

April 2001
14RP-0402A-WWEN

Prepared by:
dotCOM and Service Provider
Solutions

Compaq Computer Corporation

Contents

Introduction	5
Getting Started	5
Linux Information Resources ...	8
IP Addressing and Domain Name Registration Resources	9
General Open Source Software Resources	9
Compaq ActiveAnswers Solution Resources	9
Open Source Software Commonly Used by ISPs	10
Summary	26
Glossary	27

Open Source Service Provider Solutions on Linux

Abstract

This guide provides an overview of Service Provider (ISP/ASP) solutions based on the software available from the Open Source community. The focus of this paper is on the Linux operating system and selected Open Source applications running on Linux platforms.

The target audience includes Compaq customers, systems integrators, and Compaq sales and support personnel who need a better understanding of Open Source service provider solutions running on Linux.

Send your comments and questions on this solution or this document to activeanswers_support@compaq.com. Please include the Document Number (shown in the upper left corner of this page) in the Subject line of your message.

Notice

The information in this publication is subject to change without notice and is provided “AS IS” WITHOUT WARRANTY OF ANY KIND. THE ENTIRE RISK ARISING OUT OF THE USE OF THIS INFORMATION REMAINS WITH THE RECIPIENT. IN NO EVENT SHALL COMPAQ BE LIABLE FOR ANY DIRECT, CONSEQUENTIAL, INCIDENTAL, SPECIAL, PUNITIVE, OR OTHER DAMAGES WHATSOEVER (INCLUDING, WITHOUT LIMITATION, DAMAGES FOR LOSS OF BUSINESS PROFITS, BUSINESS INTERRUPTION, OR LOSS OF BUSINESS INFORMATION), EVEN IF COMPAQ HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

The limited warranties for Compaq products are exclusively set forth in the documentation accompanying such products. Nothing herein should be construed as constituting a further or additional warranty.

This publication does not constitute an endorsement of the product or products that were tested. The configuration or configurations tested or described may or may not be the only available solution. This test is not a determination of product quality or correctness, nor does it ensure compliance with any federal, state or local requirements.

Compaq, Compaq Insight Manager, Deskpro, FASTART, NetFlex, NonStop, PaqFax, ProLiant, Prosignia, QuickFind, Qvision, RomPaq, SmartStart, and Systempro/LT are registered with the United States Patent and Trademark Office.

ActiveAnswers are trademarks and/or service marks of Compaq Information Technologies Group, L.P.

Linux is a registered trademark of Linus Torvalds.

UNIX is a registered trademark in the United States and other countries licensed exclusively through X/Open Company Ltd.

Microsoft, Windows, and Windows NT are trademarks and/or registered trademarks of Microsoft Corporation.

Other product names mentioned herein may be trademarks and/or registered trademarks of their respective companies.

©2001 Compaq Computer Corporation.

All rights reserved. Printed in the U.S.A.

Open Source Service Provider Solutions on Linux
White Paper
Prepared by dotCOM and Service Provider Solutions

First Edition (April 2001)
14RP-0402A-WWEN

Table of Contents

1 Introduction	5
2 Getting Started	5
3 Linux Information Resources	8
4 IP Addressing and Domain Name Registration Resources	9
5 General Open Source Software Resources	9
6 Compaq <i>ActiveAnswers</i> Solution Resources	9
7 Open Source Software Commonly Used by ISPs	10
7.1 Domain Name System Software	10
7.1.1 BIND	11
7.2 Web Server Software	11
7.2.1 Apache	11
7.2.2 PHP	12
7.2.3 FrontPage 2000 Server Extensions	12
7.3 Web Log Analysis	12
7.3.1 Analog	12
7.3.2 Webalizer	12
7.4 Caching Proxy Server	12
7.4.1 Squid Proxy Server	12
7.5 High Availability Clustering	13
7.5.1 Kimberlite	13
7.5.2 Piranha	13
7.5.3 Linux-HA	13
7.6 Application Server	13
7.6.1 Enhydra	14
7.6.2 Zope	14
7.6.3 Zend	14
7.7 Mail Software	15
7.7.1 Sendmail	15
7.7.2 Qmail	15
7.7.3 POP3	15
7.7.4 IMAP	16
7.7.5 IMP	16
7.7.6 TWIG	16
7.7.7 Procmail	16
7.7.8 Fetchmail	17
7.7.9 Majordomo	17
7.7.10 LISTSERV	17
7.7.11 MAPS	17
7.8 News Software	17
7.8.1 INN	17
7.8.2 Diablo	18
7.8.3 NNTPCache	18
7.9 File Transfer Protocol Software	18
7.10 Firewall and System Access Control Software	19
7.10.1 Linux with IP Masquerading	19
7.10.2 IP Filter	19
7.10.3 Ipchains	20
7.10.4 Iptables	20
7.10.5 TCP Wrapper	20
7.11 System Security Sites	20
7.11.1 CERT	20
7.11.2 CIAC	20
7.11.3 CSRC	21
7.11.4 BugTraq	21
7.12 System Management Software	21
7.12.1 Virtual Network Computing (VNC)	21
7.13 Remote Desktop Protocol Client	21

7.13.1 Rdesktop	21
7.14 Routing Software	21
7.15 Directory Service Software	22
7.15.1 OpenLDAP	22
7.15.2 LDAP	22
7.16 Chat Software	22
7.16.1 IRC (Internet Relay Chat)	22
7.17 Shared File and Print Services	23
7.17.1 Samba File and Print Server	23
7.18 Development Tools	23
7.18.1 Perl	23
7.18.2 CPAN	23
7.18.3 Python	23
7.18.4 GNU Development Tools	23
7.18.5 Tcl	24
7.18.6 expect	24
7.18.7 TclTk	24
7.18.8 TclX	24
7.19 Client Application Software	25
7.19.1 Applications	25
7.19.2 Lynx	25
7.19.3 Pine and Pico	25
7.19.4 TIN	25
7.20 Database Software	25
7.20.1 MySQL	25
7.20.2 PostgreSQL	26
8 Summary	26
9 Glossary	27

1 Introduction

This guide offers a basic introduction to the steps involved in establishing an Internet Service Provider (ISP). It introduces Internet applications typically used by ISPs and Application Service Providers (ASPs), and describes bandwidth requirements, equipment needed, registration processes for Internet Protocol (IP) address space, domain names, and security needs. Most importantly, it will point you to further sources of information for specific details on particular Open Source Solutions used by Service Providers (SPs) on Linux platforms.

The past few years have seen explosive growth in the Open Source community, in both the quality and quantity of applications available. The public is beginning to understand what the SP community has known for several years. Open source software provides qualities essential to an SP. In layman's terms, the software is offered at no cost, source included, and includes the right to modify the software.

This lowers the entry barriers for a new SP, and most importantly gives control to a SP. Having access to the source allows SPs to understand the inner workings of an application, and provides the ability to diagnose and fix a problem in a timely manner.

