



Journal of Information, Law and Technology

Improving stability and performance of an international network of free access legal information systems

Andrew Mowbary
andrew@austlii.edu.au

Graham Greenleaf
graham@austlii.edu.au

Philip Chung
philip@austlii.edu.au

and **Daniel Autin**

Australian Legal Information Institute (AustLII)

This is a **Refereed** Article published on **22nd November 2007**

Mowbary, A *et al*, "Improving stability and performance of an international network of free access legal information systems", [JILT 2007 \(2\)](#), <http://go.warwick.ac.uk/jilt/2007_2/mowbary_etal/>

Abstract

High quality legal research must increasingly be global and comparative. This is hindered by the limited range of countries' laws covered by the centralized systems of the multinational commercial legal publishers and by the costs of accessing their materials. Networking of online legal information by commercial legal publishers goes back to the 1970s. Over the last decade a global decentralised network of *Legal Information Institutes* (or *LIIs*) has emerged, providing free access to legal information which is comparable with and sometimes better than the commercial providers. Australia's LII - *AustLII*, has been a lead player, and created and runs the *World Legal Information Institute* (WorldLII), the principal interface into the shared LII legal data.

The ad-hoc nature of the technical networking between the dozen existing LIIs means that the effective utilisation of this shared infrastructure has previously been sub-optimal and increasingly fragile. This has been exacerbated by the network's constant expansion. This paper outlines an initiative funded by the Australian Research Council's E-Research programme to address these problems by building a flexible generic set of tools to support and enhance access to WorldLII and more generally, any network of geographically distributed set of web-based systems.

Keywords

Legal research, legal information, networks, free access, commercial providers

1. The need for global legal research

Legal research increasingly has global dimensions. International trade and the WTO, the internationalisation of environmental and human rights issues, and the complexities of the growing numbers of regional agreements and organisations are some obvious examples. There is also a growing need for law reform (including the legislative implementation of treaty obligations) to have regard for successful models from different legal systems. There are calls for a more international common law and for bridges between the world's differing legal systems (eg Lord Cooke 1996; Poulin 2003; Greenleaf, Mowbray & Chung 2004).

Whether we look from the perspective of a commercial lawyer, a government law reformer or negotiator, a legal academic, or a NGO concerned with issues of global impact, there is a need for facilities which allow legal research to have a regional or even a global dimension,

2. Commercial legal publishers and their networks

The largest commercial legal publishers, particularly LexisNexis and WestLaw, operate globally in the range of national legal materials offered by their international portals, thus facilitating internationally-oriented legal research. The world's dominant legal publishing groups (Reed-Elsevier, Thomson, and Wolters-Kluwer) now own the major legal publishing houses in a large number of countries and are rapidly expanding their holdings in other countries. Despite these developments, the range of countries their content covers is still relatively small and confined largely to countries with mature markets.

When did legal information systems first contain significant content from multiple countries? By the early 1980s Lexis (now LexisNexis) had added UK and French content to its dial-up service. EURONET Diane also contained legal database from a number of countries (Bing 1984). By the mid-80s other European content, New Zealand and Australia were available on LexisNexis. Today, this has expanded to cover content from 19 countries, under the LexisNexis 'Global Legal' section (previously less politely named 'Non-US'). Thirteen are Commonwealth countries, and 'Commonwealth' is the only multi-country search provided. Provision of caselaw and legislation varies. Caselaw and legislation cannot be searched together. *A fortiori*, it is impossible to simultaneously search the whole of the LexisNexis international site.

Westlaw International has gradually expanded since 2000 to cover caselaw and legislation from six jurisdictions: the United States, United Kingdom, the European Union, Australia, Hong Kong, and Canada (Westlaw 2007). The maximum scope of a simultaneous search is 10 databases, and you need to be able to select 13 to search everything they have. Cases and legislation can be searched simultaneously, with other content. The third potential commercial network has not yet happened. The Wolters Kluwer/CCH companies have content scattered on sites around the world (US, UK, CCH Asia, Australia, New Zealand, Canada etc) but no network or mega-site.

The LexisNexis and Westlaw International systems were based on the model of one centralized location of data from multiple countries, and remain that way. In the mid-1980s

there were attempts to establish another model: a dial-up network of collaborating commercial providers of legal information from multiple countries. For example, Australia's then-CLIRS system (later bought by LexisNexis) tried to set up a network with the UK's Eurolex and Canada's QL, but this was scuttled when LexisNexis bought Eurolex, despite litigation challenging this (see Greenleaf, Mowbray and Lewis, 1988, p37). Such networks were in any event gateways at best, not allowing for simultaneous searching. During the CD-ROM era of legal research, from about 1988-95, it seems the emphasis was on circulating plastic, not building networks (Greenleaf 2004). It was not until the commercial 'Internet era' from about 1994 that this changed.

