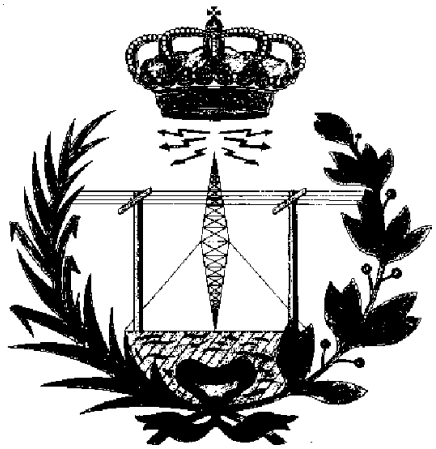
INGENIERÍA DE TELECOMUNICACIÓN

ESCUELA SUPERIOR DE INGENIEROS

UNIVERSIDAD DE SEVILLA

PROYECTO FIN DE CARRERA



Implementación de un Monitor y Analizador Gráfico de Red en el Entorno Gnome

por Juan Toledo Cota

Proyecto dirigido por Jon Tombs

Sevilla, Julio de 2001

a Mónica, por sobrellevar una
relación de muchos a muchos.
(Ella, los ordenadores y yo)

Tabla de contenidos

Lista de figuras

Agradecimientos

Jorge Chávez. Las horas interminales en su despacho no tienen precio. No hay mejor manera de aprender sobre ordenadores y tecnología que divertirse en su tecno parque temático.

Jon Tombs, mi gurú local de Unix. ¿Cuánta gente puede presumir de haber aprendido Linux de alguien que figura en el archivo AUTHORS del kernel?

Daniel López Ridruejo, por ser la fuente de inspiración para iniciar este proyecto.

A mis padres, por enseñarme a aprender.

A todos mis amigos, por no abandonarme cuando no modulo.

La Free Software Foundation, por el mágnifico conjunto de código que donan a la humanidad (y por supuesto también por el ordenador que me subvencionaron para seguir desarrollando EtherApe)

Miguel de Icaza, por iniciar el proyecto Gnome y dar un toque de español a un mundo tan dominado por el inglés.

A tantos que han contribuido a hacer del software libre lo que es: Linus Torvalds, Richard Stallman, Peter Mattis, Spencer Kimball, Josh MacDonald, Federico Mena Quintero, Damon Chaplin, Gerald Combs, Laurent Deniel, ESR, etc.

A todos los usuarios de EtherApe, y particularmente a los que han dedicado tantas horas para ayudar a depurar el programa: David Pollard, Jim Howard, etc.

```
#include “README.thanks”
```

Y a Mónica no sólo se lo dedico. También se lo agradezco.

Introducción

En los últimos años hemos vivido el desarrollo explosivo de la informática. Pero más importante si cabe que la universalización del acceso a los ordenadores ha sido el desarrollo de las redes de ordenadores, y en particular la fulminante penetración que Internet ha protagonizado a todos los niveles de la sociedad.

En el aspecto tecnológico, el hecho de que el acceso a todo tipo de datos remotos sea una realidad ubicua ha dado lugar al surgimiento casi diario de nuevas tecnologías que hacen uso de las nuevas posibilidades. Esta aceleración del progreso tecnológico ha traído consigo su propio vocablo: se habla del “Internet Time”, dando a entender que las cosas cambian mucho más rápido de lo que la industria estaba acostumbrada. Las empresas han tenido que aprender a reinventarse a sí mismas casi cada año, a riesgo de quedarse en la cuneta desbancadas por nuevos competidores surgidos de la nada.

Afortunadamente este progreso hacia un medio de comunicación global extremadamente barato no sólo beneficia a las empresas. Los ciudadanos de a pie pueden ahora ponerse en contacto con gentes del otro lado del planeta de manera casi inmediata. Esto posibilita la creación de “comunidades virtuales”, grupos de gente con aficiones comunes por raras que sean que no están limitados por su esparcimiento geográfico.

Las nuevas tecnologías que van surgiendo traen también otra consecuencia. En muchos casos los sistemas que se van poniendo en marcha crecen en complejidad y es difícil mantener una idea clara del conjunto que se está construyendo. Se echa en falta personal cualificado que pueda mantener la maquinaria en marcha, y muchas de las herramientas que facilitarían el trabajo están aún por diseñarse.

Aquí es donde las consecuencias sociales de Internet entran para echar una mano. Algunas de las comunidades virtuales que se han creado integran a profesionales de la tecnología que disfrutan de su trabajo. Pueden crear de manera independiente a las grandes empresas, y altruistamente generan soluciones a diferentes problemas que afectan a la sociedad.

Esa es la semilla de la que parte EtherApe. A partir de un problema, la necesidad de una herramienta de diagnóstico rápido de red, la sinergia del movimiento del software libre es capaz de generar una solución en un corto espacio de tiempo que puede competir de tú a tú con otros productos comerciales.

El trabajo no surge solo: sigue haciendo falta el esfuerzo de individuos dedicados. La diferencia está en la repercusión que ese esfuerzo individual tiene en el conjunto de la sociedad.

Monitorización de redes

Para cualquier actividad que se vaya a realizar son necesarias herramientas que ayuden a depurar el proceso, o que ayuden al mantenimiento una vez que esté puesto en marcha. En los trabajos que involucran redes de ordenadores, estas herramientas son los monitores de red, también conocidos como sniffers.

1 Objetivos de la monitorización de redes

Hoy en día las redes de ordenadores están muy extendidas. Desde el PC familiar conectado por módem a un proveedor gratuito de servicios de Internet, hasta el ejército de ordenadores de grandes empresas bancarias, todos son ejemplos de sistemas en red.

Potencialmente cada uno de los ordenadores de una red puede establecer una comunicación con otro. El gran número de ordenadores que intervienen es uno de los niveles de complejidad del problema, pero sólo el primero. Cada nodo puede mantener más de una conversación a la vez, y cada una de estas conversaciones puede tener un objetivo distinto.

El objetivo de estas conversaciones es variado. Algunas tendrán como finalidad el ayudar a la consecución final del producto o servicio por el que se montó la red, por ejemplo, la consulta de la página web de una agencia de viajes para la compra de un billete de avión. Otras, sin embargo, tienen lugar únicamente con el objetivo de mantener el sistema en marcha, como el intercambio constante de información que los encaminadores de las distintas redes tienen entre sí.

En un sistema tan complejo como puede ser una red informática muchas cosas pueden ir mal. Un servidor web mal configurado podría estar redirigiendo a un navegador a una página no deseada, un nodo defectuoso de la red puede generar tráfico hasta anular la capacidad de transporte de la red, o un encaminador mal

programado puede estar “dejando caer” información, o enviándola al lugar erróneo.

Además de todo lo que pueda ocurrir fortuitamente, cualquier recurso que se pone a disposición de unos usuarios puede acabar siendo abusado. Un empleado que monopoliza el ancho de banda de una empresa, tanto si es para usos lícitos como si lo es por motivos lúdicos, puede reducir de manera importante la funcionalidad de la red. Además están los peligros externos. Individuos que por diversión o por motivos económicos intentan introducirse en los sistemas de una empresa, o una universidad, saltándose las posibles medidas de seguridad que se hayan establecido.

Es por todo esto que el gestor de la red necesita de una herramienta que le ayude a analizar lo que está ocurriendo. De por sí un cable no presenta un aspecto diferente si está cargado hasta el límite de capacidad o si no está siendo utilizado. Inferir que hay un problema usando las aplicaciones diarias de la red puede no ser obvio a primera vista. Y desde luego un intruso que no quiera ser detectado no lo será a menos que hagamos una búsqueda activa.

Así pues, los objetivos que persiguen los monitores de red pueden resumirse en:

- Análisis de la eficiencia del sistema
- Diagnosis de posibles problemas
- Identificación de amenazas de seguridad

Para poder estudiar más a fondo cómo trabajan estas herramientas, necesitamos ver con más detalle cuál es el objeto de su estudio.

2 Sistemas de comunicación informáticos

2.1.1 Nivel físico

2.1.2 Punto a punto

2.2 Medio compartido

2.3 Nivel lógico

3 Algunas herramientas de monitorización disponibles

3.1 Tcpdump

3.2 Ethereal

3.3 Ntop

3.4 Etherman

Software libre

EtherApe ha sido desarrollado y distribuido como software libre. Software libre no sólo hace referencia a gratuidad, *el concepto de* libertad *es más importante. En este capítulo trataremos de hacer ver por qué el software libre no es sólo una opción a considerar por sus beneficios sociales, sino porque también conduce a la producción de programas de más calidad desde muchos puntos de vista.*

"If I have seen farther than others, it is because I was standing on the shoulders of giants." scientist Sir Isaac Newton, in a letter to his colleague Robert Hooke, February 1676

1 Definición de software libre

2 Ventajas técnicas del software libre

Velocidad de desarrollo

Arquitectura de EtherApe

EtherApe es la solución al problema del diagnóstico rápido de una red. El principal objetivo de la arquitectura adoptada es el de obtener resultados lo más rápidamente posible. La tecnología en que se basa (Gnome) y el método de trabajo (software libre) han sido de gran ayuda para la consecución del objetivo marcado. Pero no basta con eso. Se ha tomado un esfuerzo considerable en separar las funciones del programa en módulos casi independientes. De esta manera también se ayuda a mejorar la mantenibilidad y depurabilidad del código.

1 Objetivos de diseño

La idea que da lugar a EtherApe es poder reproducir los resultados de Etherman en las plataformas más accesibles hoy día. Esta afirmación genérica se traduce en detalle en:

- Representación gráfica de la red

  Se requiere una representación intuitiva de lo que sucede en la red. Los nodos se identifican con círculos y las conversaciones con líneas que unen aquellos. Se introduce información adicional en el tamaño y el color de los elementos

- Representación en tiempo real

  EtherApe pretende ser un primer recurso en la diagnosis de problemas. La información se presenta en tiempo real y la persona responsable puede decidir qué pasos ulteriores tomar tras un vistazo rápido.

- Portabilidad

  Etherman estaba limitado a las plataformas a las que los desarrolladores originales tenían acceso.

- Perdurabilidad del código

  En contraposición a Etherman, EtherApe debe poder estar disponible por largo tiempo, independientemente de que cambien las plataformas subyacentes o que el desarrollador principal abandone su puesto. La licencia GPL asegura este punto.

Estos objetivos condicionan la elección entre las posibles alternativas que estén disponibles para la implementación. Después de presentar la solución adoptada se discutirán las posibles opciones y la razón que motiva tal decisión.

2 Esquema general

EtherApe es un programa diseñado para correr bajo plataformas Unix genéricas. El código fuente en C es portable a diversas implementaciones que cumplan con la norma POSIX.

EtherApe ha sido compilado y ejecutado con éxito al menos en las siguientes combinaciones de sistemas operativos y procesadores:

- Linux

  Debian, RedHat, Mandrake, Slackware, Suse, etc.

  Probado con procesadores x86, Alpha, Motorola 680x0 y PowerPC

- FreeBSD
- NetBSD
- Solaris

  Al menos con las versiones 7 y 8

Para conseguir su objetivo, EtherApe saca partido de un conjunto de bibliotecas de libre distribución, que son las que verdaderamente definen el límite de la portabilidad del programa.

En el siguiente esquema está representada la arquitectura de EtherApe, con el conjunto de bloques funcionales que la componen.

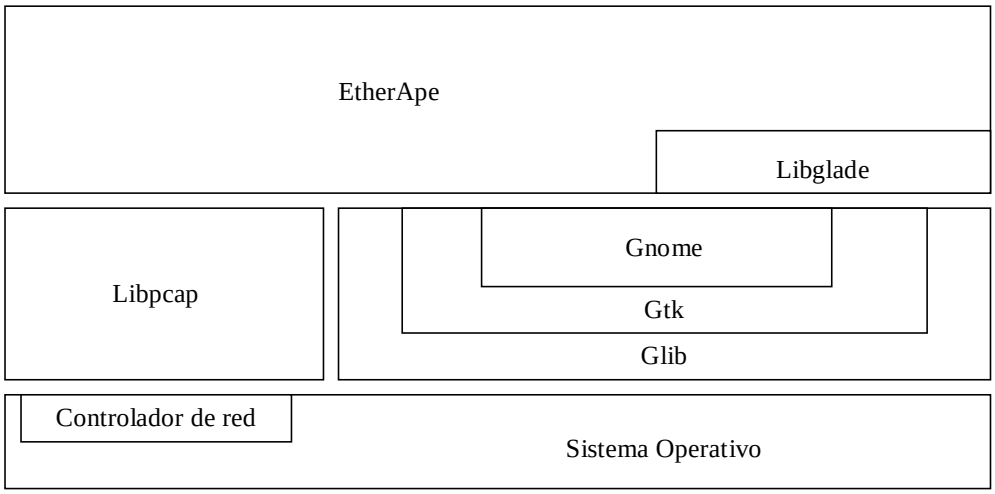


Figura 1 – Diagrama de bloques de EtherApe

Tal y como puede apreciarse en el esquema, el programa no interactúa directamente con el sistema operativo en ningún caso, sino que deja que sean las bibliotecas las que cumplan esa función.

También salta a la vista el hecho de que la interacción con el sistema operativo (y por ende, con el hardware) se realiza por dos caminos distintos que están bien diferenciados. Por un lado el sistema que se encargará de leer el tráfico de la red, y por otro el encargado de presentar los datos y el resto de la interfaz de usuario.

Nótese que en cualquier caso este es un esquema simplificado, y que cada uno de los elementos trae consigo sus propias dependencias. Ausencias notables son la biblioteca estándar de C, y las bibliotecas del sistema X-Windows.

3 Dependencias externas

Una de las máximas siempre presentes en el desarrollo de software libre (del software en general, pero como veremos más adelante es un aspecto fundamental del software libre) es el esfuerzo por no reinventar la rueda.

Este concepto alude al hecho de que demasiado a menudo un desarrollador tenderá a rechazar código no realizado por él mismo bien por desconfianza, bien por pereza a la hora de estudiar APIs (application program interface, interfaz de programación de aplicaciones) ajenos.

Sin embargo utilizar adecuadamente los recursos disponibles trae consigo un conjunto de beneficios:

- Disminución del tiempo de desarrollo

  Si bien será necesario adaptar el programa para asegurar que se ajusta a la interfaz de la biblioteca que vayamos a usar, ese tiempo siempre es menor que el que se necesita para programar desde el principio esa funcionalidad

- Aumento de la generalidad

  La biblioteca que vayamos a utilizar es un proyecto en sí mismo y sus creadores probablemente se habrán encargado de hacerlo funcionar en más entornos de los que nosotros tengamos inicialmente planeados

- Disminución del tiempo de depuración

  Por el mismo motivo que el punto anterior, los encargados de mantener esa biblioteca habrán tenido oportunidad de probar más casos patológicos, aumentando la fiabilidad del código.

  Por otro lado es cierto que no siempre es apropiado usar bibliotecas. Puede ocurrir que no haya bibliotecas existentes que se ajusten exactamente a nuestras necesidades. Aunque este no es el caso de EtherApe, una ventaja de usar bibliotecas que sean software libre cuando ocurre esto es que podemos partir del código existente, y añadir a la biblioteca sólo lo que falta hasta servir nuestras necesidades.

  Por todos estos motivos a lo largo del desarrollo de EtherApe se ha hecho un esfuerzo consciente para investigar qué código estaba ya disponible que pudiera ayudar en la implementación, y de usarlo en lo posible.

3.1 Libpcap

Libpcap es la biblioteca que nos permite obtener una copia de los paquetes que circulan por la red. Libpcap empezó siendo parte del monitor de red Tcpdump, pero con el paso del tiempo sus creadores la segregaron y la distribuyeron como una biblioteca independiente para que otros proyectos pudieran sacarle partido fácilmente.

Algunas de las características que hacen Libpcap enormemente interesante son:

- Independiente de plataforma

  A lo largo de los años en los que Tcpdump se ha ido desarrollando (la versión inicial de este programa data de 1991), el número de plataformas que soportaba ha ido creciendo sin cesar hasta convertirse prácticamente en una herramienta universal.

Libpcap ha heredado esta característica, y permite a los programas que enlazan con esta biblioteca usar un único API para extraer paquetes de la red desde casi cualquier sistema operativo: Ultrix, SunOS, VMS, Linux, HP-UX, FreeBSD, NetBSD, etc. Recientemente incluso el entorno Win32 está soportado, ampliando aún más si cabe sus dominios.

- Independiente de medio físico

  Libpcap no sólo es capaz de entender las especificidades de muchos sistemas operativos. También es capaz de abstraer las diferencias entre los distintos medios de transporte y proporciona la trama que se haya recibido sea cual sea el caso. Además, usando la función adecuada, informa a la aplicación que está por encima de cuál es el medio físico que se está usando, de modo que pueda prever cómo serán que formato tendrán las tramas que capture

- Lectura y escritura de archivos de captura

  Además de recoger paquetes “en vivo” directamente de un dispositivo de red, Libpcap es capaz de recuperar un volcado de tráfico hecho en un archivo, probablemente con otra herramienta que haya use Libpcap. En el caso de EtherApe, esto nos permite depurar problemas de usuarios utilizando como datos de ejemplo exactamente el mismo tráfico que ven ellos en su red.

  Otra ventaja adicional para la depuración es que cuando se lee de un archivo de datos pasamos del entorno aleatorio de una red en uso real a una entrada de datos que es siempre la misma. Si el programa comete un error al intentar procesarla podemos repetir el proceso cuantas veces sea necesario hasta identificar la causa del problema.

- Filtrado de tráfico

  Cuando una aplicación pide a la biblioteca Libpcap que inicie una captura, puede también pedir que sólo le muestre los paquetes que cumplan unos determinados criterios. Esta funcionalidad es extremadamente útil para aislar problemas en redes que están muy cargadas, donde pueden estar manteniéndose docenas de “conversaciones” diferentes en un momento dado.

  La sintaxis que definen los filtros es muy extensa, y entre otras cosas es posible seleccionar tráfico usando criterios de nodo de origen o destino, protocolo que se esté usando, puerto al que va destinado un paquete, etc.