2 Getting Started

Service Providers come in a variety of flavors. Some key examples are:

- NSP – Network Service Providers sells bandwidth and network connectivity to ISPs, ASPs, HSPs, and MSPs.
- ISP – Internet Service Providers are organizations that provide consumers and businesses with access to the Internet. ISPs typically provide bandwidth and sometimes additional services to their customer base. Some of these services might be electronic mail, news service, and space for Web pages.
- ASP – Application Service Providers offer access to software applications and support services via the Internet. This kind of service can save a business the cost of installing, managing, and supporting applications. Some typical examples include customer resource management tools, payroll, expense, human resources, and electronic mail.
- HSP – Hosting Service Providers are companies that provide space and services for hosting Web-based content. There are three main types of hosting services: shared, dedicated, and co-location. Shared services occur when multiple sites are hosted on a single server. Dedicated services are the hosting of a single customer website on one or more dedicated servers. With a co-location service, you provide your own server hardware and software and place them at the HSP's physical site. Some of the advantages of HSPs are network geography, backup power, and a dedicated secure facility.
- MSP – Management Service Providers most commonly offer network and application monitoring services for e-businesses.

Have you decided what type of services you'd like to offer? Do you want to be an ISP or would you like to utilize an existing Service Provider to host your service solutions? If you want to be an ISP, will you offer consumer dial-up accounts with unlimited access, for a fixed fee? Will you include free email and Web storage? Would you like to offer business customers services such as Web page design, Domain Name System (DNS) registration, host business sites, or maintain

business applications or equipment? Will you host websites for your customers? Will you offer electronic commerce solutions to your customers? Hint: the money is in the business side.

Compaq recommends deploying these solutions in a multi-tiered, highly available architecture. This is the Compaq DISA (Distributed Internet Server Array) architecture. A Linux DISA solution is shown in Figure 1 and an example Internet Service Provider architecture is displayed in Figure 2.

Figure 1. Compaq DISA Architecture on Linux

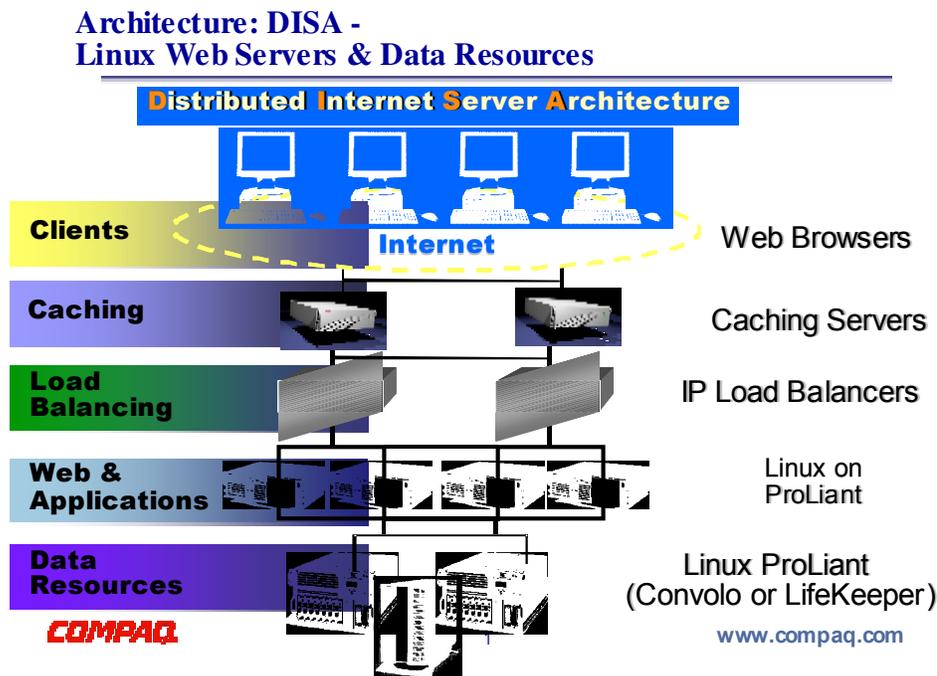
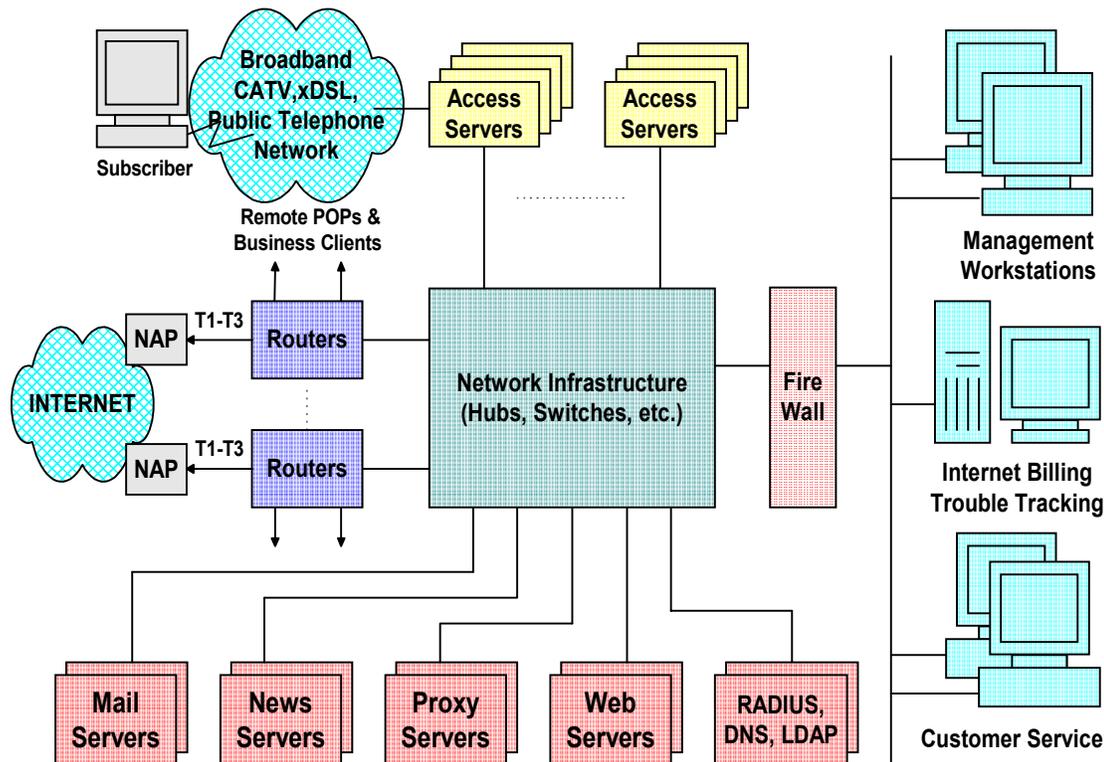


Figure 2. Architectural Overview of an Internet Service Provider



First, identify the applications that will be used to provide the services. The most popular Open Source solutions are listed later in this brief, along with a description of what they do. This is the time to choose a domain name (such as `ispconnect.com`). This name will uniquely identify a Service Provider business on the Internet.