3. The free access alternatives

The alternative source for legal information is provided by the free access providers. Since the popularisation of the World Wide Web in the mid-1990s, the Internet's provision of a relatively inexpensive but sophisticated means of access to information has provided for the first time the prospect of effective general access to at least a country's 'essential' legal information (legislation, cases, treaties, and other information that should be publicly available). In most countries, government agencies, Courts, NGOs, Universities, law firms, and others have now created numerous web sites. A side-effect of this domestic free access is that researchers in other countries also have such access. A country's provision of free access to its laws is reciprocal.

However, the mere proliferation of legal web sites has not provided a satisfactory means to conduct systematic and comprehensive legal research, either at a national level or internationally. The existence of these numerous websites may not be known to users, and they are likely to have inconsistent means of both browsing and searching (if searching is even possible). After over ten years of free access to law on the Internet, what we usually find is a multiplicity of 'silos' of what are usually small parts of national law, with no means of linking them into a comprehensive research facility. Irrespective of the level of economic development of a country, it seems that more often than not, the sources of a country's legislation do not spontaneously cooperate with its numerous providers of court decisions, its agency responsible for treaties, or law schools that produce law journals, in order to produce a comprehensive legal portal for a jurisdiction. There are exceptions, from Legifrance to Sri Lanka's Lawnet, but the rule is rivalry and lack of cooperation between the sources of law. The problem is of course exacerbated if you need to find the law from more than one jurisdiction: inter-jurisdictional official cooperation is largely unknown.

What solutions are possible at both national and international levels? The use of web spidering and search engines is part of the answer, but as yet even Google provides nothing approaching a serious legal research facility. Another possibility is the development of standards for legal information adhered to by primary providers (eg legislatures and courts), and possibly the use of distributed searches over their sites. No successful example is known.

4. Legal Information Institutes – the best alternative so far

The most successful solution to date has been the creation of comprehensive law sites for all the essential law from a jurisdiction, country or region, through aggregation of data from multiple sources. There are some governmental attempts to create such comprehensive sites such as Legifrance, but successful systems have at least as often been developed by independent, often University-based, Legal Information Institutes (LIIs). We use the term ‘Legal Information Institute’ to refer to a provider of legal information that is independent of government and which provides free access on a non-profit basis to multiple sources of legal information (generally legislation and case law). They are therefore essentially aggregators of public legal information at a national or sometimes regional level. The primary motivation and rationale of the LIIs is to promote access to law as matter of good public policy (Greenleaf 1995a, Greenleaf 1995b, Greenleaf 1995c, Greenleaf 1997 and Greenleaf 2005).

The current LIIs include: the *Legal Information Institute* (LII (Cornell) – Cornell University - 1992) for US federal law; the *Australasian Legal Information Institute* (AustLII – UNSW and UTS -1995); the *British & Irish Legal Information Institute* (BAILII – Institute of Advanced Legal Studies, London and University of Cork - 2000) for all the jurisdictions of the UK and Ireland; the *Canadian Legal Information Institute* (CanLII University of Montreal - 2000); the *Pacific Islands Legal Information Institute* (PacLII – University of the South Pacific - 2001) originally for fourteen island countries of the Pacific, now for twenty, the *Hong Kong Legal Information Institute* (HKLII – HKU - 2002) for Hong Kong and potentially other parts of China; the *Southern African Legal Information Institute* (SAFLII – Wits University– 2003; now operated by the South African Constitutional Court) for Southern and Eastern Africa, *JuriBurkina* (2004) for Burkina Faso, *New Zealand Legal Information Institute* (NZLII – VUW and Otago University - 2004), and *CyLaw* (2004) for Cyprus.

A linguistic rather than regional focus is taken by *Droit Francophone* (2003) with initial content concentrating on West and Central Africa, but some databases from across the francophonie. It is a single LII but now includes content derived from Juri Burkina. It is operated by LexUM at the University of Montreal, who operate CanLII. AustLII has developed two multi-LII sub-portals. In 2005 it developed the *Commonwealth Legal Information Institute* (CommonLII, 2005) for Commonwealth and Common Law countries (‘droit Anglophone’ is its nickname), which added databases from 20 new countries, but also relied upon the content of existing LIIs: AustLII, BAILII, CanLII, PacLII, HKLII, SAFLII and CyLaw. The *Asian Legal Information Institute* (AsianLII, 2006), drew on CommonLII, PacLII (for PNG) and HKLII, but also databases from 12 additional Asian countries. These three developments took place after the establishment of WorldLII.

LAWPHIL (Arellano Law School – 2001) provides its Philippines databases via AsianLII, but is not yet part of the network technically. The Global Legal Information Network (GLIN) at the US Law Library of Congress has also now added its databases.

