inter-Group Cooperation

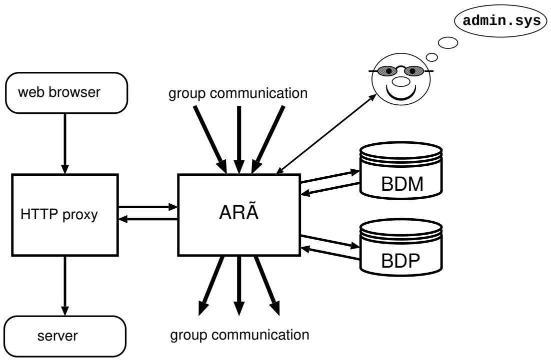

Figure 2: *ARÃ*'s architecture inside an organization

3.1 Transparently Locating Mirrors

For mirrors to be systematically accessed by users, first and foremost, we need to make the location process transparent and "intelligent"; such a service would access the most appropriate server- be it a source server or a mirror- given the source's URL or any one of its replicas. In figure 2, we sketch the way transparent location works :

- the user sends a request via a proxy;
- the proxy asks the location service : *ARÃ*;
- *ARÃ* searches in its mirroring database (BDM[2] returns a list of URLs- be it the source's URL or one of its mirrors'- corresponding to the client's specifications
- the proxy redirects this request to this URL.

This work is based on the redirection of HTTP [BL95] requests. The service uses a mirroring database, BDM, and a proxy which invokes the BDM for the most appropriate matching URL, then redirects the request to the latter[3]. Therefore, filling the mirroring database with the appropriate information is by far the most important building block of this location service : A mechanism for advertising mirrors amongst cooperating groups is used for this purpose.

[2] the french acronym for "Base de Données de Miroirs"

[3] it could well be the same URL if it does not exist in the BDM or if it happens to be the most appropriate

3.2 Group Cooperation Protocol for Discovering and Advertising Mirrors

The group cooperation protocol gives each participating organization the possibility to publicize mirrors it discovered[4], and therefore to benefit from mirroring information released by its peers.

Mirrors' properties are appropriate in large scale systems :

- the lifetime of a mirror is usually long, making asynchronous and lazy cooperation among cooperating sites feasible,
- all of a mirror's content can be identified with just one URL prefix [WO86], making the volume of location data to be stored and exchanged very small

In *ARÃ*, cooperation is done asynchronously. This is possible as incomplete mirroring information on one site does not lead to a bad access; the source server is always there to answer a request. Inter-group cooperation is achieved using point-to-point epidemic communication [PSTT95]. Such mechanism could create cycles;

[4] In fact, mirrors discovered by persons having the right to add such information to its local mirroring database

this problem is treated in the section about filtering mirroring information in section 3.3.

3.2.1 Mirroring Database (BDM)

It is the BDM which holds all the necessary information on mirrors' location; it consists of an associative table linking a URL with it's possible mirrors' URLs. This emulates a unique naming system for all replicas of some source data.

3.2.2 Cooperating Sites' Database (BDP)

Each organization, wanting to cooperate with other organizations, has to know the members of its group. Therefore, a small database consisting of all organizations with which it is cooperating should be created and constantly updated (BDP[5]). The database' content is specific to each organization, allowing each to have a different group of cooperation (see figure 1).

3.3 Filtering Cooperation Information

Information sharing by cooperation has the advantage of naturally selecting the information on the basis of cooperating groups. In *ARÃ*, all sites in a cooperating group share their location informations on mirrors; hence, in such a model with intersections between groups, information may be disseminated indefinitely [6]. Besides, not all the information is interesting to all groups, and it could also be redundant. A solution to these problems would be to filter the information received at the level of each group. Hereafter, we present two interesting solutions :

- **Simple solution :** We use a *flood degree* to determine the number of indirection levels a location information can travel,
- **Elaborated Solution :** We take into account the interest of each organization; this could be done by annotating the contents of each replica (mirror), e.g. using keywords [BDM98].

[5] BDP : the french acronym for "Base de Données des Participants"

[6] in case there are cycles

3.4 Interaction with Caches

In order to benefit from all the available replicas in the system, we are bound to deal with the replicas found in caches. In that respect, it is wise to see that future location systems will have to consider replicas in caches and mirrors.

Therefore, in our system, clients' requests are first handled by the local cache. If it is capable of satisfying the request, the client is served. If not, the cache redirects the request to *ARÃ* which returns a sorted list [7] of the URL's aliases to the cache. The latter inspects its contents to see if one of these URLs, by order of preference, is already in the cache. If it is, the client is served, otherwise, the cache redirects this request to a newly returned URL and caches the response before sending it back to the client.

4 Measurements

As a proof of concept, we present preliminary measurements aiming to show the benefits made from sharing location information among organizations with common interests. Analysing collected traces over a period of time, we will answer three questions : (1) is it interesting to mirror ?, (2) what are the best mirror candidates ?, and (3) is it interesting to cooperate ?

We used traces from three french institutions : INRIA[8], ENST[9] and INT[10]. Traces characteristics are shown in table 1.

	Requests	Start	End
INRIA	3440847	12/8/1998	3/1/1999
INT	2593281	23/11/1998	15/2/1999
ENST	236436	13/9/1998	8/10/1998
Total	6270564		

Table 1: Collected Traces' Characteristics

To answer the first question, we measured the locality of references for each of the collected traces; the results (see figures 3, 4) show that a large number of requests (more than 60%) and traffic (more than 70%) go

[7] based on the client's criteria

[8] Institut National de Recherche en Informatique et Automatique

[9] Ecole Nationale Supérieure de Télécommunications

[10] Institut National de Télécommunications

to less than 10% of the servers for all three traces. This implies that these corresponding servers, for each trace, are potential candidates for mirroring. Locality by itself is not sufficient as some highly accessed servers may be in the vicinity of the clients, and therefore might not need to be replicated. So, there need to be a complementary study on the access time of all these potential mirror candidates in order to pick the "right" mirrors and place them judiciously. This part of the study will not be discussed in this paper.