Next, decide on how much bandwidth you need. Perhaps even more importantly, you will need to decide how much bandwidth you can afford for your connection to the Internet. Will you need a fractional T1¹ or E1, a full T1 or E1, multiple T1s or E1s, T3, or E3 lines? Upstream Internet access is available from a variety of sources. It can be offered by backbone providers, national providers, or regional providers. In the preceding diagram, these providers are all listed under the generic name of Network Access Providers (NAP). In general, the higher up the food chain, the more expensive the connectivity and the better the Internet topology. Interpret this as a major factor in a customer's website response times. Decide on the equipment required (servers, disk storage arrays, backup equipment, routers, switches, hubs, dial-up access servers, and modems).

¹ The T-carrier system, introduced by the Bell System in the U.S. in the 1960s, was the first successful system that supported digitized voice transmission. The original transmission rate (1.544 Mbps) in the T-1 line is in common use today in Internet Service Provider (ISP) connections to the Internet. Another level, the T-3 line, providing 44.736 Mbps, is also commonly used by ISPs. Another commonly installed service is a fractional T-1 line, which is the rental of some portion of the 24 channels in a T-1 line, with the other channels going unused.

The T-carrier system is entirely digital, using pulse code modulation and time-division multiplexing. This definition is provided by the <http://whatis.techtarget.com/> site.

E1 (or E-1) is a European digital transmission format devised by the ITU-TS and given the name by the Conference of European Postal and Telecommunication Administration (CEPT). It is the equivalent of the North American T-carrier system format. E2 through E5 are carriers in increasing multiples of the E1 format. The E1 signal format carries data at a rate of 2.048 million bits per second and can carry 32 channels of 64 Kbps each. E1 carries at a somewhat higher data rate than T-1 (which carries 1.544 million bits per second) because, unlike T-1, it does not do bit-robbing and all eight bits per channel are used to code the signal. E1 and T-1 can be interconnected for international use.

Compaq offers qualified and tested equipment in all these areas, either directly or via our solutions partners.

After deciding Internet upstream connectivity, you will need to get a block of IP addresses. You will use these for your own equipment, and potentially offer addresses to business and maybe even consumer customers. You will need to register your Service Provider's domain name (ispconnect.com in the example above). Allotment of an address pool and even registration of a domain name can be handled by your upstream provider. Even better, go directly to the address assignment and the network naming authorities for your country. They are listed below under the IP Addressing, and Domain Name Registration Resources section. In the U.S., it is the Internet Assigned Numbers (IANA) and InterNIC. Before doing this, establish a Domain Name Service (DNS), and have assigned primary and secondary name servers to provide this service.

The Linux Documentation project (<http://www.linuxdoc.org>) is a great source of information for things like setting up networking, DNS, creating Linux routers, and creating Linux firewalls to name a few. Information is available in the form of HOW-TO articles.

Service Providers also need to think about security needs – how to protect systems and critical customer data (such as credit card information and customer lists).

Finally, the Service Provider solution is still not complete until you also take into account the elimination of single points of failure, and the need for redundant equipment.

There is detailed information at numerous external sources that are referenced in this brief, and Compaq recommends these sources when additional software detail is needed. Information about the software and organizations used in this solution is provided below. They are categorized as Linux Information Resources, Internet Protocol Addressing, Domain Name Registration Resources, and General Open Source Software Resources.

3 Linux Information Resources

A large amount of information is available on the World Wide Web. The following websites provide information on Linux, Linux distributors, organizations controlling domain name registration, and IP address assignment:

- <http://www.freshmeat.net/> – Index of Linux and Open Source software.
- <http://www.slashdot.org/> – News/stories source for solutions.
- <http://www.linuxdoc.org/> – Linux Documentation Project. Collection of HOWTO documents and guides.
- <http://www.linux.org/> – Linux Online.
- <http://www.li.org/> – Linux International.
- <http://www.tux.org/> – East coast (U.S.) Linux Users groups.
- <http://www.redhat.com/> – Red Hat Linux distribution.
- <http://www.calderasystems.com/> – Caldera Linux distribution.
- <http://www.suse.com/> – SuSE Linux distribution.
- <http://www.turbolinux.com/> – TurboLinux distribution.
- <http://www.debian.org/> – Debian Linux distribution.

4 IP Addressing and Domain Name Registration Resources

The following websites provide IP addressing, and domain name registration resources:

- <http://www.icann.net/> – The Internet Corporation for Assigned Names and Numbers.
- <http://www.iana.org/> – The Internet Assigned Numbers Authority.
- <http://www.internic.net/> – Assigns network names in top-level domains (for example, COM, NET, ORG) and points to other accredited registrars (accredited by ICANN).
- <http://www.apnic.net/> – Asia-Pacific Network Information Center.
- <http://www.arin.net/> – American Registry for Internet Numbers.
- <http://www.ripe.net/> – Réseaux IP Européens (RIPE).

5 General Open Source Software Resources

These applications are typically available in packaged form with the previously listed Linux distributions.

<http://www.opensource.org/> – Open Source software site.

<http://www.gnu.org/> – The GNU (acronym for GNU's Not UNIX) project.

<http://www.fsf.org/> – The Free Software Foundation (GNU).

<http://www.freshmeat.net/> – Great site for open source and Linux solutions.

<http://www.sourceforge.net/> – Free service for Open Source development.

6 Compaq ActiveAnswers Solution Resources

The Compaq *ActiveAnswers* site contains a growing list of documents that address the planning, deployment, and operations of Linux solutions on Compaq platforms. Go to <http://www.compaq.com/activeanswers> and access the Service Provider Solutions link for complete information.

Some examples of documents contained within the *ActiveAnswers* Linux solutions are:

- *Linux Web Solution with Apache, PHP, MySQL, and ht://Dig*
- *Installation and Configuration Guide for Linux and Apache Web Server on Intel*
- *Preventing Unsolicited Bulk Email on Linux*
- *FrontPage Server Extensions using the Apache Web Server on Linux*
- *Virtual Web Hosting Using the Apache Web Server on Linux*
- *Ensim ServerXchange and Compaq Platforms – a Comprehensive Application Hosting Solution*

7 Open Source Software Commonly Used by ISPs

The following sections contain lists of websites and other information sources that are relevant to Open Source solutions typically used by ISPs in providing their services. Most of the applications discussed here are available in package form with the Linux distributions listed in preceding Section 3 .

The goal of this section is to give an overview of each software component. It also provides a definitive source for additional information on that component. The Open Source solutions are grouped under Domain Name System (DNS) Software, Worldwide Web Server (WWW) Software, Mail Software, News Software, File Transfer Protocol (FTP) Software, Firewall Software, System Management Software, Routing Software, Dial-up Authentication Software, Directory Service Software, Chat Servers, Shared File/Print Services Software, Database Software, Development Tools, and finally pointers to Client Side Application Software.

