

Distribuciones especializadas de GNU/Linux para la Administración Pública

Sergio Talens-Oliag
sto@debian.org

Enero 2004

Resumen

En esta ponencia hablaremos de la generación y mantenimiento de distribuciones *especializadas* de GNU/Linux para su uso en las Administraciones Públicas, justificando su existencia y explicando cómo creemos que deben ser desde un punto de vista técnico

Introducción

En los últimos tiempos han aparecido varias distribuciones de GNU/Linux promovidas por las Administraciones como *LinEx* en Extremadura o *Guadalinux* en Andalucía que muy probablemente tendrán continuación en otras Comunidades.

En esta ponencia expondremos las razones por las que consideramos que es deseable el uso del *software libre* y el desarrollo de estas distribuciones específicas de GNU/Linux a partir de otras existentes.

Conceptos previos

Comenzaremos nuestra presentación comentando algunos de los conceptos que vamos a emplear a lo largo de toda la ponencia.

¿Qué es el Software Libre?

El concepto de *software libre* se refiere al derecho de ejecutar, copiar, distribuir, estudiar, cambiar y mejorar el software sin restricciones y comprende las siguientes libertades:

- Libertad de usar el programa con cualquier propósito sin ningún tipo de restricción.
- Libertad de estudiar cómo funciona el programa y adaptarlo a las necesidades propias.
- Libertad de redistribuir copias del programa y de ese modo ayudar a otros.
- La libertad de mejorar el programa y liberar esas mejoras al público.

Cuando hablemos del *software libre* en la Administración comentaremos cuales son sus ventajas y las razones por las que pensamos que vale la pena emplearlo en este tipo de entornos.

Distribución de software: paquetes y sistemas de gestión de paquetes

Desde los inicios del *movimiento del software libre* el método tradicional para distribuir los programas ha sido en forma de código fuente, generalmente empaquetado en un archivo `tar` comprimido que contiene todos los ficheros necesarios para compilar e instalar el programa.

Para instalar un programa es necesario descargar el fichero con el código, desempaquetarlo, compilarlo y copiar los ficheros ejecutables a los directorios adecuados del sistema. Generalmente la compilación implica leer la documentación, modificar algún fichero y en ocasiones instalar programas o bibliotecas adicionales.

El problema de este sistema es que es largo y tedioso, en el proceso se pueden cometer multitud de errores y para detectarlos y corregirlos hay que tener un nivel de conocimientos técnicos del que no disponen todos los potenciales usuarios del programa.

Para solventar estos problemas algunos programas comenzaron a distribuirse *precompilados* dentro de un fichero **tar** con la estructura de directorios adecuada para instalarlos en nuestro sistema.

Aunque este último sistema es muy simple, todavía tiene varios problemas:

- No hay un método claro para saber que programas hay instalados en el sistema,
- No es posible detectar los conflictos o incompatibilidades con programas o bibliotecas ya instaladas cuando instalamos un nuevo programa o biblioteca.
- Cuando un programa depende de otros programas o bibliotecas debemos verificar manualmente las dependencias, lo que puede ser tedioso en algunos casos.

Para resolver estos problemas se introdujo el concepto de *paquete* y el de *sistema de gestión de paquetes*.

Básicamente un *paquete* es un archivo que contiene binarios precompilados y *meta-datos* que describen su contenido (versión del programa, autor, origen, etc.), su relación con otros paquetes (dependencias, conflictos, etc.) e incluso las operaciones a realizar al instalar, configurar o borrar el paquete.

A lo largo del tiempo han aparecido varios formatos de paquete distintos, los más extendidos hoy en día son el del *RedHat Package Manager* (**rpm**) y el formato de archivo de Debian (**deb**).

Se puede consultar una comparativa de diversos formatos en *Comparing Linux/UNIX Binary Package Formats* [[Pkg-comp](#)]

Un *sistema de gestión de paquetes* es un sistema que mantiene información sobre los paquetes instalados e instalables e incluye herramientas para instalar, actualizar y eliminar esos mismos paquetes de un equipo concreto.

