



UNIVERSIDAD AUTÓNOMA
DE SAN LUIS POTOSÍ



BIBLIOTECA
VIRTUAL

Desarrollo de herramientas y servicios con tecnologías abiertas

MTE. Rosalina Vázquez Tapia

I.E. Jaime Virgilio Gómez Negrete

CONTENIDO

- Iniciativa Open Source y Free Software
- Proyecto GNU y licencia GPL
- Lenguajes y aplicaciones más populares
- Aplicaciones para Bibliotecas Digitales
- La Biblioteca Virtual CREATIVA: una plataforma 100% Open Source
- Conclusiones

Open Source Software (OSS) implica:

- Libre redistribución.
- Código fuente.
- Trabajos derivados.
- Integridad del código fuente del autor.
- No discriminación en contra de personas, grupos o áreas de aplicación.
- Distribución con licencia para su uso.
- La licencia no depende de programas o establece restricciones, debe ser tecnológicamente neutral.
- Caracterizado como *Free Software*

Referencia: <http://www.opensource.org/docs/osd>

Free Software o FLOSS

Free Software implica:

- Libertad de ejecutar el programa para cualquier propósito.
- Libertad de estudiar cómo funciona el programa y adaptarlo a las necesidades. Se requiere acceso al código fuente.
- Libertad para redistribuir copias.
- Libertad de mejorar el programa y compartir las mejoras para beneficio de toda la comunidad.

Referencia: <http://www.gnu.org/philosophy/free-sw.html>

Proyecto GNU y licencia GPL

- Es un acrónimo recursivo que significa *GNU is Not Unix*.
- Nace en 1983 con el propósito inicial de desarrollar un sistema operativo compatible con UNIX bajo la filosofía *free software*.
- La combinación de Linux y el sistema GNU origina un sistema operativo completo: GNU/Linux
- GNU GPL (General Public License) es una de las licencias más populares para OSS.
- GPL implementa el concepto de “*copyleft*” que especifica los términos de uso y distribución para la mayoría del software bajo esta licencia.

Referencia: <http://www.gnu.org/gnu/gnu-history.html>

Sistemas y aplicaciones Linux

- Prácticamente toda distribución Linux contiene una cantidad muy importante de aplicaciones *Open Source*.
- Los principales lenguajes para el desarrollo sobre Linux:
 - C , C++, Ada, Java y Fortran, Perl, Ruby y Python.
- Las principales distribuciones Linux:
 - Debian [<http://www.debian.org>]
 - Ubuntu [<http://www.ubuntu.com>]
 - Fedora [<http://www.fedoraproject.org>]
 - Mandriva [<http://www.mandriva.com>]

Aplicaciones Open Source

- **Sistemas de bases de datos relacionales:**
 - PostgreSQL
 - MySQL
 - SAP DB
- **Desarrollo de aplicaciones web server:**
 - Servidor HTTP Apache
 - Lenguaje interpretado PHP
- **Lenguajes de programación:**
 - Perl, PHP, Python

Aplicaciones Open Source

- **Protocolos de intercambio de información:**

- HTTP
- Z39.50
- OAI-PMH
- SOAP

- **Lenguajes de marcado y estándares W3:**

- HTML
- XML
- CSS (cascading style sheets)
- DOM (document object model)

- **Metadatos:**

- Dublin Core
- RDF

Combinaciones populares

- En el mundo Linux es común encontrar una serie de “combinaciones” de tecnologías *Open Source* para el desarrollo:
 - Apache + PHP + MySQL.
 - Apache + PERL + PostgreSQL.
 - JavaScript + XML (Ajax, para desarrollo del lado cliente).
- Es recomendable utilizar la que sea más adecuada para el proyecto y no mezclarlas pues haría más compleja la implementación.

Aplicaciones para Bibliotecas Digitales

• Desarrollo de colecciones:

- *Greenstone* [<http://www.greenstone.org/>]: Es un conjunto de aplicaciones para construir y distribuir multiples colecciones digitales. Requiere Apache, JAVA , GCC y GDBM. Dificultad de implementación: Regular.
- *DSpace* [<http://www.dspace.org/>]: Para construir repositorios digitales. Requiere JAVA, Apache Maven, Apache Ant, PostgreSQL, Jackarta Tomcat y Perl. Dificultad de implementación: Alta.