7.1 Domain Name System Software

This software provides the domain name services for the Internet. DNS is a critical component for all of the other services that an ISP provides such as Web Browsing, E-mail, News, and E-Commerce. DNS is the way that Internet domain names are located and translated into IP addresses. DNS is a distributed database that takes care of mapping a machine name (for example, a fully qualified domain name like mail.ispconnect.com) to or from its Internet IP address (for example, 192.255.16.1). The lists of domain names and IP addresses are distributed throughout the Internet in a hierarchy of authority. The original top-level domains are:

- .com (for commercial organizations, like Compaq, compaq.com)
- .edu (for educational organizations, like Harvard, harvard.edu)
- .gov (for Government organizations, like NASA, nasa.gov)
- .mil (for military organizations, like the U.S. Army, army.mil)
- .net (for networking organizations, like uunet.net, or ispconnect.net)
- .org (for noncommercial organizations, like National Public Radio, npr.org)
- .int (for international organizations, like NATO, nato.int).
- There are also countrywide top-level domains, like .us (for the United States), .jp (for Japan), .de (Germany).
- New domains are being added in 2001: .name, .biz, .info, .pro, .museum, .aero, and .coop.

ISPs typically use the .net domain for their own infrastructure (for example ispconnect.net) and utilize the .com domain for their business operations (for example ispconnect.com). ISPs need to run at least two DNS servers. One will function as a primary (definitive name/address information for the domain). The second will function as a secondary (backup for name/address information), in case the primary is unavailable. These servers also function as the central query location for all Internet domain information. They will resolve information not in their domain space by querying the appropriate DNS root name servers. It is a good idea to run these DNS servers on systems that are not located on the same Local Area Network (LAN).

7.1.1 BIND

Web Address: <http://www.isc.org/>

The Internet Software Consortium is in the final stages of becoming a nonprofit corporation dedicated to production-quality software engineering for key Internet standards. They provide a DNS software package known as BIND (the Berkeley Internet Name Domain). BIND, the most popular software for providing DNS services, is currently running the DNS services for the DNS root servers (definitive DNS information for the top-level domains). *DNS and BIND* by Paul Albitz & Cricket Liu (ISBN: 1-56592-236-0) is an excellent printed source of information.

7.2 Web Server Software

Web servers support the Hypertext Transfer Protocol (HTTP). The list below covers both WWW Servers and WWW Proxy/Cache Servers. Also included are useful publishing and log analysis tools.

7.2.1 Apache

Web Address: <http://www.apache.org/>

The Apache Web Server Project homepage contains versions of the Apache Web Server that you can download and documentation on each version. The October 2000 WWW server site survey by Netcraft (<http://www.netcraft.com/survey/>) found that over 59% of the websites on the Internet are using Apache, thus making it more widely used than all other Web servers combined.

The Apache project is an effort to develop and maintain an Open Source HTTP server for various modern desktop and server operating systems, such as UNIX and Windows NT. The goal of this project is to provide a secure, efficient and extensible server that provides HTTP services in sync with the current HTTP standards.

The Apache Web Server provides some very nice tools supporting virtual Web hosting, server side extensions, as well as bandwidth limiting capabilities for the virtual websites you are hosting. Detailed information on virtual Web hosting can be found in the Compaq [ActiveAnswers](#) document, *Virtual Web Hosting Using the Apache Web Server on Linux*.

The Apache Week site at <http://www.apacheweek.com> has excellent information on Apache configurations, and modules available for the server. Go to this site for a definitive list of Apache modules. You will find features such as blocking access, counters, server-side scripting, limiting bandwidth, and URL rewriting.

Popular Apache modules that are shipped with Apache can be found at:

<http://www.apache.org/docs/mod/index.html>

Popular Apache modules that are not shipped with Apache can be found at:

<http://modules.apache.org>

7.2.2 PHP

Web Address: <http://www.php.net>

PHP (Personal Home Pages) is the most popular method for server-side scripting. Our *Linux Web Solution with Apache, PHP, MySQL, and ht://Dig Technical Guide* details how to build server-side scripting into the Apache Web Server. See the solution document at the Compaq [ActiveAnswers](#) site.

7.2.3 FrontPage 2000 Server Extensions

Web Addresses: <http://www.microsoft.com/frontpage>

<http://www.rtr.com> (location for Linux kits from Ready-to-Run Software)

The FrontPage 2000 Server Extensions Resource Kit facilitates document publishing on a Linux Web server from remote PC clients running Microsoft FrontPage 2000. These extensions are not open source, but add nice publishing capabilities to your open source Web server solution.

7.3 Web Log Analysis

This section describes Web log analysis tools.

7.3.1 Analog

Web Address: <http://www.statslab.cam.ac.uk/~sret1/analog/>

Analog is a tool for analyzing log files. This site provides a description of Analog, sample reports that can be generated by analog, and access to sites for downloading the latest version of the software.

7.3.2 Webalizer

Web Address: <http://www.mrunix.net/webalizer>

Webalizer provides highly detailed, easily configurable usage reports in HTML format, for viewing with a standard Web browser.

7.4 Caching Proxy Server

7.4.1 Squid Proxy Server

Web Address: <http://squid.nlanr.net/Squid/>

Squid is a high-performance proxy/caching server. It supports FTP, Gopher, and HTTP data objects. A proxy server can save you on upstream bandwidth costs. It can also be utilized to implement HTTP screening capabilities for blocking “objectionable” material from customers, or a subset of your customer base. Squid tends to be faster than other proxy/caching servers because it maintains linked objects in RAM, caches DNS lookups, and supports non-blocking I/O, negative caching of objects, and DNS lookups. Squid documentation and FAQs are available from the Squid homepage.

7.5 High Availability Clustering

These solutions range from simple heart beat monitoring and Internet Protocol (IP), failover (Web failover where content is easily replicated), to shared storage solutions (appropriate for services like mail, database, and e-commerce solutions). Components of a high availability solution include membership services, communication services, cluster management, resource monitoring, storage sharing/replication, and resource fencing. There are many things to consider when designing a high availability solution. They range from network infrastructure redundancy, system redundancy features (such as redundant power, multiple network interface cards (NICs), and fans), disk file system high availability solutions such as software or hardware RAID, and clustering solutions.

File system technologies are also important to a cluster solution. Journal file systems provide for a quick takeover of shared file system resources. Cluster file systems allow disks to be shared between members of a cluster. Reiserfs, Ext3, and JFS are the top choices for journal file system solutions. GFS is both a cluster file system and a journal file system.

7.5.1 Kimberlite

Web Address: <http://oss.missioncriticallinux.com/kimberlite>

Kimberlite is an Open Source shared storage, two-node high availability cluster solution. The two nodes can be connected to a shared SCSI or Fibre Channel storage subsystem.

7.5.2 Piranha

Web Address: <http://ha.redhat.com/software.html>

Piranha is an Open Source solution for Linux high availability, IP service monitoring and fail over, and load balancing. It supports two-node failover and multimode load-balancing clusters.

7.5.3 Linux-HA

Web Address: <http://linux-ha.org/>

Linux-HA provides a high availability clustering solution for Linux. Linux-HA provides the heartbeat code (that is, it monitors nodes and takes over the IP address of the failed node).