Por todo lo anterior el uso de Libpcap es una decisión obvia a la hora de implementar cualquier monitor de red. De hecho, el número de paquetes basados en esta biblioteca es bastante alto. A modo de ejemplo podemos citar Ethereal, ya discutido más arriba; Nessus un programa auditor de seguridad; divine, que escucha los mensajes de ARP para sugerir configuraciones de red para ordenadores portátiles; Sniffit, otro analizador de red, pero con funciones específicas para volcar el contenido textual de las conversaciones (muy práctico por tanto para los piratas informáticos), etc.

El único inconveniente que se le podría achacar a Libpcap es que el hecho de que sea capaz de filtrar el tráfico usando tal diversidad de criterios implica que la misma biblioteca está de hecho analizando la pila de protocolos de cada paquete que recibe. En muchos casos eso se traduce en que se duplica el trabajo que hace el programa: la pila de protocolos se analiza dos veces, una en la biblioteca, y otra en la aplicación que soporta.

Sin embargo durante el desarrollo y el uso de EtherApe no se ha apreciado que esto suponga ningún inconveniente grave.

3.2 Glib

Glib es una biblioteca de portabilidad y de utilidades para sistemas Unix y Windows. Al igual que ocurre con Libpcap, Glib se escindió como biblioteca independiente a partir de Gtk+. Podríamos decir que Glib es un intento de estandarizar una extensión de la biblioteca estándar de C. Veámoslo con más detalle.

3.2.1 Portabilidad

Gtk+ se creó para ser un sustituto adecuado de la biblioteca de componentes gráficos Motif. Una de los objetivos a alcanzar era que el sistema fuera portable a muchas arquitecturas, salvando las distancias que separan a cada una de las implementaciones de C en cada plataforma. Hoy día es Glib la biblioteca que consigue este objetivo, y no sólo da soporte a Gtk+, sino a una gran variedad de otras bibliotecas y aplicaciones.

Veamos una lista de características que hacen que las aplicaciones que usan Glib sean más portables.

- Definición de tipos portables.

  Un problema del estándar C es que los tipos `int`, o `short int`, por ejemplo, no tienen exactamente la misma dimensión en todas las

plataformas. Depende de cómo esté definido para la arquitectura subyacente.

Por ese motivo, cuando se pretende hacer un programa portable el programador a de crear código condicional según la plataforma en la que se vaya a compilar, de manera que pueda tener confianza en el tamaño de los tipos que está utilizando.

Si se usa Glib, se deja ese trabajo a la biblioteca. Glib define una serie de tipos inambiguos, tales como gint8, guint8, gint16, guint16, gint32, guint32, gint64, guint64. Por otro lado también se añaden tipos que no están en el estándar a pesar de ser muy comunes como gboolean, gsize, gssize; y otros para simplificar el lenguaje como gpointer, gconstpointer, guchar, guint, gushort, gulong.

EtherApe saca partido de estos tipos en muchas situaciones. Por ejemplo, a la hora de definir variables que han de contener la una dirección IP (cuatro octetos) se usa el tipo `guint32`; o cuando se necesita un puntero a una zona de memoria que alberga una trama (es decir, un conjunto arbitrario de octetos), se utiliza `guint8 *`.

- Carga de código en tiempo de ejecución.

  Tanto Windows como muchas versiones de Unix soportan la incorporación de objetos compilados en tiempo de ejecución (las conocidas DLL, o de manera genérica, objetos compilados en Unix).

  Sin embargo el procedimiento para hacer uso de esa funcionalidad es dependiente de cada sistema operativo. Glib alivia el problema introduciendo un API portable, tanto para que el programa pueda averiguar si esa función está presente en la plataforma, como para ponerla en marcha.

- Canales de entrada y salida

  En C básico se habla de entrada y salida estándar, del error estándar y de descriptores genéricos de archivo. Sin embargo los sistemas operativos tienen características que van más allá de estos, como los *pipes* y los *sockets*, y trabajar con ellos requiere añadir complejidad al programa, además de utilizar mecanismos que no son portables.

  En particular, hacer lecturas o escrituras asíncronas implica diseñar el ciclo principal del programa alrededor de ellas, y es incómodo tener que estar pendiente en cada momento cómo se comportará tal o cual sistema operativo.

Mediante Glib se simplifica el proceso. Cómo veremos más adelante Glib aporta una implementación estandarizada del ciclo principal de un programa, y de esta manera puede incorporar también de manera natural en procesado de canales de entrada y salida, en sus múltiples encarnaciones, y tanto en su modo síncrono como asíncrono.

Esta característica es fundamental para EtherApe, que funciona a partir de dos flujos asíncronos independientes: la aparición de tramas de datos en la interfaz de red que se está leyendo y la interacción de usuario con la interfaz del programa.

3.2.2 Utilidades

Casi a la vez que se desarrollan los lenguajes de programación van apareciendo técnicas de programación que se comprueba son útiles. A pesar de que muchas de estas técnicas no son realmente parte del lenguaje en cuestión son tan básicas y tan necesarias que en muchas ocasiones su enseñanza se hace de manera conjunta.

Estructuras de datos como las listas enlazadas, los árboles binarios balanceados, o las tablas de hash son básicas para cualquier programa que deba almacenar información, y sin embargo por lo común cada programador debe hacer un reimplementación para cada programa que vaya a realizar, si bien lo más común era que al cabo del tiempo se creara su propia biblioteca de utilidades.

Por otro lado procedimientos como un bucle principal con soporte para eventos asíncronos, funciones que ayuden a mostrar mensajes de depuración o de registro, o que ayuden a minimizar los problemas de asignación y corrupción de memoria en C son también extremadamente comunes, si bien en muchos casos puesto que su uso ya no es tan perentorio acaba por ocurrir que los programadores lo dan de lado y da como resultados programas menos robustos y flexibles.

Glib pretende dar una solución a estos dos problemas. Estandarizando una solución el programador puede olvidarse de los detalles de estas tareas “mundanas” y pensar solamente en lo que es específico a su programa. Un conjunto de especialistas mantiene Glib, y se encarga de asegurarse de que la implementación de estos algoritmos genéricos es lo más eficiente y robusta posible.

Algunos de las utilidades que Glib aporta y de las que EtherApe hace uso extensivo son:

- El ciclo principal de eventos

  La mayoría de los programas que se hacen hoy en día no se corresponden con el modelo de antaño. En lugar de ejecutar un programa que realizará un único cálculo o función, hoy los programas se acercan más a las arquitecturas cliente/servidor.

  Los servidores de red se mantienen en segundo plano hasta que un cliente hace una llamada, y sólo entonces se despierta el proceso para realizar algún trabajo. Pero también ocurre lo mismo en las interfaces de usuario. Un programa como un procesador de texto está esencialmente parado hasta que un usuario pulsa una tecla o activa un mando con el ratón.

  Si bien las bibliotecas de componentes gráficos suelen proveer de algún mecanismo para permitir que el programa que deba soportar se comporte de esta manera, Glib ofrece la oportunidad de sustentar este esquema en cualquier tipo de programa, no sólo los que está orientados a una IGU. Así, pueden usarlo por ejemplo también servidores de red o bien bibliotecas de mayor nivel que se encarguen del sistema gráfico.

  Como ya se comentó antes EtherApe hace uso del ciclo principal de eventos para gestionar tanto la llegada de tramas de red como del tratamiento de la interfaz de usuario.

- Estructuras complejas de datos

  Glib proporciona un amplio abanico de estructuras de datos complejas, que se ajustan a cualquier uso. Tablas, listas simple y doblemente enlazadas, árboles balanceados binarios, tablas de hash, e incluso estructuras de uso mucho menos común como árboles n-arios, quarks (asociación bidireccional entre una cadena y un número entero) y tablas indizadas (tablas que pueden ser indizadas en un número arbitrario de columnas)

  Puesto que la biblioteca no puede saber qué uso exactamente se le va a dar a cada estructura, cada elemento almacena un puntero, en lugar de una estructura específica. EtherApe, por ejemplo, debe almacenar el conjunto de nodos que ha ido escuchando en la red. Para ello se crea una estructura que guarda información sobre un nodo, y se almacena un puntero a la clave que identifica a ese nodo en un árbol balanceado binario junto con el puntero a la estructura creada.

  Glib también proporciona funciones específicas para gestionar estas estructuras complejas de datos, ahorrando al programador de tener que estar pendiente de reservas de memoria, por ejemplo. Además puesto que Glib se

encarga de una buena parte de la gestión memoria, es capaz de optimizar su uso para minimizar el número real de peticiones de asignación que se hacen al sistema operativo, y de este modo acelerar la ejecución del programa.

- Mensajes de depuración y registro

  Glib incorpora un sistema flexible para el tratamiento de mensaje de depuración y registro. Usando funciones como `g_debug`, `g_info`, `g_warning`, o `g_critical` el programador separa la información que el programa debe registrar de la implementación del registro o la presentación de esa información.

  Por defecto la salida se hace por el estándar error dependiendo por ejemplo del valor de la variable de entorno DEBUG (tal y como hace EtherApe), pero también es posible derivar su procesado a otras funciones, de manera que se pueda hacer que la salida se presente en una ventana de la interfaz gráfica de usuario, o enviarlo al sistema de registro de mensajes del sistema operativo (syslogd, en los sistemas GNU/Linux)

El número de utilidades que incorpora Glib es aún mayor: funciones para la gestión de hilos, la gestión de cachés de memoria, el tratamiento robusto de cadenas o la medida de tiempos son algunas de ellas. Aquí sólo se han presentado algunos de los ejemplos de los que EtherApe se beneficia.

3.3 Gtk+

Gtk+ son las siglas de Gimp Toolkit. En 1996 dos estudiantes de Berkeley publicaron la primera versión de un programa de retoque fotográfico como software libre, llamado Gimp. En su primera encarnación Gimp usaba la biblioteca de componentes gráficos propietaria Motif, el estándar de la época, y pronto estuvo clara la necesidad de desarrollar una biblioteca libre que estuviera a la par de las necesidades técnicas.

Así surgió Gtk+, una biblioteca de componentes gráficos que desde entonces ha evolucionado técnicamente y se ha hecho muy popular, superando su objetivo inicial de dar soporte únicamente a Gimp.

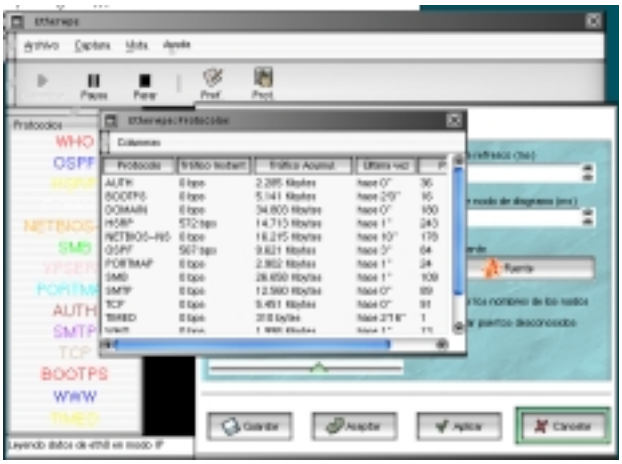
A pesar de que Gtk+ es una biblioteca escrita en C y preparada para ser usada por programas en C, Gtk+ incorpora el concepto de objetos mediante un ingenioso uso de las estructuras. Así pues los componentes se derivan unos de otros, mantienen las propiedades de sus padres y pueden ser tratados como cualquiera de sus ancestros.

De hecho la implementación interfaz de usuario como un conjunto de objetos es el paradigma que prevalece hoy en día y por tanto Gtk+ ha debido adoptar una estructura que lo soportara.

Gtk+ incorpora una larguísima lista de componentes gráficos de entre los cuales el programador de la aplicación puede elegir: botones, entradas de texto, ventanas de diálogo, árboles jerárquicos, etiquetas, menús, barras de progreso, etcétera. En este sentido Gtk+ es un gran paso adelante con respecto al estado anterior de la situación, en la que los programas que quisieran tener una presentación gráfica debían elegir entre pagar una licencia de Motif para tener un sistema moderno y eficaz o programar directamente usando las bibliotecas de bajo nivel del sistema X-Windows.

Gtk+ posee una característica que lo hace muy atractivo para los usuarios modernos. Una vez diseñado una interfaz de usuario, el aspecto final en pantalla no está totalmente delimitado, sino que el usuario tiene libertad para escoger la apariencia de los controles. Esto que parece banal hoy en día tiene una gran importancia entre el público a la hora de decidir qué programa van a usar. El programa no sólo ha de ser funcional. También ha de ser bonito.

Si bien no ha sido este el motivo fundamental para la elección como la biblioteca de componentes gráficos de EtherApe, sí es un aliciente añadido bastante interesante.



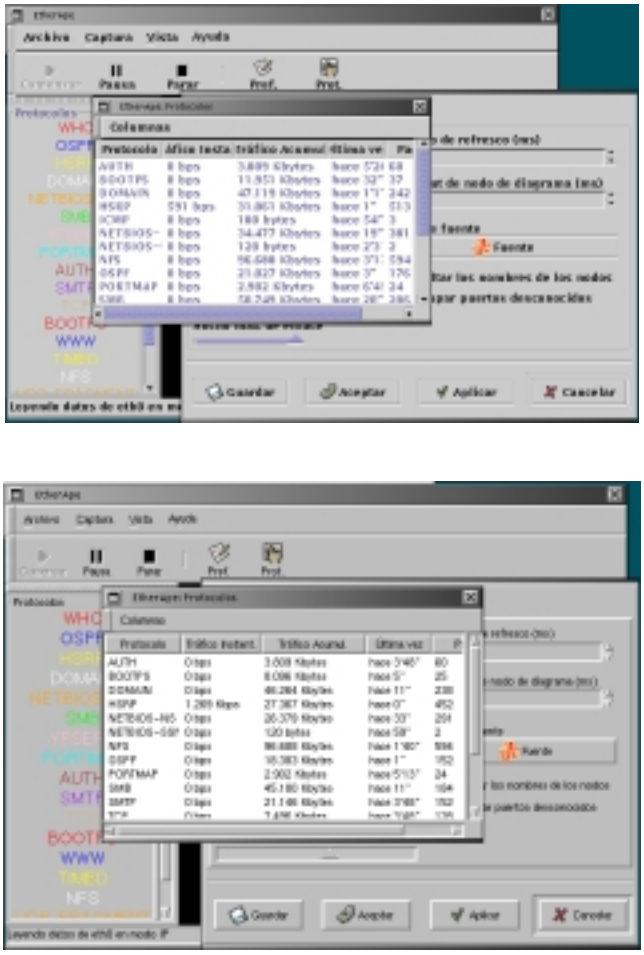


Figura 2 – Estilos de Gtk+ aplicados a EtherApe

Gtk+ tiene una estructura que lo hace fácilmente portable. Si bien en su encarnación original Gtk+ tenía como objetivo servir de puente entre la aplicación y las bibliotecas del sistema de ventanas de X, en realidad en grueso de la biblioteca no depende de X en absoluto, sino que se basa en otra biblioteca propia que se encarga de acceder a las primitivas gráficas: la biblioteca Gdk.

Puesto que los puntos de contacto con el sistema X-Windows están tan bien definidos, en los últimos tiempo ha sido posible hacer una versión de Gdk (y por tanto de Gtk+) que funciona de manera nativa en los entornos Win32. A pesar de que la versión de Gtk+ para Windows todavía no es estable, el hecho de que no sea técnicamente imposible implica que a no muy largo plazo las aplicaciones originalmente programadas para correr en entornos Unix usando Gtk+ podrían ser también aplicaciones de Windows con muy pocos cambios.

Actualmente Gtk+ es junto con QT una de las dos bibliotecas de componentes gráficos distribuidas como software libre con más aceptación. La diferencia

técnica fundamental entre las dos es que mientras que Gtk+ es una biblioteca en C, QT es una biblioteca de objetos de C++.

3.4 Gnome

3.4.1 Nota histórica

Además de la diferencia entre los lenguajes de Gtk+ y QT, hay otra distinción histórica que mantiene a los usuarios de cada una de las bibliotecas a bastante distancia.

En 1998, más o menos a la vez que la versión 1.0 de Gtk+ hacía su aparición, un conjunto de programadores en Alemania inicia un intento de desarrollar un entorno integrado de usuario (un equivalente en software libre a CDE para plataformas Unix, o la interfaz de usuario de Windows) basado en una reciente biblioteca de componentes que en aquel momento era técnicamente más avanzada que Gtk+, pero que adolecía de “problemas de licencia”. Era (y es) el entorno KDE.

Si bien QT era una biblioteca gratuita, no era software libre en sí mismo, y por tanto muchas figuras del software libre argüían que KDE hipotecaba su futuro a los designios de la compañía que desarrollaba la biblioteca QT.

Surgió entonces un movimiento alternativo para desarrollar otro entorno integrado de usuario, pero éste había de estar basado enteramente en software libre, y en particular se eligió Gtk+ como la biblioteca de componentes gráficos que habría de sustentarlos. Este segundo entorno es el conocido como Gnome.

Con el paso de los años la disputa original acabó por carecer de sentido cuando Trolltech, la compañía creadora de QT distribuyó su biblioteca bajo la licencia GPL. Sin embargo para entonces tanto KDE como Gnome habían alcanzado ya una masa crítica de desarrolladores y hoy en día siguen evolucionando como dos alternativas independientes.