Aplicaciones para Bibliotecas Digitales

- **Portales con servicios bibliotecarios y motor de búsqueda:**

- *Keystone DLS* [<http://www.indexdata.dk/keystone/>]:
Conjunto de aplicaciones para la implementación de servicios bibliotecarios, incluyendo la administración de contenidos y de portales. Requiere Apache, PHP, MySQL, XML y YAZ. Actualmente ya no se ofrece desarrollo de este producto. Dificultad de implementación: Alta.

Aplicaciones para Bibliotecas Digitales

- **Portales con motor de búsqueda:**

- *OpenSiteSearch*

- [<http://opensitesearch.sourceforge.net/>]: Desarrollo original de la compañía OCLC, actualmente es un proyecto Open Source dedicado a la implementación de portales para bibliotecas digitales con énfasis en la implementación del protocolo Z39.50. Requiere Apache, JAVA y YAZ. Dificultad de implementación: Fácil.

Aplicaciones para Bibliotecas Digitales

- **Administrador de publicaciones seriadas:**

- *CUFTS* [<http://researcher.sfu.ca/cufts>]: Desarrollado en la biblioteca de la Universidad Simon Fraser en Canadá; provee servicios de administración de contenidos electrónicos, una base de datos *A to Z* integrada, analizador de hipervínculos (resolver) y control de registros MARC.

Aplicaciones para Bibliotecas Digitales

• Motores de búsqueda:

– *Apache Lucene*

[<http://lucene.apache.org/java/docs/index.html>]: localiza recursos en HTML, XML, PDF y RTF. Permite operadores booleanos y compresión de índices. Requiere JAVA. Dificultad de implementación: Regular.

– *Zebra* [<http://indexdata.dk/zebra/>]: Localiza recursos en SGML, XML, MARC y ASCII. Permite operadores booleanos, igualación por aproximación. Es cliente Z39.50. Dificultad de implementación: Fácil.

Aplicaciones para Bibliotecas Digitales

- **Meta buscador (Búsquedas federadas):**
 - dbWiz [<http://researcher.sfu.ca/dbwiz>]: Desarrollado por la biblioteca de la Universidad Simon Fraser de Canadá; ofrece en una sola interfaz un motor de búsqueda federada a bases de datos y OPAC's.
 - Pazpar2 [<http://www.indexdata.com/pazpar2/>], Index Data. Orientado a Z39.50, utiliza Ajax, XML y Webservice API.
 - OA-Hermes [<http://oa-hermes.unam.mx/oa-hermes.html>]: Es un metabuscador desarrollado por la Bidi UNAM que integra fuentes Open Access.

Aplicaciones para Bibliotecas Digitales

• Servidores proxy:

- *EZProxy* [<http://www.oclc.org/ezproxy/>]: Permite la autenticación y el acceso a recursos de manera controlada por IP desde ubicaciones remotas. No es Open Source pero se integra perfectamente en este tipo de desarrollos. Es barato y muy eficiente.
- *Pound* [<http://www.apsis.ch/pound/>]: Proxy inverso, balanceador de carga y redireccionador de peticiones basadas en URL. Indispensable para la administración de múltiples servidores back-end y la sanitización de peticiones en HTTP/HTTPS. Considérese por motivos de seguridad.

Otras Aplicaciones

- **RABID** [<http://ict.udlap.mx/rabid/>]: *XmLibris, El Pescador, PDLib y otras.*
- **Google:** *iGoogle, Google Doc, etc.*
- **Herramientas de la Web 2.0:** Blogs, RSS, Wikis, Podcast, comunidades virtuales, etc.