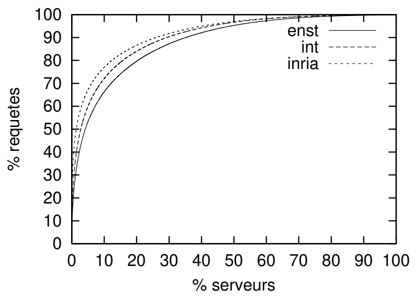

Figure 3: Number of requests per server

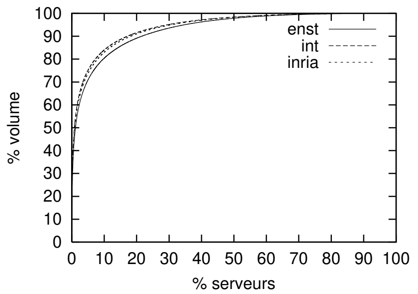

Figure 4: Traffic per server

		Mirror candidates		
		ENST	INRIA	INT
Mirror candidates	ENST	100%	48%	50%
	INRIA	9%	100%	34%
	INT	11%	38%	100%

Figure 5: Percentage of Common Mirror Candidates

Traffic generated toward the mirror candidates		Traffic		
		ENST	INRIA	INT
Mirror Candidates	ENST		80%	85%
	INRIA	52%		81%
	INT	50%	83%	

Figure 6: Locality of Requests to Potential Mirrors

But what about cooperation ? To answer this question, we first measure the percentage of common potential candidate mirrors between the three sites (table 5) and the amount of traffic the latters are responsible for (table 6). Results show that the three sites have common candidate mirrors (up to 50%) and that these common mirrors are responsible for a large amount of requests (up to 85%). Hence, we infer the potential benefits of cooperation among these organizations for locating mirrors. But, for cooperation to work, these mirrors have to exist somewhere. This is the part that distinguishes *ARÃ* from other approaches, as it gives to the users (organizations) the right to create their own mirrors and to list them in the mirror database, making them available to all other cooperating organizations; other approaches consider that, in order to be able to locate mirrors, the latters should be erected by some global "responsible" organization.

5 Conclusion

In this paper, we presented *ARÃ*, a system for discovering and advertising mirrors, as well as managing their aliases. These services are provided using an asynchronous communication protocol among groups of organizations exchanging location information about their respective knowledge of existing mirrors.

The mirroring location system presented is a non global solution which can be deployed in large scale environments. We mean by "global" solution, one that tries to locate all existing mirrors on the net. A solution of such scale has to deal with enormous problems for storing, searching and updating the mirroring database. Besides, such an approach leaves the decision for creating and advertising mirrors to the servers, leaving the users aside. As for *ARÃ*, it focuses on the needs of users in organizations with common interests, thus willing to cooperate in exchanging location information about mutually interesting mirrors. Preliminary results, obtained by analysing real collected traces from three organizations with common interests, show that exchanging the mutually "interesting" location information (probably of locally mirrored sites) leads to a big improvement of access latency, as it can be responsible for amount of clients' requests (up to 85% in the analysed traces). We have already implemented a version of *ARÃ* which we intend to use in real world situations for a more realistic study of its impact on web accesses.

References

[Aka] Akamai. http://www.akamai.com/.

[APS98] Yair Amir, Alec Peterson, and David Shaw. Seamlessly selecting the best copy from internet-wide replicated web servers. In *DISC'98 - 12th International Symposium on DIStributed Computing*, Andros, Greece, Sep 1998.

[BDM98] Vincent Bouthors, Olivier Dedieu, and Mesaac Makpangou. Recommendations sharing forum for internet communities. Url, presentation, Dyade, Bull, INRIA, 1998. `http://webtools.dyade.fr/pharos/`.

[BL95] T. Berners-Lee. Hypertext transfer protocol http/1.0. " ", HTTP working Group Internet draft, October 1995.

[KRR99] Jussi Kangasharju, Keith W. Ross, and James W. Roberts. Locating copies of objects using the domain name system. In *Proceedings of the 4th International Web Caching Workshop*, San Diego, California, April 1999.

[Moc87] Paul Mochapetris. Dns : The domain name system. RFC 1034, Network Working Group, ISI, November 1987.

[PSTT95] Karin Petersen, Mike Spreitzer, Douglas Terry, and Marvin Theimer. Bayou: Replicated database services for world-wide applications. In *Proceedings of the 15th ACM Symposium on Operating Systems Principles (SOSP-15)*, pages 3–6, Copper Mountain Resort, Colorado, December 1995. ACM.

[Pul98] Giray Pultar. Automatically selecting a close mirror based on network topology. In *12th Systems Administration Conference (LISA '98)*, Boston, Massachusetts, USA, Dec 6-11 1998.

[VtYB97] P. Vixie, S. thomson, Y.Rekhter, and J. Bound. Dynamic updates in the domain name system (dns update). RFC 2136, Network Working Group, ISI, April 1997.

[WO86] B. Welch and J. Ousterhout. Prefix tables: A simple mechanism for locating files in a distributed system. In *Proceedings of the 6th International Conference on Distributed Computing Systems (ICDCS)*, pages 184–189, Washington, DC, 1986. IEEE Computer Society.