Capítulo 2

Análisis de requisitos y descripción funcional del sistema

2.1. Introducción

Presentamos en este capítulo un estudio más detallado de la problemática que pretende solucionar este Proyecto de Fin de Carrera planteando:

- Un catálogo de requisitos a cumplir por el sistema a desarrollar.
- Un esquema con las funcionalidades de la solución adoptada.

2.2. Catálogo de requisitos

Mediante reuniones del equipo de trabajo y el estudio de las necesidades a las que se pretende dar solución con este proyecto, se identifican los requisitos listados a continuación:

REQ1: Será un sistema de monitorización remota en el que la información de monitorización será generada en los sistemas remotos y enviada a otros sistemas para su almacenamiento y procesado. Con esto se busca la alteración mímina, en términos de instalación y configuración, de los equipos de una red ya existente en estado «productivo» así como afectar lo menos posible a su rendimiento, dejando las tareas más necesitadas de recursos a otros sistemas diferentes.

REQ2: Será un sistema centralizado, lo que acarreará sus ventajas e inconvenientes clásicos. Por una parte dispondremos de una centralización de la recogida y procesado de la información, mayor facilidad de configuración y uso y menor gasto de instalación, mientras que por otro lado careceremos de redundancia y exponemos nuestro sistema centralizado a mayores problemas de seguridad (como, por ejemplo, frente a ataques de denegación de servicio).

REQ3: Será un sistema escalable y adaptable a las necesidades y al tamaño de la red a la que se desee monitorizar así como a los cambios futuros que en la misma se deseen realizar, de manera que se permita una rápida instalación y configuración en los nuevos sistemas incorporados a la red y en el sistema centralizado de monitorización. En lo que sigue en el desarrollo del Proyecto no nos ceñiremos a ninguna estructura de red concreta sino que plantearemos topologías específicas en cada caso según la necesidad y a modo de ejemplo.

REQ4: La información de monitorización será, en caso de necesidad, capaz de atravesar distintas redes sin alterar su contenido para llegar hasta el sistema centralizado de monitorización.

REQ5: Será un sistema que en *software* se basará en **GNU/Linux**. Esto nos permitirá poder escoger dentro del abanico de utilidades y herramientas ya existentes para estas plataformas buscando además la gratuidad del *software* escogido.

REQ6: Se buscará la mayor la seguridad posible tanto de las utilidades seleccionadas como de la información de monitorización.

A modo de ejemplo se muestra en la Figura 2.1 en la que se ilustra la hipotética red de una empresa en la que a través de una *DMZ* intermedia se permite el acceso a sus servidores desde Internet a la vez que se protegen los equipos de los empleados o servidores internos de la propia empresa.

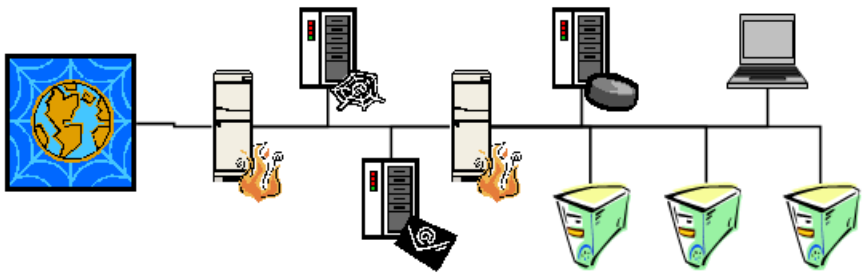

Figura 2.1: Ejemplo de red con *DMZ* intermedia

En dicha Figura 2.1 se pueden identificar, de izquierda a derecha, los siguientes elementos:

- Un acceso a Internet.
- El *firewall* exterior de la *DMZ*.
- Dentro de la *DMZ* se pueden localizar:
  - Un servidor *Web*
  - Un servidor de correo.

- Tras ello encontramos el *firewall* más interno de la *DMZ*.
- Ya en la red interna de la empresa, podemos encontrar:
  - Un servidor de archivos.
  - Distintos equipos de usuario.

Sobre este ejemplo ubicaremos en el siguiente punto 2.3 las funcionalidades propuestas en la solución planteada en este Proyecto, de manera que podamos abstraernos de esta red de ejemplo para implementar en cualquier otra topología de red nuestra solución.