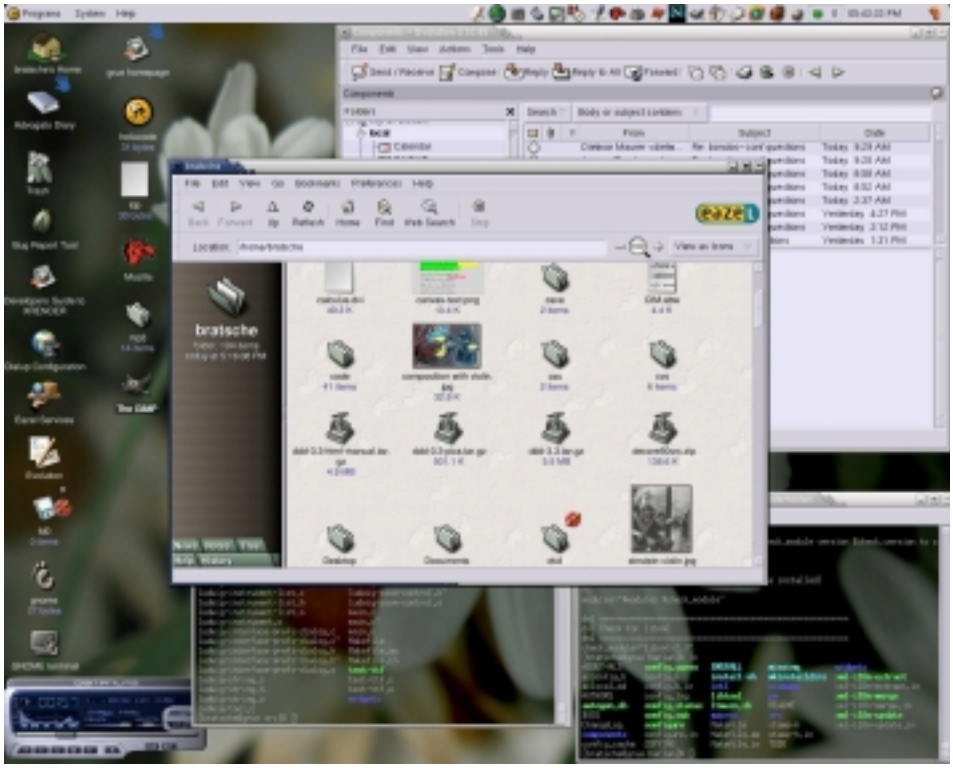


Figura 3 – El escritorio Gnome

3.4.2 Definición de Gnome

A primera vista es difícil aclarar qué es exactamente Gnome, y esto es así porque Gnome es en realidad un conjunto de cosas agrupadas bajo un mismo nombre.

- Un entorno integrado de usuario

  También conocido como el escritorio Gnome. Son un conjunto de aplicaciones que hacen que el contacto inicial de un usuario no técnico con el ordenador sea una experiencia relativamente amigable.

  Aquí se incluyen los menús de acceso a los programas, las barras de tareas, los pequeños programas que los usuarios han aprendido a esperar que estén siempre disponibles, como una calculadora, etc.

  El objetivo es lograr un escritorio atractivo y que las aplicaciones tengan un funcionamiento similar entre todas ellas para hacer más intuitivo el trabajo del usuario.

- Aplicaciones de productividad

  El núcleo de desarrolladores de Gnome no sólo pretende programar las pequeñas aplicaciones (por comparación, digamos las aplicaciones que vendrían por defecto al instalar Windows 98), sino que también están las

aplicaciones más complejas que se necesitan hoy en día en cualquier entorno productivo.

Por ejemplo, Gnumeric es la respuesta del entorno Gnome al programa de hoja de cálculo Excel, pero también hay programas para hacer procesamiento de textos, presentaciones, tratamiento de imágenes, etc.

- Una plataforma de desarrollo

  Para alcanzar los objetivos que se pretenden, Gnome se ayuda de una serie de bibliotecas que añaden funcionalidad a los programas. Una de estas bibliotecas se construye sobre Gtk+ para crear componentes gráficos más complejos (por ejemplo un selector de archivos que mantiene un historial), pero muchas más son independientes de la interfaz gráfica.

  Hay, por ejemplo, funciones para procesar los argumentos de la línea de comandos, para gestionar los documentos de ayuda en línea, soporte para una arquitectura de componentes sobre CORBA, y una capa de abstracción de sistemas de archivos (útil para leer imágenes de una cámara fotográfica como si fuera un sistema local, por ejemplo).

Si bien EtherApe no pertenece de manera formal al proyecto Gnome (una distinción que puede definirse como meramente administrativa), si que intenta cumplir todas las guías de diseño para programas de ese entorno, y desde luego hace uso extensivo de las facilidades que proporciona la plataforma de desarrollo.

3.4.3 Características de interés

En muchos casos no es indispensable hacer una aplicación dependiente de Gnome para que pueda cumplir su objetivo fundamental. De hecho cualquier aplicación que use las bibliotecas de Gnome provoca una cascada de dependencias que en muchos casos puede complicar su uso a los usuarios finales. Por este motivo no es raro encontrar aplicaciones que tengan dos versiones diferentes. Una más ligera que sólo dependa de Gtk+, y otra más amigable para el usuario pero también con un bagaje de dependencias mucho más amplio que sí usan las bibliotecas que proporciona Gnome.

En el caso de EtherApe una característica en particular decidió la balanza a favor de la inclusión de Gnome: el componente *canvas*. Este componente simplifica en sobremanera la presentación del diagrama de la red, e incluirlo ayudaba a conseguir un producto funcional en un lapso de tiempo mucho más corto.

Una vez que se decide usar Gnome, se aprovecha la circunstancia para sacar partido de otras características que ofrece el entorno. Así pues, algunas de estas son:

- El componente *canvas*.

  Como ya se ha dicho este componente es fundamental a la hora de la presentación del diagrama. La versión inicial de EtherApe que no utilizaba Gnome debía "pintar" en un componente de dibujo cada uno de los elementos diagrama en cada paso. Igualmente era necesario utilizar una técnica de doble memoria intermedia para evitar el parpadeo de la pantalla, además de necesitar código adicional para asegurar un refresco cada vez que cualquier otro elemento del sistema de ventanas obstruía parcialmente el diagrama.

  El canvas evita todos estos problemas, porque no es simplemente un área de dibujo, sino un objeto que es capaz de almacenar primitivas de dibujo y sus características aliviando al programador de las complicadas tareas gráficas. Básicamente al canvas se le ordena mantener en pantalla un círculo de un radio y color determinados en un punto determinado y el componente se encarga de todo lo demás. Las características pueden ser modificadas con posterioridad y el canvas se actualizará automáticamente.

- Documentación en línea

  La segunda característica más visible que hace evidente que EtherApe es una aplicación Gnome es la presencia de documentación en línea.

  Gnome define una serie de estándares sobre cómo debe organizarse la documentación de una aplicación, y provee además de una interfaz de programación que servirá luego para lanzar el visor de la ayuda.

  La documentación se escribe utilizando un lenguaje de estructurado de marcadores (SGML), la implementación dedicada a la documentación denominada Docbook. A partir de un único archivo de texto utilizando este lenguaje es posible generar múltiples formatos de salida: postscript, HTML, XML, texto llano, TeX, etc.

```
<!-- ============= Introduction ============================== -->
  <sect1 id="intro">
    <title>Introduction</title>

    <para>

    <application>EtherApe</application> is a graphical network monitor
      for Unix modeled after etherman. Featuring ether, ip and tcp
      modes, it displays network activity graphically. Hosts and links
      change in size with traffic. Protocols are color coded. It
      supports ethernet, fddi, ppp and slip devices. It can filter
      traffic to be shown, and can read traffic from a file as well as
      live from the network.
    </para>

    <para>
      <application>EtherApe</application> is also a tool for gathering
      network statistics, and the data can be presented in a number of
      different ways.
    </para>

    <para>
      <application>EtherApe</application>
    </para>

    <para>
      To run <application>EtherApe</application>, select
      <menuchoice>
          <guisubmenu>Applications</guisubmenu>
          <guimenuitem>EtherApe</guimenuitem>
      </menuchoice>
      from the <guimenu>Main Menu</guimenu>, or type
      <command>MYGNOMEAPP</command> on the command line.
  </para>

<!--
    <para>
      <application>EtherApe</application> is included in the
      <filename>GNOME-PACKAGE</filename> package, which is part of the
      GNOME desktop environment. This document describes version
      &version; of <application>EtherApe</application>.
    </para>
-->
  </sect1>
```

Figura 4 – Ejemplo de código Docbook del fichero de ayuda

Para la presentación en línea de la ayuda se usa el formato HTML, utilizando un visor de HTML hecho a medida. En el futuro se espera cambiar a un formato XML

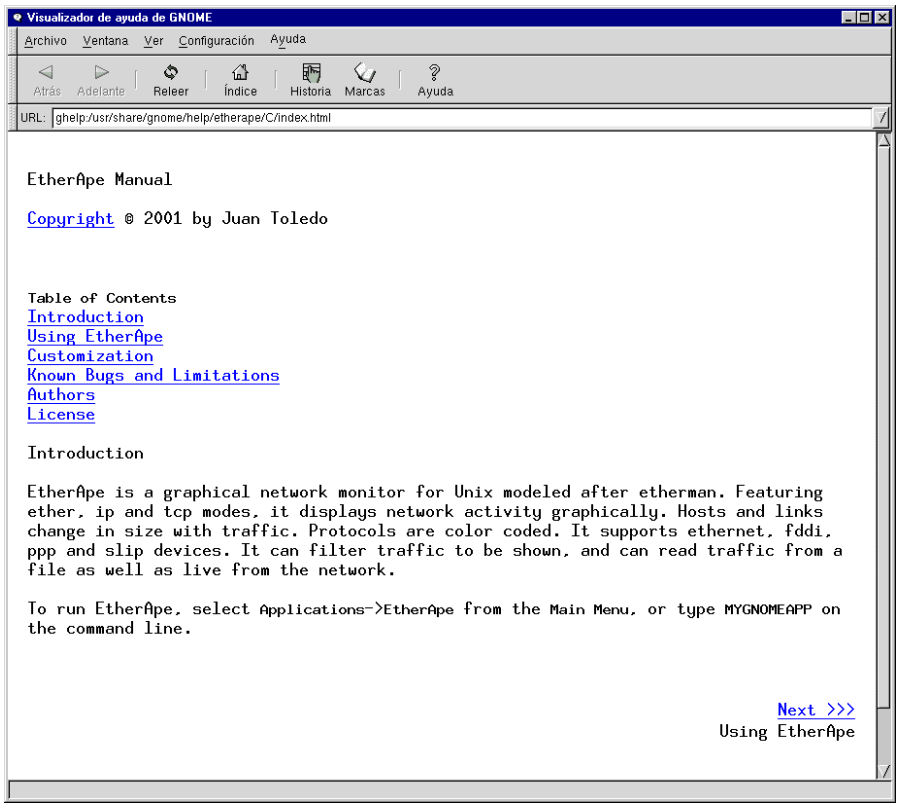


Figura 5 – Visor de ayuda de Gnome

- Almacenamiento de preferencias

  Gnome incluye un conjunto de funciones de biblioteca que permite almacenar información en un archivo específico para a aplicación.

  Su uso más inmediato es el guardar el conjunto de variables que componen las preferencias de la aplicación, de modo que una vez que el usuario ha configurado el programa a su gusto no tenga que volver a hacerlo la siguiente vez que arranque el programa.

  El funcionamiento es parecido al del registro de un sistema Windows, salvo que la información se guarda con formato ASCII en un archivo de texto específico para cada usuario en lugar de hacerlo en binario en un archivo genérico del sistema.

  De esta manera es más fácil detectar errores y administrar el sistema, a la vez que es más robusto ante la corrupción de archivos.

- Almacenamiento de historiales

  Uno de los componentes gráficos avanzados de Gnome es la entrada de texto con almacenamiento de historial. Cada vez que se introduce una entrada nueva se añade a la lista desplegable que incorpora el componente.

Además, estas entradas se van almacenado en el mismo archivo de configuración que se mencionó en el punto anterior, pero mientras que guardar las preferencias hay que hacer una llamada explícita a la biblioteca, en el caso de los historiales es un proceso automático.

EtherApe utiliza entradas de este tipo en el diálogo de carga de ficheros de captura y en la entrada de filtros de captura. La capacidad de un programa de mantener cierta memoria sobre el uso que se hace de él es altamente valorada por los usuarios, ya que obvia la necesidad de repetir pasos que ya se hayan dado en el pasado.

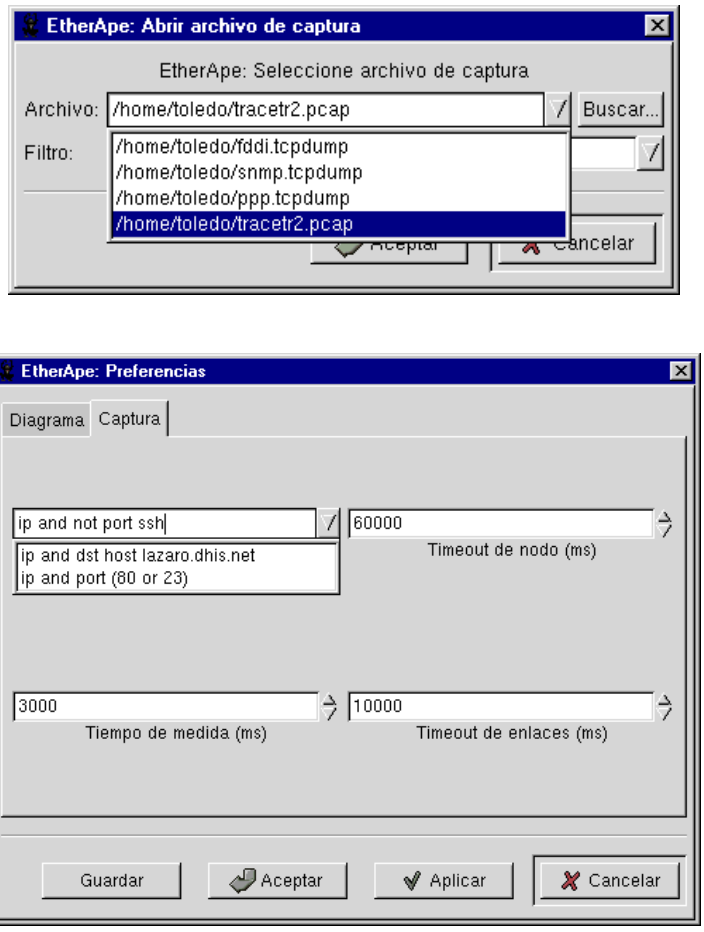


Figura 6 – Entradas de texto con historial

3.5 Glade/Libglade

Puesto que Gtk+ y Gnome son bibliotecas de C, la manera de diseñar y poner en marcha la interfaz de usuario es mediante llamadas a funciones. Este es un sistema muy poco intuitivo, tedioso y que induce fácilmente a errores.

Por cada elemento de la interfaz gráfica hay que especificar dónde se coloca, cuáles son sus atributos por defecto y cuáles y cómo son los atributos que ha de gestionar

En los Entornos Integrados de Desarrollo (IDE) modernos, el problema se resuelve mediante programas accesorios que permiten “dibujar” una interfaz de usuario y asignar propiedades a los elementos. En un paso posterior el IDE escribe el código que implementa esa interfaz, para que el programador pueda integrarlo a su programa.

```
diagram_only_toggle =
    gtk_toggle_button_new_with_label (_("Click to toggle"));
  gtk_widget_ref (diagram_only_toggle);
  gtk_object_set_data_full (GTK_OBJECT (diag_pref), "diagram_only_toggle",
                                diagram_only_toggle,
                                (GtkDestroyNotify) gtk_widget_unref);
  gtk_widget_show (diagram_only_toggle);
  gtk_box_pack_start (GTK_BOX (vbox11), diagram_only_toggle, FALSE, FALSE, 2);
  gtk_tooltips_set_tip (tooltips, diagram_only_toggle,
                            _("Toggle whether text is shown on the diagram"),
                            NULL);

  label30 = gtk_label_new (_("No text"));
  gtk_widget_ref (label30);
  gtk_object_set_data_full (GTK_OBJECT (diag_pref), "label30", label30,
                                (GtkDestroyNotify) gtk_widget_unref);
  gtk_widget_show (label30);
  gtk_box_pack_start (GTK_BOX (vbox11), label30, FALSE, FALSE, 0);
  gtk_label_set_justify (GTK_LABEL (label30), GTK_JUSTIFY_LEFT);

  vbox5 = gtk_vbox_new (FALSE, 0);
  gtk_widget_ref (vbox5);
  gtk_object_set_data_full (GTK_OBJECT (diag_pref), "vbox5", vbox5,
                                (GtkDestroyNotify) gtk_widget_unref);
  gtk_widget_show (vbox5);
```

Figura 7 – Ejemplo de código C generado por Glade

3.5.1 Glade

El programa que cumple esta función en el entorno Gnome es Glade. Mediante Glade es posible diseñar prácticamente toda la interfaz de usuario de un programa: las ventanas, los diálogos, los menús... Como si se tratara de un programa de dibujo, Glade presentan una caja de herramientas donde se encuentran todos los componentes gráficos que se pueden utilizar, tanto de Gtk+ como de Gnome.

Una vez especificado diseñada la interfaz, el programa procede a escribir los archivos en C. Por un lado están las funciones que harán que aparezca en pantalla la IGU. Por otro lado el esqueleto de las funciones que se habrán de ejecutar cuando se active alguno de los procedimientos de los componentes gráficos.

Realmente Glade va más allá, y de hecho puede escribir un programa completo que puede ser compilado y ejecutado. Incluso si el programa no necesita de procedimientos no relacionados con IGU (algo excepcionalmente raro, pero un programa que deba simplemente plantar en pantalla una serie de avisos), el programa estará completo sin necesidad de haber escrito una sola línea de código.

Sin embargo Glade también tiene algunos problemas. Conforme el programa que se está desarrollando evoluciona es más que probable que se habrán de añadir nuevas ventanas y controles. Glade fuerza a que todo el código que genera se almacene únicamente en dos archivos: interface.c para la generación de la IGU y callbacks.c para que el programador complete las funciones de rellamada de la interfaz.

Para un programa que va adquiriendo cierta envergadura este es un sistema engorroso, porque impide establecer cierto orden en la estructura del programa, separando las distintas partes que componen la interfaz. Por otro lado el algoritmo que debe detectar cuando se ha implementado ya una de las funciones de rellamada no es perfecto, y en algunos casos nos encontramos con un callbacks.c que define dos veces la misma función.

