



EUFIRELAB
EVR1-CT-2002-40028

D-06-01

<http://eufirelab.org>



EUFIRELAB:

**Euro-Mediterranean Wildland Fire Laboratory,
a “wall-less” Laboratory
for Wildland Fire Sciences and Technologies
in the Euro-Mediterranean Region**

Deliverable D-06-01

Networking Structure: Description of its Architecture

Franck GUARNIERI, Denis OVERAL

February 2004

CONTENT LIST

Summary.....	1
List of associated documents	1
1 Introduction.....	2
2 Linux.....	2
2.1.1 A general overview.....	2
2.1.2 Main features of Linux.....	3
3 Apache HTTP server.....	4
3.1 Apache HTTP server, a general overview	4
3.2 Web application development with Apache	4
4 MySQL.....	6
4.1 MySQL, a general overview	6
4.2 MySQL features	6
5 PHP.....	7
5.1 PHP, a general overview.....	7
5.2 Main features.....	7
6 References	8
7 Acronyms.....	8

SUMMARY

This document presents the major specifications and the structure of the EUFIRELAB web site.

It describes the technologies developed in the frame of the project, based on articulations between Linux, Apache HTTP server, MySQL and PHP.

For each of these technologies, the document explains by a general overview the specificity of each of these technologies and the main features used, implemented under a working web site <http://www.eufirelab.org>.

LIST OF ASSOCIATED DOCUMENTS

None.

1 INTRODUCTION

The EUFIRELAB applications rely on a computer architecture centred around Internet technologies.

The model that is used is often referred to as a *LAMP* architecture, where LAMP is an acronym for a *Linux-Apache-MySQL-PHP* system.

Linux is the name of the Operating System (OS), which is the basic layer to make a computer work.

It is the part of the architecture in charge of driving microprocessors, network interfaces, memory, and such

Apache is an HTTP server.

It is a program that handles HTTP requests, and creates answers streams that creates a web dialog with a client browser.

MySQL is a relational database management system.

It is a program that organises and accesses data in a very effective way.

PHP is a scripting language.

Among its many features, it allows easy linking of a HTTP server and a Database management system.

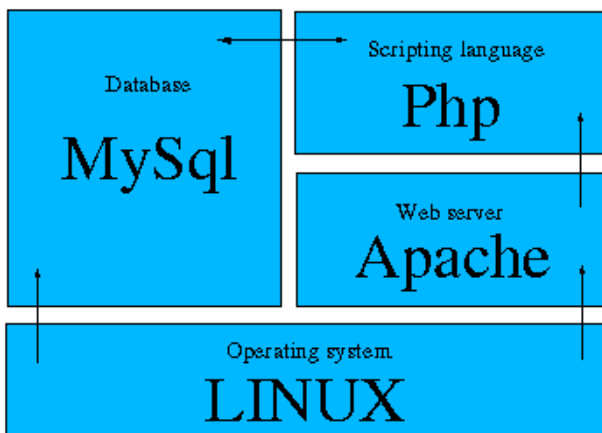


Figure 1-1: A LAMP architecture

In the following chapters, we'll see an introduction to each of these concepts used in the EUFIRELAB architecture.

2 LINUX

2.1.1 A general overview

Linux is an operating system that was initially created as a hobby by a young student, Linus Torvalds, at the University of Helsinki in Finland.

Linus had an interest in Minix, a small UNIX system, and decided to develop a system that exceeded the Minix standards.

He began his work in 1991 when he released version 0.02 and worked steadily until 1994 when version 1.0 of the Linux Kernel was released.

The current full-featured version is 2.4 (released January 2001) and development continues.

Linux is developed under the GNU General Public License and its source code is freely available to everyone.

This however, doesn't mean that Linux and its assorted distributions are free -- companies and developers may charge money for it as long as the source code remains available.

Linux may be used for a wide variety of purposes including networking, software development, and as an end-user platform.

Linux is often considered an excellent, low-cost alternative to other more expensive operating systems.

Due to the intrinsic nature of Linux's functionality and availability, it has become quite popular worldwide and a vast number of software programmers have taken Linux's source code and adapted it to meet their individual needs.

