

Introducción

El mantenimiento se puede definir como aquel conjunto de técnicas que aseguran el continuo funcionamiento del ordenador personal, se puede ver como el cuidado que se le da al ordenador para prevenir posibles fallos y para prolongarle la vida, teniendo en cuenta la ubicación física del equipo ya sea en la oficina o en casa.

Tipos de Mantenimiento

Hay tres tipos de mantenimiento:

- Mantenimiento preventivo.
- Mantenimiento correctivo.
- Mantenimiento perfectivo.

Mantenimiento preventivo de ordenadores

El mantenimiento preventivo consiste en crear un ambiente favorable para el sistema y conservar limpias todas las partes que componen un computador, por ejemplo, basta decir que la mayor parte de fallos que presentan los equipos se debe al exceso de polvo en los componentes internos, ya que éste actúa como aislante térmico. En definitiva se trata de conocer aquellos elementos que pueden disminuir la vida útil del equipo.

Otro aspecto que junto al polvo perjudica a la vida del ordenador es que el calor generado por los componentes no pueda dispersarse adecuadamente porque es atrapado en la capa de polvo.

Ni que decir tiene el hecho de junto con el polvo se mezclen las partículas de grasa y aceite que pueda contener el aire del ambiente, ya que crean una espesa capa aislante que refleja el calor hacia los demás componentes, con lo cual se reduce la vida útil del sistema en general.

Por otro lado, el polvo contiene elementos conductores que pueden generar cortocircuitos entre las trayectorias de los circuitos impresos y tarjetas de periféricos.

Si se quiere prolongar la vida útil del equipo y hacer que permanezca libre de reparaciones por muchos años se debe realizar una limpieza con frecuencia del entorno, o por lo menos evitar en la medida de lo posible una serie de factores como son:

- El agua.
- La electricidad.
- Los golpes.
- El frío.
- El calor.
- etc.

Mantenimiento correctivo y perfectivo para PCs

El mantenimiento correctivo consiste en la reparación de alguno de los componentes del ordenador, que abarcan desde una soldadura pequeña, el cambio total de una tarjeta (sonido, video, memoria), o el cambio total de algún dispositivo periférico como el ratón, teclado, monitor, disco duro, etc.

El mantenimiento perfectivo consiste en la ampliación de algún componente del equipo para aumentar su capacidad en base a una serie de requerimientos, por ejemplo, ampliar la capacidad del disco duro con uno nuevo, aumentar un módulo de memoria, etc.

Generalmente resulta mucho más barato cambiar algún dispositivo que al tratar de repararlo pues muchas veces la complejidad del dispositivo nos limita ya no solo por la necesidad de tener aparatos especiales, sino por la complejidad técnica del componente en si mismo. Por ejemplo intentar reparar una placa base de una marca X modelo XXX puede ser completamente distinto de reparar una placa base de la misma marca X pero el modelo XXY. En estos casos es más sencillo sustituir la pieza y enviar la defectuosa al servicio técnico de la casa X.

Asimismo, para realizar el mantenimiento debe considerarse lo siguiente:

- Revisión de los recursos del sistema, memoria, procesador y disco duro.
- Optimización de la velocidad del micro.
- Revisión de la instalación eléctrica (sólo para especialistas).
- Un informe de los problemas de cada equipo junto con el mantenimiento realizado a cada equipo.
- Observaciones que puedan mejorar el ambiente de funcionamiento.

Criterios que se deben considerar para el mantenimiento del PC

A la hora de realizar el mantenimiento basta con seguir dos criterios muy sencillos de aplicar:

- La periodicidad: debería de hacerse, con una periodicidad semestral, una limpieza física del ordenador para evitar tener que realizar un mantenimiento correctivo.
- La ubicación del equipo: afectará o beneficiará al ordenador, deben tenerse en cuenta varios factores:
 - En casa:
 - Alejar el ordenador de las ventanas para evitar que los rayos del sol dañen al equipo y para evitar que el polvo se acumule con mayor rapidez.
 - Colocar el equipo en un mueble que se pueda limpiar con facilidad.
 - Comprar enchufes protectores de corriente.
 - Aspirar con frecuencia la alfombra o moqueta en la que se encuentra el equipo para evitar que se acumule el polvo.
 - No colocar objetos sobre el monitor que le tapen la ventilación y por tanto no disipen bien el calor.
 - No colocar el equipo muy pegado a la pared, dejando que el ventilado de la fuente de alimentación disipe bien el calor.
 - Oficina, en el lugar de trabajo, el mantenimiento es más tedioso debido a que se genera más polvo que en casa, hay más vibraciones y seguramente descargas eléctricas, habrá aparatos que produzcan magnetismo y pueden provocar pérdidas de datos, hay más humo, etc.

Consideraciones finales con respecto al equipo:

- No exponerlo a los rayos del sol.
- No colocarlo en lugares húmedos.
- Mantenerlo alejado de equipos electrónicos que produzcan campos magnéticos ya que pueden dañarlo.
- Limpiar con frecuencia.
- No fumar.
- Evitar comer y beber cuando se esté usando.

- Utilizar sistemas de alimentación ininterrumpida en los servidores para que si la energía se corta tener tiempo para guardar la información.
- Revisión de la instalación eléctrica de la casa u oficina, pero esto lo debe de hacer un especialista.

Herramientas para Mantener el PC

Antes de empezar a realizar un mantenimiento del PC, o a ensamblar uno nuevo a partir de componentes sueltos, es muy conveniente tener en cuenta una serie de recomendaciones en cuanto a las herramientas que posiblemente se vayan a precisar y alguna advertencia sobre la manipulación de los componentes y la electricidad estática.

Para mantener y ampliar un ordenador, es necesario tener algunas herramientas, las cuales en un principio deberían de ser de una calidad aceptable para que duren más y ayuden en la facilidad de mantener.

Por lo menos se tiene que quitar la carcasa de la torre e insertar o quitar tarjetas de expansión, ventiladores, fuente de alimentación, memoria, reemplazar un lector de discos de CDROM, un disco duro, etc.

Se debe también ser capaz de configurar los jumpers y los conmutadores que se pueden encontrar en los dispositivos IDE y en la placa base.

Hay dispositivos a los que se les conectan cables de corriente y por lo tanto hay que probar la continuidad en fusibles y cables.

Para ejecutar tales tareas es necesario un equipo de herramienta básico que contenga los elementos siguientes:

- Destornillador de estrella de un tamaño mediano.
- Destornillador de cabeza plana de un tamaño mediano.
- Destornillador Torx (especial para tornillos con una cabeza en forma de estrella de 6 puntas). Por norma general, estos tornillos son típicos de PC de marca o se encuentran en lugares a donde el fabricante del PC no desea que se acceda.
- Pinza extractora: Permite recuperar un tornillo cuando se cae entre los componentes de la placa base, y también se utiliza para insertar y sacar jumpers.
- Alicates terminados en punta: Útiles para, por ejemplo, extraer los separadores de la placa base.
- Destornillador buscapolos: Para comprobar, por ejemplo, el interruptor de encendido del ordenador.
- Bridas: Para no dejar cables sueltos.
- Insertor / Extrator de chips: necesarios para llevar a cabo ampliaciones de memoria (las famosas tarjetas gráficas que se les podía expandir la memoria) y reemplazar chips defectuosos.
- Cortador de cables.
- Pelacables.
- Pulsera antiestática: Para evitar la electricidad estática.
- Un polímetro: Permite comprobar la continuidad de corriente.
- Tenacillas.
- Un disquete y una cinta para limpiar las unidades correspondientes.
- Piezas de recambio.

La siguiente lista de elementos no son imprescindibles para realizar las tareas de mantenimiento de PC's, pero sí que son útiles para determinadas tareas:

- Un soldador y su soporte.

- Un bote de aire comprimido: Quita el polvo que se acumula en el interior de la caja del ordenador y en el ventilador de la fuente de alimentación.
- Un pequeño limpiador de vacío: para las placas y el teclado.
- Materiales de limpieza. Elimina la suciedad que se “pega” en el exterior del ordenador, incluido el monitor.

Polímetros

Un polímetro es una pieza extremadamente útil en un equipo de comprobación y debe formar parte del equipo de herramientas para realizar mantenimiento. Para mantener un equipo no es necesario comprar el mejor polímetro del mercado, simplemente es bueno comprar aquel que incluya protección a las sobrecargas, ya que de haberlas, un polímetro de protección de sobrecarga es más resistente a estos abusos y el costo extra que pueda suponer se amortiza rápidamente.

Existen en el mercado polímetros analógicos y digitales, siendo estos últimos más caros y más precisos pero en el uso que se le da para el mantenimiento del equipo no son recomendables por su alto coste.

Con un polímetro se comprueban los fusibles y la continuidad de los conectores, utilizando los rangos de resistencias para medir continuidad y determinar si hay alguna resistencia alta. Se puede usar en los rangos de voltios DC, entre otras cosas, para comprobar las salidas de la fuente de alimentación.

Para comprobar un fusible, se utiliza el rango de resistencias. Primero se cortocircuitan las puntas de pruebas y luego se ajusta a cero la medida que da, utilizando el control que hay en el polímetro. A continuación se conectan las puntas una a cada extremo del fusible. Si no se obtiene una lectura de cero ohmios, el fusible está defectuoso.

Para comprobar una batería se pone el polímetro en el rango de voltaje DC y en la escala más alta en la que se espera que se encuentre el voltaje de la batería. Luego se conectan las puntas de prueba a los contactos de la batería, teniendo cuidado de que la polaridad sea correcta observando el voltaje en la escala.

Extracción y Cambio de chips

No es normal que en las labores de mantenimiento se tengan que reparar las placas base y de las tarjetas de un PC cambiando algún chip por dos motivos:

- Es una tarea de los técnicos de las respectivas casas.
- No es habitual encontrar en el mercado un componente de esas características.

Pero si que puede llegar a dar el caso de tener que quitar y volver a instalar algún que otro chip. Por ejemplo, a las tarjetas gráficas de hace 3 años (que aún se encuentran en numerosos puntos de nuestra geografía), se les podía ampliar la memoria, y esta ampliación se realizaba insertándole un chip en un slot de la tarjeta.

Piezas de Recambio

Para realizar un buen mantenimiento, aparte de las herramientas también hace falta disponer de ciertas piezas de recambio, exactamente las piezas que más hacen falta en todo sistema informático son:

- Fuentes de alimentación.
- Discos Duros.
- Disqueteras. Son estándar, de 3'5".
- Ratones. No ocupan mucho espacio y es habitual tener que cambiarlos.

- Teclados.
- Monitores: Tienen el problema de que son voluminosos para tenerlos almacenados.

Estas piezas se pueden obtener por diversos cauces:

- Comprándolas: La compra de fuentes de alimentación, disqueteras, ratones y teclados son triviales y no son excesivamente caras a no ser que se deseen componentes especiales como p.ej. un ratón inalámbrico. Sin embargo los discos duros y los monitores suelen ser caros, y los monitores son voluminosos para tenerlos de recambios.
- Reutilizando: Cuando se retira un equipo se pueden aprovechar los componentes como la fuente, el teclado, etc, para en un momento dado volverlo a utilizar.

Mantenimiento Preventivo: Elementos que afectan a la vida del equipo

Hoy en día los ordenadores (clónicos y de marca) junto con los periféricos son muy fiables. Funcionan durante mucho tiempo sin el más mínimo fallo, pero es prudente no dar ciertas cosas por supuestas. Si se acondiciona el entorno en el que se va a instalar un PC y se cuida durante el uso, se evitarán muchos problemas.

En este apartado se hace un repaso a aquellos elementos que a veces parecen insignificantes, pero que tanto pueden afectar a la vida útil de un equipo. Conocerlos ayudará a evitar que se produzcan problemas o, al menos, reducir su importancia. Se verán los factores que hay que tener en cuenta a la hora de elegir un entorno de trabajo adecuado. También se comentarán los problemas causados por los picos de tensión y cómo solucionarlos. Y una serie de consejos a seguir para prolongar la vida de los ordenadores y de los cuidados que se deben tener durante el trabajo, para conseguirlo.

Ubicación del Equipo

De todos es sabido que los mainframes requieren un entorno refrigerado con temperatura constante ya que sus complejos circuitos generan una gran cantidad de calor. El exceso de humedad y el polvo pueden reducir la fiabilidad de su gran número de interconexiones además los mainframes también necesitan fuentes de alimentación con filtros especiales y altamente fiables.

Los ordenadores actuales tienen menos componentes y generan menos calor, siendo más silenciosos que sus predecesores (sin tener en cuenta el ruido que generan los ventiladores).

No obstante, lo mejor sería cuidar el ambiente en el que está ubicado el equipo, en concreto sería correcto tener en cuenta las siguientes consideraciones:

- Tenerlo correctamente ventilado.
- No tenerlo con un exceso de calor ni de frío.
- Evitar los humos.
- Evitar colocarlo cerca de materiales magnéticos.
- Evitar ubicarlo cerca de fuentes de vibración (impresoras) pues pueden dañar los discos duros.
- No conectarlo en líneas eléctricas a la que estén conectados aparatos con gran consumo de energía como aparatos de aire acondicionado.
- Mantener el entorno limpio, libre de polvo.