7.6 Application Server

Server side scripting and database connectivity is commonly accomplished via Apache/PHP3, Perl, or Python. An application server is a general-purpose software server supporting the interaction between a client browser and the actual application. An application server allows an n-tier architecture where the server connects to a database in addition to serving many applications or services to clients.

Features that should be provided include:

- State management for HTTP transactions
- Pooling of connections to databases
- Scaling to handle thousands of simultaneous users by use of multithreading
- Load balancing across multiple machines

- High availability
- Management of transactions across multiple distributed systems
- Secure communications

7.6.1 Enhydra

Web Address: <http://www.enhydra.org>

Enhydra is an Open Source Java/XML application server run-time and development environment. Enhydra offers commercial professional services support via Lutris Technologies. The major components of the Enhydra based solution include:

- Web Servers – Apache, Netscape, or Enhydra itself for HTTP
- Application framework – Servlet run-time environment of common services (session, presentation, and database connectivity for supporting n-tier applications).
- Enhydra Multiserver – Servlet running environment with embedded Servlet management, monitoring, and debugging.
- Enhydra XMLC – XML Compiler designed to support separation between designer/developer co-development (that is, separation between presentation and business layers.)
- Enhydra JDDI – embedded Java for dynamic HTML
- Enhydra DODS – graphical tool for object to relational database mapping
- Enhydra application wizard – quick and automatic way to build a project tree
- Enhydra Director – provides clustering and failover of multiple application servers

7.6.2 Zope

Web Address: <http://www.zope.org>

Zope is an Open Source application server based on Python, a powerful object-oriented programming language. It is popular for portal site deployments.

7.6.3 Zend

Web Address: <http://www.zend.org>

The Zend engine is an Open Source scripting engine that drives PHP. Zend brings performance, reliability, and an easy-to-use interface to PHP.

7.7 Mail Software

This software supports the SMTP, POP3, and IMAP4² protocols. Also listed in this section are tools for managing mailing lists, setting up mail filtering, and setting up the process for subscriber vacation messages. Two of the most popular Web-based email solutions are also listed (IMP and TWIG). Included at the end is a resource for understanding the problems of SPAM mail, and how to deal with them.

7.7.1 Sendmail

Web Address: <http://www.sendmail.org/>

Sendmail is the BSD Mail Transport Agent supporting mail transport by means of TCP/IP using Simple Mail Transfer Protocol (SMTP). Sendmail V8 is the most popular mail transport agent being used on the Internet today. Sendmail has support for mail masquerading (the ability to make all outgoing mail appear to come from one mail domain), and for virtual mail domains (supporting many mail domains on a single system). It also has good anti-spam features for rejecting inbound mail from known spammers (senders of unsolicited mail), validating domains, keywords and pattern recognition (“make money fast”), and limiting the number of recipients in outgoing or incoming mail.

Note: Sendmail is heavily dependent on DNS. It is in DNS that you define mail exchange records. These records define the system or systems that handle mail for a specific domain.

7.7.2 Qmail

Web Address: <http://www.qmail.org/>

Qmail is considered by many as a replacement for sendmail. It also supports SMTP, and provides performance features not found in sendmail.

7.7.3 POP3

Web Addresses: <ftp://ftp.qualcomm.com/eudora/servers/unix/popper/>
<http://www.eudora.com/free/servers.html>

POP (Post Office Protocol) allows single-user hosts to read electronic mail from a server. The foregoing sites provide information on the QUALCOMM POP3 server (Qpopper).

² SMTP (Simple Mail Transfer Protocol) is a TCP/IP protocol used in sending and receiving e-mail. However, since it is limited in its ability to queue messages at the receiving end, it is usually used with one of two other protocols, POP3 or IMAP, that let the user save messages in a server mailbox and download them periodically from the server. In other words, users typically use a program that uses SMTP for sending e-mail and either POP3 or IMAP for receiving messages that have been received for them at their local server. Most mail programs such as Eudora let you specify both an SMTP server and a POP server. On UNIX-based systems, sendmail is the most widely-used SMTP server for e-mail. SMTP usually is implemented to operate over TCP port 25. The details of SMTP are in RFC 821 of the Internet Engineering Task Force (IETF). This definition is provided by the <http://whatis.techtarget.com> site.

7.7.4 IMAP

Web Addresses: <http://www.washington.edu/imap/> – University of Washington IMAP Information Center

<http://asg.web.cmu.edu/cyrus/> – Cyrus IMAP server by Carnegie Mellon University

The Internet Message Access Protocol (IMAP) allows a client to access and manipulate electronic mail messages on a server. The current version 4 of the protocol is described in RFC 1730.

7.7.5 IMP

Web Address: <http://www.horde.org/imp/>

The Internet Messaging Program (IMP) is written in the PHP scripting language. IMP provides webmail access to IMAP and POP3 accounts and contact management. IMP uses PHPLib for session management, handles MIME attachments, and works with both PHP3 and PHP4.

7.7.6 TWIG

Web Address: <http://twig.screwdriver.net/>

TWIG is a Web-based IMAP client written in PHP3. Its features include IMAP e-mail, contact management, scheduling, Usenet newsgroups, to-do manager, and bookmark management. Any piece of information can be shared with other users on the system using the built-in groupware functionality of TWIG.

7.7.7 Procmal

Web Address: <http://www.procmal.org>

Procmal is the mail-processing utility language written by Stephen van den Berg of Germany. Using procmal, you can filter hundreds or thousands of incoming mail messages per day according to a predefined set of rules. Since the procmal language understands details about most UNIX mail transport and delivery agents, it is the tool of choice for writing custom mail filtering scripts. The procmal filtering engine can be invoked by sendmail or by a user's .forward file. Procmal is also the basis of high performance mailing list software like Smartlist.

See the manual pages, procmal(1), procmalrc(5), and procmalrc(5), for more information on using procmal. Extensive examples are provided in procmalex(5).

The site for procmal tips at (<http://www.uwasa.fi/~ts/info/proctips.html>) contains pointers to documentation, examples, and FAQs, and a link to the FTP site where you can obtain the procmal kit.

7.7.8 Fetchmail

Web Address: <http://www.tuxedo.org/~esr/fetchmail/>

Fetchmail is a full-featured, robust, well-documented, remote mail retrieval and forwarding utility.

7.7.9 Majordomo

Web Address: <http://www.greatcircle.com/majordomo/>

Majordomo is a set of programs that automate operation and maintenance of Internet mailing lists.

7.7.10 LISTSERV

Web Address: <http://www.lsoft.com/>

LISTSERV is an e-mail list management software. A lightweight freeware version is available.

7.7.11 MAPS

Web Address: <http://maps.vix.com/rbl/>

The Mail Abuse Protection System (MAPS) Realtime Blackhole List (RBL) is a system for creating intentional network outages for the purpose of limiting the transmission of known unsolicited bulk e-mail (also known as spam). This site also explains the spam problem, how to test your mail system for vulnerability to unsolicited bulk e-mail, and how to prevent your mail system from being used for unsolicited bulk emailing. The consequences of spam and the methods to prevent it are addressed in the [ActiveAnswers](#) paper, *Preventing Unsolicited Bulk E-mail on Linux Systems*.