At this time, there are dozens of ongoing projects for porting Linux to various hardware configurations and purposes.

About the place of Linux today, one can say that the world of Operating Systems nowadays is mainly two-fold.

Microsoft Windows (R) reigns over the world of Personal Computers, even if seriously challenged and losing ground to Linux.

In the field of servers, Unix is the OS of choice.

Amongst the many Unix flavors that exist, Linux represents the biggest share, making it "the next generation of Operating Systems".

The most serious estimates amount to 18 million computers nowadays running under the Linux Operating System worldwide, and many of them are internet servers.

In tune with the power-of-choice tradition of Linux, many companies and communities now offer Linux along with lots of applications.

Though the OS is the same, the bundled software do vary from one distribution to another.

Red Hat, SuSE, Mandrake, and many other firms sell Linux this way, and their CDs are called distributions.

Usually the distributors charge a nominal fee for the media and for technical support in the future.

Debian is a high-quality non-commercial distribution of Linux, funded and run by about 900 volunteers.

When talking about Linux, the term "Open Source" operating system is often found.

Any software, at heart, contains binary instructions -- a continuous stream of 0's and 1's -- to the processor so that it works the way it is supposed to.

Encoding software directly in binary is no doubt a tedious, if not impossible, proposition.

So a programmer uses an English-like higher-level language to write software.

This program text is the source code; it is compiled or assembled to obtain the final executable.

Microsoft give away only the executable; Linux also gives the source code -- hence it is open-source.

2.1.2 Main features of Linux

All applications in Linux are pre-emptively multitasked.

The OS handles all scheduling of processes (and kernel threads).

No application can hog the resources unless the system administrator specifically defines it as high-priority.

This leads to smoother performance and better load-balancing.

All applications in Linux run in their own private memory space.

This means that a poorly-written application cannot tamper with the memory of another application (or the kernel).

If an application does try to access memory it doesn't own, it is immediately halted by the operating system, without disturbing any other process on the system.

This makes Linux very stable against virus attacks.

UNIX was designed with the notion that multiple people would be sharing use of the system at the same time.

Several people can log into a Linux machine and each of them can run whatever programs he likes.

UNIX applications are written from scratch with the idea that multiple people will be using them, and each may have entirely different preferences.

Many people can use the same application at the same time according to their liking!

Because of the fundamental design of UNIX, every application can run on one machine and display its interface on another.

This is extremely useful for remote administration.

Linux is outstanding in the area of memory management.

It will use every scrap of memory in a system to its full potential.

The Linux kernel occupies just 2 MB, whereas Windows NT takes 16 MB!

Linux uses a copy-on-write scheme.

If two or more programs are using the same block of memory, only one copy is actually in RAM, and all the programs read the same block.

If one program writes to that block, then a copy is made for just that program.

All other programs still share the same memory.

Demand-loading is very useful as well. Linux only loads into RAM the portions of a program that are actually being used, which reduces overall RAM requirements significantly.

At the same time, when swapping is necessary, only portions of programs are swapped out to disc, not entire processes.

This helps to greatly enhance multiprocessing performance.

Finally, any RAM not being used by the kernel or applications is automatically used as a disc cache.

This speeds access to the disc so long as there is unused memory.

On the other hand, memory management is poorest in Microsoft Windows!

Linux normally uses its own high-performance file system, which uses disc space much more efficiently, optimises for speed on reading and writing, and automatically prevents fragmentation.

The Linux file-system literally does not need a defragmenter, though one is available.

It also sees when programs make errors writing to the disc and automatically prevents them, so there is usually no need to run a disc checker unless you notice a problem

Linux enforces a strict separation between the kernel and other applications.

Most services like mail, file and print serving, web serving and so on are applications, and can usually be changed dynamically.

At worst, a specific application may need to be restarted, and not the whole system.

Reboots are only for kernel updates and hardware changes.

UNIX systems have uptimes in terms of years!

TCP/IP networking and the Internet was originally developed on UNIX systems, and most of the high-power networking in the world is done on UNIX.