Frío y Calor

Los componentes electrónicos consumen corriente y, por tanto, terminan por calentarse, pero hay que tener muy en cuenta que el exceso de calor perjudica. Para comprobar este consumo de energía basta con abrir la caja del equipo después de apagarlo y comprobar si componentes como el disco duro, los chips, etc, están calientes.

Los componentes electrónicos tienen una vida limitada, que se reduce en la medida que el calor se incrementa, por lo que es muy importante tener en cuenta la temperatura ambiente a la que el PC funciona. Dentro del ordenador, se disipa una gran cantidad de calor gracias a la corriente de aire que el ventilador hace circular a través de la caja. Además de incrementar el riesgo de fallo de los componentes, el calor excesivo puede producir otros problemas.

Si la temperatura se eleva considerablemente, habrá muchas posibilidades de que algún componente deje de funcionar bien o que funcione mal. Algunos dispositivos, como por ejemplo las fuentes de alimentación, en la actualidad disponen de diseños de circuitos protegidos contra el calor, de tal forma que si la temperatura de funcionamiento excede de ciertos niveles, se produce una parada en el funcionamiento.

Todos los componentes se calientan cuando se conectan y se enfrian cuando se desconectan, lo que produce dilataciones y contracciones, que a su vez, producen tensiones mecánicas. El problema es más grave con los chips instalados sobre zócalos que con los soldados en la placa. Con el paso del tiempo, las dilataciones y las contracciones hacen que las patillas se salgan fuera del zócalo produciéndose un fallo. Este tipo de fallos es muy común en las tarjetas gráficas, de hecho en los nuevos zócalos AGP traen incorporado una especie de "amarra tarjetas AGP", para evitar que se muevan.

De igual forma se debe evitar un exceso de frío, porque fomenta la formación de condensación, lo que puede causar la oxidación de las superficies metálicas tanto en los conectores como en los zócalos de los chips. Con el tiempo, esto produce contactos eléctricos poco fiables.

Si se ha dejado la máquina en un medio extremadamente frío durante un tiempo, no se debe conectarla hasta que pase el tiempo suficiente como para que alcance la temperatura de la habitación. La resistencia de los componentes eléctricos disminuye con la temperatura, lo que hace que por ellos circule más corriente, de este modo, al conectar el ordenador, se podría estropear algún fusible o causar daños más serios.

Los ordenadores no habrían llegado a ser tan populares si hubiesen necesitado aire acondicionado, pero no deben estar sometidos a temperaturas extremas; por ejemplo, no se deben poner junto a una ventana que a veces reciba directamente la luz solar, sobre todo en verano, cuando la temperatura puede pasar de 40°C. También se debe evitar poner junto a calentadores o fogones.

Un punto importante que no se debe pasar por alto es permitir la obstrucción de las ranuras de ventilación de las cajas de la unidad y del monitor. El ventilador no funcionará de forma totalmente eficiente si no hay suficiente espacio detrás de la unidad. No se debe poner la máquina contra una pared u otra barrera, ni apilar papeles, manuales o discos en lo alto del ordenador o del monitor.

Humedad

Un exceso de humedad puede dañar al ordenador. Este problema puede ser particularmente serio a altas temperaturas. La condensación puede tener muchos efectos, como pueden ser la oxidación de las superficies de las partes metálicas, o de los contactos

eléctricos, o la rotura de los materiales de aislamiento de la fuente de alimentación y del monitor. Este fenómeno es frecuente en zonas costeras españolas, donde el problema se agrava aún más debido a la salinidad del agua del mar que se mezcla con la humedad en las viviendas.

La ausencia de humedad, por calor, o por tiempo seco, puede crear problemas con la electricidad estática.

Polvo y Suciedad

Hay poco que se pueda hacer cuando se utiliza el ordenador en un entorno sucio y polvoriento. Sin embargo, se ha de tener en cuenta que la acumulación de polvo reduce la fiabilidad del ordenador. Una limpieza periódica podría ser necesaria para evitar problemas.

El monitor es un buen indicador de la cantidad de polvo que hay en el entorno. Los monitores de ordenador, atraen el polvo, debido a la electricidad estática que hay en la pantalla. Si se acerca la mano a la pantalla se nota como los pelos del dorso son atraídos hacia el CTR. Se pueden incluso notar chasquidos de la electricidad estática. La pantalla actúa como un imán para el polvo, que se pega a ella o se mete en la unidad del sistema.

Es asombrosa la cantidad de polvo que se puede acumular dentro de la unidad. Esto se debe a la corriente de aire del ventilador que actúa como una bomba de vacío. La mayor parte de los ventiladores funcionan soplando aire por la parte trasera de la unidad. Este aire entra a través de las ranuras de ventilación, así como por otras aberturas, como las ranuras de las unidades de disco.

El polvo puede causar varios problemas en los ordenadores, ya que se amontona en las unidades de disquetes, ensuciando sus mecanismos y lo que es peor, pasándose a los discos, causando errores de lectura y otros fallos. El polvo se adhiere también a los componentes de la placa, reduciendo la capacidad que tienen de disipar calor. Se acumula, de la misma forma, en enchufes y zócalos, en los que las sustancias químicas que transporta pueden provocar que los contactos se enmohezcan, pudiendo producir conexiones poco fiables. Por último, puede introducirse dentro del motor, que en caso extremo, podría quemarse.

Si hay que trabajar con el ordenador en un medio polvoriento, lo único y más importante que se puede hacer es aumentar la frecuencia del mantenimiento preventivo, por lo que se debería limpiar más a menudo. Sin embargo, se puede reducir la cantidad de polvo que puede entrar en el ordenador si se es cuidadoso al elegir su colocación.

Las cajas tipo torre que se colocan en el suelo aumentan el nivel del polvo ya que los humanos al andar levantan gran cantidad de polvo.

Es muy recomendable utilizar un aspirador o una mopa en las salas de ordenadores, ya que las escobas lo único que hacen es levantar el polvo, sin embargo el aspirador lo absorbe.

El problema del polvo se vuelve peliagudo si se habla de los típicos servidores ubicados en una esquina y encendidos las 24 horas del día los 365 días del año, ya que suelen atraer gran cantidad de polvo, que puede repercutir en la disponibilidad.

Los teclados son también vulnerables a los efectos de la suciedad y del polvo, ya que por su naturaleza, son propensos a recoger otras cosas como migajas y ceniza de cigarrillos. Con el paso del tiempo, estas cosas se acumulan en su interior y pueden impedir, por un atasco, el funcionamiento de las teclas.

Se deben limpiar los teclados de forma periódica, como parte del mantenimiento preventivo. Sin embargo, en un medio particularmente sucio, se debe emplear un mayor nivel de protección, por lo que sería una buena idea comprar una membrana de plástico transparente y flexible que se adapta sobre las teclas y las protege de la suciedad, del polvo y de los líquidos.

Esto es irrelevante en oficinas, pues se supone que se limpia todos los días, pero es un factor muy importante en las industrias o almacenes.

Golpes y Vibraciones

Los ordenadores son bastante resistentes, pero hay un límite al que ellos pueden resistir. Las vibraciones constantes pueden ser la causa de que los chips se salgan de sus zócalos y de que los conectores se aflojen.

Los golpes y los impactos repentinos pueden ser dañinos, sobre todo para los discos duros, que son especialmente vulnerables. Si se mueve un disco duro o se produce una vibración mientras está funcionando, se pueden producir serios daños y pérdida de datos. También hay que asegurarse que el ordenador no se instala en un lugar fácilmente desplazable.

Si hay que instalar un PC donde los choques y las vibraciones son inevitables, sería conveniente disponer de una máquina especial, resistente a dichos golpes.

Electricidad Estática

Al moverse por una habitación, se genera una carga eléctrica cuya cantidad depende de varios factores: de las ropas que se tengan puestas, del tipo de cobertura del suelo, del nivel de humedad de la atmósfera y de la conductividad de los zapatos.

Cuando se está cargado eléctricamente y se toca algo conectado a tierra seguramente se recibirá una descarga eléctrica. Estas descargas eléctricas son inofensivas para el individuo. Sin embargo, pueden ser la causa de un deterioro de la memoria del ordenador, para el micro o para la placa base (comúnmente se denomina quemar la memoria, el micro y la placa).

Los generadores de electricidad estática son:

- La piel humana.
- El vidrio.
- El nylon.
- La lana.
- El pelo.
- El plomo.
- La seda.
- El aluminio.
- El algodón.
- El acero.
- El poliéster.
- El teflón.

Se pueden utilizar pulverizadores antiestáticos para las alfombras y reducir la generación de electricidad estática. También, se pueden colocar felpudos antiestáticos, conectados a tierra, debajo del ordenador y de las sillas. Esto impedirá que se genere electricidad estática al moverse mientras trabaja en el ordenador. En climas secos será necesario un humidificador.

La electricidad estática puede ser un serio problema para los empleados de servicio técnico y para todos aquellos que trabajan en ordenadores abiertos ya que al tocar un componente electrónico, puede quedar inutilizado fácilmente. Se puede evitar esto trabajando con pulseras antiestáticas conectadas a tierra a través del chasis del equipo, pero, a veces, es posible que no se pueda evitar ni usándolos.

A continuación se dan una serie de indicaciones para eliminar el riesgo de dañar los componentes con descargas de electricidad estática:

- Tocar la caja del ordenador, o cualquier otra cosa que esté conectada a tierra, antes de tocar una placa del circuito o cualquier otra parte de la máquina. Esto dará la seguridad de que la electricidad estática se descarga inofensivamente a tierra.
- Evitar tocar los componentes electrónicos o los conectores laterales cuando se proceda a instalar, quitar o configurar placas del circuito.
- No tocar los pines de los módulos de memoria cuando se hagan ampliaciones.
- Las placas, tarjetas y memorias se deben de colocar en un embalaje antiestático.
- Utilizar calzado deportivo por ser un buen aislante.

Problemas relacionados con la corriente

Los ordenadores están diseñados para funcionar con salidas normales de corriente, sin embargo, la calidad del voltaje de la línea de corriente alterna puede variar en momentos puntuales e incluso en algunos lugares, el suministro puede ser tan poco fiable como para impedir al ordenador funcionar correctamente.

Los ordenadores mainframes son tan complejos que un leve fallo podría hacer caer a todo el sistema, destruyendo el trabajo de cientos de usuarios. En estos sistemas se emplean circuitos especiales de filtro para proteger el hardware de los efectos de las pequeñas fluctuaciones de la línea de voltaje.

Los ordenadores personales son mucho más tolerantes con las irregularidades de la corriente. Sin embargo, si la calidad del suministro es pobre, también se producirán problemas.

Hay dos tipos principales de perturbaciones que se pueden dar en cualquier edificio y por cualquier motivo que son: los picos de alto voltaje y las fluctuaciones de la línea de corriente.

Picos de Voltaje

Los picos de voltaje pueden causar el deterioro de la memoria y la caída del sistema. Ocurren con frecuencia cuando aparatos como los aires acondicionados, los calentadores y cafeteras se conectan y desconectan. Otros elementos del equipo de una oficina, que consumen mucha corriente, como los radiadores, son también culpables de estas perturbaciones.

Cuando un dispositivo se conecta y desconecta, en vez de un incremento o disminución de la corriente a través de los cables, lo que se produce es una oscilación amortiguada de alta frecuencia.

La oscilación es lo que comúnmente se llama *pico*. Algunas veces, ese pico no se elimina en los circuitos de filtro de la fuente de alimentación y aparece en las líneas de voltaje que alimentan los circuitos electrónicos de un PC, pudiendo causar deterioros en la memoria y la caída del sistema. Los picos pueden ser tan grandes como para causar el fallo de los componentes de la fuente de alimentación, e incluso llegar a afectar al micro y a la placa.

Las fluctuaciones de voltaje, producidas por la conexión y desconexión de elementos del equipo, están presentes en cualquier línea eléctrica. Las condiciones que determinan si

estas fluctuaciones son suficientemente fuertes como para hacer que el equipo funcione mal, son complejas y difíciles de analizar. Sin embargo, los pisos superiores de los edificios altos son particularmente propensos a ello.

Fluctuaciones de la Línea de Corriente

La fuente de alimentación de un ordenador está diseñada para funcionar a un cierto voltaje nominal de entrada. Las diseñadas para usarse en Estados Unidos y Canadá tienen una entrada de 110 V de corriente alterna, mientras que para la mayor parte de Europa, la entrada estándar es de 240 V de corriente alterna. Otros países usan 220 V de corriente alterna.

Cuando se compre cualquier componente para el equipo (incluida la fuente de alimentación) hay que asegurarse de que están diseñadas para ser utilizadas en España y que han sido fabricados para adaptarse a las normas de seguridad eléctricas.

