MIRA: Software para el análisis de tráfico IP sobre ATM

Carles Veciana, Josep Solé Pareta, Sergi Sales, Jordi Domingo.
Universitat Politècnica de Catalunya (UPC),
Jordi Girona 1-3, 08034 Barcelona.
{carlosv,pareta,ssales,jordid}@ac.upc.es

Arturo Azcorra, Alberto García.
Universidad Carlos III de Madrid (UC3M),
Butarque 15, 28911 Leganés (Madrid).
{azcorra,alberto}@it.uc3m.es

Manuel Alvarez-Campana, Ana B. García.
Universidad Politécnica de Madrid (UPM),
ETSI Telecomunicación, Ciudad Universitaria, 28040 Madrid.
{mac,abgarcia}@dit.upm.es

Pedro Andrés Aranda Gutiérrez
Telefónica Investigación y Desarrollo.
paag@tid.es

RESUMEN

El objetivo del proyecto MIRA es el desarrollo de una plataforma avanzada de supervisión y control de tráfico para redes académicas y de investigación. El punto de partida es el prototipo experimental MEHARI, desarrollado por los mismos grupos investigadores que participan en el proyecto MIRA. Se trata, por un lado de complementar las funcionalidades básicas del sistema MEHARI, mejorando robustez y operatividad, y por otro de ampliar sus prestaciones orientándolo a facilitar la supervisión y gestión de políticas de uso aceptable (AUP: Acceptable Use Policy) sobre redes académicas y de investigación convencionales (ej. RedIRIS) y de nueva generación (ej. RedIRIS2 y TEN-155). En concreto, en lo concerniente a la consolidación de la plataforma existente, se está desarrollando una Interfaz Gráfica de Usuario, se están adaptando los módulos de captura para que se puedan utilizar en segmentos Ethernet y para el soporte de IPv6, se está preparando la plataforma para poder tomar medidas en varios puntos de la red para mejorar las tasas de captura actuales, etc. Referente a la ampliación de funcionalidades, el objetivo es añadir módulos de medida de consumos para tarificación, supervisión de la calidad de servicio ofrecida por la red, generación de alarmas ante determinadas situaciones y gestión de mecanismos de encaminamiento avanzados.

Este artículo describe la arquitectura del sistema MIRA como consolidación del sistema MEHARI y repasa alguna de las nuevas funcionalidades en las que ya se está trabajando.

1. INTRODUCCIÓN

Conocer el uso que se da a las redes es sumamente importante para el diseño y gestión de las mismas. El entorno de red actual, dominado por la arquitectura TCP/IP (Internet) es muy dinámico y en expansión. Aparecen nuevas aplicaciones día a día que no se restringen a los servicios básicos como web, correo electrónico o transferencia de ficheros, sino que traspasan al ámbito del comercio electrónico, los servicios multimedia, etc. Los usuarios, a su vez, modifican sus hábitos (número y tipo de peticiones, duración de las sesiones, etc.) aprovechando nuevos tipos de contratos (SLA, Service Level Agreement) que ofrecen distintos niveles de calidad de servicio. Finalmente, la comercialización de Internet ha puesto de manifiesto el replanteamiento de los métodos de tarificación (tarifa plana, etc.), ya que la tarificación aplicada en los servicios de telecomunicaciones existentes hasta el momento no parece satisfacer los requisitos del nuevo medio de comunicación de masas que es Internet.

Esto mismo ha supuesto un impacto importante en las redes de ámbito académico, a través de las cuales hoy en día se puede acceder a servicios no estrictamente académicos ni de

investigación. Por consiguiente, las redes académicas deben adecuarse a políticas de uso aceptable (AUP) que impidan que dichas redes supongan una competencia desleal con las redes comerciales y por otro lado, empezar a aplicar una corresponsabilización de los gastos sin esperar una subvención completa de las instituciones públicas. En ambos casos es necesaria la monitorización del tráfico.

MIRA[1] es un sistema de captura y análisis de tráfico para entornos de red basados en tecnología ATM. Los analizadores de tráfico basados en tecnología ATM son caros y reducen su rendimiento a la hora de analizar los protocolos transportados en ATM [1], [2]. Varios proyectos han abordado el problema de la reducción de costes en plataformas de análisis de tráfico IP sobre ATM, con distintas orientaciones [3], [4], [5]. Debido al elevado volumen de datos que puede llegar a transmitir por un enlace ATM (OC3, OC4, OC12), se hace necesario reducir rápidamente la información a analizar para que sea computacionalmente tratable. Este es el caso de las plataformas de captura OCxMON (OC3 y OC12) que analizan el tráfico reportando resúmenes basados en flujos. Para poder realizar una clasificación del tráfico más rica, no únicamente en función de atributos de red y transporte, se hace necesario disponer de todo el contenido del paquete (cabeceras de aplicación y datos de usuario). Al aumentar la cantidad de información a capturar, se reduce considerablemente el tráfico analizado, pero se aumenta la riqueza con la que se puede caracterizar. De este modo, es mucho más fácil conocer el tipo de aplicaciones y la naturaleza del tráfico. Esta es la finalidad de la plataforma MIRA, clasificar el tráfico bajo el mayor número de parámetros posibles para que el gestor de la red (en nuestro caso la Red Académica Española RedIRIS) pueda determinar el correcto uso de sus recursos (uso académico y de investigación) por parte de los usuarios y al mismo tiempo justificar la existencia de una red de I+D en un mercado liberalizado, donde existe una amplia oferta de servicios Internet comerciales.