3.5.2 Libglade

Libglade aporta una ingeniosa solución al problema anterior, y representa un buen ejemplo de innovación aportado por el mundo del software libre.

Los programas que usan Libglade separan radicalmente la interfaz de usuario del resto del programa. Ni una sola línea de código se dedica a la implementación de la IGU, sino que se genera automáticamente en tiempo de ejecución a partir de un archivo de definición en formato XML generado por Glade.

Así pues, cuando EtherApe hace la llamada apropiada a Libglade, esta biblioteca carga el archivo etherape.glade y presenta la interfaz de usuario. Cuando se activa alguna función de la IGU que requiera de una función de rellamada es Libglade la que se encarga de efectuarla, sabiendo cuál es la función a la que debe llamar puesto que está especificada en etherape.glade.

Por otro lado, existen funciones que permiten al programa principal acceder a los componentes gráficos que ha generado Libglade, de modo que se puedan modificar en detalle sus propiedades en tiempo de ejecución, si es que eso es necesario.

Mediante este sistema se resuelve el problema de la estructura del código, puesto que lo que antes iba en interface.c se hace innecesario; y dado que las funciones

que antes iban en callbacks.c pueden estar definidas en cualquier parte del programa con tal de que sean funciones globables accesibles por la biblioteca en tiempo de ejecución.

Pero además Libglade hace posible algo verdaderamente sorprendente: ya no es necesario volver a compilar la aplicación para cambiar la interfaz de usuario, basta con editar con Glade el archivo de definición de la interfaz. Esto da lugar a ciclos de desarrollo aún más rápidos de los que permite Glade por sí mismo, pero también permite que sean personas diferentes las que trabajen en la interfaz y el código, o que los usuarios personalicen a su gusto la apariencia del programa con una increíble flexibilidad sin tener que cambiar en absoluto el programa.

Y por si fuera poco todo eso se consigue sin tener que hacer uso de ningún interprete. Todo sigue siendo código C compilado, con lo cual no se pierde nada en velocidad de ejecución

```
<widget>
          <class>GtkMenuBar</class>
          <name>menubar1</name>
          <shadow_type>GTK_SHADOW_NONE</shadow_type>

          <widget>
            <class>GtkMenuItem</class>
            <name>file1</name>
            <stock_item>GNOMEUIINFO_MENU_FILE_TREE</stock_item>

            <widget>
              <class>GtkMenu</class>
              <name>file1_menu</name>

              <widget>
                <class>GtkPixmapMenuItem</class>
                <name>open</name>
                <signal>
                    <name>activate</name>
                    <handler>on_open_activate</handler>
                </signal>
                <stock_item>GNOMEUIINFO_MENU_OPEN_ITEM</stock_item>
              </widget>

…
```

Figura 8 – Ejemplo de código XML del fichero etherape.glade

3.6 Ethereal

Hasta ahora se han tratado las bibliotecas más importantes con las que enlaza EtherApe para poder ser compilado. Habíamos comentado que utilizar bibliotecas externas ayuda a crear el programa más rápidamente, además de obtener un programa más robusto y fiable (siempre que las bibliotecas están bien hechas, claro está).

Sin embargo el hecho de que EtherApe sea software libre abre las puertas a un extensísimo conjunto de recursos de programación que no necesariamente tiene por qué estar disponible en forma de una biblioteca: hablamos del mismo código fuente de cualquier otro programa que sea software libre.

A pesar de que EtherApe es un programa relativamente novedoso (es la primera implementación de software libre de un monitor gráfico de red), no todos los algoritmos que debe implementar son un campo nuevo. En particular, al igual que ocurre con EtherApe, Ethereal debe identificar cuál es la pila de protocolos que transporta cada paquete para cumplir su objetivo de un análisis detallado.

Así pues, puesto que el código de Ethereal está expuesto a todo aquel que quiera detenerse a examinarlo, constituye una referencia casi más útil que los estándares que definen a los protocolos, puesto que expresa algorítmicamente el modo de identificarlos.

EtherApe obtiene un provechoso partido de esta situación, incorporando directamente a su código algunos de los procedimientos desarrollados originalmente para este otro programa. Si bien en la mayor parte de los casos no es posible copiar y pegar directamente puesto que las estructuras de datos que se manejan no son idénticas, sí que se convierte en una tarea mucho más sencilla añadir funcionalidad en este sentido.

No sólo se ha aprovechado la identificación de protocolos, también problemas menores como la representación textual de direcciones ethernet o IP a partir de su representación numérica son problemas comunes por los que resultaría inútil perder el tiempo en reimplementar.

Por otro lado hay que destacar que este comportamiento dista mucho de ser parasitismo. Muy al contrario es una simbiosis por la que todo el mundo gana. EtherApe reduce su tiempo desarrollo y aumenta en funcionalidad, y a su vez el conjunto de los usuarios de EtherApe (entre los que se incluyen los desarrolladores de Ethereal) disponen de una nueva herramienta con la que hacer su trabajo.

Puesto que la funcionalidad de Ethereal está embebida directamente en el código EtherApe, este programa no es una dependencia propiamente dicha. Es decir, no es necesario disponer de éste para poder utilizar aquel. Sin embargo, puesto que su contribución es fundamental para el éxito del proyecto se ha decidido incluirlo aquí.

3.7 Licencias del código externo

Siempre que se vaya a usar código producido externamente se ha de estar en cumplimiento de las licencias correspondientes para encontrarnos dentro de la legalidad.

En el caso de proyectos comerciales tradicionales esto implica pagar una cantidad acordada para adquirir el derecho de uso y poder enlazar la correspondiente biblioteca.

Sin embargo EtherApe se ha producido usando enteramente software libre, y por tanto las licencias de sus dependencias lo reflejan. Libpcap se distribuye bajo la licencia BSD, Glib, Gtk+, las bibliotecas de Gnome y la biblioteca estándar de C se encuentran bajo la licencia LGPL, y Ethereal se distribuye con licencia GPL.

La licencia BSD se resume básicamente en que no hay límites en el uso del software que se distribuye con ella, con tal de que en la documentación del trabajo que se produzca contenga una mención al producto en el que se basa.

La licencia LGPL (“Lesser General Public License”) es común a muchas bibliotecas de software libre. Lo que indica es que la biblioteca es en sí misma software libre, pero no se limita el uso a cualquier tipo de programa que quiera usar la funcionalidad que aporta.

Como ya se ha indicado más arriba, la licencia GPL es un tipo de licencia que permite el uso ilimitado del software, pero obliga a que trabajos derivados estén asimismo también recogidos bajo la misma licencia.

Así pues queda claro que sólo el uso de código de Ethereal fuerza a que EtherApe sea también software libre bajo licencia GPL. Sin embargo a juicio del autor hay poco que ganar manteniendo el programa como software propietario, y muchas ventajas que perder.

4 Estructura interna

EtherApe es una aplicación, que como casi todas las aplicaciones basadas en una IGU no siguen un camino lineal, sino que están controlada por eventos.

Tres son los eventos principales a los que EtherApe debe responder

- Llegada de una trama a una interfaz de red

Cuando esto ocurre el programa analiza el paquete en cuestión y actualiza todas las estructuras de datos que estén relacionadas con dicho paquete.

- Interacción con la interfaz de usuario

  En cualquier momento el usuario puede parar o iniciar una captura, cambiar la interfaz de red de donde leer los datos o finalizar el programa totalmente. En cualquiera de estos casos se debe actuar de acuerdo con el comando que se ha ejecutado.

- Vencimiento del temporizador de refresco

  El diagrama principal de red se actualiza cada cierto tiempo, fijado por el usuario. Las ventanas de información estadísticas también deben ser refrescadas, aunque lo hacen con menos frecuencia.

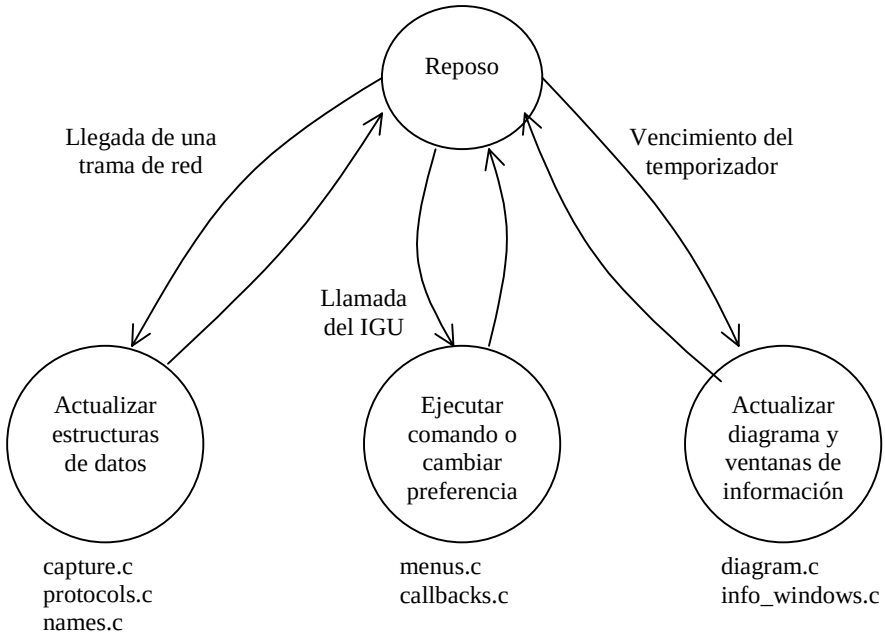


Figura 9 – Diagrama de eventos de EtherApe

En la figura se refleja también en qué parte del código se implementa la funcionalidad que requiere cada uno de los eventos.

4.1 Motor de captura

4.2 Presentación del diagrama

4.3 Información estadística

Motor de captura

Para cumplir los objetivos que se han planteado EtherApe debe ser capaz de escuchar *la red y* entender *lo que se está diciendo. La información acumulada se almacenará para su posterior presentación.*

1 Captura portable de paquetes

1.1 Adquisición de tramas

Lo mínimo indispensable para poder hacer cualquier análisis del tráfico en la red es poder adquirir información de lo que allí está ocurriendo. Se trata de conseguir que el programa obtenga una copia de cada paquete que llega a la interfaz de red.

Sin embargo por lo general las tarjetas de red están configuradas para atender exclusivamente el tráfico que le pueda concernir, esto es, aquél que le está dirigido directamente y el tráfico de broadcast. Esto se hace para minimizar la carga que se pudiera provocar en el nodo al tratar de procesar tramas que serán, en la práctica, inútiles. Si se quiere escuchar todo el tráfico, no sólo el descrito antes, se ha de configurar la interfaz de red en lo que se llama “modo promiscuo”. Eso permite al sistema operativo tener acceso a todas las tramas recibidas, y así también, por tanto, a las aplicaciones que las reclamen.

Como se puede imaginar este proceso es característico del sistema operativo que se vaya a utilizar, y también depende de la interfaz de red. La manera de resolver este problema de modo portable es utilizando la biblioteca Libpcap

Utilizando las funciones adecuadas se logra que Libpcap haga una copia de cada trama y nos proporciones un puntero a su copia. EtherApe hace un tratamiento de ese paquete y almacena sólo la información que le es interesante.

1.2 Integración en el bucle de eventos

Para que el bucle principal haga una llamada a las funciones de procesado de paquetes cuando llega uno, se ha de registrar de alguna manera en el bucle de eventos.

De manera general, Gdk (y través de él, Gtk+) permiten registrar funciones que serán llamadas cuando se cumplan ciertas condiciones en un descriptor de archivo.

En particular, cuando inicializamos libpcap para hacer una captura podemos pedir un descriptor de archivo que usaremos en la llamada de `gdk_input_add` para poder hacer el tratamiento de cada nuevo paquete que llegue.

2 Identificación de nodos

La primera tarea a realizar cuando se analiza una trama es averiguar quienes son los nodos origen y destino de la comunicación. De esta manera el tráfico que transporte esta trama será agregado a los nodos correspondientes y quedará reflejado en el gráfico.

2.1 Definición de identidad

Lo interesante en este punto es que no hay una única definición para nodo origen y destino en una trama en particular. Usando la terminología OSI, cada nivel de la pila de un nodo está manteniendo su propia conversación con otro nivel equivalente en la pila de otro nodo, que ni siquiera tiene que ser el mismo.

Así, por ejemplo, tomemos el ejemplo de un navegador de internet que está recibiendo una página web.

Si lo miramos al nivel de área local, todo el tráfico se intercambia entre el PC cliente y el encaminador de esa red local. Desde el punto de vista de redes IP, se está estableciendo una comunicación punto a punto ente el PC cliente y el nodo remoto que alberga el servidor web. Pero se puede ir más allá. En la práctica, para ver una página web es necesario hacer más de una consulta. Al menos una para descargar el documento HTML y otra por cada imagen que tenga la página. Para cada una de estas comunicaciones se establece una conexión TCP distinta, todas ellas dirigidas al puerto 80 del servidor, pero cada una iniciada desde un puerto distinto del PC cliente.

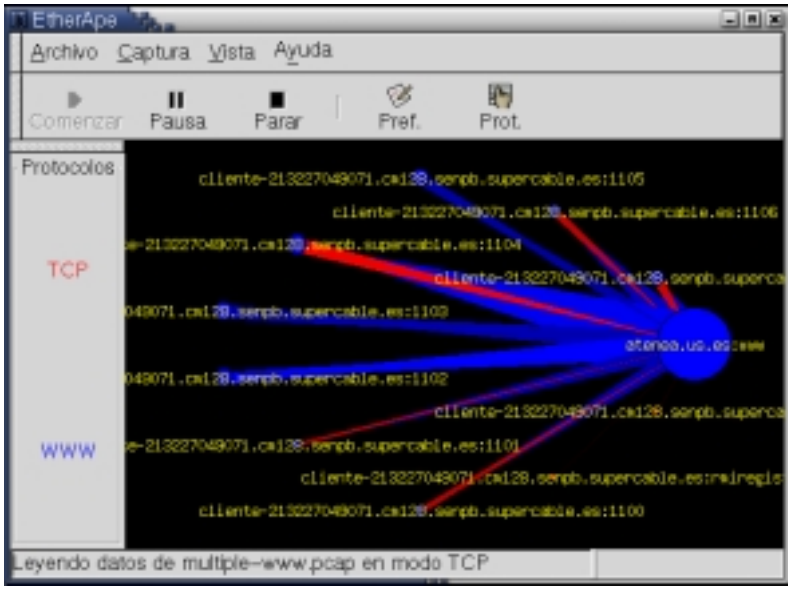


Figura 10 – Modo TCP: cada puerto es un nodo

Queda claro por tanto que en cada momento hay que tomar una decisión con respecto a como se define la identidad de los nodos. Es posible ordenar a EtherApe que actúe de una manera u otra, a través de lo que se han denominado "modos". Así pues, se puede elegir entre los modos Ethernet, IP, o TCP, para el caso anterior. Pero también existen los modos FDDI y Token Ring para redes que no están basadas en tecnología Ethernet.

Nótese que en ciertos tipos de topología, como en las redes punto a punto PPP no existe el concepto de identidades al nivel de capa de enlace, puesto que sólo puede haber un origen y un destino y cada uno sabe quién es el otro. Por eso EtherApe inhabilita los modos de capa de enlace cuando está leyendo datos de una interfaz PPP o SLIP.

2.2 Casos soportados

Ahora veremos más específicamente cómo se definen los nodos en cada modo en particular

- Ethernet

  El identificador de nodo se corresponde con la dirección hardware de la interfaz de red. En la cabecera de los tres tipos posibles de comunicación Ethernet (802.3, ETH-II y 802.2) siempre aparece la pareja de direcciones que identifican el origen y el destino. Cada una de esas direcciones es un conjunto de seis octetos.

- FDDI y Token Ring

  Si bien en estos dos casos hay información adicional, igualmente se encuentra la pareja de direcciones hardware que coincide en tamaño, aunque la posición de esta información esté situada de manera diferente a como lo está en las tramas Ethernet.

- IP

  Una vez localizada la cabecera IP en la trama, no hay más que leer las direcciones de origen y de destino, que están también siempre presentes. En este caso cada una de ellas es un conjunto de 4 octectos.

  Hay que hacer notar que si bien todas las tramas que se capturen en una interfaz de red deben albergar una cabecera que se corresponde con el nivel dos de ese medio físico, no ocurre lo mismo con los protocolos de nivel superior. Este es el caso de IP y TCP.

  Antes de intentar buscar una identidad IP a un nodo hay que asegurarse de que realmente lo que tenemos es tráfico IP. Para ello lo que se hace es que cada vez que se está utilizando uno de estos modos el programa agrega automáticamente un filtrado a la interfaz que elimina las tramas que no correspondan a ese tipo de tráfico.

- TCP

  Este es un ejemplo más complejo de definición de identidad. En los casos anteriores lo único que se hacía era copiar directamente las direcciones de origen y destino que se encontraban en las cabeceras correspondientes.

  En el caso de TCP, la cabecera sólo indica los puertos de origen y de destino, pero estos por sí solos no identifican los extremos de la comunicación. Así pues el programa crea una identificación concatenando los cuatros octetos de la dirección IP con los dos octetos adicionales del puerto TCP.

  De este modo en modo TCP cada nodo está definido por una clave de seis octetos, y la heurística que se necesita para comparar nodos entre sí es ligeramente más complicada, puesto que hay que mirar información en dos cabeceras distintas.

3 Análisis de la pila de protocolos

Además de la representación gráfica de la información, la segunda característica que hace que EtherApe se destaque entre otras alternativas para monitorizar una

red es su capacidad de hacer un análisis completo de la pila de protocolos de cada paquete.

Es común que otras aplicaciones distingan entre varios tipos fundamentales de tráfico, separando por ejemplo el tráfico IP del IPX, o distinguiendo entre ICMP, UDP y TCP en el tráfico IP. EtherApe va mucho más allá tratando de distinguir completamente para qué aplicación concreta se está enviando cada paquete que circula por a red.