La pérdida de corriente puede causar mayores inconvenientes, de tal forma que, aunque dure menos de un segundo, puede hacer que el ordenador vuelva a arrancar, con lo que se perderían todos los cambios que hayan hecho desde la última vez que se salvó el trabajo. Los archivos que no fueron cerrados correctamente, pueden hacer estragos en el sistema de archivos de la mayor parte de los SO, produciendo la pérdida de sectores y otros problemas.

Para tratar los problemas de la línea de corriente se pueden tomar tres opciones:

- Se puede comprar un enchufe protector especial de onda en almacenes de componentes electrónicos o de ordenadores, que ofrecen protección limitada contra picos de corriente.
- Comprar un rectificador de corriente. su tamaño es como el de una caja de zapatos y se conecta al enchufe de la pared, teniendo, a su vez, dos o tres salidas. Se componen de circuitos que filtran la corriente y eliminan tanto los picos como los ruidos.
- Adquirir un sistema de alimentación continua o ininterrumpida SAI (UPS, Uninterruptible Power Supply). Esta fuente, al igual que el rectificador, se conecta entre el enchufe de la pared y el ordenador. Durante el funcionamiento normal, la corriente alterna pasa por ella y por algún filtro que lleve incluido. La UPS tiene unas baterías que se cargan continuamente con el voltaje de la línea. También contienen componentes electrónicos llamados inversores que convierten el bajo voltaje de la salida de la batería en alto voltaje de corriente alterna. cuando la UPS detecta una rápida caída de voltaje, un relé se conecta sobre el circuito inversor que comienza a generar un voltaje de corriente alterna. Dependiendo del consumo de corriente y de la capacidad de las baterías de la UPS, un PC puede seguir funcionando desde unos minutos hasta hora y media. Esto permite al usuario salvar su trabajo y preparar ordenadamente una salida del sistema si la corriente no se restablece. Las UPS se pueden encontrar en una amplia gama de tipos y tamaños. Algunos más baratos generan una onda cuadrada de salida, en lugar de la senoidal de la corriente de la red. Esto es aceptable para alimentar las fuentes de alimentación conmutadas que usan los ordenadores, pero no en otros equipos.

Deterioro

Los ordenadores, como cualquier componente electrónico se deterioran con el paso del tiempo. Sin embargo, el modo en que se use, también afectará a la mayor o menor probabilidad de que ocurran fallos.

Por ejemplo aunque no parezca cierto, el hecho de encender y apagar el equipo en intervalos de tiempo cortos hará que se deteriore antes.

Si se quiere alargar la vida del PC, se debe conectar y desconectar tan espaciadamente como sea posible.

Si se deja el ordenador conectado durante mucho tiempo, sin usarlo, el CRT puede sufrir daños, ya que, al estar presentando la misma imagen de forma continuada, se puede quemar el revestimiento de fósforo de la pantalla; para solucionar este problema se puede coger el hábito de habilitar el salvador de pantallas o simplemente apagar el monitor cuando uno se ausenta durante un espacio prolongado de tiempo.

Los mayores problemas pueden venir por descuidar y abusar del equipo. Hay varias reglas básicas de buen comportamiento que se deben de conocer que son:

- Evitar fumar o comer mientras se está trabajando con el ordenador, ya que esto puede conducir a que la ceniza y las migajas se introduzcan en el teclado, pudiendo atascar las teclas.
- Mantener los disquetes aislados del polvo y de la suciedad, almacenándolos, metidos en sus fundas, cuando no se usan.
- Manejar los disquetes cuidadosamente, sin tocar la superficie de grabación. Esto ayudará a preservar los datos ya que el polvo y la grasa no se introduzcan dentro de mecanismo de la unidad.
- No se debe ser brusco en el manejo del ordenador. Para introducir un disquete en la unidad, sólo es necesario presionar suavemente con la punta de los dedos. También se debe mover el cierre de la unidad con delicadeza.
- No se deben machacar las teclas como si se tratara de una vieja máquina de escribir.
- Evitar presionar los cables. cuando se intente desconectar una regleta, se debe tirar del enchufe y no de los cables.

Mantenimiento Correctivo y Perfectivo

El mantenimiento correctivo consiste en la reparación de alguno de los componentes del ordenador, que abarcan desde una soldadura pequeña, el cambio total de una tarjeta (sonido, video, memoria), o el cambio total de algún dispositivo periférico como el ratón, teclado, monitor, disco duro, etc.

El mantenimiento perfectivo consiste en la ampliación de algún componente del equipo para aumentar su capacidad en base a una serie de requerimientos, por ejemplo, ampliar la capacidad del disco duro con uno nuevo, aumentar un módulo de memoria, etc.

En ambos tipos de mantenimiento habrá que abrir el equipo y sustituir una pieza o simplemente añadir una pieza, por eso en los siguientes apartados se va a ver cuales son las piezas más importantes de un ordenador y como realizar su sustitución o como insertar una nueva pieza en el mismo.

Los componentes físicos de un ordenador que se pueden cambiar o ampliar son los siguientes:

- Placas Base.
- Micro.
- Memoria.
- Disco duro, lectores de CD, CD-RW, DVD, etc.
- Disqueteras.

Abriendo la Caja del Equipo

Se abrirá la caja del ordenador para limpiarle el polvo o sustituir o ampliar algún componente.

Para abrir la caja será necesario un destornillador plano, de estrella o de Torx.

Existen cajas que ni siquiera tienen tornillos sino que se abren pulsando alguna lengüeta y desplazando los laterales. También es posible encontrar cajas que llevan candados con lo cual antes habrá que abrir el candado con la llave respectiva.

Cuando se habla de formatos de cajas hay que tener en cuenta dos aspectos básicos. Por una parte el formato exterior, que determina su tamaño, estilo, ... Y por otra parte los formatos basados en especificaciones técnicas, que determinarán la futura distribución de los componentes internos.

- Formatos “técnicos”: Va a determinar el formado de la placa base del PC. Los elementos que más se van a ver influenciados por el formato de la caja son la placa base y la fuente de alimentación. A continuación se detallan cuales son estos formatos:
 - XT: Es el formato de los primeros PC's. Surge en 1981 con el IBM PC. Se caracterizaban por la existencia de grandes interruptores en su parte posterior y su gran tamaño y peso.
 - AT: También diseñado por IBM, surge en 1984. Similar al formato anterior, pero con grandes diferencias en el interior de la caja que determinaba otra colocación de la placa. Tenía aspecto de torre y no era compatible con el anterior. En años posteriores fueron apareciendo formatos que apenas diferían de éste, como Baby AT que presentaba unas dimensiones menor y consecuentemente menos coste.
 - ATX: Surge en 1995 y supone un gran cambio con respecto a sus predecesores. Sobre todo en el formato de la fuente de alimentación (que permite diferentes voltajes) y en que son mucho más fáciles de ampliar.
 - NLX: Formato reciente que tiene por objetivo reducir el tamaño.
- Aspecto:
 - Mini-Torre:
 - 2 bahías 5 1/4
 - 2 bahías 3 1/5
 - FA: 200 W
 - Midi-Torre:
 - 3 bahías 5 1/4
 - 2 bahías 3 1/5 (poseen itnernas)
 - FA: 250-300W
 - Torre
 - Gran-Torre
 - Semi-Torre
 - Server
 - Mini-ITX

Pasos a dar:

- Desconectar el cable de alimentación del PC ya que no llega con tener apagado el equipo.
- Descargarse de la electricidad estática.
- Localizar los tornillos o buscar las lengüetas y quitarlos.
- Desplazar los laterales de la carcasa para tener acceso al interior

Placas Base

Para colocar la placa en la caja del equipo hace falta el siguiente material:

- Placa
- Tornillos
- arpones
- Herramientas:
 - Destornillador
 - Brazalete antiestático
 - Cutter

Antes de insertar la placa en la caja hay que realizar el siguiente trabajo previo:

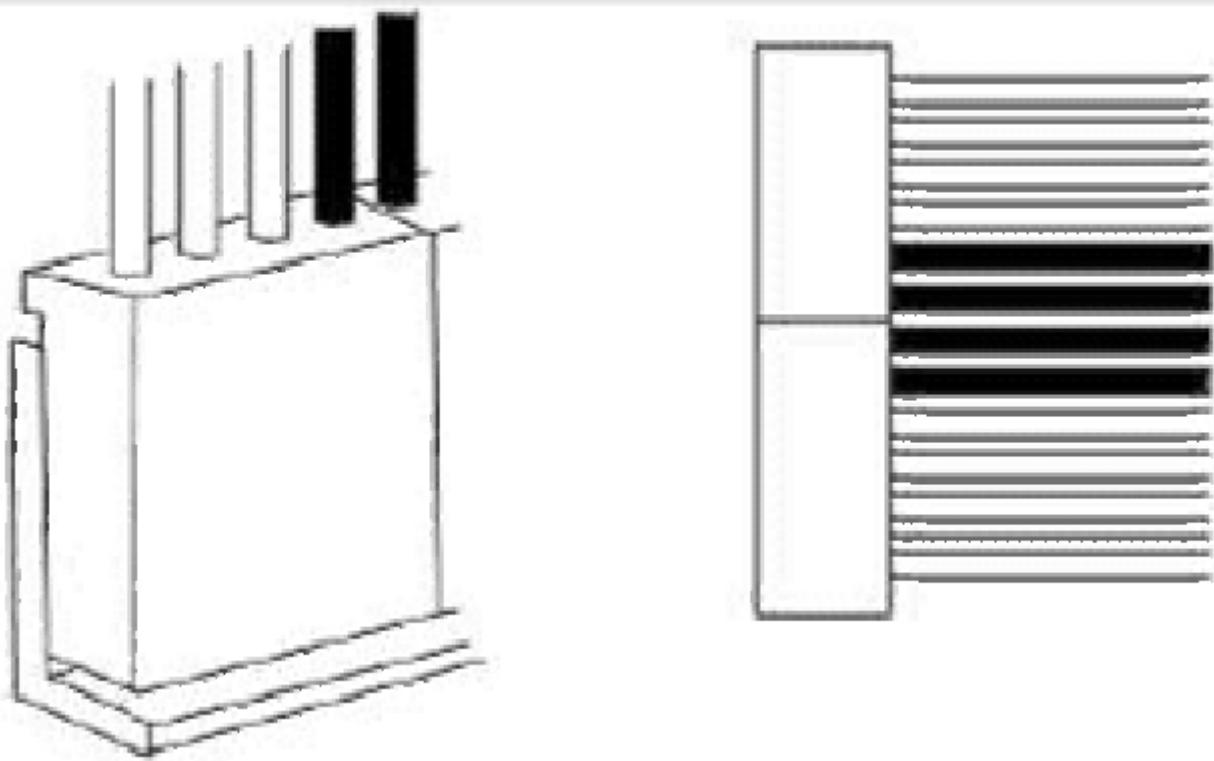
- Anclar los buses a la placa antes de instalarla, además del procesador y la memoria, ya que si la caja es pequeña, o incomoda de trabajar, más tarde será más difícil.
- Antes de colocar la placa será necesario colocar la fuente de alimentación, si bien actualmente la mayor parte de las cajas ya traen la fuente de alimentación colocada de "fábrica".
- Configurar los Jumpers ajustando el voltaje adecuado, la frecuencia de reloj del bus principal y el multiplicador.

A la hora de colocar la placa en la caja hay que seguir los siguientes consejos:

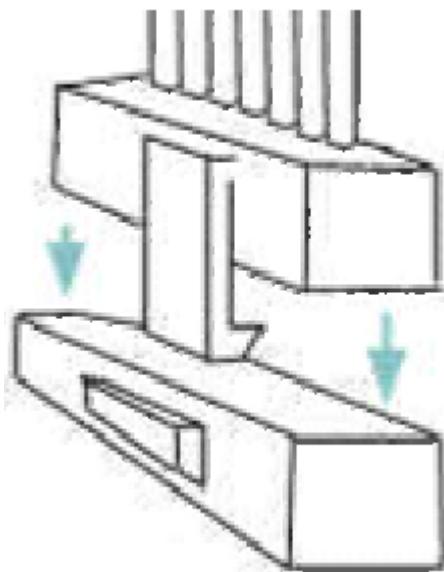
- Evitar que la placa contacte con cualquier parte metálica de la caja. Para ello se deben de utilizar los arpones suministrados con la placa.
- Si la caja no dispusiera de bahías para poder enganchar los arpones, se cortará el enganche de éstos con un cutter.
- Para evitar que la placa toque la superficie metálica de la caja también se puede recurrir a un truco muy viejo que consiste en utilizar la esponja en la que viene embalada.
- Si el procesador alcanza temperaturas muy altas habrá que recortar la superficie de esponja que se aloje debajo del microprocesador ya que se podrá llegar a quemar.
- Al atornillar los tornillos hay que cerciorarse de que la placa está bien anclada pero sin pasarse. La circuitería es muy delicada.

