



Appendix K

University of Moratuwa

The University of Moratuwa, originates from the Ceylon Technical College of Maradana, which was established in 1893. The University is overlooking the Bolgoda Lake.

LEARN (Lanka Academic and Research Network), which connects academic institutions, including Moratuwa, was established in 1990.

1. Introduction

Moratuwa is one of 20 member institutions that participate in the LEARN network (Lanka Academic and Research Network) for connection to each other and the Internet. The University of Moratuwa, University of Colombo, the Open University, University of Ruhuna and the University of Peradeniya are some of the participating institutions. The Swedish International Development Agency was involved in funding the LEARN network.

2. The Network

The hub of the LEARN network (www.ac.lk) is at the University of Moratuwa (www.mrt.ac.lk), where the LEARN Technical and Training Unit is also based. The LEARN Finance and Administration Unit is based at the University of Colombo.

Ten of the LEARN institutions (including the University of Moratuwa) are connected at 2 Mbps, while the rest connect at lower speeds. However, the total bandwidth from the LEARN network to the Internet is 3 Mbps.

The LEARN network gets its bandwidth via two ISPs, Sri Lanka Telecom (www.slt.lk) and LankaCom (www.lankacom.net). LankaCom is the largest provider of bandwidth, and is also where the Sri Lanka Internet Exchange (IX) is located. The other ISPs as well as the LEARN network therefore connects to the IX, keeping local Sri Lanka traffic within the country.

Both Sri Lanka Telecom and LankaCom have connections to the SEA-ME-WE fibre optic submarine cables, which connects them to the large SingTel IX (Internet exchange) in Singapore. They also have VSAT connectivity to the Internet.

The university itself has an optical fibre backbone.

2.1 Future plans

The LEARN network is due to be upgraded in 2004 to a 100 Mbps link between the four large universities, and to connect more educational institutions. It is hoped that the international connection can be upgraded to as much as 45 Mbps.

3. Library

The Moratuwa University library specialises in Engineering, Architecture and Information

Technology resources. On-line journals are available at:
www.lib.mrt.ac.lk/resources/journals/elecj.html

The library uses the LibSys library system (www.libsys.co.in), which provides the services such as Cataloguing, Circulation, Serials Control, an Article Indexing System, OPAC (Online Searching of Bibliographic Information) and Web OPAC (which makes the catalogue system available on-line) –see screenshow below.



On-line and other resources are explained to students at demonstration sessions. Students and researchers can also utilise library staff for conducting professional searches.

4. Optimisation

4.1 Caching

Squid proxy server is used to cache web pages that have been accessed recently. Squid is implemented as a transparent proxy (at the central cache server), therefore, users are automatically served by the proxy for all web access without any browser configuration. The University has staff with in-depth knowledge of Squid, and they have implemented some of the more advanced features of the software:

- Cache Hierarchy. The main Squid server is a parent cache to the proxies at each university. Since the links between the universities are fast compared to the Internet link, the cache hierarchy improves performance more than a single proxy at each university would have done.
- Using the Squid delay pools to allocate bandwidth per network and even per individual computer. To implement this, the Squid source code was modified.
- Using the Squid redirector to catch a download request, and present the user with a choice to download the file immediately or overnight. Using the URL redirector in original ways such as this is a popular way to extend the functionality of Squid.

4.2 The Off-line download system

The off-line download system is to encourage users to download FTP files overnight. If a user clicks on an FTP link, the Squid redirector (which is a Squid feature for URL redirection) invokes a program written by University staff. This program operates as follows:

Users accessing ftp: URLs are served by our program, which presents them with a directory listing in which each file has two links, one for normal downloading, and the other for off-line downloading. If the off-line link is selected, the specified file is queued for later download, and the user notified by e-mail when the download is complete. The system keeps a cache of recently downloaded files, and retrieves such files immediately when requested again. The download queue is sorted by file size. Therefore, small files are downloaded first. As some bandwidth is allocated to this

system even during peak hours, users requesting small files may receive them within minutes, sometimes even faster than an on-line download.

Since Squid is also an FTP caching server, files requested previously may be in cache and served immediately.

At the mirror site (www.mirror.ac.lk) A user can also supply the URL of the file to be downloaded. Then the system will automatically download the files in idle hours.

The Off-line download system will also be used to identify popular download or download sites for inclusion on the mirror site (see below).

4.3 Mirroring

The LEARN network maintains a joint mirror server (www.mirror.ac.lk) where large software sites are mirrored, helping to keep large downloads off the Internet link. Sites that are mirrored include TuCows (a popular shareware and freeware download site), CPAN (Comprehensive Perl Archive Network), MySQL, Sunfreeware, mirrors of Linux distributions and the GNU site (www.gnu.org).