The current members of the Free Access to Law Movement are as follows (as at April 2007), omitting LAWPHIL and GLIN:



The more established (ie pre-2003) LIIs are the dominant providers of free access legal in their jurisdictions. AustLII and PacLII, and possibly CanLII, are the dominant online providers *per se* for their jurisdictions.

5. International collaboration - the free access to law movement

Before 2001 there were some LIIs, but no network. Since then both a technical network and an international ‘Free Access to Law Movement’ (FATLM) have emerged from previous informal cooperation between key participants and institutions. Both developments has been centred around the University-based LIIs (but is not restricted to them) and is now expanding to include some government and other providers.

The characteristics of the FATLM have been:

- It is based on independent national or regional LIIs from both developed and developing countries;
- It involves the provision of assistance from the more established LIIs to jurisdictions that would like to establish new LIIs or otherwise improve local free access to law;
- It is creating global legal research facilities by creating networks of LIIs;
- It is articulating its goals through the Declaration on Free Access to Law (Declaration 2002), a key statement of which is ‘*Public legal information from all countries and international institutions is part of the common heritage of humanity.*’; and
- It encourages research and technical assistance through an annual Meeting of LIIs and the ‘Law via Internet’ Conference.

The methods of cooperation have developed rapidly. They started with informal discussions on cooperation in 1997 and 1999 in Sydney when AustLII hosted the first two *Law via Internet Conferences*. The catalyst for the building of an international network was the *LII Workshop on Emerging Global Public Legal Information Standards* at Cornell Law School (July 2000), involving participants from the US, Canada, Australia and South Africa, where

the name 'WorldLII' (for 'World Legal Information Institute') was used to describe the challenge of developing a global free access legal research facility. Various possible models were discussed, including a distributed search system proposed by Cornell's Tom Bruce (Bruce 2000).

In 2000-2001, AustLII provided its search engine (SINO) and hypertext mark-up software to assist the creation of BAILII, CanLII, PacLII and HKLII. By the 3rd *Law via Internet Conference* (Sydney November 2001), these developments made it possible for AustLII, with cooperation of the other LIIs, to develop and launch the first version of the World Legal Information Institute (WorldLII – <<http://www.worldlii.org>>), based on making all of the content of these five LIIs, plus some South African materials, simultaneously searchable. The initial (and current) implementation does not rely as much on distributed searches as Bruce's model, but instead uses a combination of distributed searches, regular synchronisation of databases, and centrally hosted content. This architecture is discussed later.

In 2002, the next key collaborative development was when LexUM at the University of Montreal (the operators of CanLII) hosted the 4th *Law via Internet Conference*, making it an annual event and for the first time held out of Australia. It also initiated an annual day long Meeting of LIIs coinciding with the Conference, both to discuss cooperation between existing LIIs and also to encourage the development of free access law services in developing countries. At the Montreal meeting, the *Declaration on Free Access to Law* was also drafted and adopted. In 2003, the 5th *Law via Internet Conference* returned to AustLII in Sydney, followed the Montreal precedents and added a day-long LII Technical Workshop, enabling groups of LII staff to concentrate on common technical problems. The collaborating LIIs expanded to include Droit Francophone (established by LexUM) and the Southern African Legal Information Institute (SAFLII) (established jointly by WITS Law School and AustLII).

In 2004, the 6th *Law via Internet / Internet pour le droit Conference* was the first to be held in Europe, hosted in Paris by Adij (operators of Legifrance, the free access French government system). JuriBurkina, NZLII and CyLaw joined the LIIs. PacLII hosted the 2005 Conference (with AustLII assistance) in Vanuatu. The 2007 Conference returns to Montreal.

The FATML website is being constructed at <www.freelii.org>. Membership is by invitation, with members nominating new candidates, and consensus required. The membership criteria are not fixed but involve adherence to, and support of, the Declaration and activities similar to (but not necessarily identical with) a LII. Recent members are the Institute of Theory of Techniques of Legal Information, Italy (ITTIG), LawPhil, a University-based free access provider from the Philippines, and GLIN at the US Law Library of Congress.

In the seven years since the Cornell meeting, the kernel of a global free access to law movement has therefore emerged, and is establishing principles, forums and methods of cooperation, and techniques for significantly expanding the global reach and effectiveness of free access to law. The challenge now is to engage with people in many more countries, in keeping with the goals of its *Declaration on Free Access to Law*, and to increase the technical sophistication and reliability of its networks.