El siguiente paso es colocar los conectores de la fuente de alimentación en la placa. Para realizar esto hay que tener en cuenta que existen dos tipos de placa (que coincide con los tipos de caja):

- Placas AT: Para este tipo de placas las fuentes de alimentación salen dos conectores hembras del mismo tamaño y para enchufarlos en los conectores machos de la placa hay que saber que los cables negros tienen que ir en el medio. De esta forma sólo un sentido es válido.



- Placas ATX: El conector sólo encaja perfectamente en un sentido.



Lo último que queda por realizar es conectar los LEDs luminosos y el altavoz, estos son los cables que vienen incluidos en la caja. Habrá que enchufar los cables de la placa etiquetados con el led correspondiente a la toma correspondiente de la placa.

Para realizar la correcta conexión hay que:

- Contar con el manual de la placa.
- En caso de no tenerlo, mirar las inscripciones de la placa.
- Si esto no funciona queda el método de “prueba-error”.

El Procesador

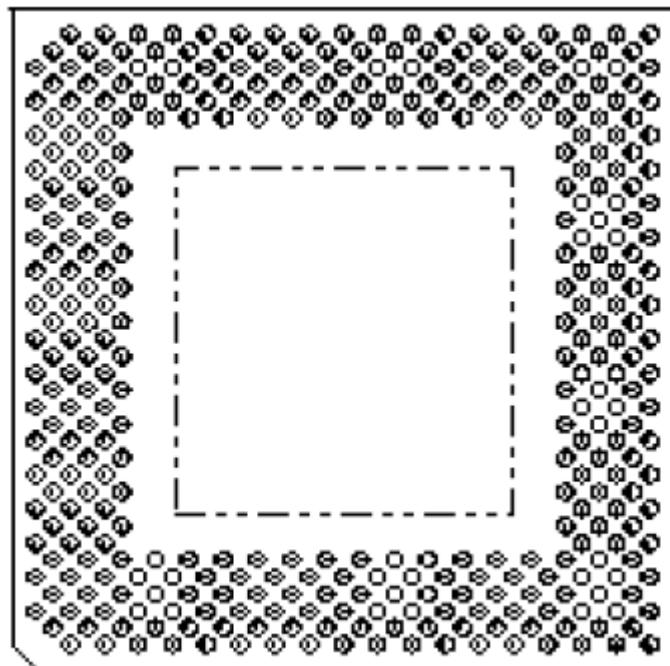
Técnicamente resulta imposible realizar un mantenimiento correctivo del procesador, por la imposibilidad de contar con las herramientas necesarias, como mucho se puede realizar una sustitución de un procesador por otro.

Realizar un mantenimiento perfectivo del procesador consiste en sustituir un procesador por otro más avanzado siempre y cuando la placa base lo soporte.

Hay que tener en cuenta que se distinguen distintos tipos de procesadores, aparte de por su fabricante y por su velocidad por el tipos de conexión. Según esta, se distinguen los siguientes:

- **ZIF** : Zero Insertion Force. Para encajar el microprocesador en la placa no hace falta hacer fuerza, simplemente basca con levantar una palanca del zócalo en la placa base, colocar el procesador encima y bajar la palanca. Es necesario decir que estos procesadores traen los pines alrededor de la carcasa de tal forma que estos fines encajan en el zócalo. En los micros los pines no se cierran en un cuadrado perfecto, sino que dejan unas esquinas que indican por qué lado encajar el micro en el zócalo.

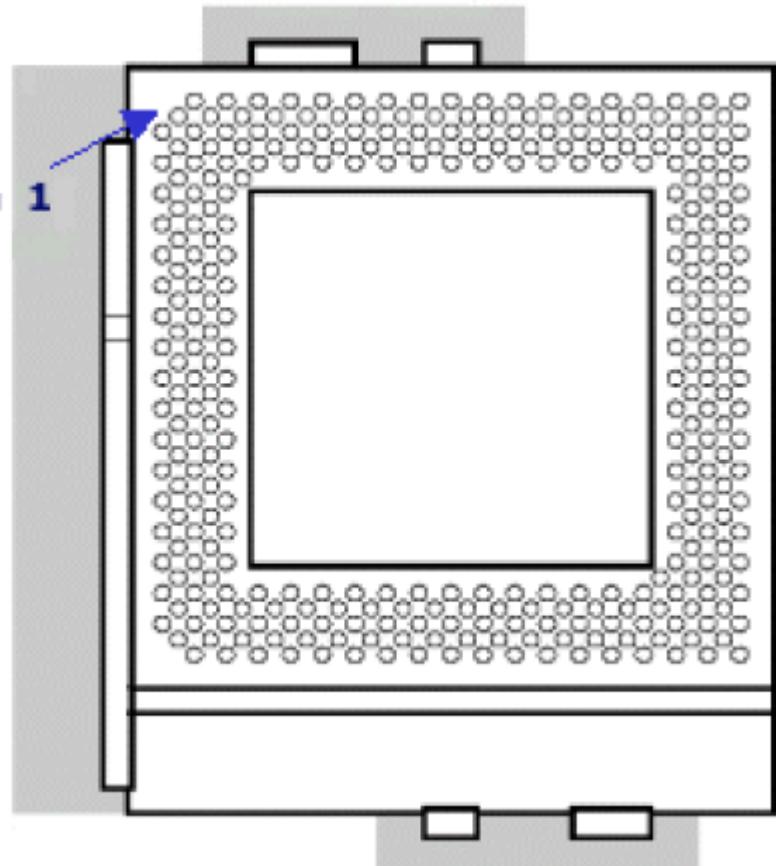
Vista de un micro



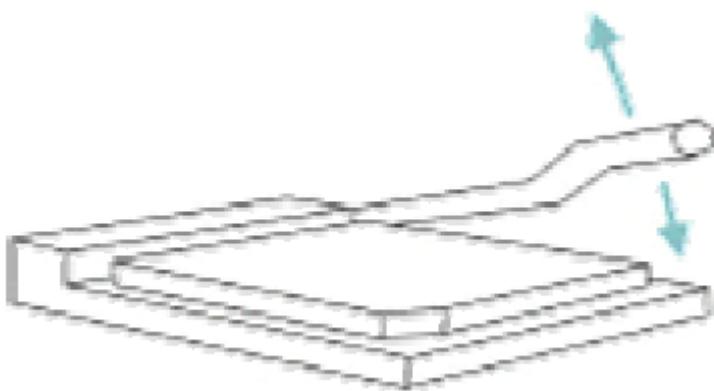
Ejemplos de estos micros son:

- Socket 7: Pentium
- Socket 8: Pentium Pro de Intel
- Socket 370 o PPGA: Pentium III más modernos y Pentium celerón más modernos
- Socket 462/Socket A: Procesadores AMD Athlon, Duron, Thunderbird, XP y MP
- Socket 423 y 478: Pentium IV de Intel

Zócalo donde se inserta el micro

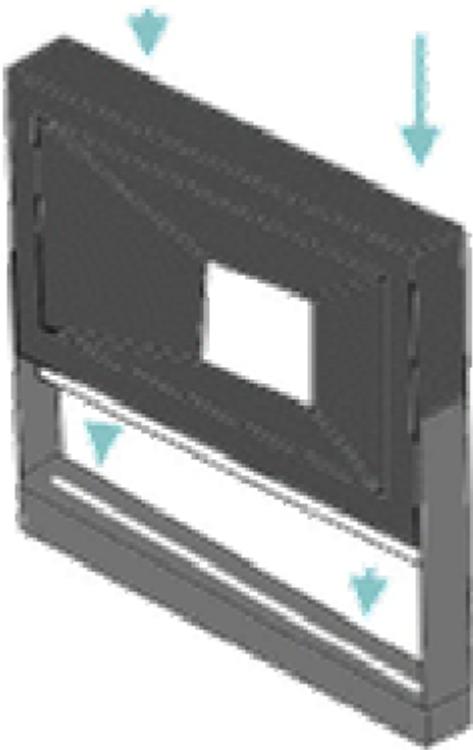


Inserción del micro en el zócalo



- **SEC** : En estos zócalos los micros se insertan como si fueran unos cartuchos y ya hay que hacer fuerza para que encaje el micro en su zócalo.
 - Slot 1: Pentium II
 - Slot II: Pentium III
 - Slot A: Athlon

Procesador SEC



- **No ZIF** : Antiguos zócalos donde se insertaban procesadores 486 y anteriores.
- **Soldados** : Antiguamente algunos procesadores estaban soldados a su placa base.

Una vez colocado el micro es necesario colocarle el disipador para evitar que se queme.

Antes de realizar este paso a la superficie del micro hay que echarle “pasta térmica” que elimina los huecos que quedan entre el disipador y el micro, consiguiendo una correcta refrigeración.

El ventilador suele necesitar alimentación, que la suele proporcionar la placa a través de unos conectores etiquetados generalmente con “CPU FAN”, que será donde hay que colocar el enchufe del disipador.

Disquetera

La disquetera se suele estropear con más frecuencia de la necesaria debido al exceso de polvo que suele acumular. Para evitar hacer un mantenimiento correctivo o perfectivo basta con realizar un mantenimiento preventivo previo.

La disquetera se conecta a la placa a través de un bus de datos que conecta la placa con la disquetera. El conector donde se enganchan los buses de datos de la disquetera a la placa se denomina FDC: Floppy Disk Controller o FDD: Floppy Disk Drive y está situado al lado de los conectores IDE's.

A continuación se exponen los pasos a seguir para realizar un mantenimiento perfectivo (añadir una nueva disquetera) o correctivo (sustituir la disquetera).

- Localizar la bahía de la caja correspondiente 3 1/4.
- Insertar la disquetera.
- Atornillarla en la bahía.
- Colocar Bus de Datos:
 - Si solo hay una disquetera, insertar el extremo del cable que contiene una doblez a la disquetera y el otro extremo del bus al conector FDC o FDD de la placa.

- Si hay ya una disquetera colocar el conector que no tiene la doblez y que está más próximo a ésta en la segunda disquetera y el otro extremo del bus al conector FDC o FDD de la placa.
 - La disquetera que tiene la doblez actuaría como maestra.
 - La disquetera que no tiene la doblez actuaría como esclava.
- Enchufar la disquetera a la fuente de alimentación utilizando el conector pequeño de la misma.

Coneutar un dispositivo IDE (Integrated Drive Electronic)

Las interfaces IDE es un estándar que define la interfaz para conectar periféricos como discos duros, cd-rom, cdrw, etc, a la placa.

EIDE (Enhanced): soporta 4 dispositivos conectados, si bien coloquialmente se le sigue llamando por el nombre IDE.

Una placa base convencional suele tener 2 canales IDE: el primario y el secundario. En la placa base se distinguen porque aparecen impresas las letras: IDE1 (primario) e IDE2 (secundario).

A cada canal se pueden conectar dos dispositivos:

- Primario (Maestro). El maestro es el primero de los dos y se suele situar al final del cable. Tiene preferencia sobre el esclavo a la hora de transmitir datos.
- Secundario (Esclavo). El esclavo es el segundo, normalmente conectado en el centro del cable entre el maestro y la controladora.

Si tenemos en cuenta el hecho de que en una placa base hay 2 canales IDE y que cada canal soporta 2 dispositivos, podemos tener en total 4 dispositivos.

Los dispositivos (discos duros, CD-ROM, DVD-ROM, etc) tienen unos “puentes” llamados jumpers, situados generalmente en la parte posterior o inferior de los mismos, que permiten seleccionar su carácter de maestro, esclavo. Las posiciones de los jumpers vienen indicadas en una calcomanía en la superficie del disco, o bien en el manual o serigrafiadas en la placa de circuito del disco rígido, con las letras *M* para designar “maestro” y *S* para designar “esclavo”.

La distribución “estándar” en un PC para un rendimiento óptimo es la siguiente:

- Canal o Puerto IDE 1:
 - Maestro: Disco Duro principal (Contiene el SO).
 - Esclavo: lector de CD-ROM/DVD.
- Canal o Puerto IDE 2:
 - Maestro: Grabadora de CD/DVD.
 - Esclavo: Segundo Disco Duro, unidad magneto óptica, ...

Los pasos para montar un dispositivo IDE son los siguientes:

- Localizar la bahía de la caja correspondiente: 3½ para discos duros y 5½ para lectores de CD, DVD, etc.
- Insertar el dispositivo.
- Atornillarla en la bahía.
- Colocar el Bus de Datos.
- Si el dispositivo está solo en el IDE, configurarlo como maestro.
- Si no está, solo se siguen las recomendaciones anteriores.

- En cuanto al Bus de Datos, se observará que uno de los extremos tiene un color rojo o rosa, este hilo tiene que encajarse con la pata del dispositivo donde está serigrafiado un 1 que suele ser el de la derecha del conector macho del dispositivo.
- Solo queda conectarle el cable de alimentación.

Memoria

Este es el componente que con más frecuencia se necesita mantener, no porque se estropee, sino porque se necesita ampliar debido a las necesidades de las aplicaciones y SO actuales.