El proyecto MIRA es fruto de la evolución de distintos proyectos orientados al análisis y caracterización de tráfico en RedIRIS. En los proyectos anteriores, se desarrollaron herramientas para la caracterización del tráfico basado en el análisis de cabeceras TCP/UDP/IP y de aplicación (CASTBA[2]), de análisis de contenidos, así como de los orígenes y destinos del tráfico con gran granularidad (MEHARI[3]). En este proyecto se integran distintos métodos de análisis de tráfico, basados en parámetros derivados de las cabeceras de los paquetes a distintos niveles (ATM, AAL5, IP, Transporte y Aplicación), así como parámetros derivados del análisis de contenidos. Se desarrollan nuevos mecanismos que nos permitirán analizar el comportamiento de las redes con topología compleja y prever sus tendencias. Se amplían también los heurísticos para la caracterización del tráfico, al mismo tiempo que se ofrecen resultados cruzados entre módulos de análisis de los proyectos anteriores. La parte software se ha estructurado como una jerarquía de módulos a cuatro niveles. Dentro de cada jerarquía se realizan funciones del mismo tipo, pero cada módulo esta especializado en una tarea determinada. La interfaz entre los distintos módulos sigue una misma especificación que permite encadenar módulos de diferentes jerarquías y combinar módulos de la misma jerarquía para obtener resultados más complejos mediante un lenguaje propio de filtrado de datos. Finalmente, también se han desarrollado herramientas de tratamiento gráfico, tanto para los resultados (generación y visualización de gráficas) como para el control de los procesos (GUI, Graphic User Interface). El resultado final es una arquitectura de procesos de análisis de tráfico abierta, que integra, además de los procesos de captura y análisis, todas las funcionalidades relacionadas con el tratamiento masivo de datos y la generación y consulta de resultados.

La estructura de este articulo es la siguiente: El apartado 2 describe la arquitectura general del sistema MIRA. El apartado 3 presenta la interfaz gráfica de generación dinámica. El apartado 4 presenta nuestra solución para la generación de gráficas. El apartado 5 describe aplicaciones que

---

[1] Metodologías para la Inspección de tráfico en Redes Avanzadas

[2] Supervisión de la CAlidad de los Servicios Telemáticos provistos por la red académica de Banda Ancha

[3] MEcanismos y Herramientas para el Análisis del uso de servicios aplicado a RedIRIS.

actualmente funcionan en el entorno MIRA. Finalmente, en el apartado 6 se muestran las conclusiones.

2. ARQUITECTURA DEL SISTEMA

Como se ha dejado entrever, los módulos que forman la plataforma MIRA se clasifican jerárquicamente en cuatro clases: *captura, preprocesado, análisis y integración*. Los procesos de captura dependen de la tecnología del enlace que es objeto de la monitorización y se encargan de capturar el tráfico sobre uno o varios enlaces y proporcionar paquetes IP completos a los módulos de preprocesado. Los módulos de preprocesado tratan el contenido de los paquetes, recopilando toda la información significativa, descartando el resto. Los módulos de análisis se encargan de elaborar los resultados y combinaciones de los mismos para obtener una información más rica (resúmenes, históricos y cruzado de datos) relativa a cada uno de los enlaces estudiados. Además, parte de los resultados obtenidos con módulos de análisis, pueden servir como configuración de otros módulos de análisis o de la jerarquía superior (preprocesado). Los módulos de integración permiten obtener una visión global en un sistema con varios enlaces distintos, mediante la correlación de la información que los módulos de análisis proporcionan acerca de cada uno de los enlaces.

La Figura 1 muestra la arquitectura funcional del sistema, tal y como ha sido aplicado a RedIRIS. En ella se observan los tres puntos de captura presentes, junto a los módulos asociados a las fases de captura, preprocesado, análisis y integración.

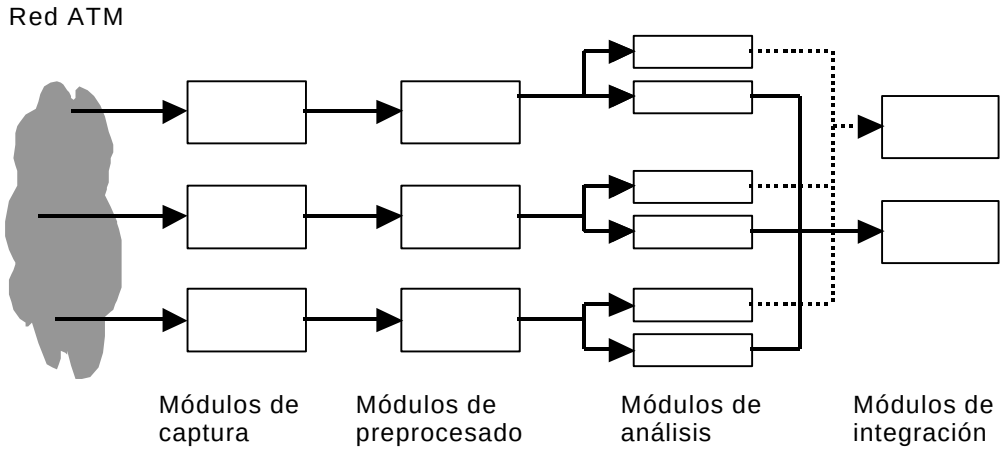

Figura 1. Arquitectura Funcional del Sistema MIRA.