There are a lot more countries with databases in WorldLII than there are members of the Free Access to Law Movement. PacLII, Droit Francophone, CommonLII and AsianLII have

incorporated into the LIIs the legal content of many organisations beyond their ‘home’ jurisdictions. How this is done varies a great deal. Sometimes the local involvement required is quite passive: a court or legislation website indicates that other are allowed to republish its content, or the copyright law of a jurisdiction states there is no copyright in legal documents (eg in Macau or India). In other cases obtaining republication rights requires requests and negotiation, but they are rarely refused if the data is already being published for free access. Otherwise, negotiating the supply of a stream of otherwise unobtainable data is sometimes very difficult particularly if a commercial provider is already obtaining it (eg Malaysian legislation). In other cases the institution concerned is pleased to obtain an opportunity to publish on the Internet (eg some Ugandan courts). PacLII and SAFLII have staff members who travel from country to country to obtain data not otherwise available. A discussion of the many complexities of data acquisition would be very lengthy. It is sufficient to say that the LII network has only scraped the surface of the valuable legal information which is potentially available.

The main constraining factor is of course funds. Every LII looks after the funding of its own system. The models on which LIIs are funded vary a great deal. AustLII has a ‘multi-stakeholder’ model, with more than 50 stakeholders in that case (see AustLII funding 2007). BAILII is similar. Most LIIs have had a considerable deal of academic funding and institutional support. CanLII is funded primarily by the Canadian legal profession. Aid agencies have made significant contributions to PacLII, SAFLII and AsianLII. Strategic alliances with some legal publishers have helped AustLII along. A small LII like CyLaw is a personal project. Yet others like NZLII live on ‘the smell of an oily rag’ and a little help from their friends, while they search for longer-term funds. There is no Daddy Warbucks (not even a George Soros) who would fund global free access to law long-term, but that doesn’t mean it can’t be done. It has been done with ever-widening scope for over a decade. There is not one formula, but as with many other aspects of open content, there are many non-business models by which numerous stakeholders can be engaged.

6. WorldLII and Droit Francophone – Network hubs

The fourteen cooperating LIIs comprising the free access to law network – AustLII, BAILII, CanLII, CyLaw, Droit Francophone, HKLII, Juri Burkina, LII (Cornell), PacLII, NZLII, SAFLII, CommonLII, AsianLII and WorldLII – have one current network hub (WorldLII) and one emerging network hub (Droit Francophone).

All fourteen LIIs, and some other legal information providers, cooperate to provide WorldLII <<http://www.worldlii.org/>>, which is operated on their behalf by AustLII. WorldLII was ‘endorsed by all the existing LII teams’ (Poulin 2004) at the LII meeting in Montreal in 2002. AustLII is a joint facility of the Faculties of Law at the University of Technology, Sydney (UTS) and the University of New South Wales (UNSW). On a day-to-day basis, AustLII runs the WorldLII facility. However, as a hub in the LII network, WorldLII only exists because of the collaboration and cooperation of the other LIIs in the network, and there is a considerable amount of consultation with other LIIs.

The following diagram summarises the current free access to law network architecture as at December 2006. Since then, the databases of the Global Legal Information Network (GLIN)

of the US Law Library of Congress have been added to the network. The architecture is described below.

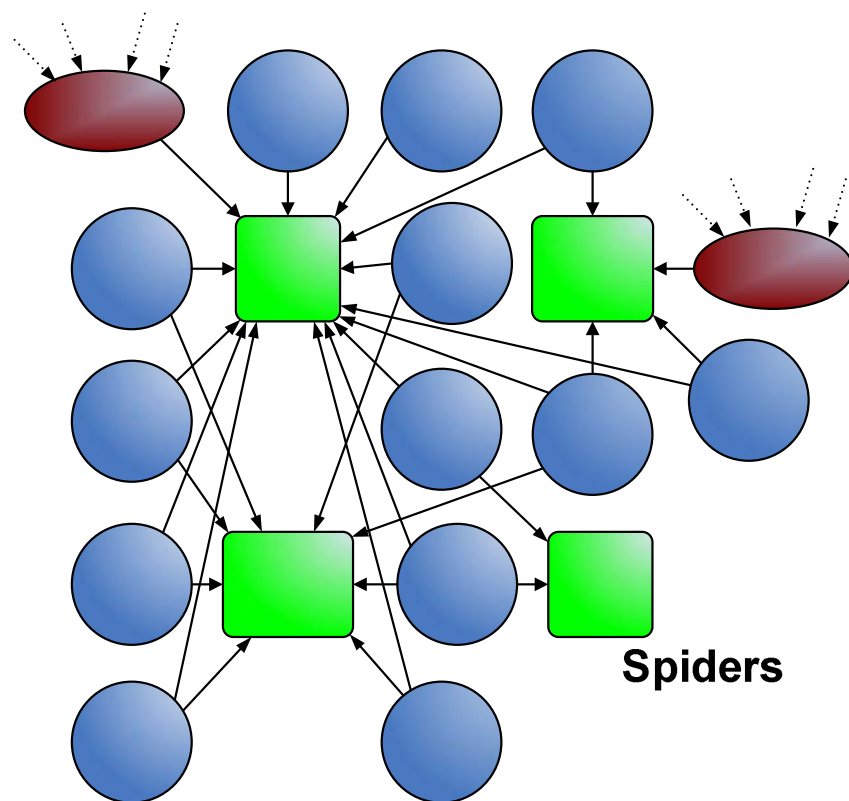


Figure 1 Free access to law network (December 2006)