Cuanta más memoria se pueda insertar es mejor.

Formatos

La memoria RAM del PC, en cualquiera de sus tipos, es físicamente un módulo o pastilla (placa de circuito impreso que agrupa varios chips de memoria) que se acaba insertando en los zócalos correspondientes de la placa. Estos módulos pueden tener diferentes formatos que a continuación se describen.

- **DIPS (Cápsula Dual en Línea, Dual Inline Package)** . Totalmente desaparecido. Sólo mencionar que en los primeros PC's se conectaba los DIPS en zócalos libres de la placa base.
- **SIMM (Módulo de Memoria Simple, Single Inline Memory Module)** . Pequeña placa de circuito impreso que almacena chips de memoria solo por un lado de la placa, y se inserta en un zócalo SIMM en la placa madre. El primer formato tenía 30 contactos. Un formato más largo en centímetros, que usa 72 contactos y puede almacenar hasta 64MB de RAM, se popularizó con los procesadores Pentium. Las memorias en este formato presentaban la restricción de tener que instalarse siempre por pares de la misma capacidad. Es decir, para conseguir 16MB, eran necesarios 2 SIMM's de 8MB o 4 SIMM's de 4MB. Esta restricción venía impuesta porque estos módulos permitían almacenar 32 bits por ciclo, y los procesadores Pentium utilizaban un bus externo de 64 bits.
- **DIMM (Módulo de Memoria Dual, Dual Inline Memory Module)** . Pequeña placa de circuito impreso que almacena chips de memoria por ambos lados de la placa, y que se inserta en un zócalo DIMM.
 - Existen 2 tipos de módulos DIMM:
 - SDR SDRAM: 168 contactos.
 - DDR SDRAM: 184 pines.
- **RIMM (Rambus Inline Memory Module)** . Pequeña placa de circuito impreso que almacena chips de memoria por ambos lados de la placa y disponen de disipadores de calor. Tienen 184 pines.

Colocando la Memoria SIMM

Para ampliar la memoria SIMM hay que tener en cuenta que es necesario insertar pares de módulos debiendo ser los módulos iguales en velocidad y tamaño.

Para poder ampliar memoria de este tipo, es preciso tener, al menos, 2 zócalos libres en la placa.

Los pasos a dar son los siguientes:

- Descargarse de la electricidad estática.
- Orientar y presentar el módulo de forma correcta sobre el zócalo de la placa.
 - Los contactos de los módulos están numerados del 1 al 30 o 72, dependiendo del tipo de módulo.

- Solo es posible una orientación del módulo en la placa ya que en la base, el conector nº 1 del zócalo tiene una pequeña rebaba que imposibilita insertar mal el módulo.
- Para poder insertar el módulo hay que inclinarlo 45º una vez que éste está sobre el zócalo.
- Posteriormente se levanta el extremo superior sin sacarlo de la ranura hasta que quede perpendicular a la placa.

Para extraer la memoria SIMM hay que presionar las lengüetas de los extremos de los zócalos hacia el exterior.

Colocando la Memoria DIMM

Estos son los módulos de 168 contactos (SDRAM) y 184 contactos (DDR, RAMBUS).

No presentan los inconvenientes de los SIMM:

- Se pueden insertar módulos de diferente capacidad.
- No es necesario colocar los módulos por pares.

Sí es conveniente tener los módulos de la misma velocidad.

Los pasos a dar son los siguientes:

- Localizar los zócalos en la placa sobre los que se insertarán los módulos de memoria. Tener en cuenta que algunas placas soportan los dos tipos de módulos: SDRAM y DDR.
- Orientar y presentar el módulo de forma correcta sobre el zócalo de la placa. Para ello observar que en la parte de los pines, la placa de memoria tiene unos huecos (2 para el caso de SDRAM y uno para DDR) que determinan la posición exacta.
 - Es conveniente comenzar a insertar módulos de memoria por el primer zócalo del primer banco, aunque no es estrictamente necesario.
- Para anclar el módulo sobre la ranura, hay que presionar suavemente sobre el módulo una vez que esté correctamente situado.
 - Estará bien colocado cuando se escuche un pequeño clic.

Bibliografía

- [Scribd \(Ibiza Ales\)](#)

Gestión de Librerías de Programas

Cuando hablamos de librerías nos referimos a conjuntos de programas, rutinas o funciones ya preparadas y a disposición de los programadores durante el desarrollo de aplicaciones. Su practicidad reside en que evitan la reescritura de algoritmos usados con frecuencia, ya que éstos pueden ser incluidos en librerías que posteriormente podrán ser llamadas desde los distintos programas a implementar.

Gestión de Librerías en C

Propiedades de las librerías en C:

- Una librería es un archivo que agrupa a otros archivos denominados miembros de la librería.
- La estructura de las librerías hace posible que puedan extraerse sus miembros.
- Al agregar archivos a una librería, se introducirá en la misma tanto el contenido de aquellos como su información de gestión (fechas, propietarios, grupos, permisos, etc).

Librerías Estáticas y Dinámicas

Librerías Estáticas

También denominadas librerías-objeto, son agrupaciones de archivos objeto (.obj) compilados en un solo archivo de extensión **.OBJ** o **.LIB**.

Los modelos de las funciones empleadas en estas librerías, junto con algunas constantes predefinidas y macros que facilitan su uso, constituyen los denominados archivos de cabecera, debido a que suelen ser llamados desde las primeras líneas (cabeceras) de los distintos archivos fuente.

Las librerías estáticas están constituidas por uno o varios archivos **.lib** , **.obj** o **.bpi** junto con uno o varios archivos de cabecera (**.h**). Al compilar un programa, el linkador agrega al ejecutable los módulos que incluyen a las funciones utilizadas en el programa, pasando aquellos a formar parte del ejecutable. Esta forma de enlazar las librerías con los programas es la que les da el nombre de estáticas.

Librerías Dinámicas

Las librerías de enlazado dinámico (DLL) son muy utilizadas en la programación para SO Windows; sistemas que incluyen multitud de librerías de este tipo en disposición de ser utilizadas por cualquier aplicación.

Aunque las librerías dinámicas se asocian generalmente a la extensión **.DLL** , también pueden estar definidas con extensiones del tipo **.EXE** , **.BPI** , **.DRV** , **.FON** , etc.

La programación de aplicaciones Windows consiste en la concatenación de llamadas a librerías dinámicas.

Manejo de Librerías

Para la programación en lenguaje C, el manejo de librerías presenta dos aspectos:

- La utilización de librerías.
- La construcción de librerías.

La utilización es segura para cualquier programa, ya que, como mínimo, habrá que hacer uso de alguna librería perteneciente a la Librería Estándar. En cuanto a la construcción, también podría darse en cualquier programa, pero dada la gran cantidad de librerías existentes, lo normal es que sólo se necesite crear una librería cuando el programa a desarrollar sea considerablemente extenso.

Evidentemente, tanto la utilización como la construcción de librerías serán diferentes dependiendo de si se trata de librerías estáticas o dinámicas.

Librerías de Enlace Dinámico en Windows

Como mencionábamos en un epígrafe anterior, las librerías dinámicas son archivos que contienen funciones y/o recursos que pueden ser requeridos por cualquier aplicación Windows. También indicábamos que podían tener tanto la extensión **.DLL** como extensiones del tipo **.EXE** (ejecutable), **.DRV** (controlador de dispositivo), **.FON** (fuente de Windows), etc. La diferencia entre las librerías cuyo archivo tiene extensión **.DLL** y las creadas sobre archivos **.EXE**, **.DRV**, **.FON**, etc, es que, mientras que las primeras se cargan porque son solicitadas por los programas al SO, el resto se cargan porque aparecen referenciadas (por el propio Windows o por un determinado programa) en archivos de inicialización de Windows.

Ventajas e Inconvenientes del Empleo de DLL's

Ventajas:

- El contenido de una DLL puede ser usado por cualquier aplicación Windows.
- La reutilización de las DLL's implica una reducción en el tamaño de las aplicaciones.
- Reducción del tiempo de compilación y/o carga de las aplicaciones, debido a la disminución del tamaño de las mismas.
- Ahorro de espacio en disco.
- Independencia de las DLL's respecto de las aplicaciones.

Inconvenientes:

- Tienen que almacenarse en la carpeta del sistema para poder ser utilizadas.
- El tiempo que tarda la aplicación en acceder al código que necesita de la DLL es mayor del que emplearía si dicho código formara parte de la propia aplicación.

Estructura de una DLL de 32 Bits

Podríamos dividir la DLL en los siguientes elementos:

- Archivo de cabecera. Conjunto de declaraciones y/o definiciones (de variables, funciones, procedimientos, etc) a usar por la librería.
- Punto de entrada y salida de la DLL (DllEntryPoint). Función que se ocupa de la carga y descarga de la DLL en la memoria principal.
- Funciones de la DLL. Aquellas especificadas e implementadas por el programador de la librería.

Creación de una DLL de 32 Bits

Para la creación de una DLL podemos usar lenguajes del tipo Visual Basic, Delphi, Visual C++, etc.

Con cualquiera de los lenguajes deberemos crear varios archivos, cada uno de los cuales contendrá un tipo de elemento útil para la construcción de la librería. Por ejemplo, si empleásemos Visual C++, deberíamos crear:

- Un archivo con extensión **.c** que contendrá el código fuente de las funciones de la librería.
- Un archivo con extensión **.def** que contendrá la información necesaria para el linkador.
- Dos archivos con extensión **.h** que será los archivos de cabecera del archivo fuente (estos sólo serán necesarios para crear un programa que utilice la DLL, pero no para la creación de la DLL en sí).

Compilación y Linkado de la Librería

Tras la compilación de los archivos anteriores, el compilador generará un archivo **.lib** . Después del linkado, se creará un archivo **.dll** (esta sería la librería en sí).

El acceso a esta DLL podrá hacerse mediante dos tipos de llamadas:

Llamada Estática

Este tipo de llamada va a utilizar el archivo creado por el compilador (**.lib**).

Con este método, el enlace entre el programa y los recursos de la DLL tiene lugar durante el linkado del programa. Es decir, será el linkador quien, utilizando los archivos objeto (**.obj**), los archivos librerías (**.lib**) y los archivos de recursos compilados (.res), cree la aplicación Windows (**.exe**).

Con este proceso, el código de la librería queda incluido en el ejecutable.

Ventajas:

- La librería se carga junto con el ejecutable (la contiene).
- El enlace tiene lugar en tiempo de compilación.
- Las funciones de la librería pueden ser utilizadas como funciones internas de la aplicación.

Inconvenientes:

- La aplicación almacena en su interior el código de la librería, lo que hace que su tamaño sea mayor.
- La librería tiene que incluirse en cada aplicación que la necesite.
- El objetivo de la reutilización sólo se cumple en parte.
- La memoria principal contiene a la librería durante todo el tiempo de ejecución de la aplicación.
- La librería y la aplicación tienen una dependencia total.

Llamada Dinámica

La llamada dinámica emplea el archivo creado por el linkador (**.dll**).

El enlace dinámico, como su nombre indica, se producirá en tiempo de ejecución; es decir, la librería se cargará en memoria cuando la aplicación la requiera al sistema. Este proceso utilizará las funciones LoadLibrary y FreeLibrary para la carga y descarga, respectivamente, de la dll en la memoria principal.

Ventajas:

- La aplicación no almacena junto con su código a la librería, lo que reduce el tamaño de la aplicación (la librería se almacena en un archivo aparte).
- Ninguna aplicación que utilice a la librería deberá incluirla en su código.
- Se utilizan los beneficios de la reutilización en su totalidad.
- La librería sólo se carga en la memoria principal cuando va a utilizarse. Cuando deja de utilizarse podrá descargarse de la memoria.
- La librería y la aplicación son independientes.

Inconvenientes:

- Necesidad de solicitar al SO la carga de la librería.
- El enlace se produce en tiempo de ejecución, hecho que dificulta la manipulación de la librería.
- Las funciones de la librería deben ser llamadas mediante punteros.

Gestión de Medios Magnéticos

Discos Magnéticos

Propiedades de los Discos Magnéticos

Un disco magnético (rígido o flexible) consiste en un soporte de almacenamiento externo que complementa a la memoria principal (RAM) de una computadora.

Sus propiedades más significativas son:

- Capacidad de almacenamiento masivo de información en un espacio muy reducido, con el consiguiente bajo coste relativo por byte almacenado. El cliente realiza una llamada a un servicio como si fuera local.
- El memoria no volátil, ya que mantiene la información almacenada aún a falta de suministro eléctrico.
- Proporciona acceso casi directo al lugar donde se encuentra el bloque de datos a leer o escribir.
- La información almacenada en un disco se agrupa en archivos o ficheros (files) identificables por su nombre.