La arquitectura diseñada permite asignar recursos de computación a tareas de una forma extremadamente flexible, en función de la topología de la red a monitorizar, y de la demanda de recursos por parte de las operaciones implicadas. Por ejemplo, dado que el procesamiento, al igual que la captura, requiere de una gran cantidad de recursos si el tráfico analizado es elevado, se ha implementado una arquitectura en la que ambas tareas pueden tener lugar en máquinas distintas conectadas por una red de alta velocidad (Ethernet dedicada a 100 Mbps). Ésta ha sido de hecho la decisión adoptada en la aplicación a RedIRIS del sistema. Podría incluso decidirse realizar las distintas tareas de análisis en máquinas distintas, si ello se considerase oportuno, o por el contrario, realizar todas las tareas en una misma máquina.

En la Figura 2 se puede apreciar cuál es la estructura concreta de equipos que se ha utilizado en la aplicación del sistema MIRA a la captura y análisis de tráfico en la red troncal de RedIRIS.

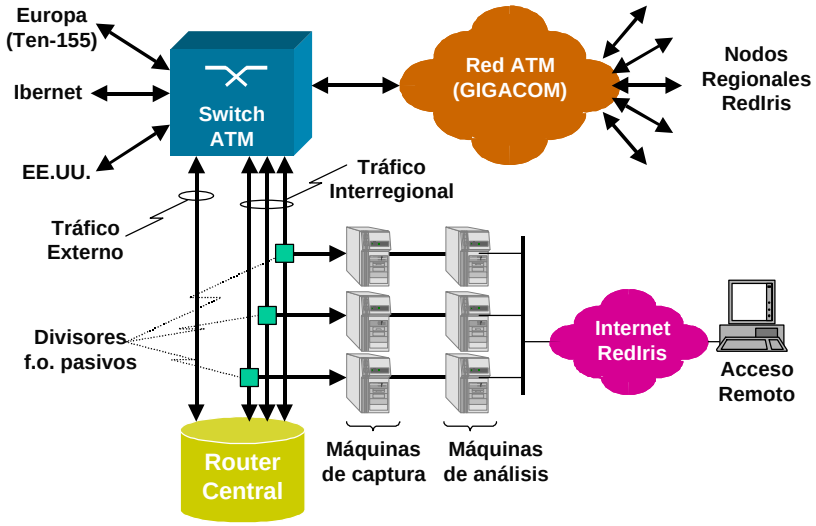

Figura 2. Estructura de equipos utilizada en RedIRIS.

A continuación describimos someramente el funcionamiento de todos los módulos que componen el sistema MIRA, sin distinguir aquellos que fueron heredados del sistema MEHARI.

2.1 Módulo de captura de tráfico

El objeto de este bloque funcional es capturar las muestras de tráfico, para que sean posteriormente analizadas por otros bloques del sistema MIRA. En la actualidad se dispone de un módulo capaz de capturar tráfico ATM sobre interfaces OC-3. Este módulo puede ser configurado para capturar cualquier célula ATM transmitida por el enlace, o sólo las pertenecientes a una lista de pares VPI/VCI. Las células capturadas son reensambladas en tramas AAL5, que son periódicamente volcadas a disco para su posterior procesado por parte del módulo de análisis.

Cada plataforma de captura comprende un PC estándar con sistema operativo FreeBSD y dos tarjetas ATM Fore (una para la captura en cada sentido de transmisión) utilizando un *firmware* especial (OC3-MON). Se precisa de una de estas plataformas por cada punto de medida que se desee instalar en la red. Obsérvese que se trata de un sistema de bajo coste comparado con otros equipos de monitorización comerciales.

2.2 Módulos de preprocesado

Los módulos de preprocesado son tres: *módulo de generación de biflujos, módulo de análisis de cabeceras de aplicación, y módulo de clasificación de flujos*. Cada uno de éstos módulos funciona de forma independiente, y caracteriza la información capturada identificando flujos y caracterizándolos con atributos directamente derivados de la información de los paquetes, deducidos mediante heurísticos o consultas a bases de datos (Internet Routing Registry; etc.). Debido al elevado volumen de tráfico que puede ser transportado sobre los enlaces de ATM monitorizados, estos módulos intentarán descartar las muestras tan pronto como sea posible, conservando únicamente los parámetros que serán de interés para el resto de módulos. Además existe un proceso de realimentación implícita hacia el sistema de captura mediante el cual este último es capaz de autorregular la tasa de captura para adaptarse a la velocidad con que las muestras son consumidas por el sistema de preprocesado.

Este bloque funcional es el responsable de un procesamiento básico de las muestras generadas por el subsistema de captura. Por cada subsistema de captura presente en el sistema se tendrá un módulo de preprocesado.

2.2.1 Módulo de generación de biflujos

Este módulo de preprocesado está basado en el concepto de flujos IP y el reconocimiento de patrones. Respecto al concepto de flujos, este bloque funcional lleva a cabo una agregación estadística de todos los paquetes pertenecientes a un mismo flujo IP durante un período de tiempo configurable. Un flujo IP se define como la agregación de los paquetes que tiene el mismo cuádruplo {dirección IP origen, puerto TCP/UPD origen, dirección IP destino, puerto TCP/UDP destino}. Además, se produce una exploración del contenido de los paquetes para tratar de determinar la presencia de patrones específicos (síntomas), configurables por el usuario. Como resultado, se genera periódicamente un fichero que contiene un resumen de los flujos IP detectados y el número de veces que un patrón ha sido detectado en él.

Posteriormente se realiza una correlación de los datos correspondientes al tráfico en uno y otro sentido capturado en el enlace. El resultado es una colección de flujos IP bidireccionales (biflujos) con una lista de síntomas ponderada para cada sentido, que constituyen la entrada para el subsistema de análisis.

