

www4mail: Helping to Bridge the Digital Divide

Clement Onime, Enrique Canessa

The Abdus Salam International Centre for Theoretical Physics – ICTP; P.O.Box 586 - 34100 Trieste, Italy

<mailto:{onime,canessa}@ictp.trieste.it>

Abstract

To some extent the exchange and transfer of knowledge and technology today is possible using off-line Web browsing. Software applications like www4mail provide Web Navigation and Database Search via e-mail by preserving as much as possible the layout of an original Web page and allow to follow hyperlinks as if one were surfing the Web via a live Internet connection. This type of approach, has proven that high-speed connectivity is not essential to help closing the information gap. In particular, www4mail has found new applications in the delivery of information to blind people as well as in the delivery of electronic Journals via the new eJDS programme of the Abdus Salam ICTP. www4mail has been adapted to incorporate Publishers restrictions and to accommodate the plethora of user's networking and computer environments found world-wide.

Overview

Since 1998, we have been developing and maintaining an open software application, named 'www4mail' which allows anyone to browse the Web, access on-line databases and any other Web based information resources via e-mail only [1]. The www4mail software is licensed under the GNU Public License [2]. The most significant impacts of this strategy are that www4mail is very low cost to establish and that it has been able to evolve fast through extensive user feedback. Most recently, www4mail has found new applications in the delivery of electronic Journals via the eJDS programme of the Abdus Salam ICTP.

While the technological revolution of the World-Wide Web continues to expand fast, there is still a large community of Internet users who only have access to e-mail, because their Internet providers do not offer full Internet connections (due to the lack of adequate infrastructure and low-bandwidth lines) or because they simply cannot afford to paid for full or long sessions Internet capabilities at, e.g., available Internet Cafes.

Many of these users live in remote areas of less-developed countries, and rely on e-mail to access essential information, e.g., medical, business and news, as well as for interpersonal communication. www4mail (logo in Fig. 1) allows to get Web pages with all of its functionality intact, with or without graphics. Users can follow links and fill out forms as if they were surfing the Web via a live connection; each action triggers an e-mail message to the www4mail server requesting the next page, or delivering keywords to a search engine, etc. When this is displayed on a Web browser, it shows up some (click-able) buttons before each highlighted line representing the hypertexts that link to other Web documents (see Fig.2).

This feature is added to facilitate navigate the Internet by simply selecting such buttons. Today several main public www4mail servers are available (e.g, at Kabissa in USA [3] and Bellanet in Canada [4]). These are operating at full capacity and deliver thousands Web pages each via e-mail to anyone, anywhere and at any time free of charge.



Fig. 1: www4mail logo



Fig.2: Example of Web page received via www4mail

Brief on Technicalities

The www4mail software is written in modular Perl (a script programming language), and it allows: retrieval of web pages, searching of arbitrary databases, filling of Web forms (GET or POST), to follow links (as in on-line Browsing), handle frames and multi-lingual Web pages, to download web pages as a PostScript or PDF files, to handle current Internet standards (MIME, HTML4.0, Javascripts ...). The software runs under Linux, it is being used actively on RedHat and Mandrake versions, some reports of usage on free BSD.

One sends an e-mail to a www4mail server and in the body of an e-mail message, one types the URL address of the Web page to be retrieved together with some of the many options available. Such options include, e.g., options to retrieve unmodified HTML sources in the body of message, options to retrieve unmodified HTML sources as attachment, help files, etc. Users can also fill out forms as if they were surfing the web via a live connection; each action will trigger an email message to the www4mail server delivering keywords to a search engine, etc.