Actualmente, la mayoría de procesos de E/S de datos utilizan en su origen o destino los discos magnéticos:

- La inmensa mayoría de las aplicaciones se encuentran almacenadas en disco (en forma de archivos 'ejecutables'). Cuando van a utilizarse estas aplicaciones, se copian (en parte) en la memoria principal y son ejecutadas desde ésta.
- Después de procesar los datos que se encuentran en la memoria principal, los resultados de este proceso se almacenarán en disco.

Por último, otra característica a indicar sobre los discos magnéticos (los 'discos duros' en este caso), es que se pueden utilizar como memoria virtual; es decir, como una extensión de la memoria principal del ordenador.

Estructura Física de los Discos Magnéticos

Físicamente, los discos magnéticos están fabricados con: mylard en el caso de los discos flexibles y aluminio o cristal cerámico en el caso de los discos rígidos.

La estructura física de un disco la forman unas superficies magnéticas denominadas **caras**, cada una de las cuales se divide en anillos concéntricos que constituyen las **pistas**, que a su vez agrupan a los **sectores** (unidades mínimas de almacenamiento cuya capacidad habitual suele ser de 512 bytes de información).

El proceso de grabación de los discos se logra, al igual que en un grabador de audio, por la acción de un campo magnético de polaridad reversible (N-S ó S-N), que imanta la pista al actuar sobre ella. Para este proceso, existe una cabeza para cada cara del disco. Los brazos que soportan a las cabezas se mueven juntos; es decir, que si la cabeza de la cara superior está sobre una determinada pista, la de la cara inferior se encontrará situada en la misma pista.

La lectura la realizan las mismas cabezas, mediante un proceso inverso al de grabación, a través del cual detectarán los campos magnéticos existentes a lo largo de la pista accedida.

En el proceso, tanto de grabación como de lectura, sólo podrá encontrarse activa una única cabeza de las existentes en el medio magnético (dos en los discos flexibles y múltiples en los rígidos).

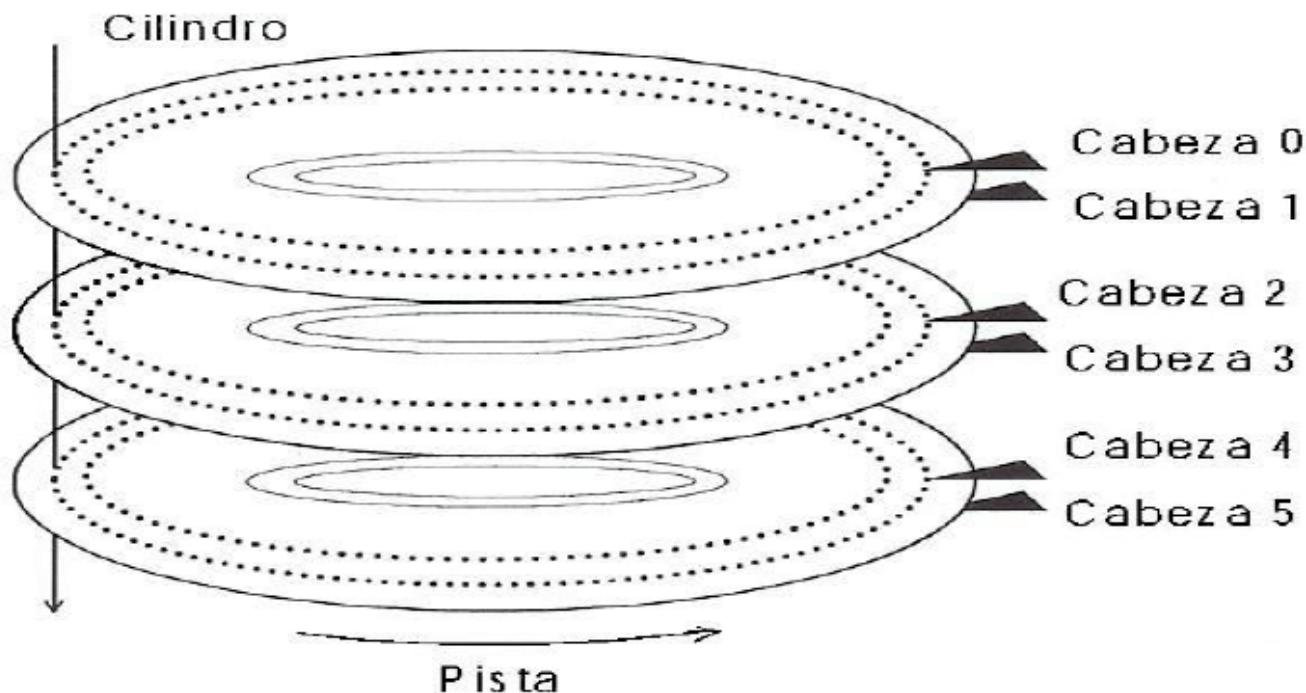
En las propiedades indicadas en el epígrafe anterior se hacía referencia al acceso casi directo de los discos magnéticos. Conociendo ya la estructura física de estos discos podemos indicar que lo de directo se refiere a la forma de acceso a las pistas y lo de casi hace referencia a la forma de acceso a los sectores una vez situados en la pista correspondiente (este último es un acceso secuencial cuyo tiempo es tan reducido que se considera despreciable). Para esta operación de localización de un sector concreto dentro del disco se emplea lo que se conoce como su dirección o CHS (número de cilindro, número de cabeza, y número de sector).

Importancia del concepto de cilindro

El hecho de que un disco rígido sea en realidad una agrupación de discos (o platos) cada uno de los cuales dispone de dos caras, además de duplicar la capacidad de almacenamiento, permite la lectura o escritura del doble de datos antes de desplazar el cabezal a otra pista, accediendo a una cara y luego a la contraria. De esto surge el concepto de cilindro que no es más que el conjunto de pistas que se sitúan bajo las cabezas de lectura/escritura en un momento determinado (o conjunto de pistas de un disco que tienen el mismo radio). Este concepto también es aplicable a los discos flexibles, aunque, al disponer estos de dos caras únicamente, se suele asociar a los discos rígidos porque en estos el concepto de cilindro es gráficamente más evidente.

De lo anterior se deduce que la mejor forma de grabar la información sobre los discos magnéticos es cilindro a cilindro acelerando con ello el proceso de escritura/lectura al minimizar los movimientos de los cabezales en búsqueda de las pistas.

El número de cilindros de un disco, por tanto, se corresponderá con el número de posiciones en las que pueden situarse los cabezales; enumerándose aquellos desde 0 (el más exterior) en forma creciente hacia el interior, correspondiendo el número mayor al más interno.



Representación gráfica de los cilindros de un disco duro

Posicionamiento, Latencia y Acceso en un Disco Rígido o Flexible

El acceso a un sector situado en una determinada cara del disco, pasa por posicionar el cabezal sobre el cilindro donde se encuentra la pista que contiene el sector, y, posteriormente, esperar a que, mediante el giro del disco, el sector deseado se sitúe debajo de la cabeza. En esta operación intervienen dos tiempos:

- **Posicionamiento** . Tiempo necesario para que el brazo con la cabeza correspondiente se coloque directamente sobre el cilindro seleccionado (pocos milisegundos).
- **Latencia (demora rotacional)** . Tiempo necesario para que el sector a localizar se sitúe bajo la cabeza lectora/escritora (en promedio es el tiempo de media vuelta).

El tiempo de acceso resulta pues, la suma de los anteriores, o lo que sería igual, el tiempo que transcurre desde que la controladora envía la orden al cabezal de posicionarse sobre un cilindro, hasta que la cabeza correspondiente accede al sector buscado.

$$T_{\text{acceso}} = t_{\text{promedio posicionamiento}} + t_{\text{promedio latencia}}$$

Estructura Lógica de los Discos Magnéticos

Desde el punto de vista de la estructura lógica de un disco duro, podríamos dividirlo en:

- **Sector de arranque (Master Boot Record)** . Es el primer sector del disco (0, 0, 0), y en él se encuentra la tabla de particiones y un pequeño programa de inicialización. Este programa se ejecuta al encender la computadora, y su función es leer la tabla de particiones y ceder el control a la partición primaria activa.
- **Espacio particionado** . Zona del disco que contiene las particiones. Una partición es cada una de las divisiones de tamaño fijo de un disco que se asocia a una unidad lógica (C:, D:, etc, en el caso de los SO Windows). Cada partición ocupa un bloque de cilindros contiguos del disco duro, pudiendo establecerse distintos sistemas de archivos (FAT, NTFS, ...) para las distintas particiones posibles.
- **Espacio NO particionado**. Se trata de la zona de disco que no ha sido particionada y que, por lo tanto, no puede ser utilizada.

La tabla de particiones del disco duro puede contener hasta 4 entradas, lo que determina el número máximo de particiones primarias que se pueden crear en el disco. No obstante, este límite de particiones puede superarse empleando una de las entradas para almacenar una partición extendida (tendríamos 3 primarias y 1 extendida). La partición extendida podrá contener tantas unidades lógicas como necesitemos.

La principal diferencia entre las particiones primarias y las extendidas es que, mientras que las primeras son arrancables y pueden ser utilizadas para contener a los SO, las extendidas no son arrancables y se utilizan normalmente para almacenar datos. Además, de entre las distintas particiones primarias, habrá que indicar cuál es la activa, es decir, la verdaderamente arrancable.

Visto lo anterior, lo primero que hay que hacer con un disco duro antes de su utilización es:

- **Crear las particiones** (utilizando herramientas del tipo FDISK).
- **Formatear las particiones creadas** . Este proceso consiste en la creación de la estructura que permita el almacenamiento de información utilizando un determinado sistema de archivos.

En el sistema de archivos FAT (MS-DOS y sistemas Windows), la estructura lógica de una partición la forman: el sector de arranque, varias copias de la tabla de asignación de

archivos, el directorio raíz y el área de datos. La FAT (tabla de asignación de archivos) es el índice del disco. En ella se indican los clusters (unidades de asignación) que utiliza cada archivo, así como los libres y los defectuosos.

La estructura lógica de una partición utilizada por un sistema UNIX tradicional está constituida: un bloque de arranque, un superbloque (contiene el número de inodos, el número de bloques, etc), un vector de inodos (similar a la FAT anterior) y los bloques de datos.

Herramientas para la Gestión de Discos Magnéticos

Considerando como herramientas de gestión las expuestas en el punto anterior (herramientas de particionado y formateo), en este punto vamos a centrarnos en otras herramientas no tan esenciales como aquellas, pero sí bastante comunes en la gestión de discos magnéticos:

- **Comprobador de errores** . Su misión es analizar el contenido del disco en búsqueda de incoherencias en el sistema de archivos. Si, por ejemplo, en un sistema FAT existen dos archivos que apuntan al mismo contenido aparecerá un error de vínculos cruzados o si aparecen datos no asociados a ningún archivo se indicará el error de cadenas perdidas. Comprobadores de errores usuales son: chkdsk /f en MS-DOS, ScanDisk en Windows y fsck en UNIX.
- **Desfragmentador de disco** . Esta herramienta busca la agrupación física (sobre el disco) de toda la información concerniente a un mismo archivo, con el fin de acelerar la lectura de datos. La fragmentación se produce por la creación, modificación y eliminación de archivos. El sistema de archivos de UNIX no precisa de desfragmentador debido a que su velocidad de trabajo no se degrada con la creación, modificación y eliminación de archivos.
- **Compresor de datos** . Se trata de un método que busca maximizar la capacidad de las particiones mediante la compresión de la información que contienen. No obstante, esta metodología ralentiza el funcionamiento general del SO, debido a que deben ejecutarse continuamente algoritmos de compresión/descompresión. Además, tiene como inconveniente la dependencia de la información del programa de compresión, circunstancia que podría provocar problemas de incompatibilidad d futuros en caso de producirse errores.

Normalmente, resulta más eficaz la compresión de ficheros de forma independiente (en lugar de particiones completas).

Los SO actuales incorporan sus propios métodos de compresión/descompresión (en UNIX: gzip -> para archivos independientes, tar -> para árboles de archivos). Además, existen herramientas ajenas a los SO para realizar estas operaciones (WinZip, WinRAR, IsoBuster, ...)

- **Copias de seguridad** . La realización de copias de seguridad del contenido del disco en otro medio de almacenamiento, es un método para garantizar la recuperación de datos destruidos por errores humanos, de situaciones imprevistas o de hardware.

Sistemas RAID (Redundant Array Of Independent Or Inexpensive Disks)

Los sistemas de matriz de discos independientes (baratos) redundantes son utilizados para el control de errores en los discos. Emplean varios discos para evitar (o minimizar) la pérdida de información en caso de que se produzca algún error. La redundancia hace referencia a la información extra que no sería necesaria si no se produjesen errores.

La gestión de los sistemas RAID no es accesible por el usuario, pudiendo ser gestionada por hardware (tarjetas RAID) o por software (SO). como suele ocurrir, el método más eficiente (pero más costoso económicamente) es el que utiliza tarjetas hardware, debido a que desocupa a la CPU de las tareas RAID.