2.2.2 Módulo de reconocimiento de cabeceras de aplicación

El módulo de reconocimiento de cabeceras realiza un análisis de las cabeceras hasta el nivel de Aplicación (HTTP, FTP, IRC, etc.) y resume la información por servidores y por flujos. El análisis que se realiza al nivel de aplicación consiste en la verificación del protocolo de aplicación indicado por el puerto del protocolo de transporte (TCP o UDP), el puerto permite identificar el diccionario de palabras clave a utilizar de la base de datos de diccionarios. La información es acumulada en registros cuyo identificador es el servidor (dirección IP del servidor y aplicación detectada), contabilizando tanto los volúmenes (paquetes y bytes) servidos como consumidos. La alta granularidad de estos resultados (dirección IP) permite disponer de informes muy detallados con todos los servidores activos a ambos lados del enlace. Pueden añadirse fácilmente nuevas aplicaciones a detectar, definiendo diccionarios de palabras clave del servicio que se desea analizar. Todo el tráfico que no pertenece a ninguna de las aplicaciones monitorizadas es reportado en forma de flujos (direcciones IP origen y destino, protocolo de transporte, puertos origen y destino, bytes y paquetes). Una descripción más detallada de este módulo en concreto puede encontrarse en [6]

2.2.3 Módulo de clasificación de flujos

El módulo de clasificación de flujos permite contabilizar distintos tipos de flujos en tiempo real. La definición de flujo es configurable, permitiendo acumular información por dirección IP (origen, destino o ambos) por aplicación, por subred y por conjuntos arbitrarios de redes. La principal utilidad de este módulo de preprocesado es monitorizar la carga que genera una entidad, aplicación, servidor, protocolo, al mismo tiempo que clasifica su tráfico en tiempo real con una granularidad alta (dirección IP origen y destino, protocolo de transporte y puertos). Con este resultado se pretende dar una visión completamente dinámica del tráfico de la red, permitiendo reconfigurar los elementos estáticos de la misma (routers, etc.). Como se verá más adelante (apartado 5.2), este módulo puede ser utilizado para alimentar las tablas de ruta de los routers que trabajan en modo “routing explícito”.

2.3 Módulos de análisis

Los módulos de análisis son el tercer eslabón de la cadena MIRA. Se alimentan de datos ya preprocesados, los enriquecen con nuevos atributos obtenidos de cruzarlos con otros datos y los acumulan para obtener los informes finales, útiles para administradores y gestores del sistema.

Dado que muchos de los procesos de análisis de tráfico se basan en acciones simples de filtrado, comparación, acumulación y truncado de ficheros de registros (flujos, servidores, bytes, etc.), al tiempo que los ficheros de resultados de cualquiera de los procesos siguen una sintaxis simple y muy parecida entre ellos, se ha desarrollado un intérprete de comandos orientado al análisis de tráfico (*filter*). *Filter* permite definir cadenas de comandos y ejecuciones de procesos de análisis

externos, coordinando las cadenas de distintos procesos de análisis y refinamiento de resultados, así como la generación de históricos. Este tipo de módulo se puede considerar como genérico, en el que sólo es necesario la especificación del tipo de fichero junto con la herramienta *filter*.

Para el tratamiento de situaciones más específicas se han desarrollado otros módulos que permiten obtener un resultado final concreto, dependiente de procesos de integración, generación de históricos complejos y formateado de informes. Ejemplos de este tipo de módulos son: el módulo de usos y el módulo de orígenes y destinos. Si un proceso específico llega a generalizarse, podría incluirse como una operación más de la herramienta *filter*.

2.3.1 Módulo de usos

Este módulo parte de los ficheros generados por el módulo de generación de biflujos. Diferentes módulos de análisis se encargan de elaborar los distintos resultados deseados (por ejemplo, tomar decisiones sobre la clasificación de un tipo de tráfico en función de los síntomas – cadenas de caracteres, protocolos, etc. – detectados en la fase de preprocesado, o detectar automáticamente servidores).

Por ejemplo, como parte de las pruebas de campo realizadas en RedIRIS, se consideró la clasificación del tráfico en cuatro categorías (académico, comercial, lúdico e indeterminado). Para ello se desarrolló un módulo que a partir de los ficheros que contenían los síntomas detectados en los flujos IP (resultado del módulo de preprocesado), y aplicando un conjunto de heurísticos fácilmente configurables, realizaba la citada clasificación y generaba ficheros periódicos con esta información.

2.3.2 Módulo de orígenes y destinos

El módulo de orígenes y destinos también parte de los ficheros de flujos generados por el Módulo de Generación de Biflujos y por cada enlace analizado acumula volúmenes (paquetes y bytes) entrantes y salientes, en función de Sistemas Autónomos y subredes. La información referente a Sistema Autónomo y red es consultada de una copia actualizada del Internet Routing Registry (IRR).

Los resultados concretos obtenidos del tráfico por enlace externo de RedIRIS son: tráfico entre RedIRIS y TEN155, tráfico entre otros destinos de RedIRIS, tráfico entre RedIRIS y la red comercial española y tráfico entre RedIRIS y USA.

Del tráfico con origen o destino TEN155 se obtienen resultados por cada uno de los sistemas autónomos que forman parte de TEN155, y del tráfico con origen o destino a la red comercial española se obtienen resultados por cada una de las redes comerciales que la forman.

2.4 Módulo de Integración

La aplicación de MIRA en topologías de red complejas, exige la integración de los datos obtenidos por distintas líneas de transmisión. En la implementación realizada, los módulos de análisis operan en una localización físicamente cercana a la fuente de los datos (la línea de transmisión), para transmitir la información elaborada, más compacta que el resultado de los procesos de captura y de preprocesado, a un lugar común en donde se consolida. La forma en la que esta integración se realiza dependerá de cada módulo de análisis concreto, y por tanto, habrá un módulo de integración por cada tipo de módulo de análisis distinto en el sistema.

