Interactive Shared Bookmark

Luigi Lancieri
France Telecom
CNET de Caen, France
E-mail luigi.lancieri@cnet.francetelecom.fr

Introduction

ISB has been designed to looking up information on Internet easier. This tool aims to be operated in conjunction with local area networks, for instance inside a company or, more generally, within Intranets. ISB's capabilities are situated between bookmark's ones and search engine's ones. So to speak, ISB allows each user to take advantages of their colleague's bookmarks. Inter activity is one of the main functionality of this shared bookmark.

As in most companies, people are connected to Internet through a Proxy-server, for reasons regarding security and cache-provided benefits. ISB will make up a database from HTML pages that are present in the cache of the proxy-server, and this database is what is first consulted by users.

Company members have to use a traditional web browser to connect to ISB (on local HTTP server). An HTML form is available to them to formulate their queries using keywords similarly to traditional search engines (e.g, Alta vista, Lycos). ISB searches its database and then sends back to the user the addresses of distant sites containing information related to a specific topic, and ranked according to its level of interest.

Information pages can be viewed (by clicking on their address) and rated (good, average, bad) by the user. Such a rating makes it possible to sort addresses according to their level of interest. ISB also elaborates statistics on HTML page access rate, such statistics being taken into account in order to classify addresses.

ISB is suitable to multi-site companies, especially in Intranet context. To do this, ISB contains slave modules, which are scattered among the various company's sites, as well as a main module that manages information sharing between slave modules. Such an architecture makes possible a large information sharing between users of various sites. Nowadays, tendency is, indeed, to distribute caches in a global network architecture (e.g Squid project). The purpose is mainly to reduce web traffic and speed up the load of HTML pages. In this context ISB seems to be an efficient way to share between users the large amount of data contained in these different caches.

Architecture

The components of such an architecture are: a Proxy-server, an HTTP-server, a specifically designed software. This software can be reached by means of the HTTP-server through the Common Gateway Interface (CGI).
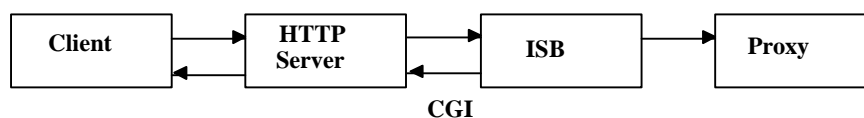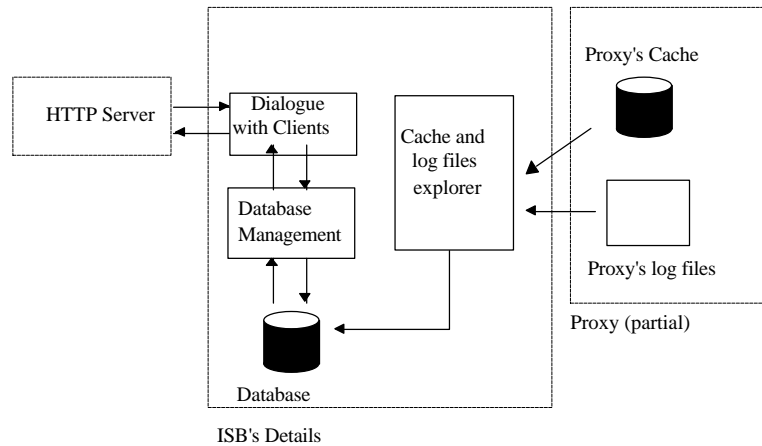

Figure 1. General architecture

ISB components implement the following main functions : the cache and proxy's log files' explorer, the dialogue with user and the building and management of the database.

Figure 2. Detailed Architecture

Advantages

As said before, ISB should be considered as a groupware tool half-way between a local bookmark and a classical search engine. Using this technique presents several advantages:

1) Users receive results ranked according to several suitable criteria : an information quality rating granted by users, an access rate to HTML pages and the content rate of keywords per page. The subjective characteristics of some of these criteria, as information quality grade granted by users are corrected by averaging each criterion. Moreover the compound use of all of these criteria allows a good estimation of the interest of a certain page.

2) Everyone in a site can take advantage in their colleagues' searches, as all viewed pages are present in proxy's cache and are used to build the search database. This database is built according to requirements of a limited user's group (firm's staff), which makes the search process more accurate, so that the database size and the updating delay can be minimised. ISB is also expected to transfer unsatisfied requests to a « general » external search engine (e.g. altavista).

4) ISB is locally implemented, so that it can be reached easier and faster than a classical external search engine would be. The proxy server is set up to make a nightly updating of its cache. The desired HTML page is downloaded (if it is present) from the in-site proxy's cache, so that users have little time to wait.

5) Database building may be cheap, as ISB needs not run all over the web, unlike a classical in site or external search engine.

A large part of this system's interest and efficiency lies in its inter-activity. Actually, the more ISB is used, the more its database enriches itself (by learning process), and the more powerful ISB becomes.