Como se comentó antes, es aquí donde el estudio del código fuente de Ethereal ha sido instrumental. Ethereal debe reconocer cada protocolo y presentar de manera legible cada uno de los campos que componen la cabecera de todos los protocolos.

Puesto que EtherApe sólo debe identificar los protocolos, no destripar su contenido, basta con replicar los algoritmos de identificación.

3.1 Sistemas de identificación

No todos los casos son iguales a la hora de identificar un protocolo. El sistema varía según la altura de la pila de protocolos en la que nos encontremos, y en algunos casos es prácticamente imposible implementar un esquema que nos dé la certeza sobre el protocolo que se está usando.

- Protocolos de nivel de enlace

  Cuando se pide a libpcap que inicie una captura sobre una interfaz de red o un archivo de captura, libpcap devuelve una estructura que contiene información sobre la captura que se va a realizar. Entre otras cosas un campo de esa estructura nos indica qué medio físico soporta la red que vamos a monitorizar.

  En la mayoría de los casos cada medio físico utiliza sólo un protocolo de nivel de enlace. De esta manera se puede identificar directamente la presencia (o ausencia en su caso) del protocolo de nivel 2.

  En las redes ethernet sí hay tres posibilidades distintas, que se corresponden con las tres variedades de tramas ethernet que existen, 802.2, 802.3 y Ethernet-II. Sin embargo hay métodos para identificar cada caso unívocamente, sin posibilidad de error, y por tanto no es un problema.

  Si es más importante identificarlos en lo que respecta al análisis de los niveles superiores, puesto que la heurística es diferente en cada caso.

- Identificación definida por estándar

  Este sistema es igualmente sencillo en los casos en que los protocolos tienen reservados un campo para identificar el protocolo de nivel superior que transportan.

  Ejemplos de este caso son las tramas Ethernet-II (que pueden transportar IP e IPX, por ejemplo) o las cabeceras de IP (ICMP, TCP, UDP, IGMP, GRE, OSPF, EGP, etc.)

- Puertos TCP y UDP asignados

  En las cabeceras de TCP y UDP no hay un estándar que defina taxativamente el tipo de tráfico que transportan. Sin embargo el IANA (Internet Assigned Numbers Authority) sí ha hecho un esfuerzo de normalización para ayudar a la interoperabilidad de las aplicaciones de red.

  Los 65536 puertos posibles están divididos en tres grupos: puertos de sistema (0-1023), puertos registrados (1024-49151) y puertos dinámicos o privados (49152-65535).

  En los dos primeros casos IANA publica una lista de servicios que pueden estar escuchando en cada puerto. En problema es que si bien es casi seguro que una conexión TCP a un puerto 80 sea de HTTP, no es forzosamente cierto, sobre todo con los puertos mayores que 1024, que pueden ser abiertos por cualquier usuario sin privilegios.

- Heurística

  Si los mecanismos anteriores no funcionan todavía es posible ir probando si la estructura del paquete se corresponde con algún protocolo conocido. Este comportamiento, que sí está implementado en Ethereal, aún no lo está en EtherApe.

3.2 Identificación de conversaciones

En algunos casos, una vez que se ha establecido una comunicación entre dos entidades en nodos distantes, ambos mantienen un estado que hace innecesario referir explícitamente en cada paquete algunos parámetros de la comunicación.

Esto ocurre por ejemplo en los mensajes de respuesta de los protocolos basados en RPC. Sólo en la petición se indica cuál es el servicio que se va a utilizar, pero en la respuesta no se hace, puesto que se considera implícito.

Igualmente ocurre con el funcionamiento pasivo del protocolo FTP. En un momento dado el servidor indica al cliente que ha abierto un puerto adicional

aleatorio (no registrado) para atender exlusivamente una transferencia de archivos con ese cliente. En este caso indentificar tramas sucesivas por el número de puerto puede dar lugar a error.

EtherApe resuelve estos problemas porque mientras que analiza los paquetes está buscando paquetes que indiquen el inicio de una conversación. Si los encuentra, almacena los parámetros de esa conversacion, y luego puede usar esos datos para tratar de indetificar el protocolo de subsiguientes tramas.

3.3 Resultado del análisis

4 Extracción de nombres

Hasta ahora hemos estado tratando las identidades de los nodos como un concepto interno destinado a localizar el origen y el destino del tráfico que se cursa. Sin embargo también es necesario informar al usuario de esos mismo detalles, pero si es posible hacerlo de una manera que sea más comprensible para una persona que un conjunto de octetos.

4.1 Definición de nombre

Para este proyecto un nombre es un pareja de cadenas alfanuméricas que se asocian a un protocolo. La primera cadena es la representación numérica de una identidad. La segunda es la representación más fácilmente reconocible para el usuario de esa misma identidad. Esta definición general es compleja, y se entenderá mejor lo que se quiere decir con un ejemplo.

Para una comunicación entre dos nodos IP, la identificación que se usa es un conjunto de cuatro octetos. EtherApe guarda esta identificación internamente en una palabra de 32 bits, digamos *0xC193A094*. La versión **numérica** del nombre para este caso es *193.147.160.148*, mientras que la versión **resuelta** es *www.esi.us.es*.

4.2 Multiplicidad de nombres

Mientras que para un modo en particular un nodo sólo tiene una única identidad, ese mismo nodo puede estar relacionado con un extenso número de nombres, por dos motivos distintos.

Por un lado es posible asignar un nombre al nodo casi a cada nivel de la pila de protocolos. Por ejemplo, para un ordenador conectado a una red ethernet hay que pensar el la dirección ethernet, en la dirección IP, y en la identificación propia de un nodo de una red Microsoft para compartir archivos.

Por otro lado, incluso para un mismo protocolo es probable que un mismo nodo se identifique con más de un nombre. Así, por ejemplo, trabajando en el modo IP, ese nodo está relacionado con múltiples nombres TCP (combinación de dirección IP y puerto TCP)

Por todo ello a la hora de diseñar las estructuras de datos se ha tenido en cuenta esta situación para dotarlas de la suficiente flexibilidad para poder almacenar toda esta información con éxito.

4.3 Resolución de nombres

La coversión entre el identificador y la pareja de nombres que le corresponde es específica para cada protocolo. Veremos algunos de los casos.

- Direcciones Ethernet

  El nombre numérico se representa utilizando el formato más común: separando cada octeto representado en hexadecimal por dos puntos

  El nombre resuelto se obtiene a partir de tablas de traducción si es que estas están instaladas en el sistema. EtherApe busca en primer lugar el archivo /etc/ethers para intentar traducir la dirección a un nombre reconocible. Si no está disponible busca a continuación /etc/manuf, que es un archívo genérico que relacion los tres primeros octectos de la dirección con el fabricante de la tarjeta.

- Direcciones IP

  El nombre numérico es el habitual. Cuatro números decimales separados por un punto.

  La resolución completa del nombre tiene sus propios problemas añadidos. Es posible pedir al sistema que nos informe del nombre que se corresponde con una direción IP, pero esta función bloquea hasta que se obtiene un resultado.

  Si se hiciera directamente, entonces el programa se quedaría congelado cada vez que apareciera una nueva dirección IP, y esto es lo contrario de lo que se pretende con el programa: la capacidad de poder echar un vistazo de manera inmediata a lo que ocurre en la red.

  El modo de resolverlo es introduciendo funciones que resuelvan asíncronamente. De ese modo EtherApe no resuelve completamente el nombre hasta que el servidor de DNS no está preparado para ello. Esta funcionalidad está implementada en dns.c, y es otro ejemplo de las ventajas

del software libre, puesto que es una adaptación directa de la implementación original que se hizo para mtr[1].

- Direcciones TCP

  Es un caso casi idéntico al anterior. A la dirección IP obtenida se le agregan dos puntos y el número de puerto. En la versión resuelta se consulta el número de puerto en el fichero de sistema /etc/services para buscar el nombre del servicio.

- Nombre NetBIOS o SMB

  En los protocolos que usan NetBIOS, el nombre de cada una de las estaciones que participan en el protocolo se explicita en cada trama, y únicamente es necesario extraerla. El único inconviente es que en el caso de NetBIOS el nombre no se transmite como texto en claro, sino que está codificado y es necesario traducirlo.

  Ese mismo nombre se transmite en ocasiones para ciertas funciones del protocolo SMB. En este caso el nombre se transmite en claro, así que es más fácil recuperarlo.

  Además del nombre, en ambos casos el último octecto es un código que se utiliza para especificar la función que tiene ese nodo dentro de la red. Puesto que en SMB un nodo puede tener más de una función (cliente, servidor, servidor de nombre, comprobación de identificaciones) el mismo nodo tiene un conjunto de nombres SMB asociadas.

  En este caso no es posible dar una representación estrictamente numérica, así que se limita al nombre sin tener en cuenta el código de función.

  En la representación resuelta sí se tiene en cuenta ese código, y se representa textualmente su significado.

4.4 Selección del nombre principal

5 Estructuras de datos

[1] El programa mtr es una versión avanzada del clásico traceroute. Mantiene un estadística constante sobre cada punto de una ruta, y no bloquea mientras está intentado resolver el nombre de cada encaminador. La página web de mtr está en http://www.bitwizard.nl/mtr/

Interfaz de usuario

La peculiaridad que caracteriza a EtherApe es la representación gráfica en tiempo real de los datos capturados. Un diagrama fácilmente reconocible es el objetivo fundamental del programa. Sin embargo no es la única información que EtherApe presenta al usuario, también hay estadísticas de datos agregados. Todo esto, junto con las opciones de configuración de EtherApe es lo que define la interfaz con el usuario.

1 El diagrama de red

El diagrama de red es el componente más destacado de la interfaz de usuario, así como el objetivo fundamental del programa. A partir de los datos capturados se presenta en pantalla un diagrama que traduce de manera gráfica el tráfico en la red.

Los círculos representan a cada uno de los nodos identificados y las líneas que los unen es el tráfico cursado entre ellos. El tamaño de los elementos varía según el ancho de banda que se esté usando, y el color de los elementos se traduce en el protocolo más utilizado, bien por el nodo, bien por un enlace en particular.

1.1 El componente canvas de Gnome

El motivo fundamental por el que EtherApe es una aplicación del entorno Gnome es por la facilidad que su componente canvas aporta a la implementación del diagrama.

The GNOME canvas is a high-level engine for creating structured graphics. This means the programmer can insert graphical items like lines, rectangles, and text into the canvas, and refer to them later for further manipulation. The programmer does not need to worry about repainting these items or generating events for them; the canvas automatically takes care of these operations.

El canvas de Gnome es un motor de alto nivel para la creación de gráficos estrucutrados. Esto quiere decir que el programador pueder insertar elementos gráficos como líneas, rectángulos y texto dentro del canvas, y referirse luego a elos para manipularlos. El programador no debe preocuparse con el redibujado de los elementos, o de la generación de eventos asociados, puesto que el canvas lo hace automáticamente.

Además de los elementos descritos, el canvas permite que se creen nuevos elementos, y agrupar un conjunto de elementos para tratarlo como uno sólo. No es de extrañar por tanto que el canvas tenga un enorme éxito entre los programadores de aplicaciones para Gnome.

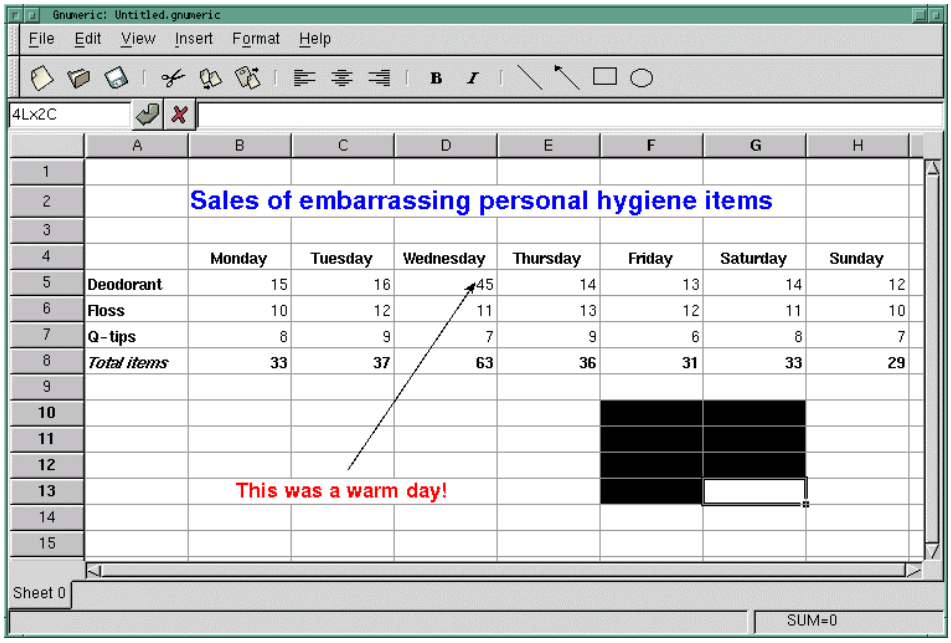


Figura 11 – Ejemplos de canvas: Calendar y Gnumeric

El canvas implementa conceptos complejos, como la multiplicidad de sistemas de referencia para localizar los elementos o la técnica del “doble buffering” para evitar parpadeos.

1.2 Elementos que componen el diagrama

Desde el punto de vista del usuario sólo son dos los tipos de elementos del diagrama: los nodos y los enlaces. Sin embargo su estructura interna es ligeramente más compleja.

1.2.1 Nodos

En realidad un nodo está formado por un elemento grupo que engloba a un elemento elipse y a un elemento de texto.

Las propiedades del conjunto que se manipulan después de su creación son la posición del conjunto, el color del círculo, el radio del círculo y el color del texto. Adicionalmente también es posible hacer invisible por completo el texto.

La utilidad del elemento grupo es la de tomarlo como origen de referencia para el posicionamiento de la elipse y el texto. De este modo basta con redefinir la posición del elemento grupo con respecto del canvas para mover el conjunto.

En cada refresco del diagrama se revisan las propiedades de cada nodo, y si alguna de ellas ha cambiado se pone en cola un refresco del diagrama (de hecho

espera a haber actualizado todos los nodos y enlaces antes de redibujar el canvas).

En cada pase se comprueba cuál es la cantidad de tráfico a medir y el protocolo principal del nodo de captura asociado para actualizar el radio y color respectivamente de cada nodo de canvas.

Asimismo se actualiza el texto identificativo con el del nombre más representativo que el nodo tenga especificado en cada momento. Se comprueba además cual es el color de texto designado en ese momento y si se ha indicado que el elemento de texto debe hacerse invisible.

1.2.2 Enlaces

Si se analiza el diagrama con anteción se observará que los enlaces no son rectángulos que unen dos nodos. En lugar de eso se verá que son triángulos, donde un nodo está colocado en el centro de la base y otro en el vértice opuesto.

Esto es así porque durante la captura se ha utilizado una estructura enlace para *cada sentido* de la comunicación. De esta manera se mantiene una estadística más detallada sobre las carácterísticas de la comunicación y se hace aparente en qué sentido se está moviendo más informacion en el diagrama.

En cada actualización la posición del triángulo se actualiza siguiendo como referencia la posición de los elementos grupo de los nodos asociados, y el color se corrige para indicar el protocolo más utilizado.

En el diagrama de ejemplo se observa claramente la distinción entre los dos sentidos de una comunicación. En este caso el cliente está recuperando una página web de un servidor remoto. Se aprecia cómo hay mucho mayor cantidad de tráfico descendente. Por otro lado el color nos indica que mientras que del servidor recibimos mayoritariamente tráfico describiendo la página web (protocolo HTTP), el cliente envía sobre todo tráfico TCP vacío de contenido, que este caso se corresponde con los asentimientos asociados a cualquier conexión TCP.

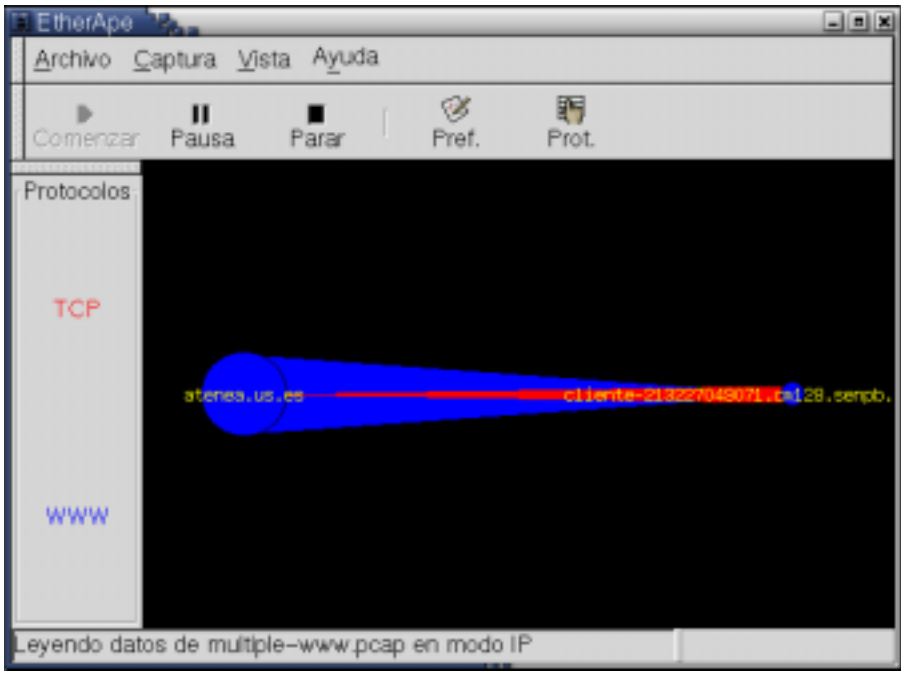


Figura 12 – Ejemplo de comunicación asimétrica

1.2.3 La Leyenda

La leyenda se genera automáticamente conforme se capturan paquetes, asignando los colores disponibles a los protocolos identificados siguiendo una estructura round-robin.

En este momento este es uno de los principales problemas de la interfaz de usuario de EtherApe, puesto que en cada ejecución un mismo protocolo puede tomar un color diferente, y pasado un tiempo se habrán identificado tantos protocolos que la asignación de colores no será unívoca.