3. INTERFAZ GRÁFICA DINÁMICA (GUI)

Una vez definidos los distintos tipos de módulos de la arquitectura pasamos a definir la interfaz entre procesos que hace que la arquitectura del sistema sea abierta. Eso se consigue mediante la especificación del formato de los datos a intercambiar (ficheros de intercambio) entre los módulos, y cómo se controlaran dichos módulos (ficheros de configuración).

Los ficheros de intercambio de datos son autoexplicativos, es decir, contienen toda la información sobre su contenido (en la cabecera) de forma que cualquier otra aplicación los pueda interpretar si reconoce la gramática.

Los ficheros de configuración con sintaxis estandarizada permitirán el uso de la interfaz gráfica de configuración dinámica para su manipulación (en la Figura 3 se muestra la interacción con los procesos y ficheros)

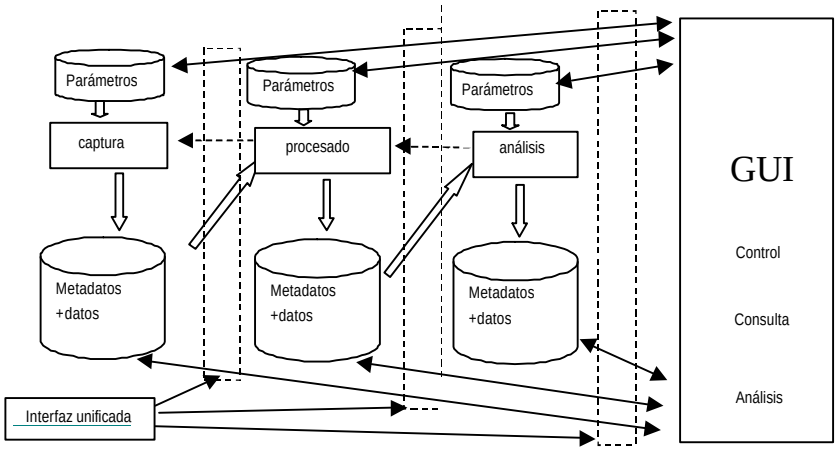

Figura 3. Interfaz de control única de los procesos.

Para simplificar la tarea de integración de los módulos bajo un mismo aspecto, y simplificar la gestión de los mismos, se ha definido una interfaz de interacción con los módulos (ejecución, configuración). Se ha definido un lenguaje en el que se especifican los tipos de parámetros que puede recibir un proceso y unas marcas que estructuran estos parámetros en grupos y les dan significado semántico. La interpretación de los ficheros de configuración por parte de la aplicación gráfica, hace posible la generación dinámica de la interfaz de usuario. Por ejemplo, se generan ventanas para cada grupo de parámetros, se muestran los parámetros en pestañas, se presentan de forma gráfica los valores de los parámetros permitiendo la edición de los valores existentes, se ejecutan los procesos asociados, se insertan líneas de ayuda y botones de navegación, se permite la edición de ficheros, etc. (Figura 4).

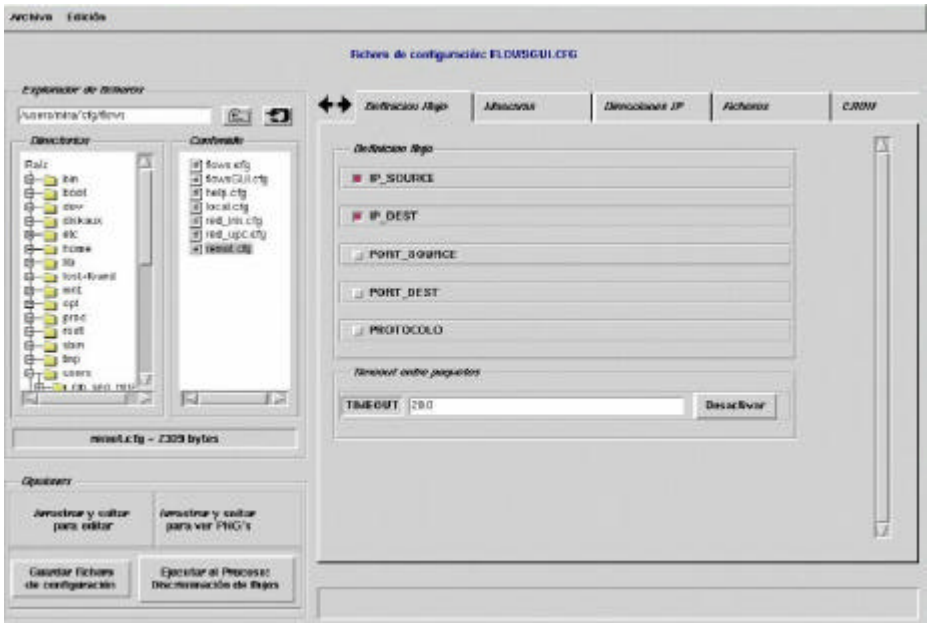

Figura 4. Aspecto de la Interfaz Gráfica

Cualquier módulo de MIRA, puede ser gestionado desde la GUI (Graphic User Interface) sin necesidad de programar en un lenguaje gráfico.

4. GRÁFICAS DE RESULTADOS

Es evidente que la mejor forma de analizar la información es de forma gráfica. Mas aún cuando los volúmenes son muy elevados o la información muy diversa. Los proyectos anteriores confiaron en aplicaciones comerciales para la generación de los gráficos finales. Las gráficas más ricas se consiguieron con aplicaciones en entorno Windows, pero esto suponía un problema de integración con las aplicaciones de captura y análisis (UNIX) además de problemas de licencia.