A continuación se enumeran los niveles RAID más habituales:

- **RAID 0** (disk striping, discos en bandas). En este nivel, la información se distribuye entre todos los discos que forman el conjunto RAID, proporcionando una mayor velocidad en las transferencias debido al trabajo conjunto de todos los discos para acceder a un mismo archivo. No obstante, si falla alguno de los discos perderemos toda la información. La implementación de RAID 0 precisa de 2 discos como mínimo.
- **RAID 1** (disk mirroring, discos en espejo). Basado en el empleo de discos para duplicar la información. Con este método, cada vez que se escriba en un disco, deberá grabarse la información de su disco copia para mantener la coherencia. A diferencia del método anterior, si en éste falla un disco, el sistema podrá continuar funcionando sin detenerse. Es habitual implementar RAID 1 con 2 discos. Este sistema permite una capacidad de almacenamiento igual a la mitad de la capacidad total de los discos de que disponemos. Pueden combinarse RAID 0 y RAID 1 para formar el sistema RAID 10. Con RAID 10, la información se distribuye en bandas por varios discos y cada disco se duplica, lo que requiere un número par de discos (4, 6, 8, ...).
- **RAID 2**. Ofrece detección y corrección de errores en los discos mediante la utilización de códigos de Hamming. Este nivel está incluido en la actualidad en los propios discos, por lo que ha dejado de ser un sistema a elegir por el usuario.
- **RAID 3**. Emplea un disco para almacenar la paridad. La información se distribuye a nivel de bits entre los distintos discos. Si un disco falla, la información se reconstruiría mediante la operación O-exclusiva (XOR) de los discos restantes. Son necesarios un mínimo de 3 discos para implementar un RAID 3. Todos los discos funcionan a la vez, lo que hace bajar el rendimiento con sistemas transaccionales (múltiples accesos sobre pequeñas cantidades de datos).
- **RAID 4**. Utiliza un disco para el almacenamiento de la paridad, al igual que el anterior; si embargo, los datos se distribuyen a nivel de bloque (en lugar de a nivel de bits) y se puede acceder a cada disco de forma individual. Este hecho mejora el rendimiento en sistemas transaccionales.
- **RAID 5**. La paridad se almacena entre todos los discos, eliminando el excesivo uso del disco de paridad que hacían los dos niveles anteriores. Este método es el más eficiente, ofreciendo la mayor tasa rendimiento/coste y el menos coste por megabyte de información. Se necesitan al menos 3 discos para su desarrollo; no obstante, el funcionamiento óptimo se alcanza a partir de los 7 discos.

Los últimos tres niveles se denominan de discos en bandas con paridad (disk striping with parity), y en ellos, podremos calcular la capacidad máxima de información que pueden almacenar sumando la capacidad de todos los discos y restándoles la capacidad de uno (redundancia).

Otros Medios Magnéticos

ZIP

Tienen un tamaño similar a los floppys de 3,5", lo que los hace fácilmente portables. Sus capacidades habituales son 100 y 250 Mb, aunque actualmente existen de 750 Mb.

JAZZ

Son discos similares a los anteriores (son compatibles) pero con capacidades de 1 y 2 Gb.

Tecnología Magneto-Óptica

LS-120 Superdisk

La tecnología Láser Servo fue desarrollada en 1996. Se trata de una tecnología mixta (magnética y óptica) compatible con la de los floppys tradicionales; es decir, un lector/grabador de este tipo puede leer y escribir sus propios discos de 120 Mb y los floppys convencionales de 1,44 Mb.

Este sistema es producto de una mezcla de tecnologías de los floppys, discos duros y CD-ROM's.

Combina una medio magnético con un método óptico utilizado para el posicionamiento de las cabezas de lectura/escritura, lo que conduce a un aumento considerable de la capacidad del medio respecto de los floppys, y a lograr velocidades de transferencia de hasta 400 kb/s (la mitad de veloces que los ZIP).

Discos Magneto Ópticos (MO)

Emplean, para la grabación del medio, un láser que calienta la superficie del disco (302º F). Existen dos variantes de funcionamiento de esta tecnología:

- El calor provoca la oxidación del metal del medio, lo que permite la orientación de su magnetismo mediante un imán (es la técnica más empleada).
- El calor cambia la estructura del medio, provocando que sea cristalino o amorfo.

Existen discos de 5,25" (650 Mb, 1.3 Gb, 2.6 Gb y 4.6 Gb) y de 3,5" (128, 230 y 640 Mb).

Cintas Magnéticas

Ofrecen una gran capacidad de información junto con velocidades de transferencia muy bajas; motivo por el cual son empleadas casi exclusivamente para realizar copias de seguridad. Suponen un coste ínfimo por Mb.

- **DIGITAL AUDIO TAPE (DAT)** ↗
 - ancho de cinta → 4 mm
 - capacidad → desde 1,2 hasta 5 Gb
- **EXABYTE** ↗
 - ancho de cinta → 8 mm
 - capacidad → desde 2,3 hasta 5 Gb
- **QUARTER INCH CARTRIDGE (QIC)** ↗
 - ancho de cinta → 6,35 mm
 - capacidad → desde 60 Mb hasta 1,35 Gb
- **DIGITAL LINEAR TAPE (DLT)** → capacidad → 10 Gb
- **MINI CARTRIDGE** → capacidad → desde 750 Mb hasta 4 Gb

Tipos de cintas magnéticas

Copias de Seguridad (BACKUP)

Las copias de seguridad o backups pueden definirse como copias de la información realizadas usando un medio de almacenamiento secundario, cuyo objetivo es salvaguardar la información ante posibles errores humanos, de hardware, etc.

Pérdida de información

Las pérdidas de datos pueden provenir de las circunstancias más variadas:

- Fallo del disco duro.
- Error humano (eliminación no deseada).
- Interrupción de una aplicación por fallo durante la grabación de la información en el disco.
- Acción de un virus o troyano.
- Accidente inevitable en el entorno del sistema informático (incendio, inundación, etc).

Frecuencia de Backups

La frecuencia de ejecución de backups dependerá tanto de la frecuencia de actualización de la información del sistema como de la información que el administrador del mismo esté dispuesto a perder.

En sistemas relativamente importantes los backups son diarios.

Medios empleados para las Backups

Las copias de seguridad pueden hacerse, entre otros medios, sobre:

- Una partición dentro del mismo disco duro que contiene la información a proteger (mínima protección).
- Un disco duro auxiliar, dentro del mismo equipo donde se encuentra el disco duro con la información a proteger.
- Un disco duro en un equipo distinto al que contiene la información a proteger (backup por red).
- Un CD-R, CD-RW, DVD-RW, DVD+RW, etc.
- Una cinta magnética (tape backup).
- Floppys, ZIP's, JAZZ's, etc. (para copias de pequeñas cantidades de información).

Tipos de Backups

- Completa. Copia de toda la información a salvaguardar.
- Progresiva o Incremental. Copia de la información nueva o modificada desde el último backup completo o progresivo (se necesitaría la última copia completa y todas las copias incrementales para restaurar la información).
- Diferencial. Copia de la información nueva o modificada desde la última copia completa (la recuperación de los datos precisa de la última copia completa y la última diferencial).

Controles de Cambio

Controlar el Proyecto y Eliminar los Retrasos

Los cambios son un pilar básico dentro de la vida del desarrollo de software. En la práctica, el trabajo requiere de una administración formal de los cambios. Si contamos con una administración de cambios del software realmente efectiva podremos conseguir que:

- Los equipos de desarrollo puedan entregar el software dentro del tiempo y presupuesto establecidos y con una calidad predecible.
- Los líderes de proyecto conozcan en todo momento el estado y avance del desarrollo del software y tengan certeza del mismo dentro del tiempo prefijado.
- Los desarrolladores utilicen y controlen con orden y seguridad sus colecciones de archivos y componentes diferentes para cada aplicación.
- Los ‘probadores’ sepan cuándo una nueva construcción de software requiere ser sometida a un paquete de pruebas y las mejoras o correcciones que debe presentar.

Las organizaciones de desarrollo exitosas consideran que el control de cambios durante todo el ciclo es la clave para asignar prioridades a las actividades del equipo, así como para controlar las dificultades que surjan durante el desarrollo. Si no implementamos dicho control, el caos de los cambios se apoderará del control del proyecto.

La Administración de la Configuración y Control de Cambios (SCM) es la disciplina de la ingeniería de software que agrupa las herramientas y técnicas de uso de las mismas que una compañía emplea para administrar los cambios de los componentes de software. Cuando la SCM se encuentra integrada en otras actividades del desarrollo (requerimientos, análisis y diseño, construcción, pruebas), se denomina Gestión de Cambio Unificada (UCM).

Existen guías que describen cómo controlar, dar seguimiento y monitorear los cambios para permitir un desarrollo iterativo exitoso; así como la forma de establecer espacios de trabajo seguros para cada desarrollador, aislándolo de los cambios realizados en otros espacios de trabajo y controlando los cambios de todos los artefactos de software (modelos, código, documentos, etc). Para llevar a cabo estas metodologías, se utilizan herramientas de ‘control de versiones y configuraciones’ y de ‘control de cambios’ que, por un lado, automatizan las metodologías, y por otro unen al equipo de desarrollo para conseguir un trabajo paralelo y coordinado. Estas herramientas permiten a cada desarrollador contar con un espacio de trabajo seguro donde puede realizar los cambios de manera independiente para que una vez probados puedan integrarse con el resto del desarrollo, garantizando de esta forma la calidad, el tiempo de entrega y la satisfacción del cliente con el producto desarrollado.

Gestión de Cambios

Podemos distinguir dos enfoques diferentes dentro de la gestión de cambios, dependiendo del mayor o menor grado de modificación del producto.

Si el cambio a realizar afecta a gran parte de los componentes del producto, podrá plantearse como un nuevo desarrollo, y aplicar un nuevo ciclo de vida desde el principio, aunque aprovechando lo ya desarrollado de la misma forma que se reutilizan los prototipos.

Si el cambio afecta a una parte bastante localizada del producto, entonces se puede organizar como simple modificación de elementos. Hay que tener en cuenta que cualquier cambio en el código del producto software siempre implicará una revisión de los elementos de documentación afectados; es decir, cambiar el código de algunos módulos puede requerir, además, modificar los documentos de diseño o incluso, en el caso de mantenimiento perfectivo, modificar el documento de especificación de requisitos.

Tomando como referencia la gestión, la realización de cambios se puede controlar mediante dos clases de documentos, que en ocasiones pueden unirse para formar un único informe:

- Informe de problema: describe una dificultad en la utilización del software que precisa de alguna modificación para subsanarla.
- Informe de cambio: describe la solución dada a un problema y el cambio realizado en el producto software.

El primer documento puede ser originado por los propios usuarios. Este informe se pasa a un grupo de ingeniería para la comprobación y codificación del problema planteado, y posteriormente a un grupo de gestión para decidir la solución a adoptar. Este grupo de gestión da comienzo al informe de cambio, que se pasa de nuevo al grupo de ingeniería para su total desarrollo y ejecución.

Comprobación de los Objetivos del Control de Cambios

Los objetivos del control de cambios son comprobados al realizar entrevistas con:

- El Director de Tecnologías de Información.
- La Administración de la Función de Servicios de Información.
- La Administración de Desarrollo de Sistemas, Aseguramiento de la Calidad de Control de Cambios, Operaciones y Seguridad.
- La Administración de Usuarios implicada en el diseño y manejo de aplicaciones de sistemas de información.

De las que se obtienen:

- Procedimientos organizacionales relacionados con la planificación de sistemas de información, el control de cambios, la seguridad y el ciclo de vida de desarrollo de sistemas.
- Procedimientos de la función de servicios de sistemas de información relacionados con la metodología del ciclo de vida de desarrollo de sistemas, el aseguramiento independiente de la calidad, los estándares de seguridad, la implementación, la distribución, el mantenimiento, la liberación del software, los cambios de emergencia y el control de versiones del sistema.
- Un plan de desarrollo de aplicaciones.
- Formato y bitácora de requisitos de control de cambios.
- Contratos con proveedores relacionados con servicios de desarrollo de aplicaciones.

Evaluación de los Controles

La evaluación de los controles de cambios deberá tener en cuenta si:

- La bitácora de control de cambios garantiza que cualquiera de los cambios mostrados han sido resueltos.
- El control de cambios es un procedimiento formal tanto para los grupos de desarrollo como para los usuarios.

- El usuario está conforme con el resultado de los cambios solicitados, el tiempo de realización de los mismos y los costes.
- Para una muestra de cambios en la bitácora de control de cambios:
 - la documentación actual refleja con exactitud el ambiente modificado,
 - los cambios hayan sido efectuados tal y como fueron documentados,
 - el cambio implicó modificaciones en los programas y operaciones.
- El proceso de cambios es controlado respecto de las mejoras en el conocimiento, la efectividad en el tiempo de respuesta y la satisfacción del usuario con el resultado del proceso.
- El mantenimiento del sistema de Intercambio de Rama Privada (PBX) se incluye en los procedimientos de control de cambios.