7.8 News Software

This software is used for handling Usenet. Usenet is a worldwide, distributed discussion system that consists of newsgroups classified hierarchically by subject. Articles or messages are posted to these newsgroups, and the articles are then fed to other interconnected computer systems via a wide variety of networks. The architecture is shown in Figure 3, which follows.

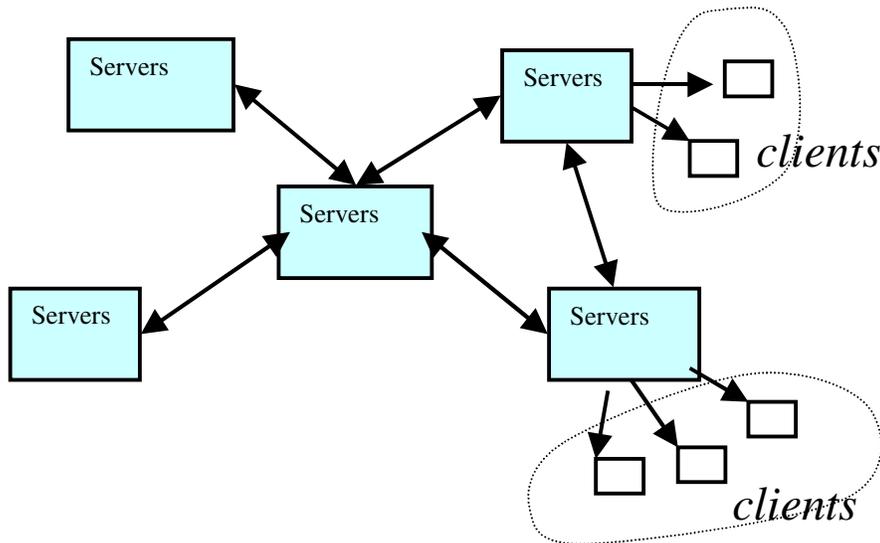
7.8.1 INN

Web Address: <http://www.isc.org/inn.html>

The Internet Software Consortium homepage for InterNetNews (INN) provides release notes and access to the latest kit. For more information on news feeds and the InterNetNews server, see the Usenet and InterNetNews document by Rich Salz, and the set of FAQs (Frequently Asked Questions). Both of these are available at the above URL.

For information on how to enable authentication of Usenet group changes using PGPverify, visit the following URL:

<ftp://ftp.uu.net/networking/news/misc/pgpcontrol/README.html>

Figure 3. Architecture of Newsgroups

7.8.2 Diablo

Web Address: <http://www.openusenet.org/diablo/>

Diablo is a backbone news transit system, designed to replace INN (InterNetNews Daemon) on backbone machines. The transit part of Diablo is well established and should work flawlessly. The transit server portion of Diablo cannot be run on a machine that needs to accept NNTP (Network News Transport Protocol) post commands or NNTP commands related to news reading. You cannot point a news reader client to a host running the transit portion of Diablo.

7.8.3 NNTPCache

Web Address: <http://www.nntpcache.org>

NNTPCache efficiently executes on the local host masquerading as a network news reading server. In fact, what it does is pass certain NNTP commands through to real (remote and possibly local) news servers based on various pattern-matching rules. NNTPCache then takes the output from those servers. That output is then cached and indexed. The next time information is asked for, or can be logically inferred from the previously collated information, it is sent directly from the cache, without consulting the remote servers.

7.9 File Transfer Protocol Software

Web Address: <http://www.wu-ftp.org/> – Washington University, St. Louis, Missouri.

<http://www.proftpd.org/> – ProFTPD is a proven, high-performance, scalable FTP server.

File Transfer Protocol (FTP) is a client/server protocol that allows a user on one computer to transfer files to and from another computer over a TCP/IP network. An anonymous FTP account on your system allows any remote user to log in to your system using an FTP or anonymous user name. Once logged in, that user has access to a special directory hierarchy that contains publicly accessible files and permission to transfer those files to another system using FTP.

7.10 Firewall and System Access Control Software

This section covers IP Masquerading, Packet Filtering, Proxy Firewalls, and additional system access control software.

Masquerading is a useful facility for translating network traffic from one network to another by hiding one or more addresses behind a single visible address. It accomplishes this by using a form of network address translation (NAT).

There are two basic types of firewalls, IP or Packet Filtering Firewalls, and Proxy Firewalls. Packet filtering firewalls block all but selected network traffic. All traffic through a TCP/IP network is sent in the form of packets. The start of each packet says where it's going, where it came from, the type of the packet, and other administrative details. This start of the packet is called the header. The rest of the packet contains the actual data being transmitted and is called the body.

A packet filter is a piece of software that looks at the header of packets as they pass through, and decides the fate of the entire packet. The filter might decide to deny the packet (that is, discard the packet as if it had never received it), accept the packet (that is, let the packet go through), or reject the packet (deny it, but tell the source of the packet that it has done so). Under Linux, packet filtering is built into the kernel.

Proxy firewalls make the network connections on your behalf. They are designed to control the flow of packets based on the source, destination, port and packet type information contained in each packet.

7.10.1 Linux with IP Masquerading

Web Addresses: <http://www.linux.org> – Linux Organization

<http://ipmasq.cjb.net> – Linux IP Masquerade Resource

IP masquerading can be implemented with Linux, and standard networking software. IP masquerading is a form of network address translation that many routers already support. The idea behind this implementation is that people running Linux can install the IP masquerading features and get the features of the high priced routers and NAT boxes without paying the high prices.

IP masquerading lets you use a single Internet-connected computer running Linux with a real IP address as a gateway for non-connected machines with unregistered (not routed out to the Internet, like the class A network 10) IP addresses. The Linux box with a real address handles mapping packets from your intranet out to the Internet, and when responses come back, it maps them back to your intranet.

7.10.2 IP Filter

Web Address: <http://coombs.anu.edu.au/~avalon/ip-filter.html>

IP Filter is a TCP/IP packet filter, suitable for use in a firewall environment. It can either be used as a loadable kernel module or incorporated into your Linux kernel. IP Filter can explicitly deny or permit any packet from passing through, distinguish between various interfaces, filter by IP networks or hosts, selectively filter any IP protocol, and act as a Network Address Translator (NAT).

If you are using a pre-version 2.2 Linux kernel, you need to use IP Filter. However, if you are using a version 2.2 or later Linux kernel, you should use Ipchains.

7.10.3 Ipchains

Web Address: <http://netfilter.filewatcher.org/ipchains/>

Linux ipchains is a rewrite of the IP Filter code. If you are using a V2.2 or later Linux kernel, you should use ipchains.

7.10.4 Iptables

Web Address: <http://netfilter.kernelnotes.org/>

Iptables is built on top of netfilter (the new packet alteration framework for Linux 2.4). It is an enhancement on ipchains, and is used to control packet filtering, and Network Address Translation (masquerading, portforwarding, and transparent proxying).