Este problema es nuevo en los programas de la clase de EtherApe, puesto que implementaciones anteriores no eran capaces de reconocer tan variedad de protocolos, y en consecuencia era sencillo hacer una asignación estática de colores.

Una posibilidad sería asignar un conjunto cerrado de colores a los protocolos que se consideren más importantes (por ejemplo, HTTP, DNS, NFS), dejando la atribución de los demás a una asignación circular. Sin embargo esto tiene el inconveniente de que cada usuario considerará principales un conjunto diferente de protocolos, con lo que probablemente no habremos arreglado nada.

Así pues la situación se mantendrá hasta que en una revisión posterior del programa se haya desarrollado una solución que permita al usuario seleccionar complentamente cómo desee que actúe el programa, leyendo sus preferencias

bien de un archivo de configuración, bien mediante un diálogo hecho expresamente a ese efecto.

1.2.4 La barra de estado

1.3 Gestión de eventos del canvas

1.4 Estructuras de datos

Lo más importante a destacar es el hecho de que las variables que se usan en el diagrama duplican en muchos casos a las variables que se utilizan durante la captura.

Esta decisión se tomó bien temprano en el diseño, desde el momento que quedó patente que el sistema de captura era lo suficientemente genérico como para ser de utilidad no solo para el sistema de representacion que utiliza EtherApe, sino para otras posibles herramientas de análisis de red.

Por este motivo, previendo que un futuro el motor de captura se convierta en una biblioteca independiente el resto del programa trata de comportarse como si de hecho ya lo fuera, de modo que se facilite una más que probable transición en el futuro.

Así pues nos encontramos con dós árboles binarios adicionales, canvas_nodes y canvas_links, que soportan el grueso de la información que controla el diagrama.

2 Las ventanas de información estadística

Conforme el programa fue evolucionando, el motor de captura se iba haciendo potente, almacenando más información sobre las características de tráfico de la red de lo que es posible representar en un diagrama.

El diagrama cumple su función de “foto” de lo más importante del estado actual mostrándonos por ejemplo el nombre más representativo o el protocolo más usado. De esta manera se ignora una gran cantidad de información que se ha ido recopilando y para la que hay que buscar algún modo de presentación.

Es este un trabajo que aún no está terminado. Aún queda información registrada a la que no es posible acceder, pero con lo que ya está implementado se sientan las bases para una sistema genérico de representación de información.

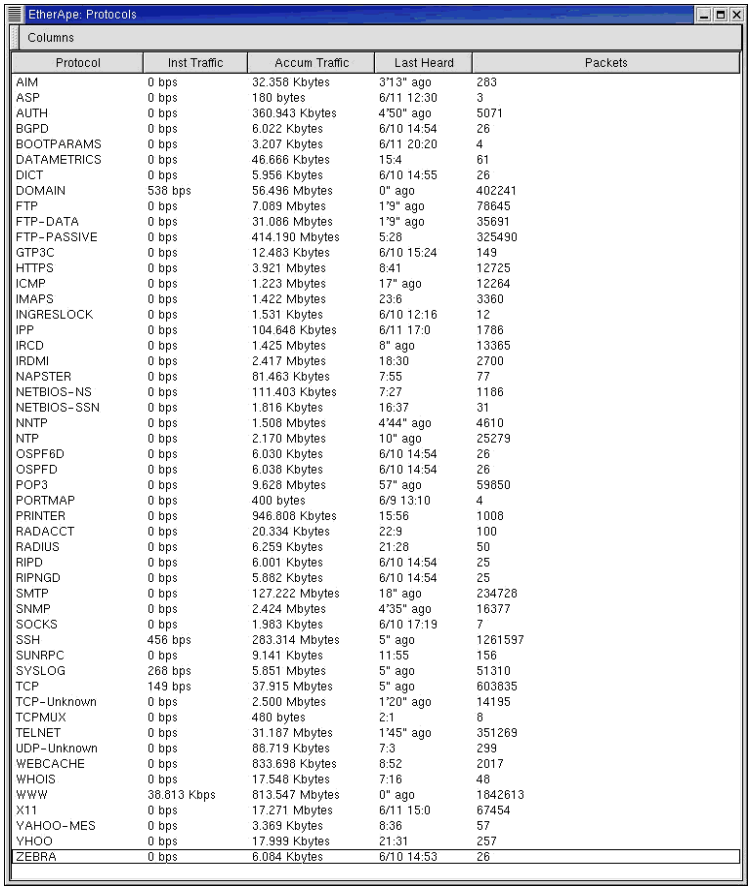
2.1 Ventana de protocolos

El primer resultado de este esfuerzo por ampliar el conjunto de información que presenta el programa es la ventana de protocolos. Aquí se presenta la informacion que no es específica de ningún nodo, sino que engloba todo el tráfico que se ha escuchado en la red.

Para cada paquete que escucha, el motor de captura analiza su pila de protocolos, y luego añade la información pertinente de ese paquete a lista global de protocolos que se va guardando.

Analizando esta información es posible presentar la tabla que se muestra en la ventana de protocolos, donde se incluyen los siguientes campos:

- *Nombre*. Es posible agregar todos los paquetes que utilizan un puerto TCP o UDP no identificado bajo sendos nombres genéricos TCP-Unknown y UDP-Unknown. La segunda opción es identificarlo con el número de puerto, e.g. TCP-2454.
- *Tráfico instantáneo*. Se calcula haciendo media a lo largo de un periodo definido por el “Tiempo de medida” (seleccionable por el usuario).
- *Tráfico acumulado*. Cantidad de tráfico acumulada desde que se inició la medida. Se resetea cuando se para una captura.
- *Última vez*. Índica en qué momento fue escuchado ese protocolo por última vez. Hasta 10 minutos da un valor relativo (hace 2’35”), luego un valor absoluto. Cuando han pasado más de 24 horas indica también el día.
- *Paquetes*. Proporciona una métrica diferente del impacto de un protocolo en la red. En muchas ocasiones un protocolo que esté generando una tormenta de broadcast no está mandando gran cantidad de tráfico, sino gran cantidad de paquetes que pueden ser pequeños. El que los paquetes sen más o menos grandes no afecta al hecho de que todas las estaciones están procesando todos los paquetes que se hayan recibido, y ese es el problema.



Protocol	Inst Traffic	Accum Traffic	Last Heard	Packets
AIM	0 bps	32.358 Kbytes	3"13" ago	283
ASP	0 bps	180 bytes	6/11 12:30	3
AUTH	0 bps	360.943 Kbytes	4"50" ago	5071
BGPD	0 bps	6.022 Kbytes	6/10 14:54	26
BOOTPARAMS	0 bps	3.207 Kbytes	6/11 20:20	4
DATAMETRICS	0 bps	46.666 Kbytes	15:4	61
DICT	0 bps	5.956 Kbytes	6/10 14:55	26
DOMAIN	536 bps	56.496 Mbytes	0" ago	402241
FTP	0 bps	7.089 Mbytes	1"9" ago	78645
FTP-DATA	0 bps	31.086 Mbytes	1"9" ago	35691
FTP-PASSIVE	0 bps	414.190 Mbytes	5:28	325490
GTP3C	0 bps	12.483 Kbytes	6/10 15:24	149
HTTPS	0 bps	3.921 Mbytes	8:41	12725
ICMP	0 bps	1.223 Mbytes	17" ago	12264
IMAPS	0 bps	1.422 Mbytes	23:6	3360
INGRESLOCK	0 bps	1.531 Kbytes	6/10 12:16	12
IPP	0 bps	104.648 Kbytes	6/11 17:0	1786
IRCD	0 bps	1.425 Mbytes	8" ago	13365
IRDMI	0 bps	2.417 Mbytes	18:30	2700
NAPSTER	0 bps	81.463 Kbytes	7:55	77
NETBIOS-NS	0 bps	111.403 Kbytes	7:27	1186
NETBIOS-SSN	0 bps	1.816 Kbytes	16:37	31
NNTP	0 bps	1.508 Mbytes	4"44" ago	4610
NTP	0 bps	2.170 Mbytes	10" ago	25279
OSPF6D	0 bps	6.030 Kbytes	6/10 14:54	26
OSPFD	0 bps	6.038 Kbytes	6/10 14:54	26
POP3	0 bps	9.628 Mbytes	57" ago	59850
PORTMAP	0 bps	400 bytes	6/9 13:10	4
PRINTER	0 bps	946.808 Kbytes	15:56	1008
RADACCT	0 bps	20.334 Kbytes	22:9	100
RADIUS	0 bps	6.259 Kbytes	21:28	50
RIPD	0 bps	6.001 Kbytes	6/10 14:54	25
RIPNGD	0 bps	5.882 Kbytes	6/10 14:54	25
SMTP	0 bps	127.222 Mbytes	18" ago	234728
SNMP	0 bps	2.424 Mbytes	4"35" ago	16377
SOCKS	0 bps	1.983 Kbytes	6/10 17:19	7
SSH	456 bps	283.314 Mbytes	5" ago	1261597
SUNRPC	0 bps	9.141 Kbytes	11:55	156
SYSLOG	268 bps	5.851 Mbytes	5" ago	51310
TCP	149 bps	37.915 Mbytes	5" ago	603835
TCP-Unknown	0 bps	2.500 Mbytes	1"20" ago	14195
TCPMUX	0 bps	480 bytes	2:1	8
TELNET	0 bps	31.187 Mbytes	1"45" ago	351269
UDP-Unknown	0 bps	88.719 Kbytes	7:3	299
WEBCACHE	0 bps	833.698 Kbytes	8:52	2017
WHOIS	0 bps	17.548 Kbytes	7:16	48
WWW	38.813 Kbps	813.547 Mbytes	0" ago	1842613
X11	0 bps	17.271 Mbytes	6/11 15:0	67454
YAHOO-MES	0 bps	3.369 Kbytes	8:36	57
YHOO	0 bps	17.999 Kbytes	21:31	257
ZEBRA	0 bps	6.084 Kbytes	6/10 14:53	26

Figura 13 – Ventana de protocolos

La tabla se implementa usando el componete GTK-Clist (de lista en columnas). Una ventaja de este componente es que permite ordenar la tabla de muchas maneras diferentes, si se proporciona en cada caso una función que compare los datos de dos filas arbitrarias.

De este modo se han programado 5 funciones de comparación diferentes, de modo que haciendo click en la cabecera de cada columna la tabla queda ordenada para ese campo en particular.

2.2 Ventana de información de nodo

Si bien es posible pedirle a EtherApe que utilice seis medidas distintas para calcular el radio del nodo, en ningún caso es posible ver toda esa información a la vez.

Por otro lado en el diagrama siempre se muestra el nombre resuelto, si es que esta disponible, y hay casos en los que resulta útil saber ese nombre no resuelto.

Para mostrar esta información debe hacerse doble click sobre un nodo en diagrama. El código de gestión de eventos de diagrama reconoce este hecho y abre la ventana correspodiente.

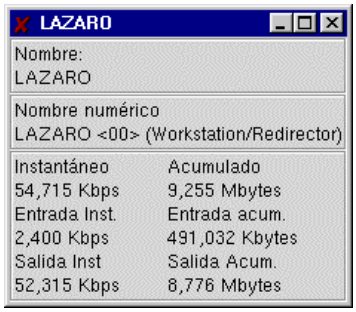


Figura 14 – Ventana de información de nodo

EtherApe sabe en realidad mucho más sobre un nodo de lo que aparece en esta ventana sencilla. En sucesivas versiones debería ser posible desplegar su árbol completo de protocolos conocidos, y mostrar asimismo todos los nombres que ha usado.

En estos momentos es posible tener acceso a partede esa información, usando los mecanismos de depuracion a través de la salida estándar de la consola. A continuación se presenta un ejemplo de sesión. La salida a partir de la línea “NODE LAZARO INFORMATION” es la información buscada, que sale como resultado de llamar la ventana de información de ese nodo desde el interfaz gráfico.

```
lazaro:~# export DEBUG=INFO
lazaro:~# /usr/bin/etherape
INFO: Dispositivos de captura disponibles: eth0 eth1 lo
INFO: Live device eth0 opened for capture. pcap_fd: 5
INFO: Link type is Ethernet
INFO: Diagrama iniciado
INFO: Reading TCP and UDP services from /etc/services
INFO: DDP protocols not supported in rtmp  1/ddp    # Routing Table Maintenance
Protocol
INFO: DDP protocols not supported in nbp  2/ddp    # Name Binding Protocol
INFO: DDP protocols not supported in echo  4/ddp    # AppleTalk Echo Protocol
INFO: DDP protocols not supported in zip  6/ddp    # Zone Information Protocol
INFO: Nodes: 23. Canvas nodes: 15
INFO: NODE LAZARO INFORMATION
INFO: Protocol level 1 information
INFO:   Protocol ETH_II
INFO:           Name: LAZARO-RJ45
INFO: Protocol level 2 information
INFO:   Protocol IP
INFO:           Name: cliente-213227048220.cm128.senpb.supercable.es
INFO: Protocol level 3 information
INFO:   Protocol TCP
INFO:           Name: cliente-213227048220.cm128.senpb.supercable.es:1372
cliente-213227048220.cm128.senpb.supercable.es:1371
cliente-213227048220.cm128.senpb.supercable.es:1370
cliente-213227048220.cm128.senpb.supercable.es:1369
cliente-213227048220.cm128.senpb.supercable.es:1368
cliente-213227048220.cm128.senpb.supercable.es:netbios-ssn
cliente-213227048220.cm128.senpb.supercable.es:1367
cliente-213227048220.cm128.senpb.supercable.es:1365
cliente-213227048220.cm128.senpb.supercable.es:1366
cliente-213227048220.cm128.senpb.supercable.es:ssh
INFO:   Protocol UDP
INFO: Protocol level 4 information
INFO:   Protocol X11
INFO:   Protocol RPC
INFO:   Protocol SSH
INFO:   Protocol UNKNOWN
INFO:   Protocol NETBIOS-SSN
INFO:           Name: LAZARO
INFO:   Protocol NTP
INFO:   Protocol DOMAIN
INFO:   Protocol POP3
INFO: Protocol level 5 information
INFO:   Protocol UNKNOWN
INFO:   Protocol X11
INFO:   Protocol SMB
INFO:   Protocol UDP-Unknown
INFO:   Protocol NTP
```

Figura 15 – Información de nodo en consola

Puede observarse cómo para cada nivel de la pila de protocolos se muestran los protocolos utilizados y cada uno de los nombres utilizados a este nivel (Ethernet: LAZARO-RJ45, IP: cliente-213227048220.cm128.senpb.supercable.es, NetBIOS: LAZARO, etc.) .

La versión final que se haga en la interfaz gráfica deberá también incluir la misma información presente en la ventana de protocolos.

2.3 Ventana de información de protocolo

A partir de la ventana de protocolos es posible presentar una ventana específica para un único protocolo.

En estos momentos su utilidad es limitada, puesto que no da más información que lo que aparece en la misma ventana de protocolos. En todo caso su valor reside pemitir al usuario ceñirse al protocolo en el que está interesado liberar espacio de pantalla para otros usos.

Figura 16 – Ventana de información de protocolo

El espacio de la parte inferior de la ventana está reservado para dar una información detallada sobre todos los nodos que han utilizado ese protocolo, de manera simétrica a lo que deberá ocurrir con la ventana de información de nodo.

3 Elementos activos de la interfaz de usuario

Son elementos activos a aquellos que están presentes para permitir que usuario. Si bien tanto el diagrama como la ventana de protocolos cumplen esta definición puesto que hacer doble click hace que aparezcan nuevas ventanas, nos centraremos en aquellos que están expresamente pensados para cumplir esta función.

3.1 La barra de menús

Los sistemas de menús son el modo más común de interactuar con una aplicación gráfica.

Gnome define un estándar de cómo se han de distribuir los menús para hacer que para un usuario novato usar aplicaciones de este entorno se más intuitivo. Ejemplos de esto es el menú *Archivo*, que ha de estar el primero a la izquierda, o el menú *Ayuda*, que es el último por la derecha, pero no está alineado a la derecha.

Veamos con detalle las distintas posibilidades.

3.1.1 El menú Archivo

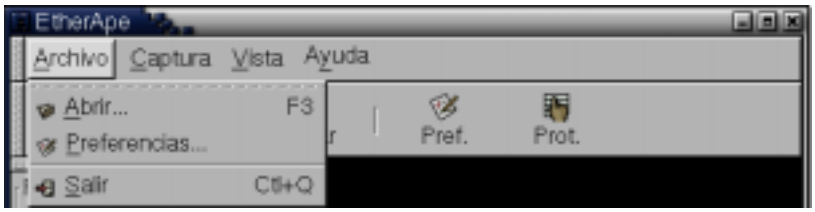


Figura 17 – El menú Archivo

- *Abrir*. Abre la ventana de selección de archivo para cargar un archivo de captura. Una vez seleccionado finaliza la captura actual y comienza leer del archivo.
- *Preferencias*. Abre el diálogo de preferencias.
- *Salir*. Finaliza la captura en curso y cierra la aplicación. Es importante asegurarse de que se han cerrado correctamente los procesos de libpcap, o se corre el riesgo de dejar la interfaz de captura en modo promíscuo.[2]

Por ser el menú más común, es el que más especificado está para las aplicaciones Gnome. El comando Salir debe estar aquí presente, y si existe el comando Abrir, también.

De hecho los tres comandos de este menú son Stock Items, pertenecen al estándar Gnome. Entre otras cosas, esto hace que los comandos estén automáticamente traducidos a cualquier idioma.

3.1.2 El menú Captura

Desde aquí se controlan los parámetros principales de la captura.

[2] Es importante asegurarse de que no se va a dejar la interfaz de red en modo promíscuo, por ello la aplicación es capaz de interceptar la señal de interrupción (que se produce cuando el usuario hace Control-C desde la consola, por ejemplo) para asegurarse de que libpcap tiene la oportunidad de cerrar de manera limpia.