Tecnologías utilizadas:

- Rack de 5 servidores DELL
- Sistema Linux, distribuciones Mandriva y Ubuntu.
- Servidor web Apache versión 1.3.nn
- PERL como lenguaje principal de programación.
- Gestor de base de datos PostgreSQL

CREATIVA: Desarrollo de colecciones digitales

ACERVOS UNIVERSITARIOS DIGITALIZADOS

Libros de la Editorial Universitaria	170
Tesis de Lic. y Posgrado	480
Número de Revistas Universitarias	36
Videos de TV universitaria	80

RECURSOS ELECTRÓNICOS

Libros electrónicos	2462
Bases de datos	43
Revistas electrónicas	18 605
Sitios web y obras de consulta	280



TOTAL:

YA TENEMOS:

22 114

CREATIVA: Digitalización de contenidos

Diseñamos nuestro modelo de libro electrónico y metodología de digitalización, considerando:

- Seguridad de la información.
- Representación fiel del documento original.
- Metadatos conforme a Dublin Core.
- Acceso al texto completo
- Facilidad de uso
- Formatos de entrada: PDF, DOC, TXT, TIF.
- Formatos de salida: PDF, HTML, JPG, TXT.
- Resoluciones: preservación (600 dpi), OCR (300 dpi), imagen para web (72 dpi).
- Concordancia con las recomendaciones de la WWW.

Desarrollamos un sistema administrativo:

- Opera a partir de dos bibliotecas de funciones de Linux programadas, que alimentan la base de datos y el portal respectivamente.
- Permite trabajar colaborativamente en la construcción de las colecciones digitales.
- Consta de nueve módulos para las correspondientes estaciones de trabajo.
- Lleva el registro de todos los recursos electrónicos que contrata la Institución.
- Administra los usuarios del servicio de acceso remoto, *Mi Cuenta*.
- Registra las estadísticas de uso de la totalidad de los recursos y servicios del portal.
- Controla el acceso e impide intromisiones no autorizadas.

CREATIVA: Sistema estadístico

- Desarrollado en Perl.
- En una base de datos global se registra: IP, estampa de tiempo, URL de acceso, detecta acceso local o remoto, e ID del usuario.
- Con una aplicación programada en el Cron de Linux, se almacena mensualmente la información y se crea una tabla por mes.
- El sistema genera las estadísticas por mes, semana o día y para cualquier recurso.
- Se lleva una bitácora por día y hora de toda la actividad que los usuarios realizan por sesión activa en el portal.

- Implementamos ***DbWiz***.
- Motor de búsquedas federadas multiprotocolo:Z39.50/SRU/SRW, SOAP, y en general programando el driver respectivo.
- Desarrollado en PERL, PostgreSQL, javascript.
- Actualización constante basado en TRAC.
- Forma parte de la suite ***reSearcher***, integrada por tres módulos adicionales: ***CUFTS*** (Administrador de publicaciones seriadas), ***GODOT*** (Resolver de hipervínculos) y ***Citation Manager*** (Administrador bibliográfico).

CREATIVA: Búsqueda y recuperación de información

- Incorporamos algoritmos de búsqueda a texto completo con apoyo del clúster de base de datos de PostgreSQL.
- Diseñamos un nuevo sistema de navegación por índices alfabéticos denominado **Alfa-Zeta**.
- Este sistema genera las listas a partir de cuatro palabras con significado semántico que estén contenidas en los metadatos de los recursos electrónicos.

CREATIVA: Servicios Virtuales

1) Para todo público :

- Consulta a los catálogos y recursos de acceso abierto
- Acceso a los acervos universitarios a texto completo
- Herramientas de búsqueda, navegación, ayuda e interacción
- Videoconferencia y teleconferencia (EDUSAT)
- Versión de texto para invidentes *
- Recorridos virtuales *

* Próximamente



CREATIVA: Servicios Virtuales

2) Para universitarios:

TODOS los anteriores +:

- Mi Cuenta (Acceso remoto)
- Portafolio Digital (Biblioteca personal)
- Referencia Virtual (Asesoría de especialistas)
- Banco de imágenes *
- Videoteca Virtual *
- Reserva Digital *

* Próximamente



3) Para profesores e investigadores

TODOS los anteriores +:

- Digitalización de materiales didácticos (Paquetes instruccionales)
- Repositorio Institucional de Artículos (RIA) *
- Guías temáticas *