Las herramientas de gestión de paquetes de bajo nivel para los formatos **rpm** y **deb** son el **rpm** y el **dpkg** respectivamente, aunque sobre ellas se han implementado multitud de herramientas de más alto nivel, llegando hoy día a una situación en la que es posible emplear el mismo sistema de alto nivel (**apt**) con formatos de paquete distintos (actualmente **deb** y **rpm**).

Las distribuciones de GNU/Linux

Después de las descripciones anteriores, diremos que una *distribución* de GNU/Linux es un conjunto de *paquetes* preparados para instalar un sistema operativo completo en un disco duro. El núcleo empleado será **Linux** y gran parte de las bibliotecas y programas del sistema serán parte del proyecto GNU (**glibc**, **gcc**, **bash**, ...); de ahí el nombre de GNU/Linux.

Las características más importantes de una distribución serán:

- El *sistema de gestión de paquetes* y el formato de *paquete* empleado.
- El *sistema de instalación* inicial.
- La cantidad, calidad y grado de actualización de los *paquetes* que incluye. En este aspecto es muy importante saber si la distribución dispone de algún sistema de control de calidad.
- El tipo de licencias de los programas empaquetados y distribuidos; ¿todos se pueden redistribuir libremente?, si no es así, ¿es fácil distinguir los que si de los que no?
- El *soporte* que de a sus usuarios, tanto en lo relativo al seguimiento de errores y problemas de seguridad como en el uso diario del sistema (documentación existente, foros de discusión en los que encontrar ayuda, existencia de soporte *de pago*, etc.).
- La frecuencia con la que aparecen actualizaciones de la distribución *oficial* y la relación de esto último con el soporte de las distintas versiones de la distribución.
- Compatibilidad con el *LSB* o *Linux Standard Base* ([\[LSB\]](#)).

Existen gran cantidad de distribuciones de GNU/Linux, las más populares actualmente son:

- **Debian GNU/Linux** (<http://www.debian.org/>), desarrollada por más de 1200 voluntarios de todo el mundo, emplea el formato de paquete **deb** y los paquetes oficiales son totalmente redistribuibles. Funciona en 11 arquitecturas diferentes.

[Pkg-comp] <http://www.kitenet.net/~joey/pkg-comp/>

[LSB] <http://www.linuxbase.org/>

- **Gentoo** (<http://www.gentoo.org/>), al igual que debian, desarrollada por voluntarios, emplea un sistema de paquetes basado en código fuente que son compilados antes de instalar. Funciona en 6 arquitecturas diferentes.
- **Fedora** (<http://fedora.redhat.com/>), proyecto lanzado por la compañía Red Hat en sustitución de la distribución denominada *Red Hat Linux*. Es un proyecto público (aunque sigue estando controlado por la empresa), emplea el formato de paquete `rpm` y los paquetes son redistribuibles libremente.
- **Red Hat Enterprise Linux** (<http://www.redhat.com/>), distribución comercial basada en Fedora.
- **SuSE Linux** (<http://www.suse.com/>), distribución comercial adquirida recientemente por Novell, emplea el formato `rpm`.
- **Mandrake** (<http://www.mandrakelinux.com/>), distribución comercial, emplea el formato `rpm`.

Los *Live CD*

Además de las distribuciones *instalables* últimamente se han puesto muy de moda los *Live CD*; aunque en la mayoría de publicaciones y webs se suelen considerar distribuciones, en realidad se trata de imágenes de un sistema de ficheros que pueden arrancar directamente desde el CD.

Este tipo de sistemas son útiles para hacer demostraciones o probar una distribución sin necesidad de instalarla (aunque con limitaciones, ya que en el sistema de archivos *live* no se puede escribir) e incluso se pueden emplear para hacer instalaciones por clonación (formateando el disco, copiando los contenidos del sistema de archivos que hay en el CD y modificando un poco el sistema para que arranque correctamente desde el disco duro).