La necesidad de tener un producto completo, independiente de aplicaciones comerciales, con tipos de gráficos básicos, con utilidades para la personalización (etiquetado, tipo de ejes, ordenaciones de los datos, porcentajes, etc.) nos llevó a la evaluación de diversas aplicaciones gráficas en entorno UNIX. Finalmente se optó por el uso de una librería de libre distribución que nos permitió definir todos los tipos de gráficas necesarios en el proyecto. Se trata de la librería GD[7], sobre la que se ha programado un proceso de nivel superior capaz de interpretar ficheros en formato MIRA. Un ejemplo de las gráficas generadas con la herramienta puede verse en el apartado siguiente.

5. FUNCIONALIDADES DEL SISTEMA MIRA

El sistema MIRA incorpora en la actualidad nuevas funcionalidades respecto a las disponibles en el sistema MEHARI, en las que se aprovecha la arquitectura descrita para la obtención de datos, su tratamiento y su consulta. A continuación se repasan algunas de estas funcionalidades.

5.1 Corresponsabilización de gastos

La finalidad de esta aplicación es ayudar a cumplir las políticas de uso aceptable (AUP) definidas. Sacar estadísticas del uso que se hace de la red, para de una forma justa responsabilizar a los usuarios de la misma.

Para poder llevar a cabo una corresponsabilización equitativa, hace falta tener en cuenta varios parámetros: el tipo de tráfico que esta pasando por la red y el tipo de enlace que utiliza. Cuando se habla de tipo de tráfico se hace referencia a la información generada por el módulo de usos desarrollado por este proyecto, y cuando se habla de tipo de enlace se refiere a los resultados obtenidos por el módulo de orígenes y destinos. Esta aplicación de MIRA cruza los dos resultados obtenidos por los módulos citados, los resultados obtenidos se pueden observar en la Figura 5.

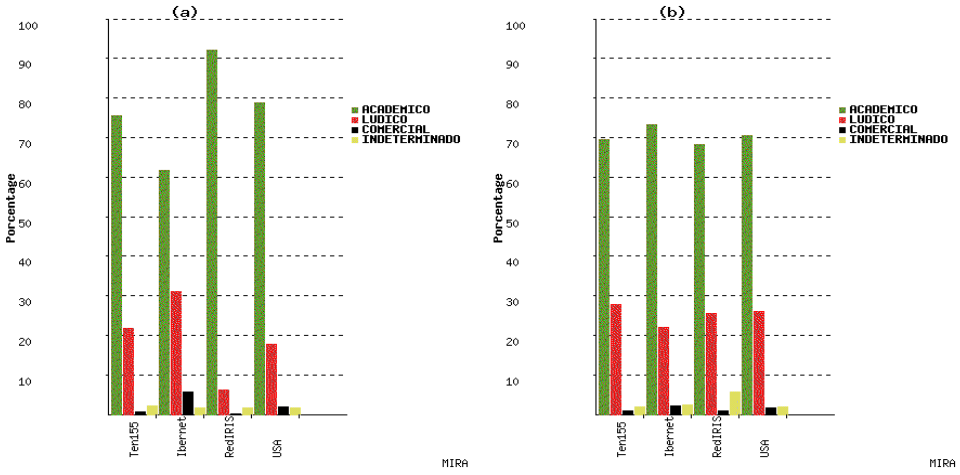

Figura 5. Clasificación de tráfico por destinos y tipo. a) Tráfico de entrada a uno de los troncales de RedIRIS. b) tráfico de salida de uno de los troncales de RedIRIS

La Figura 5 muestra en porcentajes los posibles destinos de tráfico (RedIRIS, TEN155, Ibernet y USA) y el tipo de tráfico que transportan (académico, comercial, lúdico o indeterminado), tomados de una muestra de tráfico en uno de los troncales de RedIRIS desde el 23 de Junio al 27 de Julio de este año.

Otro posible resultado que se puede obtener es el cruce entre los resultados provenientes del módulo de orígenes y destinos, que hacen referencia al tráfico entre RedIRIS y la red comercial Española (Ibernet), obteniéndose una visión de las subredes que comparten más información (ordenadas decrecientemente) con RedIRIS conjuntamente con el tipo de tráfico en porcentajes que se genera. La Figura 6, muestra este tipo de resultado para uno de los troncales de RedIRIS.