WorldLII currently (April 2007) brings together 870 databases from 123 countries, plus international institutions, that are provided by the various LIIs so as to present all of the data as an integrated whole. The number of databases made searchable via the network is expanding at about 25% per annum. This includes the ability to conduct searches across all or various cross sections of the data as well as to browse into data into a consistent way. As well as being a network hub, WorldLII is a significant LII in its own right, hosting collections of databases that do not exist on other decentralised LIIs, most notably the decisions of 26 international Courts and Tribunals (Greenleaf, Chung & Mowbray 2005). Although it has primarily an English language interface, WorldLII aims to provide the broadest and most generic interface to the content of all LIIs, irrespective of in what language they are available.

Droit Francophone is a developing network hub based on the common element of the French language. It contains many databases from Francophone countries that it hosts, but it is also developing the capacity to provide comprehensive network access to the francophone content of other LIIs, such as JuriBurkina and CanLII. There are continuing discussions about inclusion of databases from other Francophone countries and networking of legal information systems in such countries. At this stage the LII networks are therefore most extensive for the Anglophone and Francophone countries. GLIN has some coverage of most Spanish-speaking countries, but only provides a facility to search abstracts of documents (in English and sometimes Spanish) and not the full Spanish texts of the documents.

WorldLII Catalog Cylaw

WorldLII Interface AustLII

PacLII

The LIIs want to encourage other language-based hubs within the overall network, and are keen to see hubs emerge for major international languages such as Spanish (WorldLII has a ‘Spanish Project’ demonstration page with its current Spanish databases to encourage this), and Chinese (HKLII is a bilingual system, but there are no other Chinese databases in the network as yet). In short, the LII network is developing into a multi-hub network sharing a common data set, and the solutions that this initiative will develop are intended to benefit the development of all emerging hubs, not only WorldLII.

7. WorldLII hardware architecture

The current WorldLII system hub runs on AustLII hardware in Sydney. It brings together resources which are hosted on hardware located in 9 countries: Australia, Canada, Cyprus, Hong Kong, New Zealand, South Africa, United Kingdom, United States, and Vanuatu. All systems are Unix-based but the types of physical systems differ markedly in terms of their speed and capacity. At the one extreme, AustLII uses a number of Sun Microsystem Enterprise multi-processor servers and multi-terabyte network-attached storage. Some of other LIIs rely on much more modest hardware. For example, BAILII uses four Intel Linux-based servers with single processors.

Network access also varies dramatically between the different systems both in terms of speed and reliability. AustLII is connected directly to the NSW RNO wide area network with redundant 100 Mbs links. On the other hand, PacLII relies on a single satellite connection via USPnet with more limited bandwidth.

WorldLII in its current implementation relies mainly on data replication from the other LIIs to Australia. This is done via various combination of web spidering, automated transfers (using mechanisms such as rsynch over ssh) and customised protocols to support remote searching.

8. The E-Research Initiative

In late 2005, AustLII was successful in obtaining funding from the Australian Research Council as part of the E-Research Initiative to help improve the WorldLII network. Some of the problems that were identified in the application and which were to be addressed included:

- Network latency – where final search results rely upon a number of asynchronous searches, search performance will only be as fast as the slowest responding search result server.
- Reliability – given the fundamentally centralised nature of the WorldLII system, it is reliant upon a single LII (AustLII). This means that WorldLII will not be available if AustLII or its associated networks fail even though the other LIIs are operational. All LIIs have concerns in terms of providing the maximum possible up-times.

- Performance – there is currently no attempt to leverage system performance to utilise the combined processing power of all LIIs. Similarly, some LIIs rely upon networks that work well locally, but that are non-optimal in terms of serving the international community.
- Scaleability and maintainability – as the number of LIIs is increasing it is unlikely that the current approach will scale. A lot of information about remote data sets is maintained manually or semi-manually. No mechanism exists for automated exchange of meta-data.
- The primary aim of the project was to build a much more robust and tightly integrated network of LIIs making much better use of resources (both systems and data). This has involved different levels of participation by the various LIIs depending upon the local availability of equipment and personnel. The two main components of the project were:
 - to build a flexible tool to support and enhance access to WorldLII and more generally any network of geographically distributed web based systems. Some of the particular technical features needed to include:
 - Local load balancing between a number of machines offering access to the same or different web based systems
 - Fall-over access to a set of local machines
 - Real-time synchronization of data between local data sets
 - Flexible load balancing and allocation of requests to a geographically dispersed set of machines
 - Replication of data sets across a global network of machines
 - to implement this in the WorldLII production system and as possible across the LIIs. Ideally, all of the LIIs would replicate each other and would also be able to act as “WorldLII” itself. Requests for access to any individual LII (including WorldLII) could be redirected to take account of network speeds and system loads.