2.3. Funcionalidades de la solución planteada

Tras ver los requisitos que deberá cumplir la solución de este Proyecto presentamos esquemáticamente las funcionalidades y principios de la solución planteada:

FUNC1: Se utilizará el sistema operativo basado en GNU/Linux **Debian Sarge**. La elección de este sistema operativo busca la utilización de *software* libre a la vez que persigue la seguridad de las aplicaciones que se seleccionen. Esto se debe a la rigurosa política de seguridad perseguida por el proyecto Debian que es impuesta en sus aplicaciones, haciendo que la versión *stable* (o versión reconocida como oficial) ofrezca unas altas cotas de seguridad y estabilidad en todos sus aspectos.

FUNC2: Se realizará una monitorización remota del uso de la red: el sistema realizará una recogida estadística de tráfico de la red. Esta recogida de información del estado de la red se efectuará de forma continua obteniendo información constantemente del estado de la red, de la carga que la atraviesa y de los servicios que están siendo utilizados. Este sistema deberá ser poco exigente en recursos (tanto de la red como de los sistemas que intervengan en la monitorización) dado su continuo funcionamiento, y se buscará también su automatización. Se realizará un procesado y análisis de la información ya almacenada en nuestro sistema centralizado de monitorización. El procesado buscará mejorar la legibilidad y la calidad de la información de monitorización generada por la red, realizándose en el sistema centralizado a partir de la información ya almacenada para no ocupar los recursos de los sistemas de la red.

FUNC3: Se hará una monitoritorización del tráfico de la red: por las limitaciones implícitas que poseerá la monitorización estadística del punto anterior, será necesario establecer en el sistema un proceso complementario de monitorización completa del tráfico de la red. Este sistema no actuará de forma continua sino que será selectivo tanto temporalmente (se hará funcionar en los intervalos de tiempo necesarios) como con la información a recoger (ya que en estados de alta carga de la red someter a la misma a un sobreesfuerzo por la extensa información de monitorización no sería adecuado).

FUNC4: Se efectuará una monitorización remota de *logs*: el sistema realizará la recogida de informes, alarmas y registros de los sistemas de la red. Estos registros serán los generados por los sistemas *firewall* de la red así como por los distintos servidores que pudiesen existir en la misma. Esta recogida también actuará de forma continua por lo que se buscará su poca exigencia de recursos. Igualmente, se realizará un procesado y análisis de la información ya almacenada en nuestro sistema centralizado de monitorización.

FUNC5: Se desarrollará una interfaz del sistema (GUI): se confeccionará una interfaz de configuración para las utilidades escogidas que facilite su uso y plantee un manejo cómodo a un administrador humano. En el desarrollo de dicha interfaz se buscará realizarlo adaptándonos a una estructura de tres capas (capa de presentación, capa de negocio y capa de datos) y dividiéndola en módulos (un módulo para cada utilidad escogida) para facilitar su desarrollo y su integración con otros proyectos.

Así, planteamos en la Figura 2.2, sobre la red ya mostrada y a modo de ejemplo, el esquema de nuestra propuesta de un sistema de monitorización remota que pueda solventar la problemática planteada. En dicha figura se muestra la existencia del sistema centralizado de monitorización que recogerá la información adecuada desde otros sistemas de la red.

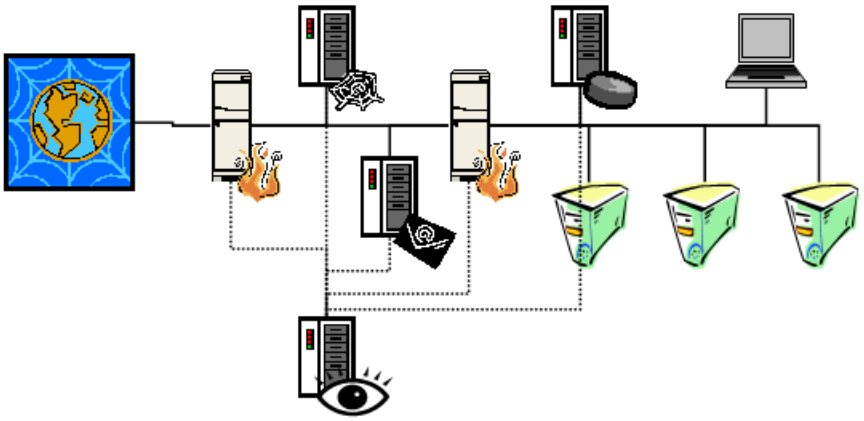

Figura 2.2: *DMZ* intermedia con sistema de monitorización

Las distintas funcionalidades que desarrollará nuestra solución podrían identificarse sobre esta red como se muestra en la Figura 2.3.