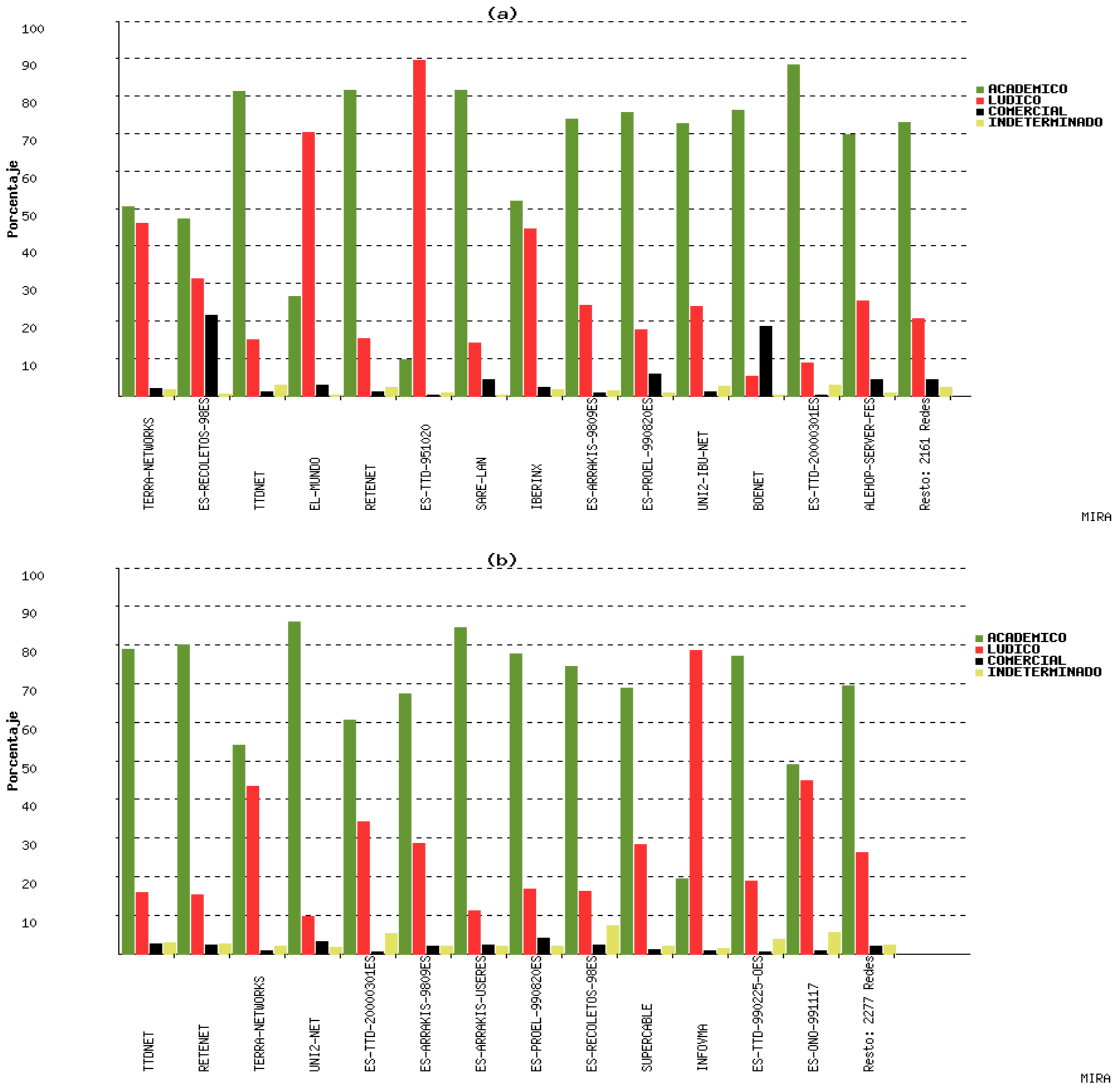

Figura 6. Clasificación de tráfico entre uno de los troncales de RedIRIS y Ibernet.
a) Tráfico de entrada a RedIRIS.
b) Tráfico de salida de RedIRIS.

Tomando como punto de partida una tabla de costes elaborada en función de los destinos y tipo de tráfico podría elaborarse una matriz para cada una de las redes que forman parte de la red académica española con su coste asociado.

5.2 Routing explícito

La necesidad de convivencia entre las redes académicas y la Internet comercial, está llevando a diseñar una segunda red para las aplicaciones académicas y de investigación que requieren altas prestaciones.

El principal problema a afrontar al desplegar una red académica de altas prestaciones es evitar la utilización para fines no académicos de la misma. En muchos entornos, la solución adoptada para asegurar que el tráfico de experimentos autorizados es el único que atraviesa la red académica es la utilización de routing explícito en los nodos de acceso a la red académica.

Esto implica la necesidad de programar conjuntamente los dos puntos de acceso de la Red Académica con informaciones cruzadas. Además hay que mantener una monitorización de los enlaces (mediante el módulo de clasificación de flujos) para comprobar que el uso es el adecuado y realizar la correspondiente coordinación entre los puntos de acceso a la Red Académica en el caso de transgresión de la política de uso aceptable de la misma. En los entornos nacionales, estos problemas se reducen al nivel organizativo. En un entorno de una Red Académica Paneuropea, esto pasa a ser difícilmente manejable y el sistema MIRA puede servir para automatizar el proceso, si se lo engrana con la aplicación de routing explícito que se describe a continuación.

La base tecnológica de la aplicación de routing explícito es la utilización del protocolo BGP-4 y de sus extensiones multiprotocolo descritas en la RFC2858 [8] para intercambiar la información de los caminos explícitos programados en cada punto de acceso. Esto se realiza mediante un router basado en el sistema operativo Linux y el demonio de routing MRTD [9] convenientemente ampliado para transportar las rutas explícitas.

El sistema MIRA vigila automáticamente los enlaces y detecta si existe un uso inapropiado del mismo, en cuyo caso produce un fichero de comandos para el router, en el que se retira la ruta explícita con tráfico ilícito del enlace académico. Este cambio se propaga por la red mediante el protocolo BGP-4 y el tráfico dirigido a la institución transgresora pasa a ser transportado por la Internet convencional.