9. LBD – The Load Balancing Daemon

A major outcome of the E-Research project was the development of a replacement DNS server. This server is referred to in-house as the *Load Balancing Daemon* (or “LBD” for short which is also an acronym for the project’s mascot – the *Little Black Duck* :-)

Whilst various commercial products exist that are capable of doing some of the things that we needed LBD to do (for example, Foundry’s ServerIron family of switches), they tend to be very expensive and often rely upon dedicated hardware with proprietary extensions to

operating systems. It was often necessary to purchase hardware and/or software for each participating geographical location. Whilst we could (and have) implemented proprietary solutions for local load balancing and failover between systems that we directly control, in the context of a cooperating group of free to air providers this is not practical, particularly when most of them have very modest financial resources.

The open source systems that are available are not stable, maintained or flexible enough for our purposes. For example, 'lbnamed' is a load balancing name server which directs users to the host with the least amount of unique logins and lower load averages (Schemers 1995). It did not incorporate the proximity of the hosts in its weighting. Another example is 'Super Sparrow' which uses the BGP routing protocol to determine the peer closest to the client and relies upon an existing DNS server Dents to handle the actual requests (Horman 2001). However, it does not take into account of the load or capacity of the hosts and is probably more difficult to deploy. More recently, the 'GeoDNS BIND patch' by Nicholas Moldavsky makes use of the concept of having multiple views of a single zone to filter whole countries without having to specify IP addresses (Moldavsky 2005). It relies upon integrating Maxmind's GeoIP (<<http://www.maxmind.com/app/country>>) to BIND. In this way, it is relatively easy to configure the server to send users from a particular country to one host and other users to another. However, this does not take into account the load and other host attributes for load balancing purposes. We have not found any existing solution to our needs and expect that this is a generic problem that extends beyond WorldLII and that the development of the software for this project should have wider application.

The LBD application is an extensible Domain Name Server that responds to DNS lookup queries with a record set containing a sorted list of the "best" matches for a given server or service. The definition of "best" can be defined by any programmer but in the current version will match first on geographic proximity between the requesting application and the target server. The algorithm will use other host attributes (for example, server response times) as secondary sorting criteria if required.

LBD is implemented in Perl and uses the Net::Server CPAN library – this takes care of most of the "infrastructure" requirements and has allowed development to focus on the specifics of the DNS requirements. The program also uses the Stanford::DNS library which handles some of the underlying DNS protocol encoding and decoding work.

LBD uses a "fork" model to handle simultaneous requests from concurrent clients. As well as accepting standard DNS requests on UDP port 53, LBD also listens for TCP connections on port 3500 – this is where control information is shared. LBD supports master / slave DNS replication so that state can be shared amongst multiple DNS servers to improve redundancy. In addition to state replication, the control port can be used for changing information on available servers, their attributes and the groups to which they belong.

In LBD, the administrator defines "groups". A group describes a set of servers that contain a particular resource. For example, in the current configuration there is a group "paclii". When a web browser looks up "paclii.lb.austlii.edu.au" the hostname being looked up is split by LBD into a group (in this case, "paclii") and a residual ("lb.austlii.edu.au"). LBD then retrieves the list of available hosts in that group. A host would be considered "unavailable" if either an administrator has set its status that way (for example, for planned maintenance) or if

a server monitoring system detects that the server is down. The list is then sorted to find a “best” match:

- Some hosts in the group will be configured to have an “affinity” for a certain geographical area. In PacLII’s case, the Vanuatu servers will have an affinity for Pacific island nation states. Other hosts in the group will have no affinity and these will be the default servers for other countries.
- All hosts in the group will have one or more performance metrics. This can be configured by administrators. In the current implementation, the key performance metric is response time to the most recent HTTP GET request sent by a monitoring system (for example, the open source Nagios system <<http://nagios.org/>> – but almost any commercial or open source system could be used). A community of web sites might choose other metrics however, such as server capacity, percentage utilisation or network bandwidth costs.
- Once the list is sorted it may need to be truncated. This is because most current BIND implementations randomize record sets containing more than one resource record. When this happens, the sort order is lost and so clients are not guaranteed to connect to the “best” server (they will, however, at least connect to an available server). The current configuration therefore limits record sets to one record (however, the most recent version of BIND may allow the resource record order to be preserved – this is being investigated).
- Finally, LBD needs to make sure all resource records are returned as A records for the standard DNS requests. LBD can return NS, TXT and CNAME records if requested.