About 75% of the web servers on the Net run a version of UNIX.

In fact, Linux has the largest market share for the entire Internet, running:

- 25.7% of the news servers,
- 26.9% of the web servers, and
- 33.7% of the FTP servers in the world.

Apart from extremely fast and reliable networking, dozens of major and minor network services are usually provided when you get Linux.

Web servers, file and print servers, ftp servers, NIS servers, IRC servers, news servers, and more are available for free or very little cost.

Linux is the Operating System that the EUFIRELAB computers are running.

3 APACHE HTTP SERVER

3.1 APACHE HTTP SERVER, A GENERAL OVERVIEW

HTTP stands for HyperText Transfer Protocol.

It is the "language" computers of the internet use to talk one another.

HTTP provides hypertext facilities such as links, images, videos, sounds and many other user-related features that permits navigation between computers linked to the Internet.

HTTP is the protocol used to transfer HTML pages between one computer (the server) and several others (the clients).

Browsers (such as *Microsoft Internet Explorer*, *Netscape Navigator* or *Mozilla*) are used to make the client connection in a HTTP relationship.

HTTP servers (such as *Apache*, later described in this chapter, or *Microsoft Internet Information Service (IIS)* among others) are used to make the server side of the HTTP relation.

The Apache HTTP Server Project is an open-source HTTP server for modern operating systems including UNIX and Windows NT.

It is a secure, efficient and extensible server that provides HTTP services in accordance with the current HTTP standards.

Apache has been the most popular web server on the Internet since April of 1996.

The August 2002 Netcraft Web Server Survey found that 63% of the web sites on the Internet are using Apache, thus making it more widely used than all other web servers combined.

Several key factors have contributed to Apache's success:

- The Apache license; it is an open source license that allows for both commercial and non-commercial uses of Apache.
- Talented community of developers with a variety of backgrounds and an open development process based on technical merits.
- Modular architecture; Apache users can easily add functionality or tailor (adapt) Apache to their specific environment.
- Portable: Apache runs on nearly all "flavours" (types) of Unix (and Linux), Windows, BeOs, mainframes...
- Robustness and security

Many commercial vendors have adopted Apache based solutions for their products, including Oracle, Red Hat and IBM.

3.2 WEB APPLICATION DEVELOPMENT WITH APACHE

There are basically two ways of providing content with Apache HTTP server.

- Static content: Apache can serve static content, like HTML files, images, etc.
- Dynamic content: For many websites, the information changes constantly and pages need to be updated periodically or generated on the fly. This is what server side programming is all about: programming languages, tools and frameworks that help developers query and modify information from different sources (databases, directory services, customer records, other websites) and deliver the content to the user. In this field, CGIs, servlets or PHP scripts are perfect tools to make dynamic contents for HTML pages.

Apache is a modular, featureful server, and many features make the difference when it comes to compare Apache HTTP server features to other web servers' :

Raw performance is only one of the factors to consider in a web server (flexibility and stability come usually first).

Having said that, there are solutions to improve performance on heavy loaded web-servers serving static content.

Apache provides ways in which you can measure and control bandwidth usage.

Throttling in this context usually means slowing down the delivery of content based on the file requested, a specific client IP address, etc.

This is done to prevent abuse and thus enhances response time.

Apache has several modules that allow distribution of HTTP requests among servers, for redundancy, increased availability, etc.

Apache provides extensive virtual hosting support, meaning that a single Apache server can host several web servers, and thus several domain names.

There are several solutions that provide secure transactions for Apache servers.

This enables Apache servers to be used for e-commerce or other scenarios where sensitive information is exchanged (like credit card numbers).

In many situations (subscription services, sensitive information, private areas), user authentication is required.

Apache includes basic authentication support.

Additional authentication modules exist that connect Apache to existing security frameworks or databases, including: NT Domain controller, Oracle, MySQL, PostgreSQL, LDAP, etc.

The Java on Apache community is a very dynamic and active one, as shows the quantity and quality of its subprojects, such as Ant, ORO and Regexp, Slide, Struts, Taglibs, Tomcat, Velocity, Watchdog, JServ, JSSI to name only a few..