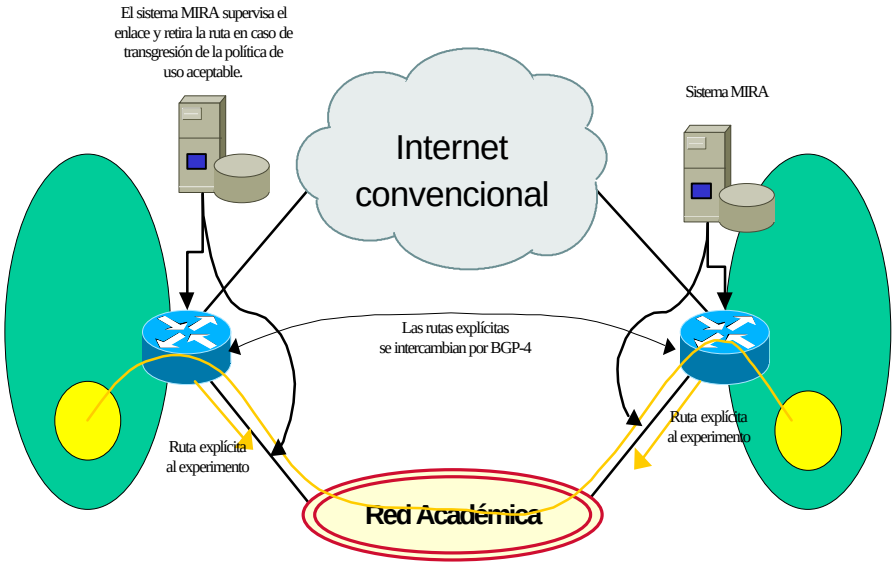

Figura 7. Esquema de funcionamiento del routing explícito.

El router MIRA tiene dos modos de comunicación: BGP-4 sin extensiones para la conexión con la Internet convencional y BGP-4 con las extensiones de routing explícito para conexiones a la Red Académica de altas prestaciones.

5.3 Consulta de resultados vía Web

Viendo la necesidad de consulta por varios usuarios de los datos generados por todos lo módulos del sistema MIRA, se ha dado como solución la publicación de resultados y su tratamiento (de forma simple) vía Web. Esta nueva interfaz, mediante un proceso de selección, ofrecerá la visualización de resultados horarios, diarios, semanales y mensuales, ya sea

numéricamente o gráficamente. Esta interfaz está internamente ligada con los módulos de análisis de modo que una vez obtenidos los datos, la generación de las páginas que los muestran es dinámica.

En un futuro, se prevé la personalización de las gráficas: elección de comunidades autónomas, datos de entrada o salida, tipo de gráfica a generar, por ejemplo: logarítmica o lineal, porcentaje o valor absoluto, etc.

6. CONCLUSIONES:

El proyecto MIRA; es la continuación de otros proyectos anteriores, todos ellos van en la línea de desarrollar un sistema de monitorización de bajo coste que dé soporte a la implantación de políticas de uso aceptable (AUP) en el entorno de redes académicas y de investigación. El proyecto MIRA está convirtiendo el prototipo proveniente de los proyectos anteriores en un proyecto acabado y abierto a la inclusión de nuevas prestaciones. El resultado es un conjunto de procesos para el análisis de tráfico con un conjunto de utilidades que permiten añadir nuevas funcionalidades y modelar las existentes. La definición de las interfaces entre módulos (ficheros de entrada y salida, ficheros de configuración, parametrización y control) permiten adaptar otras aplicaciones (Netramet[10], NetFlow[11]) de forma fácil. El lenguaje de script permite experimentar al administrador del sistema, y decidir el tipo de análisis que desea obtener automatizando el tratamiento. Los informes finales pueden ser muy heterogéneos, desde clasificaciones del tráfico por tipo de aplicación, por servidor (dirección IP), por dirección de Red, Sistema Autónomo, enlace, grupo de entidades, tipo de entidad, tipo de tráfico, etc. y de distinta granularidad: horarios, diarios, etc.

AGRADECIMIENTOS

Este trabajo ha sido financiado por la CICYT (Comisión Interdepartamental de Ciencia Y Tecnología) a través del contrato con número de referencia 2FD97-2234-C03-02.

REFERENCIAS

[1] M.Alvarez et al. “CASTBA: Medidas de Tráfico sobre la Red Académica Española de Banda Ancha”, VIII Jornadas TELECOM I+D, Madrid-Barcelona, Octubre 1998.

[2] M Álvarez-Campana et al “Caracterización de Servicios Telemáticos sobre la Red Académica de Banda Ancha”, Primer Seminario del Programa Nacional de Aplicaciones y Servicios Telemáticos (SPAST-I). Navarra, diciembre 1999.

[3] J.Aspirdof et al. “OC3MON: Flexible, Affordable, High Performance Statistics Collection”, INET’97, Lamasya, june 1997.

[4] P. Lizcano et al “MEHARI: System for Analysing the Use of the Internet Services”,Computer Networks 31 (1999), pag.2293-2307.

[5] A. Azcorra, et al “Análisis del uso de los servicios Internet a partir del sistema MEHARI: prueba de campo sobre RedIRIS”, Primer Seminario del Programa Nacional de Aplicaciones y Servicios Telemáticos (SPAST-I). Navarra, diciembre 1999.

[6] Carlos Veciana-Nogués, Jordi Domingo-Pascual y Josep Solé-Pareta. “Server Location & Verification Tool for Backbone Access Points”, ITC Specialist Seminar on IP Traffic Measurement, Modeling and Management, Monterey (USA), Septiembre 2000.

[7] GD graphic library. http://www.boutell.com/gd/.

[8] T. Bates, Y. Rekhter, R. Chandra and D. Katz, "Multiprotocol Extensions for BGP4" , RFC2858, Junio,2000

[9] MRTd : Multithreaded Routing Toolkit del MERIT http://www.mrtd.net

[10] Nevil Brownlee “Traffic Flow Measurements. Experiences with Netramet” IETF RFC 2123, March 1997

[11] Cisco Netflow White paper: http://www.cisco.com/warp/public/732/netflow/nflow_wp.htm