En muchas ocasiones las distribuciones adaptadas son simplemente un *Live CD* que contiene una instalación de una distribución estándar con paquetes adicionales, como por ejemplo tenemos **Knoppix** ([Knoppix]), que es un *LiveCD* derivado de la distribución inestable de **Debian** y **Metadistros** ([Metadistros]), que es un proyecto que tiene como objeto facilitar la creación de los *Live CD* a partir de cualquier instalación de Linux.

Adaptación y especialización de distribuciones

A partir de las distribuciones más populares han aparecido multitud de adaptaciones y especializaciones; en general lo que se suele hacer es seleccionar un subconjunto de los paquetes de la distribución base, se añaden unos pocos (que a veces son simplemente paquetes de la distribución original modificados) y se incluye un sistema de instalación modificado o incluso completamente reescrito.

Las razones por las que se generan estas distribuciones son múltiples y variadas; en algunos casos se trata de generar distribuciones comerciales a partir de otras totalmente libres (p. ej. **Libranet**, **Lindows** o **Xandros** a partir de **Debian**), generalmente simplificando la instalación, reduciendo el número de paquetes y vendiendo soporte a usuarios finales. En otros se trata de adaptarlas para usos concretos, simplificando la instalación y el mantenimiento; así tenemos distribuciones orientadas a la educación (**Skolelinux**), distribuciones que hacen especial hincapié en la seguridad (**Adamantix**, **NSA Security Enhanced Linux**, etc.), que sirven para transformar un PC en un *router* (como el *Linux Router Project*).

GNU/Linux en las Administraciones Públicas

¿Por qué usar Software Libre en la Administración Pública?

Motivos éticos

El *software libre* puede ser usado, modificado y distribuido sin restricciones, lo que promueve la participación y difusión del conocimiento entre sus usuarios, evita que se realicen copias ilegales (ya que en nuestro caso se puede hacer libremente) y a la hora

[Knoppix] <http://www.knopper.net/knoppix/>

[Metadistros] <http://metadistros.hispalinux.es/>

de trabajar con o para la Administración en temas relacionados con la informática no favorece a ninguna empresa o individuo, ya que en principio toda persona que tenga acceso a un PC puede utilizar y conocer el *software libre* en las mismas condiciones.

Motivos técnicos

En general el código fuente del software libre más empleado suele ser robusto y fiable ya que al estar disponible suele ser revisado por mucha gente, tanto para ampliar o alterar su funcionamiento como para corregir los errores que puedan ir apareciendo.

Aunque en general esto no sea importante para el usuario doméstico, en una Administración es una gran ventaja, ya que le da independencia tecnológica; al disponer o poder disponer de personal capaz de leer y escribir el código de los programas es posible modificarlos para añadir funcionalidades o corregir errores siempre que sea necesario en lugar de depender de entidades externas, que con frecuencia son empresas extranjeras.

Motivos económicos

La gran mayoría del *software libre* es gratuito, lo que reduce los costes generados por el pago de licencias. Además, las modificaciones o adaptaciones de programas pueden ser realizadas por personal propio o contratado a empresas de la industria local del software, de modo que el dinero invertido en el trabajo no acabe en otros países.

Tipos de distribución y servicios relacionados

En una organización pública puede ser interesante disponer de distintos tipos de distribuciones según el público al que vayan dirigidas:

- Administradores de sistemas,
- Personal interno de la organización,
- Ciudadanos.

y la función que vayan a cumplir:

- Sistemas servidores de propósito general,
- Sistemas servidores especializados,
- Sistemas de usuario final o de escritorio.

Una vez identificado el tipo de público al que va dirigida una distribución y la función que va a desempeñar, se puede optar por emplear una distribución ya existente tal cual está o adaptarla para generar una versión propia.

En principio, para implantar y utilizar una distribución, la Administración tendrá que disponer de infraestructura para proporcionar servicios relacionados con ella.