De esta forma:

1. El sistema centralizado de monitorización ejecutará el sistema **Debian Sarge**, presentando al Administrador de Red una interfaz del sistema (GUI) que le permitirá interactuar con el sistema de monitorización. Este interfaz permitirá a su usuario cambiar elementos de configuración ya existentes o introducir elementos nuevos en la configuración.

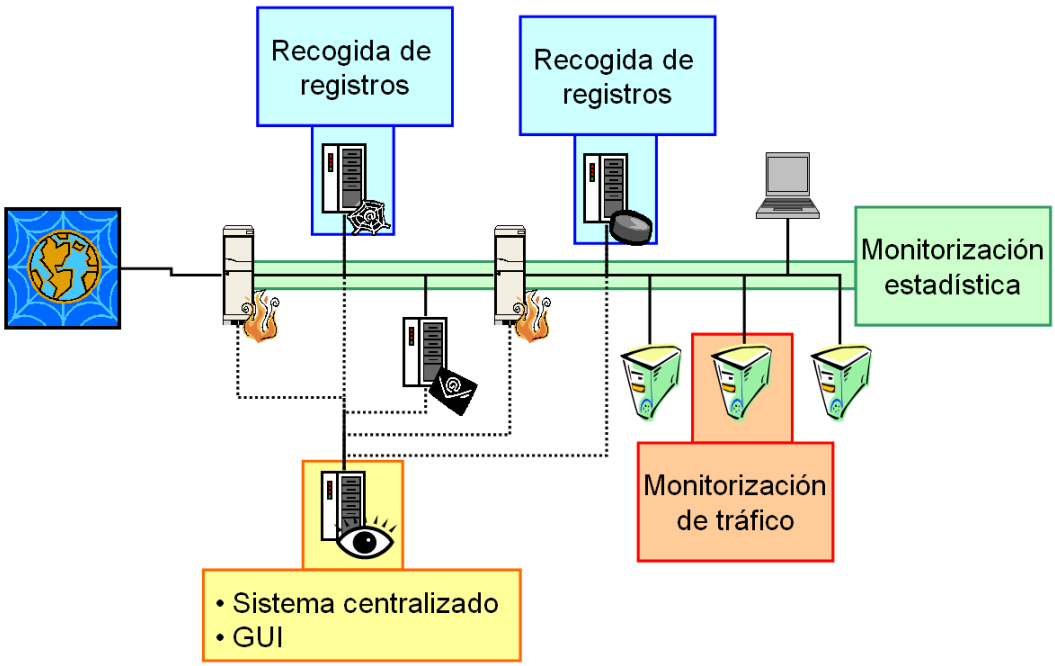

Figura 2.3: Funcionalidades y su identificación con la red de ejemplo

2. Se recogerá la información estadística del uso de la red a partir de la información que circule por la propia red. Esta información estadistica será generada de forma remota en sistemas de la red, minimizando la alteración de la red para su instalación y puesta en funcionamiento.

3. Se podrá efectuar una monitorización de tráfico de una forma selectiva, tanto en el tiempo como sobre la información en sí. Esta monitorización de tráfico también se efectuará de una forma remota en sistemas de la red, por los mismos motivos que se han indicado en el punto anterior.

4. Se recogerán los informes y registros de los sistemas de la red, tales como servidores o *firewalls*. Se buscará minimizar la complejidad de la instalación y puesta en funcionamiento de esta capacidad en los sistemas de la red.

5. El transporte de la información de monitorización se realizará atravesando las redes que sean necesarias, ya sea la propia red de datos que se monitoriza o redes independientes o exclusivas para la monitorización.

6. Se realizará un procesamiento y análisis de la información de monitorización (cualquiera de las tres citadas anteriormente) en el sistema centralizado, evitando el consumo de recursos de los sistemas de la red dedicados ya a otras labores.

A partir de este ejemplo simple puede abstraerse la implantación de las funcionalidades deseadas de nuestro sistema de monitorización en una red de

topología arbitraria.