Figura 18 – Submenú de modo

- *Submenú de modo*. Permite seleccionar cualquiera de los modos de catptura disponibles en la interfaz de captura que esté activa en ese momento. Cuando se cambia de modo se reinicializa el estado del programa, puesto que la definición de los nodos cambia y los datos ya no pueden seguir agregándose.

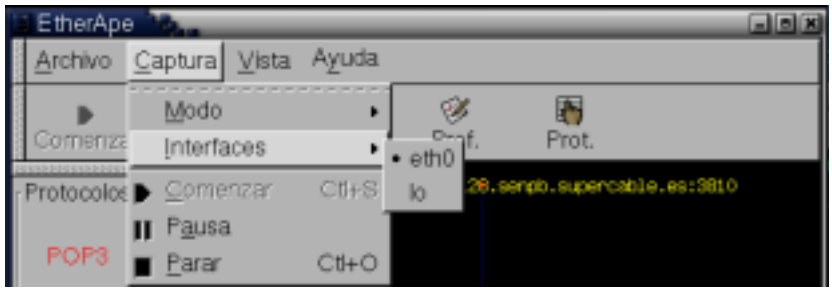


Figura 19 – Submenú de interfaces

- *Submenú de interfaces*. Mediante el uso de funciones de servicio de libpcap, EtherApe detecta todos los interfaces del sistema con los cuales es posible hacer una captura en vivo. Cambiar la interfaz también finaliza la captura anterior y resetea las estadísticas.
- *Comenzar*. Empezar a leer datos de la interfaz de lectura o de un archivo de captura.
- *Pausa*. Cuando se está leyendo de una interfaz en vivo, este comando congela el diagrama, pero se siguen procesando los paquetes que llegan de modo que cuando se siga adelante mediante el comando Comenzar los datos estadísticos se actualicen y sigan siendo válidos.

  Cuando se lee de un archivo de captura, se congelan tanto el diagrama como la captura. Al final del archivo de captura el programa se pone en pausa automáticamente para permitir analizar las ventanas de estadísticas.

- *Parar*. Finaliza la captura actual y libera la interfaz de red, si se estaba usando. Después de parar el programa mantiene su estado: la interfaz de captura o el fichero de lectura, y el modo que se esté utilizando.

3.1.3 El menú Vista

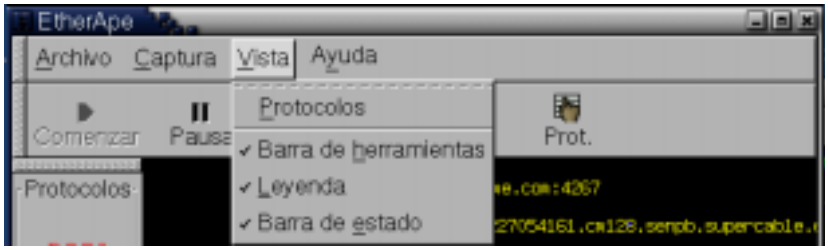


Figura 20 – Menú Vista

- *Protocolos*. Abre la ventana de protocolos
- *Barra de herramientas*. Permite seleccionar si la barra de herramientas está o no presente, de modo que se pueda hacer más hueco para el diagrama propiamente dicho.
- *Leyenda*. Ditto, para la tabla de la leyenda.
- *Barra de estado*. Ditto, para la barra de estado.

3.1.4 El menú de Ayuda

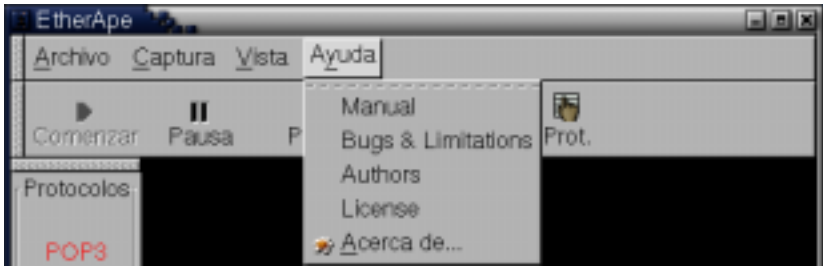


Figura 21 – Menú de Ayuda

El estándar Gnome especifican que todas las aplicaciones deben incluir un menú de ayuda, que incluya como mínimo enlaces a la ayuda en línea y el comando Acerca de...

- *Páginas de ayuda*. Estos comandos de menú no se crean de la misma manera de los demás, sino que las rutinas de inicialización de Gnome las añaden automáticamente a partir del fichero

/usr/share/gnome/help/etherape/C/topic.dat, que se distribuye junto con la aplicación.

De esta manera se facilita el que los encargados de la documentación puedan trabajar independientemente de los programadores.

- *Acerca de...* Este comando está igualmente estandarizado por Gnome, así como el cuadro diálogo que se presenta al activarlo.

3.2 El diálogo de preferencias

EtherApe permite configurar muchos de los parámetros que definen su funcionamiento. El diálogo de preferencias se divide en dos hojas, una de las cuales controla de qué manera se hace la captura, mientras que la otra configura la presentación de los datos en pantalla.

3.2.1 Parámetros del diagrama

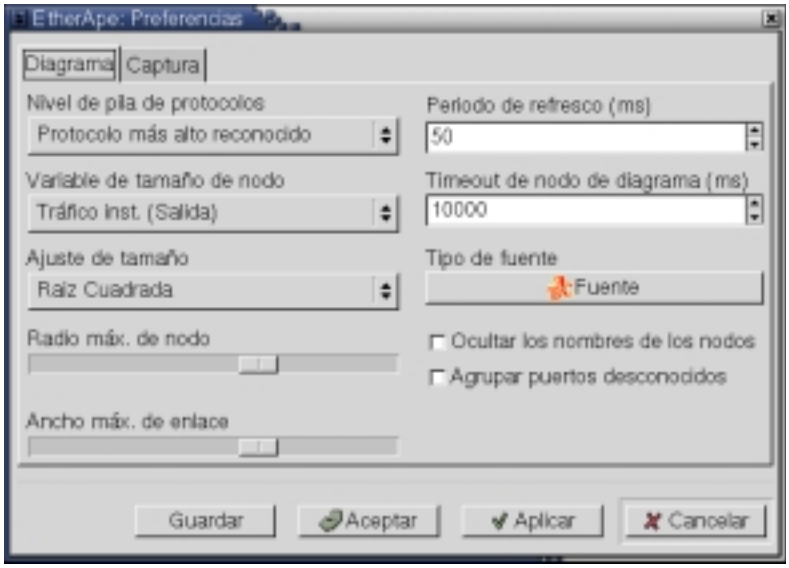


Figura 22 – Diálogo de preferencias del diagrama

- *Nivel de pila de protocolos*. Define el conjunto de protocolos que se distinguirán en pantalla. Los posibles valores son niveles fijos de la pila (nivel 2, nivel 3, nivel 4, etc.) o bien presentar siempre el protocolo de nivel más alto reconocido.

  Hay que tener en cuenta que la selección de niveles fijos no se ajusta exactamente con la definición de niveles en un sistema OSI, simplemente

con el orden en que se van encontrado para cada medio físico. Para una red Ethernet sí produce el valor esperado, pero para FDDI, por ejemplo, el nivel 2 es FDDI, pero el 3 puede ser LLC, mientras que IP va en este caso en el cuarto nivel.

Cuando se usa el ajuste de "Nivel más alto reconocido", se mezclan protocolos de varios niveles, pero suele resultar ser lo más útil, y es el valor por defecto.

- *Variable de tamaño de nodo*. El radio del nodo puede determinarse a partir de seis variables diferentes. Suma del tráfico de entrada y salida, tráfico de entrada y tráfico de salida, tomados como valores instantáneos o como acumulados.

  Para los valores instantáneos se utiliza la variable "Tiempo de medida" que se configura en las preferencias de capturas. El acumulado es a partir de la última vez que se reseteó la captura.

  Variando este parámetro pueden analizarse hechos interesantes, tales como averiguar quién ha consumido más ancho ancho de banda, o quién ha servido más tráfico, o bien quién está monopolizando la red en este momento. El valor por defecto es "tráfico instantáneo de salida"

- *Ajuste de tamaño*. Por su naturaleza, las comunicaciones de red suelen funcionar por ráfagas. Esto es, un nodo puede permanecer largo tiempo callado y repentinamente comenzar a enviar información. Por otro lado, diferentes protocolos manejan cantidades muy dispares de tráfico. Por lo general protocolos que usen puentes o routers para su configuración automática utilizan paquetes de menos de 200 bytes, mientras que típicamente un protocolo para compartir archivos en una red local intentará copar el ancho de banda máximo.

  Así pues, si el ancho de los nodos se relaciona de manera lineal con el tráfico, para un ajuste que sirva para comparar protocolos que usen poco ancho de banda provocará que cuando aparezca tráfico de otro tipo la pantalla se rellene completamente por el tamaño que han alcanzado el radio de los nodos implicads. Si se ajusta a la inversa, los protocolos de menor impacto se hacen prácticamente invisibles.

  El modo de mantener una visión de conjunto es utilizar una función de ajuste, que se controla con este parámetro. Es posible utilizar una función linea, logarítmica o de raiz cuadrada. Este ajuste afecta tanto al cálculo del radio de los nodos como al del ancho de los enlaces.

- *Radio de nodo / Ancho de enlace*. Una vez decidida la función de ajuste, este parámetro permite aplicar una corrección multiplicativa al valor calculado. De este modo tenemos aún más control sobre cómo queremos que distintos niveles de tráfico aparezcan en pantalla.

  Este ajuste es necesario puesto que no es lo mismo analizar el tráfico de un enlace punto a punto sobre un módem que el de una línea dedicada T3.

- *Periodo de refresco*. Es el tiempo que pasa entre sucesivas actualizaciones del diagrama. Permite definir al usuario el nivel de equilibrio deseadoentre validez del diagrama y uso de la CPU.

  Hay que tener en cuenta que en cada refresco no sólo se redibuja el diagrama, sino que es necesario procesar las estructuras de datos de captura para asegurar que vamos a utilizar datos válidos. Por ejemplo los paquetes o nodos demasiado antigos se descartan en este momento, puesto que hacerlo de manera contínua supone un esfuerzo de procesado innecesario.

  También aquí hay que permitir flexibilidad, puesto que el análisis de una red que transporte mucho más tráfico requerirá de mucho más tiempo de procesado y el periodo de refresco se habrá de ajustar al alza en consecuencia.

- *Expiración de nodo de diagrama*. Se utiliza para mantener en pantalla sólo los nodos que esté activos (hayan intervenido en tráfico de algún tiempo) dentro del lapso definido.

  Hay que distinguir entre tiempo tiempo de expiración de nodo de diagrama y el de captura, siendo el primero siempre menor que el segundo. Cuando un nodo expira del diagrama únicamente desaparece allí (se elimina el canvas_node correspondiente), pero el motor de captura sigue guardando la información relativa a este nodo.

  Esto se hace así por dos motivos. El más importante es el de guardar la estadística de este nodo durante más tiempo, de modo que cuando vuelva a estar activo esa información se agregue a la que ya está recogida. Así, por ejemplo, si se hubiera eliminado completamente, a la siguiente aparición el proceso de recogida de nombres tendría que comenzar desde el principio.

  Por otro lado cabe la posibilidad de que el usuario mantenga un interés especial por un nodo y mantenga abierta su ventana de información de nodo. Aunque el nodo se quede inactivo por un tiempo es necesario refrescar la información de la ventana, y eso se hace a partir de los datos guardados en el nodo de captura.

- *Tipo de fuente*. Define la fuente a utilizar para superponer el nombre de los nodos en el diagrama.
- *Ocultar los nombres de los nodos*. En algunos casos puede ser innecesario y hasta molesto identificar cada uno de los nodos del diagrama para observar el comportamiento global de la red. Esta opción permite eliminar los nombres.

  Por otro lado no se puede negar el valor lúdico de la presentación en pantalla de EtherApe, y eliminar los nombres contribuye a crear un diagrama más actractivo para los usuarios que sólo quieran entrenerse mirando el monitor.
- *Agrupar puertos desconocidos*. Si bien esta opción afecta de manera importante a la presentación, en realidad define cómo se realiza la captura, y en versiones sucesivas se trasladará al panel correspondiente.

  Mediante este parámetro se puede decidir si paquetes dirigidos a puertos TCP o UDP no reconocidos se agrupan dentro de protocolos genéricos TCP-Unknown y UDP-Unkown, o bien se crea un protocolo nuevo por cada puerto TCP-Port 3364, o UDP-Port 6734.

  Mientras que en el primer caso se está perdiendo información debido al agrupamiento, si el programa está analizando una red con protocolos que abren puertos arbitrarios y cuyo funcionamiento no está implementado, el número de protocolos reconocidos podría crecer sin control hasta hacer el digrama bastante difícil de reconocer.

  En cualquier caso es objetivo del programa la eliminación de puertos no reconocidos. Su presencia se considera como un caso todavía no implementado.

3.2.2 Parámetros de captura

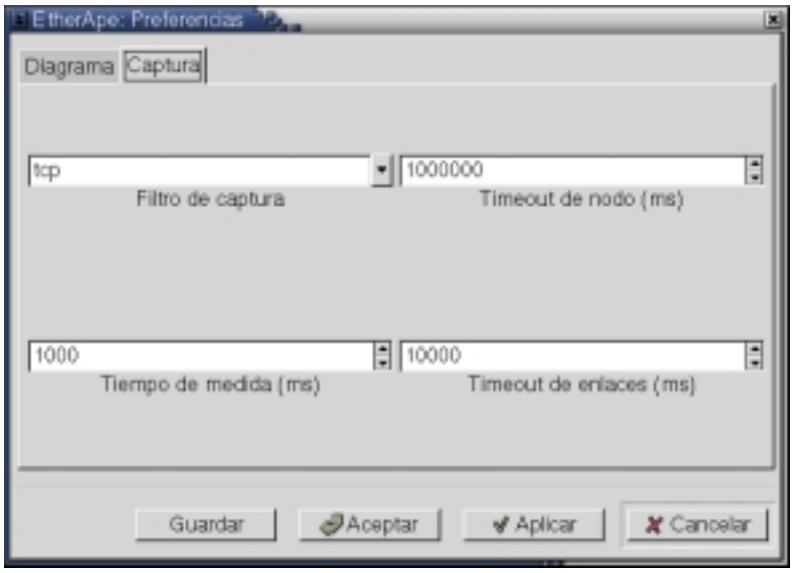


Figura 23 – Diálogo de preferencias de captura

- *Filtro de captura*. Permite limitar el tipo de tráfico que EtherApe va a procesar. Esta funcionalidad la proporciona la biblioteca libpcap, y por tanto la sintaxis a utilizar es la que esta exige.

  Hay que recordar que cuando se trabaja en un modo específico se debe asegurar que sólo se trata de procesar nodos cuyas identidades se admitan en ese modo. Para conseguirlo el programa actualiza automáticamente el filtro de captura utilizado para ajustarse al modo actual.

- *Expiración de nodo de captura*. Toda la información recogida sobre los nodos que lleven todo este tiempo inactivos será deshechada.

  Si al explicar el tiemo de expiración de nodo de diagrama justificábamos la necesidad de recordar esta información, es igualmente importante olvidarla cuando sea necesario.

  EtherApe está concebido como una herramienta de primera necesidad, y como tal no es raro encontrar administradores de sistemas que la dejan corriendo en una consola ininterrumpidamente durante semanas. Durante el funcionamiento normal de una red habrá nodos que estarán activos casi la totalidad del tiempo (un servidor local, el encaminador), y otros que podrían aparecer una sola vez y no volver a hacerlo más (el servidor de una página web que se visita ocasionalmente). Si el programa no deshechara ningún tipo de información sus requerimientos de memoria crecerían hasta agotar los recursos del ordenador donde está corriendo.

Así pues este parámetro permite que el programa siga corriendo indefinidamente, ajustado el equilibrio entre información almacenada y las necesidades de memoria.

- *Tiempo de medida*. Lapso de tiempo de utiliza el programa para calcular todos los valores instantáneos que se vayan a presentar.

  Puesto que este valor es ajustable, el usuario puede decidir si está más interesado en calcular una media a largo plazo o en observar el comportamiento más aproximado al instantáneo.

- *Expiración de enlaces*. Juega el papel que el tiempo de expiración de nodos, salvo que para los enlaces.

  Sin embargo, puesto que de los enlaces no se guarda información que pueda tener un interés adicional en el futuro es innecesario contar con el equivalente de expiración de enlaces de diagrama.

3.2.3 Aplicabilidad y permanencia de los parámetros

Una característica imprescindible para hacer que un programa sea amigable a los ojos de los usuarios es dotarles de cierta inteligencia. Eso significa por ejemplo el que sean capaz de recordar automáticamente valores anteriores del filtro de captura, o qué archivos de captura se han abierto recientemente.

Igualmente es importante permitir al usuario almacenar sus preferencias para recuperarlas en la siguiente sesión. Gnome aporta funciones de biblioteca que facilitan este trabajo, manteniendo automáticamente el archivo donde se almacenan.

Utilizando el botón de Guardar puede hacer que los cambios hechos en el diálogo de preferencias durante la sesión actual se almacenen para la siguiente.

Hay que notar que modificar los parámetros no funciona igual para todos. Accionar algunos de ellos implica un cambio inmediato en el programa (como por ejemplo cambiar el factor de multiplicación de radio de nodo), mientras que para hacer efectivos otros es necesario pulsar el botón Aplicar (por ejemplo, para ver el cambio de tipo de fuente en los nombres)

El botón Aceptar cierra el diálogo de preferencias aplicando los cambios automáticamente, mientras que Cancelar también cierra el diálogo, pero sin aplicar los cambios.

3.3 La interfaz de línea de comandos

Procedimiento de desarrollo

Una parte importante de la elaboración de EtherApe ha sido el esfuerzo dedicado a que el programa siga los estándares de desarrollo comúnmente utilizados por la comunidad del software libre, en lugar de imponer un estilo propio. A pesar de que esto requiere un esfuerzo inicial nada despreciable, esta labor no tarda en dar sus frutos.