7.10.5 TCP Wrapper

Web Address: <ftp://ftp.porcupine.org/pub/security/index.html>

TCP Wrapper intercepts an incoming network connection and verifies that the connection is allowed before passing the connection to the network daemon. TCP Wrapper is configured through the /etc/hosts.allow file. The FTP archive listed above is for the Mathematics and Computing Science Department at Eindhoven University of Technology (the Netherlands) and contains TCP Wrapper kits.

7.11 System Security Sites

This section contains information about websites concerned with system security.

7.11.1 CERT

Web Address: <http://www.cert.org/>

The Computer Emergency Response Team (CERT) is a clearinghouse for security-related events that occur in the Internet community. If you are an administrator, subscribe to the CERT mailing list and frequently check the CERT advisories.

According to its charter, CERT works with the Internet community to facilitate the community's response to security events involving hosts, takes proactive steps to improve the community's awareness of security issues, and conducts research aimed at improving the security of existing systems. CERT services include a 24-hour hotline for responding to security incidents, product vulnerability assistance, and technical documentation and tutorials.

7.11.2 CIAC

Web Address: <http://www.ciac.org/ciac/>

The Computer Incident Advisory Capability (CIAC) site is maintained by the U.S. Department of Energy. Their website offers computer security information, as well as workshops, consulting, and security incident-handling information.

7.11.3 CSRC

Web Address: <http://csrc.ncsl.nist.gov/>

The Computer Security Resource Center (CSRC) is a U.S. Government archive on security information and contacts maintained by the National Institute of Standards and Technology (NIST).

7.11.4 BugTraq

Web Address: <http://www.securityfocus.com>

BugTraq is a full-disclosure UNIX security mailing list. This is a detailed discussion of UNIX security holes – what they are, how to exploit, and how to fix them.

7.12 System Management Software

This section contains information about system management resources.

7.12.1 Virtual Network Computing (VNC)

Web Address: <http://www.uk.research.att.com/vnc>

Virtual Network Computing (AT&T labs Cambridge) develops remote access software that allows you to bring up a remote display of the system that is being managed. VNC is a client/server solution. It is composed of vncserver, vncviewer, and a Java viewer for browser access. Unlike X based tools, no state is stored in the client side viewer.

Linux/UNIX/NT/W95/W98/Macintosh systems can be remotely managed with this software.

7.13 Remote Desktop Protocol Client

This section contains information about remote desktop management.

7.13.1 Rdesktop

Web Address: <http://www.rdesktop.org>

Rdesktop is an open source client for Windows NT Terminal Server, capable of speaking its Remote Desktop Protocol (RDP) in order to present the user's NT desktop.

7.14 Routing Software

This section contains information about routing software.

Web Address: <http://www.gated.org>

Merit GateD Consortium offers GateD, the most popular software for routing. You can use your Linux box as a router. With the Linux 2.2 kernel, you can optimize the system to be used as a router. There is also support for routing with Linux for T1-interface network cards.

7.15 Directory Service Software

Directory Services provide a way to centrally administer subscriber information in a highly scalable manner. The most common way to do this is with a directory server that supports the LDAP (Lightweight Directory Access Protocol). The Lightweight Directory Access Protocol is an Internet standard directory service protocol that runs over TCP/IP. It can be used to provide a standalone directory service or to provide lightweight access to the X.500 directory. You can use these servers to centralize authentication information, message store locations, vacation messages, and access privileges. Perl integration with LDAP is available through the PerlLDAP libraries.

7.15.1 OpenLDAP

Web Address: <http://www.openldap.org/>

The OpenLDAP Project is a collaborative effort to provide a robust, commercial-grade, fully-featured, and open source suite of LDAP applications and development tools. The project is managed by a worldwide community of volunteers who use the Internet to communicate, plan, and develop the OpenLDAP suite and its related documentation.

7.15.2 LDAP

Web Address: <http://www.umich.edu/~dirsvcs/ldap/>

One of the most popular LDAP servers is provided by the University of Michigan. Documentation, patches, kits and archives for the University of Michigan LDAP server are available at this website.

7.16 Chat Software

This section describes chat software.

7.16.1 IRC (Internet Relay Chat)

Web Address: <http://www.irchelp.org/>

Internet Relay Chat (IRC) allows users to communicate with each other in real time across a network of Internet servers. IRC is an Internet-based system that allows people from all over the world to have live real-time textual "conversations." It is based around a network of interconnected servers for which the user runs an IRC client program to connect. Each user on an IRC network must have a unique nickname, and can join one or more channels. They can also send and receive private messages and even transfer files. The help archive URL above provides general information and pointers to many FAQs, clients, and servers.

7.17 Shared File and Print Services

This section describes file and print servers.

7.17.1 Samba File and Print Server

Web Addresses: <http://www.samba.org/>

The Samba File and Print Server for Windows provides file and print services to SMB clients, such as Windows for Workgroups, Windows NT, or LanManager. It also provides Netbios name serving and browsing support. At the Samba site, you can obtain documentation on the Samba server, report problems and download the latest software. Visit the comp.protocols.smb news group for information on the Samba server protocol.

7.18 Development Tools

This section describes development tools commonly used by ISPs.

7.18.1 Perl

Web Address: <http://www.perl.org>

The Practical Extraction and Report Language (Perl) is an interpreted language distributed over Usenet. See this website for more information on Perl, including FAQs, documentation, newsgroups and software updates. Perl is a favorite tool of system administrators, and is often referred to as the “duct tape of the Internet”.

7.18.2 CPAN

Web Address: <http://www.cpan.org>

Comprehensive Perl Archive Network (CPAN) contains resources for using Perl as building blocks, and integration tools for your Web, mail, news, LDAP, and other solutions.

7.18.3 Python

Web Address: <http://www.python.org>

Python is an interpreted, interactive, object-oriented programming language. It has modules, classes, exceptions, very high-level dynamic data types, and dynamic typing. For comparisons to other languages such as Tcl, Perl, Scheme, or Java, refer to <http://www.python.org/doc/Comparisons.html>.

7.18.4 GNU Development Tools

Web Address: <http://www.gnu.org>

The GNU site lists a variety of development software and tools at its site above. It is at this location you can find the GNU C and C++ compilers, version control software, and build utilities. If you are porting a software package, and need to build the source, there is a good chance that it will need to be built using the GNU tools.

7.18.5 Tcl

Web Addresses: <http://sourceforge.net/projects/tcl>

Tool Command Language (Tcl) is a string-processing language for issuing commands to interactive programs. The website at Scriptics Corporation provides information on topics related to Tcl. Their FTP site contains a collection of Tcl kits that are available for downloading.

7.18.6 expect

Web Address: <http://expect.nist.gov/>

Expect is a tool for automating and testing interactive applications, such as telnet, FTP, passwd, fsock, rlogin, and tip, to name a few. *Exploring Expect: A Tcl-Based Toolkit for Automating Interactive Applications* (ISBN 1-56592-090-2), written by Don Libes and published by O'Reilly & Associates, is an excellent source of information. The expect homepage provides access to FAQs, examples, contributed scripts and software.