Apache also provide commercial-quality standards-based XML solutions that are developed in an open and co-operative fashion

Perl and Apache make a powerful and popular combination.

There are several projects that use these two technologies.

Apache has also modules for popular server side languages such as Perl, Python,Tcl, JavaScript, Haskell, etc

The Apache HTTP server is at the heart of the EUFIRELAB web site, it is the program that handles any query make to this site.

4 MYSQL

4.1 MYSQL, A GENERAL OVERVIEW

A relational database management system (RDBMS) is a program that lets you create, update, and administer a relational database.

An RDBMS takes Structured Query Language (SQL) statements entered by a user or contained in an application program and creates, updates, or provides access to the database.

Some of the best-known RDBMS's include Oracle's database product line, Computer Associates' CA-OpenIngres, and IBM's DB2.

MySQL, as used in the EUFIRELAB project, is a Relational DataBase Management System.

A database management system (DBMS), or simply a database system (DBS), consists of

- A collection of interrelated and persistent data (usually referred to as the database (DB)).
- A set of application programs used to access, update and manage that data (which form the data management system (MS)).

The goal of a DBMS is to provide an environment that is both convenient and efficient to use in

- Retrieving information from the database.
- Storing information into the database.

Databases are usually designed to manage large bodies of information.

This involves

- Definition of structures for information storage (data modeling).
- Provision of mechanisms for the manipulation of information (file and systems structure, query processing).
- Providing for the safety of information in the database (crash recovery and security).
- Concurrency control if the system is shared by users.

The MySQL database server is the world's most popular open source database.

Its architecture makes it extremely fast and easy to customise.

Extensive reuse of code within the software and a minimalistic approach to producing functionally-rich features has resulted in a database management system unmatched in speed, compactness, stability and ease of deployment.

The unique separation of the core server from the storage engine makes it possible to run with strict transaction control or with ultra-fast transactionless disk access, whichever is most appropriate for the situation.

The MySQL database server is available without a license fee under the GNU General Public License (GPL).

4.2 MYSQL FEATURES

They are summarised here below.

ANSI SQL syntax support : the MySQL database server supports a broad subset of the ANSI SQL 99 syntax.

Cross-platform support : optimised binaries exist for a wide range of platforms, including Linux, Microsoft Windows, FreeBSD, Sun Solaris, IBM's AIX, Mac OS X, HP-UX, AIX, QNX, Novell NetWare, SCO OpenUnix, SGI Irix, and Dec OSF.

You can connect to a MySQL database server from all of the major platforms, using nearly any programming language.

Independent storage engines : MySQL database server's unique independent storage engines let you choose the type of database storage that is most appropriate for your particular needs.

Transactions : using the InnoDB or Berkeley DB (BDB) storage engines, the MySQL database server supports transactions.

The InnoDB storage engine also supports foreign key constraints.

Flexible security system, including SSL support : the MySQL database server has an advanced permissions and security system, including support for SSL transport-layer encryption.

The security system also allows you to limit server resources on a per-user basis.

Query caching : the latest version of the server includes a new query cache, which can significantly increase the performance of commonly-issued queries, without requiring any special programming.

Performance can be increased by over 200% in typical usage.

Replication : using database replication, you can have many "slave" servers running off a single "master" server for robustness and speed.

Full-text indexing and searching : full-text indexes allow you to search fields containing arbitrary text for specific words and phrases, including relevance rankings.

Several MySQL databases have been created for the EUFIRELAB project, and are described in this deliverable.

5 PHP

5.1 PHP, A GENERAL OVERVIEW

PHP (acronym for "PHP: Hypertext Preprocessor") is a project of the Apache Software Foundation.

It is a widely-used general-purpose scripting language that is especially suited for Web development and can be embedded into HTML.

When it comes to the question "What possibilities does PHP offer?", the answer is "Basically, anything!"

PHP is mainly focused on server-side scripting, so you can do anything any other CGI program can do, such as collect form data, generate dynamic page content, or send and receive cookies.