Estos servicios serán, a grandes rasgos:

- Formación para usuarios y administradores,
- Servicio de soporte ante problemas para el personal propio,
- Mantenimiento de los equipos y la distribución, incluyendo servicios de administración y actualización remota de equipos de escritorio y servidores especializados,

Motivos para generar una distribución propia

Vistos los servicios relacionados con la distribución o distribuciones que se deben dar, se ve rápidamente que adoptar como propia una distribución de GNU/Linux tiene algunos problemas:

- Número de programas: empleando una distribución generalista los usuarios tienen demasiadas opciones para escoger en casi cualquier aspecto del sistema (varios sistemas de ventanas, varios paquetes ofimáticos, etc.), por lo que se hace muy difícil poder formarlos en el uso de todos ellos y, lo que es más importante, se complica mucho el soporte y el mantenimiento, al tener que conocer un número demasiado elevado de programas.
- Programas no incluidos en la distribución original: si dentro de la Administración es necesario emplear software que no está en la distribución original habrá que disponer de una infraestructura paralela para mantenerlos y gestionarlos, salvo que trabajemos con una distribución gestionada por voluntarios que nos permitan incluirlos.

- Control de las versiones y corrección de errores: si se emplea una distribución comercial es la empresa la que tiene la última palabra a la hora de publicar las correcciones y a la de cambiar de versión.
- Modificaciones y traducción de programas: muchas Comunidades Autónomas Españolas son bilingües, por lo que será necesario traducir algunas de las aplicaciones existentes a los idiomas locales. Aunque a la larga las traducciones deberían acabar incluidas en las versiones originales de los programas, en principio se pueden generar paquetes especiales con las traducciones de algunos programas mientras se van incluyendo en las versiones originales.
- Preconfiguración de los programas: para muchas aplicaciones puede ser interesante disponer de una configuración distinta de la estándar a la hora de instalarlas, ya que eso simplifica mucho el soporte. Como ejemplo de esto podríamos tener preconfigurados los servidores de nombres, de correo, etc. para las instalaciones internas de la Administración o la selección de idiomas en una versión para el público en general.

Por estas y otras razones, parece que lo más adecuado para una Administración es montar un equipo dentro de su organización para que se encargue de generar y mantener una versión adaptada de una distribución de propósito general cuyas tareas serían:

- decidir qué tipos de distribuciones son necesarias,
- seleccionar los paquetes que se deben incluir en cada una,
- preparar el sistema de instalación para que sea lo más automático posible,
- preparar las preconfiguraciones que parezcan más adecuadas,
- preparar el sistema para que funcione con el idioma o los idiomas empleados en la Comunidad Autónoma, incluyendo traducciones propias si es necesario,
- crear una infraestructura para que pueda existir una comunidad local de usuarios y desarrolladores,
- montar un sistema de gestión de errores para los usuarios locales, que funcione en el idioma o los idiomas propios,
- establecer canales para relacionar la distribución propia con la original, fundamentalmente se trataría de que algunos de los integrantes del equipo participen en el desarrollo y las comunidades relacionadas con la distribución original y hagan de puente (y de filtro) entre el sistema de gestión de errores local y el de la distribución base.
- generar y mantener las sucesivas versiones de la distribución y crear versiones en forma de *LiveCD* para permitir que sea probada sin necesidad de instalarla.

Elección de Debian como distribución base

Suponiendo que se acepte nuestra idea inicial de que es adecuado el generar una distribución adaptada, proponemos emplear *Debian GNU/Linux* como distribución base.

Puede que Debian sea uno de los mayores proyectos de *software libre* del mundo si lo medimos por el número de desarrolladores (más de 1000), número de paquetes (más de 12500 en la versión oficial) o el número de informes de error procesados (más de 200.000 en 10 años).