7.18.7 TclTk

Web Address: <http://www.tcltk.com>

TclTk is a programming toolkit developed by John Ousterhout at the University of California, Berkeley. The TclTk toolkit is similar to other GUI toolkits (such as Xlib, Xview, and Motif). Unlike those toolkits, TclTk does not require that you use C or C++ in order to manipulate the widgets, and you can build useful applications very rapidly after you gain some expertise.

The TclTk website at Scriptics Corporation provides pointers to TclTk kits. The Tcl WWW Info page at SCO provides FAQs, introductory and reference material, and sample applications.

7.18.8 TclX

Web Address: <http://www.neosoft.com/tclx/>

TclX (Extended Tcl), a superset of standard Tcl, was created by Karl Lehenbauer (karl@neosoft.com) and Mark Diekhans (markd@grizzly.com). The kit can be freely distributed for any use without license or fee. Available over the Internet since 1989, TclX adds capabilities to Tcl and is the source of many of the capabilities of the baseline Tcl release, including arrays, files, sockets, file events, and date and time handling. TclX has three basic functional areas: a set of new commands, a Tcl shell (UNIX shell-style command line and interactive environment), and a user-extensible library of useful Tcl procedures.

The TclX homepage provides pointers to documentation and contains a link to download the source code.

7.19 Client Application Software

The following sections describe client application software.

7.19.1 Applications

Web Address: <http://www.tucows.com/>

The Ultimate Collection of Winsock Software website contains access to numerous Internet applications for personal computer users.

7.19.2 Lynx

Web Address: <ftp://ftp2.cc.ukans.edu/pub/lynx/>

Lynx is a fully featured Web browser for users connected to a system via cursor-addressable, character-cell terminals or terminal emulators. Lynx is a product of the Distributed Computing Group within Academic Computing Services of the University of Kansas.

7.19.3 Pine and Pico

Web Address: <http://www.washington.edu/pine/>

Pine is an IMAP e-mail client for terminals or terminal emulators. Pine is easy to learn; command choices are presented at the bottom of each screen. Pico is the text editor that Pine uses, and is available as a separate, standalone program. This website at the University of Washington provides complete documentation, FAQs, and access to the software.

7.19.4 TIN

Web Address: <http://www.tin.org>

TIN is a full-screen news reader for terminals or terminal emulators on UNIX systems. The TIN home page allows you to download the latest version, access TIN mailing lists, and submit bug reports.

7.20 Database Software

The following sections describe database software.

7.20.1 MySQL

Web Address: <http://www.mysql.com>

Multi-user, multi-threaded SQL database server. An excellent tool for centrally managing your solutions. There are also nice modules for the Apache Web Server that allows you to utilize your SQL database. MySQL is very popular for Web solutions with backend database connectivity. The *Apache Linux Web solution with Apache, PHP, MySQL, and ht://Dig Technical Guide* in the Compaq [ActiveAnswers](#) solution website is a great resource.

7.20.2 PostgreSQL

Web Address: <http://www.postgresql.org>

PostgreSQL is an object-relational database management system (DBMS). It supports most SQL constructs, including subselects, transactions, and user-defined types and functions.

8 Summary

This is just the tip of the Open Source Iceberg. Every tool you need is either available, or can be constructed from an existing Open Source tool. Open Source gives you, the ISP, control over your solution.

9 Glossary

Anonymous FTP

Using the Internet's File Transfer Protocol (FTP), anonymous FTP is a method for giving users access to files so that they do not need to identify themselves to the server. Using an FTP program or the FTP command interface, the user enters "anonymous" as a user ID. Usually, the password is defaulted or furnished by the FTP server. Anonymous FTP is a common way to get access to a server to view or download files that are publicly available.

.forward file

On a UNIX system, this file (unique to each user and located in their home directory) is used to reroute mail or perform some processing on that mail message.

IMAP

Internet Message Access Protocol is a protocol for retrieving e-mail messages. The latest version, IMAP4, is similar to POP3 but supports some additional features. You can store and retrieve messages from folders that you maintain on the mail server. With IMAP4, you can search through your e-mail messages for keywords while the messages are still on the mail server. You can then choose which messages to download to your machine.

INN

The InterNetNews package is a complete Usenet news system. It includes innd, an NNTP server, and nnrpd, a news reading server. INN separates hosts that feed you news from those that have users reading news.

Innd

The main process that implements the network news transport protocol (NNTP) for InterNetNews. The 'd' notation at the end of 'inn' denotes daemon, which means it is a standalone process, that is, a process that runs in the background and performs a specified operation at predefined times or in response to certain events.

InterNIC

This is a global registrar for the top-level domains of .com, .net, and .org.

NNTP

Network News Transport Protocol, a protocol used for Usenet News. It consists of a set of client-server and server-server commands that are utilized to send or receive Usenet news messages across the Internet. For example, the NNTP POST command is used by client programs wishing to post a Usenet news article to a particular discussion group.

POP

Post Office Protocol is a protocol used to retrieve e-mail from a mail server. Most e-mail applications (sometimes called an e-mail client) use the POP protocol.

Procmal

An autonomous mail processor. It is usually invoked via the .forward file mechanism, or from within the mail server.

procmailrc

Procmail resource file, which contains a mixture of environment variable assignments, and recipes for processing mail.

Procmailsc

Procmail weighted scoring technique. In addition to the traditional true or false that you can specify in a recipe, you can use a weighted scoring technique to decide if a certain recipe matches or not.

Procmailex

Example procmailrc files.

RFC

The Requests for Comments (RFCs) form a series of notes, started in 1969, about the Internet (originally the ARPANET). The notes discuss many aspects of computer communication, focusing on networking protocols, procedures, programs, and concepts but also including meeting notes, opinion, and sometimes humor. Some RFCs become standards.

More information on RFC can be found at this site: <http://www.rfc-editor.org/>

Root Name Servers

The root name servers know the location of the name servers authoritative for all the top-level domains. Given a query about any domain name, the root name servers can provide the names and addresses of the name servers authoritative for the top-level domain of that domain name.

Server side extensions

In the context of the World Wide Web client-server model, server side extensions are extensions to the Web server that allow for additional processing or functionality on the server side.

SMTP

Simple Mail Transfer Protocol, a protocol for sending e-mail messages between servers. Most e-mail systems that send mail over the Internet use SMTP to send messages from one server to another.

Spam

A term used to describe unsolicited junk e-mail. They typically consist of commercial advertising, get-rich-quick schemes, and the sale of dubious products.

Unsolicited Bulk E-mail (UBE)

Unsolicited Bulk E-mail is the targeting of large numbers of Internet mail users with unsolicited direct mail messages. Commonly referred to as SPAM.

X.500

An ISO and ITU standard that defines how global directories should be structured. X.500 directories are hierarchical with different levels for each category of information, such as country, state, and city.