PHP scripts can be used in several different ways, like other programming language, but for the EUFIRELAB application, we are using it as HTTP server-side scripting.

This is the traditional and main target field for PHP.

You need three things to make this work:

- The PHP parser (CGI or server module),
- a webserver and a web browser.

You need to run the webserver, with a connected PHP installation.

You can access the PHP program output with a web browser, viewing the PHP page through the server.

PHP can be used on all major operating systems, including Linux, many Unix variants (including HP-UX, Solaris and OpenBSD), Microsoft Windows, Mac OS X, RISC OS, and probably others.

PHP has also support for most of the web servers.

This includes Apache, Microsoft Internet Information Server, Personal Web Server, Netscape and iPlanet servers, O'Reilly Website Pro server, Caudium, Xitami, OmniHTTPd, and many others.

For the majority of the servers PHP has a module, for the others supporting the CGI standard, PHP can work as a CGI processor.

So with PHP, you have the freedom of choosing an operating system and a web server.

This makes for a very wide and very active community working with PHP.

With PHP, you also have the choice of using procedural programming or object oriented programming, or a mixture of them.

Although not every standard OOP feature is realized in the current version of PHP, many code libraries and large applications (including the PEAR library) are written only using OOP code.

With PHP you are not limited to output HTML.

PHP's abilities includes outputting images, PDF files and even Flash movies (using libswf and Ming) generated on the fly.

You can also output easily any text, such as XHTML and any other XML file.

PHP can autogenerate these files, and save them in the file system, instead of printing it out, forming a server-side cache for your dynamic content.

5.2 MAIN FEATURES

One of the strongest and most significant feature in PHP is its support for a wide range of databases.

Writing a database-enabled web page is incredibly simple.

The following databases are currently supported:

- | | | |
|-----------------------|---------------|------------------------|
| - Adabas D | Ingres | Oracle (OCI7 and OCI8) |
| - dBase | InterBase | Ovrimos |
| - Empress | FrontBase | PostgreSQL |
| - FilePro (read-only) | mSQL | Solid |
| - Hyperwave | Direct MS-SQL | Sybase |
| - IBM DB2 | MySQL | Velocis |
| - Informix | ODBC | Unix dbm |

PHP also has a DBX database abstraction extension allowing the user to transparently use any database supported by that extension.

Additionally PHP supports ODBC, the Open Database Connection standard, so you can connect to any other database supporting this world standard.

PHP also has support for talking to other services using protocols such as LDAP, IMAP, SNMP, NNTP, POP3, HTTP, COM (on Windows) and countless others.

You can also open raw network sockets and interact using any other protocol.

PHP has support for the WDDX complex data exchange between virtually all Web programming languages.

Talking about interconnection, PHP has support for instantiation of Java objects and using them transparently as PHP objects.

You can also use CORBA extension to access remote objects.

PHP has extremely useful text processing features, from the POSIX Extended or Perl regular expressions to parsing XML documents.

For parsing and accessing XML documents, we support the SAX and DOM standards.

You can use our XSLT extension to transform XML documents.

While using PHP in the e-commerce field, you can use the Cybercash payment, CyberMUT, VeriSign Payflow Pro and C CVS functions useful to create online payment programs.

At last but not least, PHP has many other interesting extensions, such as the mnoGoSearch search engine functions, the IRC Gateway functions, many compression utilities (gzip, bz2), calendar conversion, translation, and much much more...

The use of PHP allowed the creation of the EUFIRELAB web site.

6 REFERENCES

Linux home page	http://www.linux.org/
Linux counter	http://counter.li.org/estimates.php
Apache home page	http://www.apache.org/
Apache usage	http://www.netcraft.com/Survey/
MySQL home page	http://www.mysql.com/
PHP home page	http://www.php.net/

7 ACRONYMS

CGI	Common Gateway Interface, a server-side scripting language
HTTP	HyperText Transfer Protocol
PHP	recursive acronym for "PHP: Hypertext Preprocessor", a multi-purpose scripting language
RDBMS	Relational DataBase Management System
SQL	Structured Query Language