Entre otras nos proporciona las siguientes ventajas:

- Es gestionada por una comunidad de desarrolladores *voluntarios* de todo el mundo organizados de forma democrática y que se comprometen a respetar el *Contrato Social de Debian*, que garantiza que:
 1. La distribución será y se mantendrá libre al 100%,
 2. Los integrantes del proyecto devolverán a la comunidad del software libre; los nuevos desarrollos tendrán licencias libres y se colaborará en todo lo posible con los autores originales de los programas empaquetados para Debian.
 3. No se esconderán los problemas; se empleará un sistema de gestión de errores público.
 4. Las prioridades son los usuarios y el *software libre*.
- Dispone de gran cantidad de paquetes de gran calidad técnica, gracias al uso de un sistema de seguimiento de errores público y a la definición dentro de la organización del *Debian Policy Manual* que describe la política de requisitos de la distribución, que incluye la estructura y contenidos del archivo de Debian, algunos aspectos del diseño del sistema operativo y el conjunto de requisitos técnicos que los paquetes deben cumplir para poder ser incluidos en la distribución oficial.

- Se revisan las licencias de todos los programas que se quieren incluir en la sección principal de la distribución para verificar que cumplen con las *Debian Free Software Guidelines*, lo que nos garantiza entre otras cosas que no vamos a incumplir ninguna ley al redistribuir los paquetes de esta sección.
- La distribución define varios estándares propios que facilitan enormemente el trabajo de un adaptador, como por ejemplo el uso de un sistema denominado *DebConf* para gestionar las configuraciones de los paquetes de modo automático o el uso del *Debian Menu System*, que permite localizar fácilmente las aplicaciones desde cualquier gestor de ventanas.
- Es una distribución multiplataforma (soporte de once arquitecturas) y con soporte multilingüe.
- Los cambios de la versión estable son poco frecuentes (quizás demasiado poco frecuentes); entre versiones sólo se realizan actualizaciones por motivos de seguridad.

Para cada cambio de versión se suelen establecer una serie de objetivos y una vez conseguidos se sigue un proceso de validación riguroso basado en el sistema de gestión de errores.

Tal y como hemos mencionado antes, no es recomendable emplear Debian *directamente* por varias razones:

- Excesivo número de paquetes.
- No permite personalización ni preconfiguración de paquetes.
- Herramientas de gestión de errores y comunicación con desarrolladores en inglés.
- Lentitud a la hora de incorporar traducciones.

En la actualidad existen varios proyectos en marcha para simplificar la generación de distribuciones basadas en Debian, una de las tendencias más interesantes es la descrita en la página de las *Custom Debian Distributions* ([CDD]). Un buen ejemplo de este tipo de proyectos es Skolelinux ([Skolelinux]), que es un intento de crear una versión de Debian orientada al sistema educativo.

Una de las ventajas de este tipo de aproximaciones es que las versiones adaptadas se integran casi totalmente con el proyecto original y por tanto no se duplican esfuerzos y las aportaciones de unos y otros pueden ser aprovechadas inmediatamente por todos.

Además, con este modelo, la distribución especializada es un subconjunto de la de propósito general, por lo que pasar de la primera a la última es prácticamente inmediato: basta con actualizar la lista de paquetes a partir del almacén original e instalar.

Conclusiones y trabajos futuros

El uso del *software libre* dentro de las Administraciones Públicas está justificado por múltiples motivos éticos, técnicos y económicos.

Por las características de este tipo de organizaciones es necesario proporcionar servicios de formación, soporte y mantenimiento del software.

Esto lleva a que sea recomendable para los organismos públicos crear una versión propia de una distribución de GNU/Linux, siempre y cuando se haga partiendo de proyectos ya existentes y se intente evitar la duplicación de esfuerzos.

Por último, decir que a nuestro modo de ver, la línea a seguir es la del proyecto **Debian**, creando las distribuciones propias a partir de su distribución de GNU/Linux y tratando de seguir la línea de las *Custom Debian Distributions*.

Para próximos trabajos nos planteamos el desarrollo de un documento más técnico sobre cómo generar una *Custom Debian Distribution* en el que se presente la infraestructura necesaria para hacerlo y defina mecanismos para enlazar lo que se haga en esas distribuciones con el proyecto Debian original.

[CDD] <http://wiki.debian.net/index.cgi?CustomDebian>

[Skolelinux] <http://developer.skolelinux.no/>