10. Deployment

The project has involved close cooperation with a number of LIIs and one of the side effects of the project has been to intensify the collaboration between all of the international network of Legal Information Institutes that comprise the Free Access to Law Movement.

Apart from introducing a new DNS server, we aimed to retain as far as possible, the simple WorldLII system architecture. In particular, we did not want to introduce additional points of failure that would result from adopting a truly distributed model (such as suggested by Bruce 2000). Accordingly, WorldLII continues to rely upon an expanded replication / synchronisation model. A complete copy of all LII data is held in Sydney and is mirrored locally both at UTS and UNSW.

As part of the project, LexUM/CanLII and AustLII bought and installed new servers on a knock for knock basis. This has meant that LexUM now has control of a server in Sydney and AustLII has control of one in Montreal. Data and system setup is synchronised on both systems using rsync on a regular basis. The Montreal system also serves as a North American home for a number of other systems including PacLII, SAFLII, NZLII, CommonLII, AsianLII as well as WorldLII itself.

The PacLII system is maintained in Vanuatu, but is mirrored in Sydney. Traffic originating from the Pacific (or more accurately, the USP network) is directed to the Vanuatu server, with the rest of the world defaulting to Sydney. Similarly, the SAFLII system is maintained in South Africa but is again mirrored in Sydney. At the moment, all traffic is directed to Sydney, but as larger servers are installed in Johannesburg, load balancing / failover similar to PacLII will be deployed.

In its current configuration, LBD has been built to work with BIND (Berkeley Internet Name Domain) – an implementation of the Domain Name System (DNS) protocols supported by the Internet Systems Consortium (ISC). BIND has been configured to delegate requests to “lb.austlii.edu.au” to the LBD server. BIND therefore fields all DNS requests in the first instance and returns the answers from LBD to clients. This method of integrating a load balancing system into DNS is recommended by the ISC but does have some drawbacks, for example, the aforementioned problem with record set re-ordering by the BIND server (ISC 2004). LBD is currently being deployed on an experimental basis. It is being gradually introduced across the free access to law network.

11. Future networking directions

As a result of the E-Research project, we now have in place a much more flexible and robust set of mechanisms for managing the WorldLII network and there is a lot of scope for expanding the extent of system replication, load balancing and failover.

There are a number of research directions that have been opened up by the LBD being so easy to extend, such as adding support for SRV record types which are arguably better suited to geo-centric load balancing. However, the complexity of the domain name system and the availability of a first-class open source implementation (BIND) do suggest the possibility of incorporating the LBD algorithms directly into BIND (despite the ISC’s reservations about such an approach) (ISC 2004).

Future research directions could include adding support for SRV records. SRV records are a relatively new DNS type that supports priorities in record sets. In this way, DNS servers could indicate clearly to web browsers which servers are preferable and in what order to try them. Combined with the geo-centric algorithms in LBD this could improve the resilience of geographically distributed load balancing. This would require extending LBD to support the SRV record type (not too difficult given LBD’s architecture) but also adding SRV support to a client web browser such as the open source Firefox. The ISC has written about this as being their preferred approach to DNS load balancing but notes that browser authors are waiting for domain name server authors and administrators to add SRV support, whilst that group is waiting for browser authors to add support (a classic “chicken and egg” problem).

Integrating the LBD algorithms into BIND would yield a geo-centric load balancing domain name server that also supported DNS-Sec, delegation, zone transfers and all the other features of BIND. However, BIND is a poor platform on which to try new techniques for load balancing so this would not be attempted until the algorithms had been fine tuned to work in real world settings.

12. Do networks matter?

Does networking matter in a post-Google world? Can't search engines just spider everything? Well, no. Most LIIs use the Robot Exclusion Protocol (Robots Exclusion 2007) to exclude spiders from at least their case-law, on privacy grounds (you should not find an old friend's divorce or criminal case when arranging a school reunion). Networking LIIs can also add many forms of organisation of the data shared between LIIs that general search engines don't yet provide (Greenleaf 2005). An on-the-horizon problem is the relationship between the LIIs and those who wish to build law-oriented services which attract large advertising revenues partly by spidering LII content, but do not bear any of the ongoing costs of building the quality legal information that the LIIs provide. The attitude of LIIs to spidering is an ongoing question, not only on privacy grounds. The future may be different, but then it will be a question of who pays for the LII value-adding from which search engines will profit.

All LIIs have challenges in maintaining their own operations, the most common one being that free access cannot be provided at no cost. The main issue for the Free Access to Law Movement as a whole, and for LII networking, is to find the best ways to expand both to make them more inclusive of organisations with similar goals, and a more global range of quality legal information. So far, both are largely Anglophone plus Francophone.