The server will fetch the page(s) in question and will deliver them upon request. URLs can be of the form: http://, https://, ftp://, news:// For the www4mail system administrator, there are over 50 options which can be configured in the server to set parameters such as: maximum quota per user, gateway administrators, maximum size of each request, size to split large files, access control lists, deny some sites, etc

The installation procedure of the server is also simple but O.S. experience is needed. To optimize its configuration some preliminary monitoring is necessary.

www4mail has broken new ground in flexibility for the Web-to-email service provider offering comprehensive management facilities: Web server style logging of access which allows usage of www4mail to be monitored and graphed by standard Web statistics programs, access control which can be used to restrict access to particular user groups, i.e. non-profits or particular developing countries regions. Besides, it allows for content filters which can be used to prevent abuse of the www4mail program by users attempting to use www4mail to access inappropriate material.

All of these features can be enabled or disabled as desired by the www4mail service provider. www4mail is designed to replicate, as closely as possible, the experience of browsing the Web via a full Internet connection including the search of on-line databases. The software is designed to be hardware efficient and can be configured to whatever hardware is available. However periodic software upgrades have been necessary to cope with increasing www4mail usage demand and also to fix bugs. Further technical information on www4mail programme can be found in [5].

eJDS application

The eJournals Delivery Service (eJDS) of the Abdus Salam ICTP in Trieste is geared to facilitate the access to current scientific literature [6]. Researchers in the world's poorest nations can now receive current and/or old scientific papers free of charge by e-mail by the new application of the www4mail software. The eJDS operates since January 2002. Several publishing companies have joined the eJDS programme. Under this agreement, developing world scientists can today have access to a much wider range of current scientific information and findings than ever before. And the publishers are able to reach researchers who would otherwise not have either the technical or financial means to read articles in their journals in a timely fashion. The Abdus Salam ICTP can broaden its vital role to meet its mandate and transfer knowledge to scientists from the South.

Although eJDS is a free service, it will not give access to Web sites or on-line Documents that are not in the list of allowed eJournals or to those eJournals in which an eJDS user has not been granted access. Using www4mail Access Control Lists (ACL), eJDS implements 2 levels of control, one to deny/allow access to some particular Copyrighted Web sites, and a second to give access to authorized e-mail addresses only (subscribers list). Another important www4mail feature is the multi-quota system, as Publishers have set limits to the number of articles per journal that can be accessed each day (three), month (12) and year (100), www4mail keeps now a complete log file per user. If a eJDS user have exceed his (daily, monthly, yearly) quota, the server will inform with a '503 Error Message - Service Unavailable (quota finished)' and will stop answering requests.

The eJDS procedure to follow is similar to that used when connected in real time, direct mode to any Web server. However the www4mail software need still to be supported to meet the demands of the dynamical Web technologies specially used by Publishers to prevent misuse and identify their customers.

Besides the navigation and search of the entire Internet via e-mail, using any standard web browser and a MIME (Multipurpose Internet Mail Exchange)-aware e-mail program,

most recently new additions to www4mail (from eJDS) is to allow the use of cgi scripts for obtaining the replies from users, that www4mail sends the user an e-mail but the user's reply returns via a CGI interface, this has the benefit of reducing the user's waiting time.

Remarks

www4mail as service, and the software behind it, directly addresses the problem of the digital divide by providing vital and equal access to the wealth of Web information to Internet users and in particular to eJournals to scientists in remote areas. In addition, www4mail preserves Web contents, language, style, etc, and it can also help blind Internet users surf the Web when using it together with text-audio software.

In all likelihood, the demand for Web-to-email technologies will keep pace with the expansion of the Internet. It would be then desired that governments in industrialized countries, non-for-profit organizations, etc could look at bridging the digital divide from this perspective too.

The most important lesson learn from www4mail is that high-bandwidth access to the Internet is not essential in order to bridge the digital divide.

References

- [1] www4mail Project - <http://www.www4mail.org>
- [2] GNU Licensing - <http://www.gnu.org>
- [3] Kabissa Space for Africa - <http://www.kabissa.org>
- [4] Bellanet Supporting Collaboration in the Development Community - <http://www.bellanet.org>
- [5] InformIT web site: <http://www.informit.com>
- [6] eJDS project: <http://www.ejds.org>