Durante el resto de este documento iremos estudiando en mayor detalle cada una de las funcionalidades presentadas para el desarrollo de la solución de este Proyecto en una serie de capítulos autocontenidos. Primero, en el capítulo 4, hablaremos sobre la recogida de información estadística del uso de la red. En segundo lugar veremos en el capítulo 5 las posibilidades de recogida de información completa de tráfico desde un sistema remoto. En el capítulo 6 aspectos sobre la recogida de registros y de informes de los sistemas de nuestra red. Tras ello, en el capítulo 7 veremos algunas formas de procesar la información obtenida en nuestro sistema centralizado de monitorización. En el capítulo 8 trataremos algunos aspectos acerca del desarrollo de la interfaz de configuración.

2.3.1. Justificación

Pasamos en este punto a hacer una justificación más extensa de las funcionalidades que se desarrollarán en este Proyecto de Fin de Carrera.

1. Selección del sistema operativo **Debian Sarge**: para solventar los requisitos de utilizar un sistema basado en *software* libre se podía haber escogido otra distribución del sistema operativo GNU/Linux. Distribuciones como **Red Hat** o **Fedora** son comunes en sistemas servidores y bien podría haberse escogido una de ellas para el desarrollo de este Proyecto. En cambio, en nuestra opinión, la distribución **Debian Sarge** ofrece las siguientes ventajas:

   - El proyecto Debian persigue unos objetivos de seguridad y estabilidad en sus aplicaciones superiores a las del resto de las distribuciones, haciéndola una distribución más robusta que las demás.
   - Debian organiza su *software* en un sistema de paquetes muy adecuado para la instalación cómoda y limpia. Además ofrece una cantidad de programas y utilidades bajo licencias de *software* libre enorme que facilitará encontrar el *software* adecuado para nuestras necesidades.
   - El gran impulso que la distribución **Ubuntu** está teniendo dentro de la comunidad de GNU/Linux impulsa indirectamente a la distribución Debian, al basarse Ubuntu en esta última. La grán penetración del proyecto Ubuntu facilita la introducción de los usuarios al mundo Debian, y está planteando a muchos administradores el cambio a este sistema operativo cada vez más manejable y a la vez más seguro que el utilizado en sus servidores.

2. Selección del enfoque de sistema de monitorización remoto: hasta el momento hemos justificado el uso de sistemas de monitorización remota para:

   - Reducir gasto de recursos en sistemas de la red.
   - Minimizar tanto el tiempo como la complejidad de la instalación del sistema en un entorno «en producción».
   - Permitir la escalabilidad del sistema, dotándolo de mayor flexibilidad.

Si bien esto es cierto, queremos ahora hacer la siguiente reflexión acerca de esta *monitorización remota*:

A día de hoy, el sistema tradicional de monitorización empleado consiste en un sistema instalado en la red (en el lugar adecuado para ello) que captura y forma la información de monitorización. Ese sistema realizará las labores tanto de captura como de procesado de la información y se accederá a esta información de monitorización accediendo al sistema en sí, y este acceso podrá ser un acceso remoto[1].

En cambio, la monitorización remota dispone de un funcionamiento diferente desde la base: la información se capturará en la red (igualmente en el lugar adecuado) pero es enviada a un sistema que será el que realice el procesado. Hay una separación de las labores de captura y procesado. Por decirlo de algún modo, el sistema de monitorización *accede de forma remota* a la información que captura, e igualmente a este sistema de monitorización se podrá acceder de forma remota.

Esto, a nuestro parecer, ofrece las ventajas ya citadas de reducción de gasto de recursos, complejidad de instalación y flexibilidad del sistema. Es fácil entender esto si nos replanteamos en este punto la pregunta que introdujimos en la introducción de este documento:

> Si administramos una red grande de una empresa corporativa con distintas subredes para distintos departamentos cada una de ellos con sus propios *firewalls*, distintas *DMZs* para distintos accesos a Internet, ¿cómo monitorizar las distintas subredes? ¿Cómo analizar las necesidades de tráfico y ancho de banda de los departamentos? ¿Cómo comprobar el estado o las alarmas de los distintos *firewalls* y sistemas alojados en distintos segmentos de la red? ¿Cómo centralizar esa información? ¿Cómo hacerlo de forma eficiente?