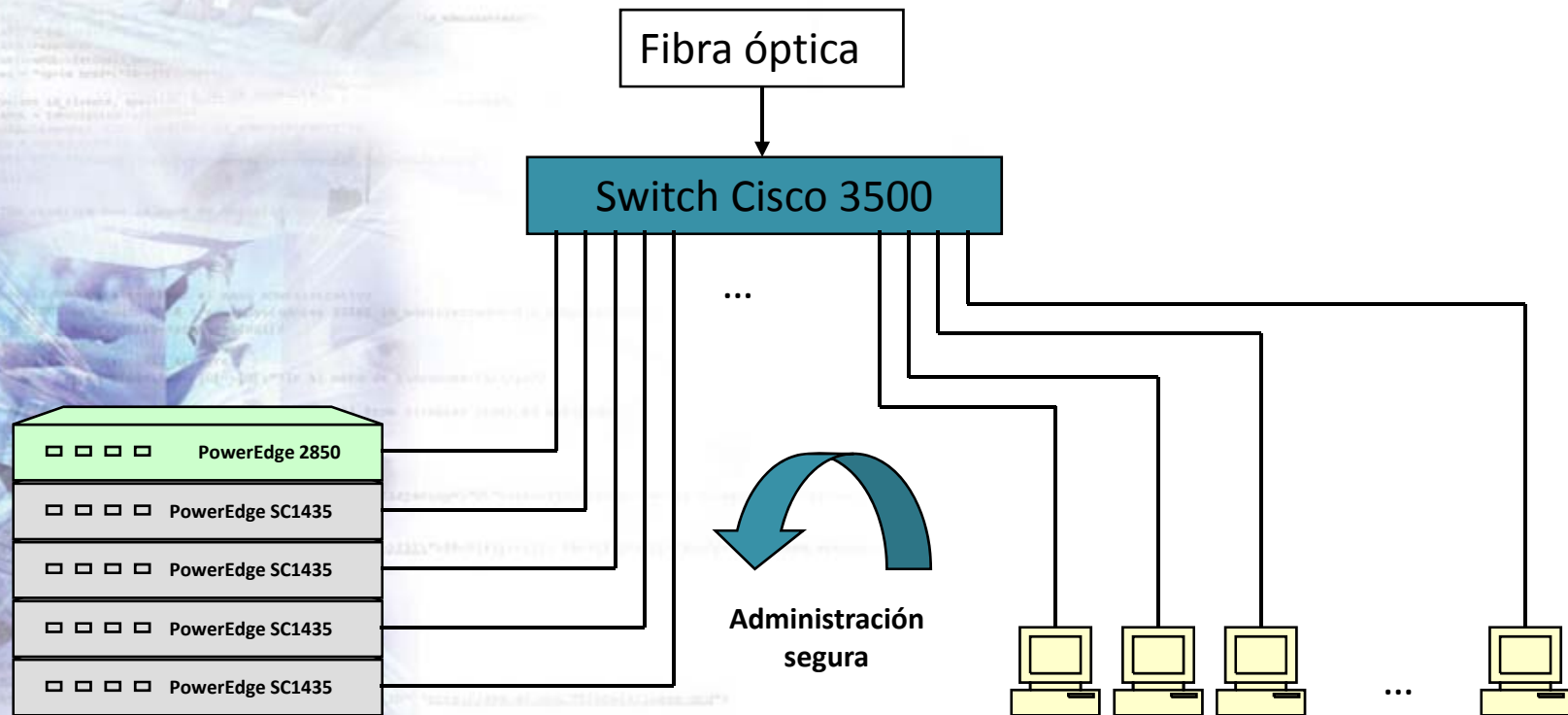
4) Para alumnos

- TODOS los anteriores +:
- Mis Cursos

* Próximamente



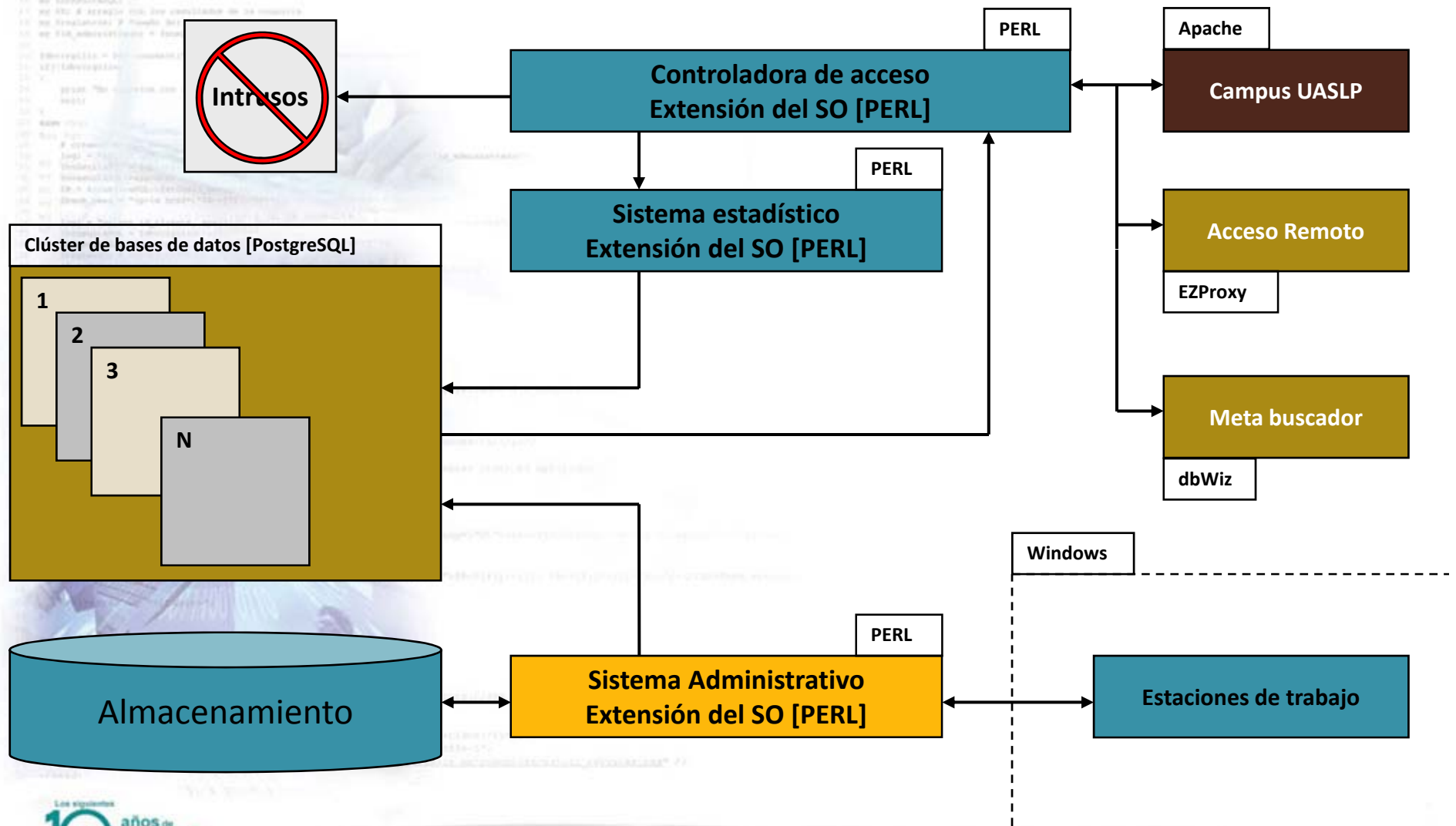
CREATIVA: Arquitectura en hardware



Granja de servidores

Equipo CREATIVA

CREATIVA: Arquitectura en software



CONCLUSIONES

- Nuestra experiencia nos ha enseñado que no hay una aplicación 100% acorde a nuestras necesidades.
- Los sistemas *Open Source* nos permiten llevar la **creatividad** a la práctica y desarrollar un producto con calidad a un bajo costo.
- Existen muchas alternativas de software *Open Source*, lo importante es identificarlas, evaluarlas e implementar las que más se adecuen a nuestras necesidades, pero sobre todo, colaborar y contribuir a esta iniciativa.

Gracias por su atención

alinavn@uaslp.mx

<http://creativa.uaslp.mx>