The technical sophistication of the network, and corresponding collaboration between LIIs, has increased a lot recently, and augurs well for long-term robustness. A related challenge is to increase the interconnections between data on different LIIs, beyond common search results to direct links between data. CanLII's Reflex (CanLII 2007) and work underway at other LIIs recognises the importance of citations to adding value both locally and across the LII network.

The future for the LII network, its interfaces, and the Free Access to Law Movement is to increasingly provide a global alternative to the expanding global reach of both the current commercial legal publishing duopoly and the Internet-wide search engines typified by Google. In doing so, it can help to provide better access to law in many countries, developed and developing, and can encourage them to join in a global project that supports economic progress, the rule of law and democracy.

References

Bruce T 2000, 'WORL DLII: A sketch for a distributed search system', LII Workshop on Emerging Global Public Legal Information Standards, Cornell Law School, Ithica, NY, July 2000, <<http://barratry.law.cornell.edu/Summit/worldlii.htm>>.

Cooke 1996, 'The Dream of an International Common Law', in *Courts of Final Jurisdiction: The Mason court in Australia*, C Saunders (Ed), Federation Press, 1996.

Declaration on Free Access to Law 2002 (as amended)
<<http://www.worldlii.org/worldlii/declaration/>>.

Greenleaf G, Chung P and Mowbray A 2005, 'Responding to the fragmentation of international law – WorldLII's International Courts & Tribunals Project', *Canadian Law Library Review*, 2005, vol 30 (1), pp 13- 21.

Greenleaf G, Mowbray A, King G and van Dijk P 1995a, 'Public access to law via internet: the Australasian Legal Information Institute', *Journal of Law & Information Science*, (1995) Vol 6 No 1.

Greenleaf G and Mowbray A 1995b, 'Law via the internet: the Australasian Legal Information Institute', *Australian Law Journal*, vol 69, pp 581 (ISSN 1321-3563).

Greenleaf G, Mowbray A and King G 1995c, 'Public legal information via Internet: AustLII's first six months', *Law Technology Journal*, CTI Law Technology Centre, Vol 4 No 2 November 1995, 5-10, ISSN 0961-6902 (also in *Australian Law Librarian*, Vol 3, No 4-5, 1995, 144-153, ISSN 1039-6616).

Greenleaf G, Mowbray A and King G 1997, 'New directions in law via the internet – The AustLII Papers', *Journal of Information, Law and Technology (JILT)*, vol 2 (ISSN 1361-4169).

Greenleaf G, Chung P and Mowbray A 2004, 'A new home online for Commonwealth law: A proposal for a CommonLII', *Journal of Information, Law and Technology (JILT)*, 2004 (2).

Greenleaf, G 2004 'Jon Bing and the History of Computerised Legal Research – Some Missing Links' in Olav Torvund and Lee Bygrave (Eds) *Et tilbakeblikk på fremtiden* ("Looking back at the future") 61-75, Unipub, Oslo, 2004

Greenleaf G. 'Global Legal Research: WorldLII and the Future' *Internet Newsletter for Lawyers*, Jan/Feb 2005

Greenleaf G, Chung P, and Mowbray A 2005 'Emerging Global Networks for Free Access to Law: WorldLII's Strategies' *Journal of Electronic Resources in Law Libraries*, Vol 1

Issue 1, Haworth Press (accepted 9 July 2005); available at [2007] UNSWLRS 16 (UNSW Law Research Paper No. 2007-16)

Horman S 2001, 'Globally Distributed Content: Using BGP to Take Over the World', <http://www.supersparrow.org/ss_paper/stuff/ss_paper.pdf>.

Internet Systems Consortium (ISC) 2004, 'BIND and Load Balancing', <<http://www.isc.org/products/BIND/docs/bind-load-bal.html>>.

LexisNexis (2007) 'Company History' <<http://www.lexisnexis.com/presscenter/mediakit/history.asp>>

Moldavsky N 2005, 'GeoDNS BIND patch', <<http://www.caraytech.com/geodns/>>.

Poulin D 2003, 'Open Access to Law in Developing Countries', *Proceedings of the 5th Law via Internet Conference*, AustLII, Sydney, November 2003.

Robots Exclusion (2007) – The Web Robots Pages at <<http://www.robotstxt.org/wc/exclusion.html>>

Schemers R J 1995, 'lbnamed: A Load Balancing Name Server in Perl', 1995 LISA IX, Monterey, CA, 17-22 September 1995, <<http://www.stanford.edu/~riepel/lbnamed/lisa95-paper.ps>>.

Westlaw (2007) Westlaw International 'Welcome' page <<http://www.westlawinternational.com/>>