Traduciendo esta pregunta en términos de número de sistemas necesarios para realizar la labor de monitorización, recursos económicos, coste de instalación y configuración, etc., del sistema clásico no remoto de monitorización están claras las ventajas del sistema de monitorización remota que se propone en este Proyecto frente a ese enfoque más clásico, y por ello hemos elegido el enfoque de la monitorización remota para este Proyecto.

3. Monitorización estadística del uso de la red: tener una buena contabilidad del uso de una red es una gran herramienta para la auditoría de la misma. Mediante sistemas de supervisión del tráfico de la red, podemos hacer un estudio detallado de los servicios más utilizados así como detectar los malos usos que se hagan de la misma.

   Esto obliga a mantener un sistema permanente de recogida de la información que atraviese nuestra red, y como ya hemos dicho nuestro sistema se basará en la monitorización remota. Se dispondrá de algún mecanismo que capturará la información de la red y la hará llegar al sistema centralizado de monitorización.

[1] Con este acceso remoto queremos referirnos a que no necesitará hacerse dicho acceso *in situ*, sino que podría hacerse a traves de una red.

Un aspecto a tener en cuenta es el lugar donde deberá ubicarse ese mecanismo capturador de la información. Se ha dicho que se monitoriza la red pero, obviamente, la red en sí no podrá realizar la captura sino que deberá hacerse en alguno de los sistemas de la red. Esto obliga a buscar los mejores sitios posibles para establecer dichos mecanismos, y estos sitios por norma general serán los interfaces de las distintas redes (como el interfaz entre nuestra red de área local e Internet) ya que son lugares donde el traspaso de información será notable. Esta norma general siempre podrá verse alterada en función de las necesidades concretas de la red que se desee monitorizar si se da el caso de tener unas necesidades más específicas.

4. Monitorización del tráfico de la red: el sistema de monitorización estadística anterior estará limitado por la tecnología utilizada, como se verá en el capítulo 4, "*Monitorización estadística de tráfico*". De esta problemática surgirá la necesidad de suplir las carencias que acusará la monitorización estadística.

   La monitorización *total* del tráfico o *retransmisión* del mismo no será un sistema capaz de sustituir a la monitorización estadística a pesar de suplir sus carencias. El hecho de retransmitir totalmente el tráfico que circula por la red hacia el sistema de monitorización podría, según la topología de la red, someterla a un sobreesfuerzo que no pueda soportar, causando más problemas por la congestión producida que los solucionados. Por ello, la monitorización estadística y la retransmisión deberán coexistir en el sistema de monitorización remota, y esta segunda solo funcionará de manera selectiva tanto en el tiempo como con la información que se desea retransmitir (por el ya mencionado problema de la congestión).

   A esto hay que añadir que, al igual que ocurre con la monitorización estadística, hará falta algún tipo de mecanismo que nos permita realizar la retransmisión. Dicho mecanismo debe de estar emplazado en el lugar adecuado para funcionar correctamente, y esas ubicaciones serán por norma general o los interfaces entre distintas redes o algún servidor concreto del que dispongamos en nuestra red.

5. Monitorización remota de *logs*: mientras que los mecanismos anteriores de monitorización estadística y retransmisión del tráfico se centraban en la propia red, la monitorización remota de *logs* o recogida de registros se centra en los sistemas de la red.

   El objetivo de este mecanismo es obtener los registros de incidencias, informes de estado, alarmas, etc., de los distintos sistemas que compongan nuestra red: los distintos servidores, los *firewalls*... De esta forma podremos tener en detalle cuál es el estado de la red, y de esta forma se podrán detectar de forma anticipada los posibles malfuncionamientos erráticos de los sistemas.

6. Interfaz del sistema (GUI): para facilitar el proceso de configuración de las aplicaciones seleccionadas que solventarán las funcionalidades anteriores se propone el desarrollo de una interfaz del sistema. Dicha interfaz estará ideada para la configuración y control de las aplicaciones que se ejecuten del lado del sistema de monitorización. Si bien podría haberse

planteado el control sobre alguna de las aplicaciones que se manejarán en los sistemas remotos esto añadiría una complejidad alta al desarrollo de la interfaz, y tal desarrollo ha sido descartado en principio para el desarrollo de este Proyecto. Se podrá consultar con mucha más profundidad el concepto de esta interfaz de configuración en el capítulo 8, "*Interfaz de configuración*".