5. Bandwidth management

The LEARN network uses a modified Squid proxy to manage bandwidth to users and networks. The Delay Pools dynamically changes the bandwidth available to a user between a minimum and maximum value, depending upon the load on the cache server.

The section below is quoted from Dr Dias' article 'Managing Internet Bandwidth: The LEARN Experience', which is available on the Internet at

www.isoc.org/isoc/conferences/inet/01/CD_proceedings/T86/ManagingInternetBandwidth.htm.

The technique is also described in a presentation by Dr Dias at

www.learn.ac.lk/workshops/feb02/bw-squid.ppt

Delay Pools

Squid's delay pools feature was configured to allocate bandwidths to each site. As cache hits fall outside this allocation, sites effectively receive a bonus when accessing popular pages. This method benefited smaller sites, which did not exceed their allocation, and thus obtained fast access times. The load offered by larger sites, on the other hand, exceeded their bandwidth allocation, thus resulting in their traffic being delayed. This method was adopted as it was considered equitable. It was observed that a significant proportion of the traffic comprised file downloads. If these were eliminated or reduced, the performance of interactive web access would improve. Our primary strategy for downloads was to encourage off-line downloading (described below), and restrict downloading during peak hours. Accordingly, we configured the delay-pools (initially at U. Moratuwa) to allocate bandwidth per-user (i.e., IP address). After some experimentation an average bandwidth of 1 kB/s was decided on. This was found to be acceptable (though by no means ideal) for most sites. However, waiting for a page to download at this speed was p-a-i-n-f-u-l-l-y s-l-o-w. Therefore, the pool size (which specifies how much data can be accessed before bandwidth limitation is enforced) was set to allow a "reasonable" page to be accessed quickly. After experimentation, this value was set to 90kB, which allows approx. two average web pages to be viewed without delay. File downloads, on the other hand, are quickly limited to 1kB/s. It was noted that the parent-children cache configuration, although recommended by Squid, resulted in significantly more delay than a single cache. Therefore, we modified the cache configuration so that each site-cache accessed the Internet directly, but used Squid's cache digest feature to check if a file was available in another cache, and if so, to obtain it from there.

Dynamic Delay Pools

One of the major limitations of Squid's delay-pools implementation (and the other bandwidth allocation methods we examined) was that even if a user accesses the

web at night, when the overall usage is minimal, he only receives the allocated bandwidth. Therefore, the system is inefficient, and users have no incentive to use the system during off-peak hours. Therefore, we modified Squid to allocate delay pools dynamically. Each cache is allocated an overall bandwidth, and the bandwidth allocation of each user and user group is dynamically varied to keep the overall usage near the overall allocation. During low-usage periods, each user gets an essentially unlimited bandwidth. With this feature, we can provide much better access speeds to off-peak users, who are thus encouraged to use the system during off-peak times for high-volume applications.

User Groups

In the initial implementation, all users had equal status. However, LEARN's requirement was for lecturers and researchers to have good performance, while undergraduates need not have the same performance. Also, some universities, faculties, and research groups were willing to pay for more performance, while others were not. Our modified Squid provides for multiple user (i.e., IP address) groups to be defined, and for each group to be assigned a per-user bandwidth and a per-group bandwidth. Users within a group are allocated any unused bandwidth for that group, and unused overall bandwidth is allocated to groups which require it. Currently, only two groups are defined: general users, who are allocated 1kb/s, and researchers, who are allocated 3kb/s. With 3kb/s we find that most web sites can be accessed with no delaying.

6. Monitoring

Staff at Moratuwa university have developed their own monitoring tool named LEARNStat. It analyses traffic flow data collected by NeTraMet and uses RRDTOOL for plotting graphs. LEARNStat is used for analysing total bandwidth usage, to identify congestion and traffic trends, and is sufficient for analysing bandwidth usage.

LEARNStat requires NeTraMet, RRDTOOL and Perl to be installed.

See www.ac.lk/LEARNStat for more details. This tool is now Open Source.

Squid logs are analysed for web and FTP usage.

7. Conclusion

The University of Moratuwa is fortunate in having experts at bandwidth optimisation involved in the network, and in the LEARN network. The university staff have developed several solutions, and published papers on optimisation.

The joint academic network LEARN, is also a good model, because in pooling skills and financial resources a country can obtain better connectivity for its academic institutions than when each institution is responsible for its own connectivity.

Case study contact

Ruwan Gamage,
Assistant Librarian,
University of Moratuwa Library,
Sri Lanka
Email: ruwan at lib.mrt.ac.lk
Tel: 94 1 650301 ext. 379