1 Introducción

El éxito de un proyecto depende de muchos factores. No basta con contar con un desarrollador de cierta calidad técnica. Es necesario mirar más allá para poder decidir cuáles son las herramientas adecuadas para realizar el trabajo, por ejemplo.

Pero cuando se trata de software de libre distribución hay que tener en cuenta que para lograr que se materialicen algunas de sus ventajas además hay que seguir sus propias reglas. Pongamos un ejemplo: Si pretendes que otros usuarios que utilizan una plataforma diferente a la tuya comprueben si tu código funciona en su entorno, debes al menos asegurarte de una lista de cosas:

- Que el usuario pueda conocer de la existencia de tu programa
- Que le sea posible obtener una copia con facilidad
- Que el programa haga al menos un esfuerzo razonable de funcionar en sistemas no probados
- Que el usuario disponga de medios eficaces para comunicarse con el programador e informar de su experiencia.

Asi pues, es importante recordar que tan importante como crear el programa es asegurar un procedimiento de desarrollo que le permita integrarse con la comunidad del software libre y sacar partido de ella.

2 Estándar GNU de programación

A lo largo de los años en que se ha ido distribuyendo software libre, el formato en que cada paquete se distribuye se ha ido formalizando hasta asentarse en un estándar.

Este estándar describe un conjunto de procedimientos que si son implementados permiten a un usuario o programador que acaba de acceder al programa conocer a priori de qué manera trabajar con él.

Algunos aspectos que pueden parecer triviales como la nomenclatura del paquete de distribución, la estructura de directorios en que se descomprime, y algunos archivos de presencia indispensable; realmente contribuyen a hacer más fácil el trabajo a las personas que no han intervenido directamente en el desarrollo del programa, puesto que se ven ante un esquema conocido.

Desde el punto de vista del usuario, si tiene delante un paquete que sigue el estándar GNU, sabe que con un alto grado de probabilidad todo va a ir bien con más que ejecutar una serie aprendida de comandos. En el caso de etherape, sería

```
lazaro$ tar xvfz etherape-0.7.8.tar.gz
lazaro$ cd etherape-0.7.8
lazaro$ ./configure
lazaro$ make install
```

Primer se descomprime el paquete. Luego se ejecuta el script llamado configure, que analiza el entorno de compilación actual, y por último se compila el paquete completo y se procede a la instalación.

Cualquiera que haya tratado de compilar un programa que tenga un cierto número de dependencias de software y que no siga este esquema sabrá valorar en su justa medida su simplicidad y su potencia.

2.1 Autoconf

En este sentido Autoconf es una herramienta indispensable. Su objetivo es facilitar la compilación de un mismo programa en plataformas diferentes, así como normalizar el proceso de ajuste de la compilación en esas arquitecturas.

El resultado final de utilizar esta herramienta será el script configure, que se encargará de realizar esas tareas.

Para producir este script el programador dispone de un lenguaje específico con el que puede definir los requerimientos de su programa. Se puede, por ejemplo, requerir que se diponga de la bibliotecas de resolución de nombres, o de ciertas ficheros de cabecera estándar.

Autoconf generará un script que será capaz de descubrir en qué tipo de plataforma está corriendo y adaptarse, de forma que tenga mayores probabilidades de averiguar si se cumplen los requerimientos, y de ser así, adaptar el proceso de compilación a las especificidades de la plataforma.

Por otro lado, es común que paquetes de bibliotecas que el programador vaya a utilizar, pongan a su disposición macros para ser utilizadas junto con autoconf, de modo que automáticamente se pueda verificar si están instaladas en esa plataforma.

2.2 Automake

La manera más habitual de guiar el proceso de compilación de un programa es utilizando un Makefile, un archivo que define los objetivos a construir y las dependencias que deben cumplirse antes de poder se construidos.

Automake es una herramienta que permite facilitar enormemente el proceso de generación de estos archivos, a la vez que se integra con autoconf para mejorar el resultado conjunto.

Así, cuando se usa esta herramienta el programador sólo debe indicar, por ejemplo, cuáles son los archivos que contienen el código fuente, sin ni siquiera preocuparse por las bibliotecas de las que depende el programa.

El script configure se encargará de descubrir esas dependencias, y, con ayuda de la información generada por automake, construir un archivo Makefile específico para la plataforma en que se esté ejecutando.

Además, al utilizar automake el archivo Makefile generado resulta ser muy funcional, integrando automáticamente multitud de objetivos que son a la vez muy útiles y esperados por los usuarios.

Así, por ejemplo, el usuario sabe que *make install* compilará el ejecutable y copiará todos los archivos necesarios en el sistema. Pero también dispone de *make uninstall*, que automáticamente borrará los archivos que se hubieran copiado, en caso de que quisiera deshacerse del programa, o *make clean*, para borrar los archivos generados en la compilacón.

El programador también saca partido, al disponer de objetivos que le són tan útiles como *make dist*, que genera automáticamente un archivo de distribución que sigue todas las normas GNU.

3 Traducciones: Gettext

Cuando se distribuye un programa de software libre a través de internet, el mercado potencial de usuarios no está restringido a ningún área geográfica.

El software puede serle de utilidad a cualquiera, y es bueno asegurar el mayor número de usuarios posibles para contribuir a que los mecanismos beneficiosos del software libre hagan efecto.

En este sentido hay que tener en cuenta que no todo el mundo habla el idioma que el programador ha elegido para su programa (que no tiene por qué ser el suyo propio), y por tanto es de interés encontrar un mecanismo que permita de manera sencilla la traducción de los mensajes al usuario.

Este mecanismo se llama gettext. Gettext cumple dos funciones: por un lado es el código que ajusta los mensajes a un idioma específico en tiempo de ejecución, y por otro son herramientas y esquemas de trabajo para facilitar la traducción de los programas.

Así, los programas se distribuyen junto con los catálogos de mensajes en diferentes idiomas, y se compilan junto con el código que los utiliza si es que la plataforma de destino no incorpora ya la biblioteca que cumple esa función.

El trabajo de la traducción se divide en dos frentes separados. Por un lado el programador marca cuáles son las cadenas que se presentan al usuario y que por tanto será necesario traducir. El marcaje supone un trabajo mínimo, puesto solo que consiste en pasar las cadenas por la función de biblioteca _().

Por otro lado los traductores ejecutan una serie de programas que generan archivos de traducción para cada idioma que deberán editar. A partir de estos archivos de texto que ya están traducidos se generarán archivos binarios con los catálogos de mensajes.

Gettext no es el único sistema que existe para la traducción de programas, pero es el estándar de facto para los programas de software libre. Esto implica que al implementar este sistema dentro de EtherApe, cualquier persona con experiencia puede contribuir una traducción en muy poco tiempo.

Así ha ocurrido de hecho, y es gracias a la colaboración desinteresada de dos usuarios que EtherApe está disponible en francés y holandés, además de las versiones en inglés y español aportadas por el autor.

4 Gestión de codigo: CVS

Cuando se desarrolla un programa de cierta entidad es necesario utilizar una herramienta que permita mantener un historial de su evolución. Esta es la función principal de CVS (siglas Concurrent Versioning System).

Utilizando CVS es posible seguir los cambios que se van introduciendo, y deshacerlos también, si es necesario. También permite marcar un momento en el tiempo del código, de modo que sea posible recuperar el proyecto completo tal y como estaba en un momento particular, o cuando se distribuyó una versión en particular.

Cuando se usa a fondo, es posible construir un árbol de desarrollo con varias ramas paralelas. Esto permite, por ejemplo, iniciar un cambio importante en el código sin perder la estabilidad de la versión anterior. En esta situación aún es posible incorporar correciones de errores en la versión estable, y más adelante integrar la versión inestable con el tronco principal.

Pero además CVS cumple una función fundamental en el esquema de desarrollo de software libre. CVS es un sistema concurrente que soporta autenticación. De esta manera es posible mantener un equipo de personas trabajando sobre la misma base de código, sin necesidad de estar haciendo llegar parches a una persona encargada específicamente del mantenimiento de la versión oficial.

Esto tiene un enorme valor para un proyecto de software libre. Cualquier herramienta que facilite el trabajo a contribuyentes de código en perspectiva es una ventaja que merece la pena poner en marcha.

En el caso de EtherApe, por ejemplo, los traductores tienen acceso directo a los archivos que le corresponden, y la persona encargada de preparar el paquete debian también edita directamente los archivos de configuración específicos.

Cuando llega el momento de sacar una nueva distribución, todos los cambios están ya incorporados, y el proceso se simplifica.

Si además se permite acceder al código en el servidor de CVS de forma anónima, se da la oportunidad a los usuarios que así lo deseen de probar lo antes posible las últimas modificaciones.

5 Promoción del código

Uno de los aspectos más importantes del software libre es asegurar el máximo de usuarios posibles.

Cuantos más usuarios haya más rápidamente se localizarán los errores. Cuanto antes se corrijan los errores mejor reputación tendrá el programa y más personas lo usarán. A mayor número de usuarios, mayor el número de ellos con conocimientos y el interés suficientes para mejorar el programa. Y así sigue el ciclo.

Es un ciclo de realimentación positiva, pero los beneficios son mayores haciendo un esfuerzo por hacer que la maquinaria funcione suavemente.

El primer paso es disponer de un sitio en el WWW para distribuir el programa e incluir toda la información posible para ayudar a los usuarios o a las personas que puedan contribuir de alguna manera al desarrollo del código.

En este punto pretendemos “vender” el producto, y así es como se debe trabajar. Cuanto más profesional sea la página web, mejores expectativas tendrán los posibles usuarios y más probable será que se lancen a probarlo.

De cualquier manera, no basta con poner una página web. Para que el público dirija su navegador a un sitio en concreto primero debe conocer de su existencia, así que también hay que hacer publicidad.

La manera más efectiva de hacer publicidad es hacer saber las páginas que se encargan de seguir todo el software disponible de la existencia de nuestro programa. El sitio por excelencia es Freshmeat (www.freshmeat.net), que a diario consultan muchos miles de personas para comprobar qué novedades hay.

Puesto que nuestra aplicación utiliza Gnome, también se ha utilizado su página web al efecto para promocionar EtherApe.

“Release early, release often”, es una máxima del software libre. Cuanto más se avance el programa y más versiones intermedias se publiquen, mayor será su exposición, y el número de personas que habrán tenido oportunidad de hablar de él.

Figura 24 – Página web de EtherApe

6 Sourceforge

Algunos de los procedimientos de los que se han hablado requieren una cierta cantidad de recursos. Un sitio web requiere contratar espacio en un servidor y pagar por el ancho de banda utilizado.

Encontrar una máquina accesible a internet donde poder instalar y administrar un servidor de CVS es todavía más difícil, y lo mismo ocurre con las listas de correo.

Afortunadamente la iniciativa empresarial hace tiempo que empezó a ver los beneficios que tiene para todo el mudo el software libre. Una de las empresas más grandes dedicadas a hacer negocios exclusivamente con Linux, VA, puso a disposición de los desarrolladores de software libre el material necesario para contrarestar la carestía de recursos. Este proyecto se llama Sourceforge, y a pesar de no contar aún ni con dos años de antigüedad se ha convertido en un pilar reconocido del fomento del software libre.

Sourceforge cede gratuitamente una extensa lista de recursos y servicios, entre los que se encuentran.

- Sitio web

  Espacio en disco duro y ancho de banda suficientes para cubrir las necesidades del proyecto más exitoso.

- Gestión de listas de correo

  No sólo proporcionan las máquinas que hacen que el sistema funcione. Además liberan al desarrollador de la carga de su mantenimiento, proveyéndole de un sencillo interfaz para su gestión.

- Servidor de CVS

  Al igual que ocurre con las listas de correo, mantener un servidor de CVS no es ninguna tarea trivial. Los técnicos de Sourceforge lo hacen posible para miles de proyectos diferentes. Además también permite hojear el código que se guarda en el servidor a través de un navegador de WWW.

- Acceso shell y granjas de compilación

  El desarrollador puede acceder directamente a una multitud de máquinas para satisfacer cualquier necesidad que tenga. Un grupo de sistemas heterogéneos está disponible para compilar código, ya sea porque la potencia de los sistemas personales de los desarrolladores no es suficiente, como para poder comprobar directamente la portabilidad de su código en las distintas plataformas.

- Sistemas de seguimientos de errores.

  El contacto con los usuarios es fundamental para poder sacar partido al primer hecho básico del software libre: la facilidad con que el software es probado en multitud de entornos y circunstancias diferentes.

  Sourceforge proporciona un sistema a través de sus páginas web para permitir hacer un seguimiento de los errores. El seguimiento es público, de manera que se disminuye la duplicidad en los informes y un mayor número de usuarios se beneficia de los consejos que pueda dar el desarrollador.

En definitiva Sourceforge es una herramienta fantástica, que hace posible aumentar la productividad del desarrollador liberándole de las tareas que podríamos llamar administrativas.

La página principal nos confirma que en el momento de escribir estas líneas EtherApe no es sino uno de los 23,501 proyectos registrados. No todos acabarán teniendo éxito, pero indudablemente lo tienen mucho más fácil gracias a este servicio.

Usos de EtherApe

1 Análisis remoto

LPI como parte del certificado

Futuras líneas de trabajo

Como ocurre con cualquier proyecto de software libre, no es posible marcar el final de la vida de un programa. Cualquier persona que tenga interés puede seguir añadiéndole caracterísiticas o eliminado errores de programación.

Sin embargo el autor no ha abandonado todavía y hay varias líneas en las que todavía queda por hacer, algunas de las cuales ya se han comentado.

- Selección de colores

  Crear una interfaz gráfica adecuada para que el usuario pueda decidir cómo se asignan los colores a los diferentes programas. Las preferencias deben poder ser grabadas.

- Completar las ventanas de estadísticas

  Permitir analizar los nodos que usan un protocolo determinado, y al mismo tiempo ver qué protocolos usa un nodo específico.

- Fijar preferencias relativas a nodos individuales

  Cada nodo tendría su juego de preferencias, en principio heredadas de las globales para nodos nuevos, pero que luego podrían ser cambiadas por el usuario y guardadas para una sesión posterior.

  Se podría fijar entonces individualmente el tiempo de medida, por ejemplo.

- Selección manual de nombre preferente

  Permitir al usuario seleccionar qué tipo de nombre prefiere en pantalla para los nodos, a ser posible de manera individual para cada uno.

- Volcado a base de datos

  Cierto tipo de usuarios puede estar interesado en utilizar los datos recogidos por EtherApe en otros estudios. Debería ser posible volcar esa información, a ser posible con una interfaz genérica a sistemas de bases de datos.

- Convertir el motor de captura en un proyecto independiente

  No todos los usuarios quieren o pueden instalar Gnome en sus sistemas para utilizar EtherApe. Convirtiendo el motor de captura en una biblioteca que se

distribuyera independientemente se facilitaría mucho el desarrollo de versiones alternativas. Por ejemplo una versión que sólo dependa de GTK+, u otro que funcione en modo texto.

Todas estas mejoras irán llegando con el tiempo, lo que no se puede asegurar es el orden en que lo harán, puesto que en muchas ocasiones son los usuarios los que fijan las prioridades, y estan no tienen por qué coincidir con las del desarrollador.

Conclusión

El proyecto que comenzó como ejercicio para probar las virtudes del software libre ha acabado teniendo mucho más éxito del esperado.

Despues el punto de vista técnico EtherApe supera en muchos aspectos a Etherman, en gran medida gracias a la enorme ayuda que supone poder basarse en tanta cantidad de software de calidad.

Pero en el camino EtherApe ha cosechado un reconocimiento que no era esperado.

Tan sólo dos meses después de comenzar el proyecto el autor fue invitado a dar una de la sesiones en el congreso de Expo Linux 2000, celebrado en Madrid. A raiz de aquella presentación la Free Software Foundation decidión patrocinar el desarrollo financionado un nuevo equipo informático.

Durante semanas el programa estuvo entre los 10 más descargados de los miles de proyectos disponibles en Sourceforge. En los 427 días que el proyecto lleva en línea la página web ha tenido 379.885 visitas, y el programa ha sido descargado 62.757 veces.

Estas cifras no son indicativas de el número real de usuarios del programa, puesto que paulatinamente un mayor número de sitios web hacían copias del programa para ofrecerlo directamente a sus usuarios, tales como Tucows o Icewalkers.

Pero probablemente mucho más importante haya sido el efecto que ha tenido el que la mayoría de las compañías dedicadas a distribuir software para Linux en CD hayan incluido EtherApe entre los programas ofertados. Debian, RedHat, Mandrake, Conectiva y Linux/PPC son sólo algunas de ellas. El sistema operativo FreeBSD también da la opción de instalarlo directamente.

El autor ha tenido la oportunidad de escribirse directamente con más de 200 usuarios distintos. Algunos tan distinguidos como la NASA, que utilizó EtherApe para monitorizar el uso de red de equipos destinados a volar en la Estacion Espacial Internacional, o una compañía dedicada a diseñar sistemas de aviónica.

No cabe duda que mantener un proyecto que es utilizado 24 horas al día, 7 días a la semana en entornos de producción fuerza al desarrollador a cuidar los detalles y que el resultado sea bueno. Pero el hecho adicional de que el código está libremente disponible, y de que de hecho se pretende hacerlo legible para fomentar las contribuciones externas implica que no sólo se cuida el resultado, sino también la maquinaria.

El objetivo de reproducir el funcionamiento de Etherman no sólo se ha alcanzado. La licencia GPL bajo la que se distribuye EtherApe implica que el programa ya quedará abandonado, y que en el futuro sólo le cabe mejorar.

Bibliografía

- Radia Perlman

  Interconnections: Bridges and Routers. Addison-Wesley Publishing Company

- Havoc Pennington

  GTK+/Gnome Application Development. New Riders Publishing. http://developer.gnome.org/doc/GGAD/

- Eric S. Raymond

  The Cathedral & the Bazaar. O'Reilly. http://www.tuxedo.org/~esr/writings/cathedral-bazaar/

- Internet Assigned Numbers Authority

  Port numbers. http://www.iana.org/assignments/port-numbers

Índice