Evaluación de la Suficiencia

La comprobación de la suficiencia del control se realizará probando que:

- Para una selección de cambios, la administración ha aprobado los siguientes puntos:
 - petición del cambio,
 - descripción del cambio,
 - acceso al programa fuente,
 - terminación del cambio por parte del programador,
 - solicitud para trasladar el programa fuente al entorno de prueba,
 - finalización de las pruebas de aceptación,
 - solicitud de compilación y traslado al grupo de producción,
 - identificación y aceptación del impacto general y específico,
 - elaboración de un proceso de distribución,

y revisando el control de cambios en cuanto a la inclusión de:

- fecha del cambio solicitado,
- persona o grupo que lo solicita,
- petición aprobada de cambios,
- aprobación del cambio efectuado (servicios de información),
- aprobación del cambio efectuado (usuarios),
- fecha de actualización de la documentación,
- fecha del traslado al grupo de producción,
- aprobación del cambio por parte del grupo de aseguramiento de la calidad,
- aceptación por parte del grupo de operaciones.

También se tendrán en cuenta, a la hora de evaluar esta suficiencia:

- Los tipos de análisis de cambios realizados sobre el sistema para la determinación de tendencias.
- La valoración de la adecuación de las librerías de la función de servicios de información y la identificación de la existencia de niveles de código base para advertir y prevenir la ocurrencia de errores.
- Si existen procedimientos de E/S (“check in/check out”) para cambios.
- Si la totalidad de los cambios en la bitácora fueron resueltos con la conformidad de los usuarios y si no se llevaron a cabo cambios que no hayan sido anteriormente especificados en la bitácora.
- Si los usuarios tienen conocimiento de la necesidad de procedimientos formales de control de cambios.
- El proceso de reforzamiento del personal garantiza el cumplimiento de cada uno de los procedimientos de control de cambios.

Evaluación del Riesgo de los Objetivos de Control NO Alcanzados

Esta evaluación se realizará llevando a cabo mediciones (“benchmarking”) de la administración del control de cambios contra organizaciones similares o estándares internacionales de buenas prácticas reconocidas en la industria correspondiente.

Para sistemas seleccionados de la función de servicios de información, se ejecutará:

- Una verificación para comprobar si la documentación determina el requerimiento o si el cambio del sistema ha sido aprobado y priorizado por parte de la administración de las áreas usuarias afectadas y el proveedor de servicios.
- La confirmación de la existencia y adecuación de evaluación del impacto en formas de control de cambios.
- La obtención del conocimiento del cambio mediante una comunicación de la función de servicios de información.
- La asignación del cambio a los correspondientes recursos de desarrollo.
- La adecuación de los sistemas y los planes de prueba de los usuarios.
- La migración formal de prueba a producción a través del grupo de aseguramiento de la calidad.
- La puesta al día de los manuales de usuario y de operación para mostrar el cambio efectuado.
- El reparto de la nueva versión a los usuarios correspondientes.

Además, la evaluación de este riesgo concluirá determinando, para una selección de cambios de información, que:

- Sólo se efectuaron cambios que hayan sido aprobados por la función de servicios de información.
- Todos los cambios han sido tenidos en cuenta.
- Las librerías actuales (fuente y objeto) muestran los últimos cambios llevados a cabo.
- Las modificaciones en el procedimiento de control de cambios de:
 - aplicaciones internas y adquiridas,
 - software de sistemas y de aplicación,
 - gestión del control de cambios por parte del proveedor.

Introducción

Los detalles de administración de una red dependen en gran medida del SO de red utilizado. Se ha optado por desarrollar la mayor parte de los apartados desde la perspectiva de dos SO: Windows 2000 y Red Hat Linux 6.0.

La estructura general del presente tema será la siguiente. En primer lugar se realiza la introducción a las funciones del administrador de redes. A continuación se presentan algunos conceptos básicos de TCP/IP necesarios para comprender la mayoría de las herramientas de administración de red. En tercer lugar se presentan algunas de las facetas de administración de redes bajo el prisma de los dos SO elegidos. Por último se hace un breve repaso a las herramientas de monitorización de tráfico y supervisión de red.

Funciones del Administrador de la Red

Las funciones de un administrador de red son tan amplias y variadas como desconocidas en muchos ámbitos. Frente a la idea preconcebida y bastante generalizada de que un administrador es la persona encargada de dar de alta a los usuarios y administrar el

acceso a los recursos más habituales como programas, archivos e impresoras, encontramos una realidad mucho más amplia.

Para que una red funcione correctamente y esté preparada para crecer en un entorno de alta demanda de recursos es necesario realizar un importante trabajo que no siempre es conocido.

Las funciones de un administrador de red pueden dividirse en varias categorías principales:

- Administración de usuarios.
- Mantenimiento del hardware y software.
- Configuración de nodos y recursos de red.
- Configuración de servicios y aplicaciones internet.
- Seguridad de red.
- Supervisión y optimización.

Administración de usuarios

Desde el momento en que se crea una cuenta de usuario para su acceso a la red es necesario estar preparado para darle el soporte necesario con el fin de facilitarle el acceso a los recursos del sistema.

Algunas de las tareas más comunes relacionadas con la gestión de usuarios son:

- Creación y borrado de cuentas de usuario.
- Creación de grupos.
- Formación de usuarios.
- Definición de políticas de uso.
- Configuración de sistema de ayuda para gestionar el soporte a los usuarios.
- Asignación de espacio de almacenamiento y ficheros en el sistema de archivos compartido.
- Creación de cuentas de correo electrónico.

Mantenimiento del hardware y software

Fundamentalmente este punto cubre la necesidad de establecer y mantener los recursos de infraestructura de la red. En este punto podemos destacar las siguientes tareas:

- Instalación y configuración del SO.
- Administración de discos y del sistema de archivos.
- Administración RAID.
- Instalación de controladores de dispositivos.
- Recuperación y diagnóstico de fallos del sistema.
- Adición y eliminación de hardware del sistema.

Configuración de nodos y recursos de red

Para que la red funcione como un sistema coherente con entidad propia, y no como una serie de nodos aislados, es necesario llevar a cabo una serie de tareas:

- Instalación y configuración de tarjetas de red.
- Conexión de host a concentradores y conmutadores.
- Conexión de host a routers.
- Configuración de routers.
- Configuración de impresión en red y compartición de archivos.
- Administración de DNS.
- Integración de diferentes SO de red.

Configuración de servicios y aplicaciones internet

Actualmente son pocos las redes locales que pueden concebirse sin pensar en una interconexión con la red WAN más extendida: la red internet. Por ello, dentro de las funciones de un administrador de red debemos incluir:

- Configuración de servidores y clientes de correo electrónico.
- Instalación y configuración de servidores web.
- Instalación y configuración de servidores de aplicaciones.
- Instalación y configuración de servidores de BD.

Seguridad de red

El impresionante auge de Internet en los últimos años ha provocado un crecimiento exponencial de los usuarios de esta red. Este hecho proporciona una serie de ventajas derivadas de la diversidad de los servicios e información que podemos encontrar en Internet, pero también supone un gran potencial para el uso indebido y la actividad criminal. Por otra parte, es necesario contemplar también el aspecto de seguridad interna, dentro de la propia red local. Para ello, deben considerarse las siguientes actividades:

- Configuración de firewalls.
- Acceso remoto a la red.
- Creación de directivas de seguridad.

Supervisión y Optimización

Es necesario controlar los problemas derivados de un malfuncionamiento de la red, de un mal uso de la misma o de debilidades en la estrategia de seguridad:

- Afinar el rendimiento de la red.
- Supervisión y registro.
- Gestión de cambios.
- Auditoria y pruebas.

Conceptos básicos de TCP/IP para Administradores

Redes TCP/IP

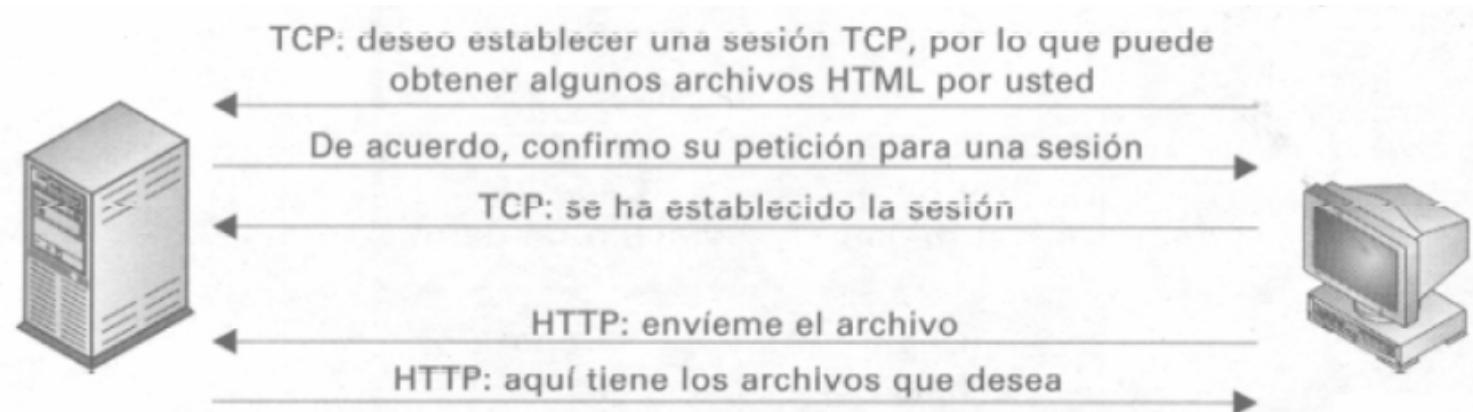
El propósito de TCP/IP es proporcionar a los equipos un método para transmitir datos entre sí. TCP/IP proporciona los medios para que las aplicaciones envíen datos entre redes y para que las redes entreguen esos datos a las aplicaciones de otros equipos o *host*.

En TCP/IP normalmente a los *host* se les asignan nombres que corresponden a sus direcciones IP, para facilitar la tarea de identificación de las máquinas.

La transmisión de datos en TCP/IP se realiza con paquetes de datos. Los paquetes son bloques diferenciados de datos que se transmiten de forma independiente de la red. Cada paquete contiene los datos que cada *host* desea compartir, así como la información necesaria para encaminar el paquete al destino correcto a través de la red.

Cada protocolo agrega sus propios datos en el encabezado para establecer la información que le permita mantener una conversación lógica con el protocolo homónimo en el host con el que transfiere información.

La siguiente figura muestra un ejemplo simplificado del intercambio de información que se produce entre una estación de trabajo y un *host* para intercambiar una página HTML a través del protocolo HTTP sobre TCP/IP:



Como puede verse, el protocolo TCP proporciona a los dos nodos que se comunican los recursos necesarios para establecer sesiones de diálogo. Una vez establecida dicha sesión, se transfieren paquetes de datos con un protocolo http, situado en un nivel superior de la pila de protocolos.

TCP/IP frente al Modelo OSI

TCP/IP es una familia de protocolos y aplicaciones que realizan funciones diferenciadas que se corresponden con capas específicas del modelo OSI (Open System Interconnection). En este punto es importante establecer un matiz importante: TCP/IP no es una familia de protocolos del modelo OSI.

De hecho, el modelo OSI nació de la necesidad de establecer un estándar que garantizara la interconexión de sistemas y equipos de diferentes fabricantes, cuando el estándar de facto que se ha impuesto mayoritariamente en el mercado es TCP/IP. Sin embargo, si existe una clara similitud entre TCP/IP y OSI en cuanto a la separación de responsabilidades a través del establecimiento de capas de protocolos, tal como muestra la siguiente tabla:

MODELO TCP/IP	MODELO DE REFERENCIA OSI
Datos de aplicación	Capa de aplicación
	Capa de presentación
	Capa de sesión
TCP	Capa de transporte
IP	Capa de red
Control de acceso al medio	Capa de vínculo de datos
Capa física	Capa Física

IP es el protocolo de capa de red más utilizado en todo el mundo. La capa de red tiene dos misiones fundamentales:

- Proporcionar una dirección única a cada host de la red: los host de cada red concreta tienen direcciones que los identifican únicamente como miembros de dicha red.
- Encaminar los paquetes a través de la red, seleccionando el camino más adecuado.

TCP es un protocolo orientado a conexión de nivel de transporte que asigna a las aplicaciones y protocolos de nivel superior un número de puerto. Estos números se transmiten junto con los datos a través de la red y son utilizados por el receptor para determinar a qué proceso debe entregar los datos recibidos.

TCP/IP no implementa estrictamente los niveles 5, 6 y 7 del modelo OSI como componentes independientes y diferenciados. Las aplicaciones que utilizan TCP/IP normalmente implementan alguno de estos niveles, o todos, dentro de la aplicación según sea necesario.

Direccionamiento IP

Cada dirección IP está formada por un conjunto de 32 bits, organizados en 4 grupos de 8 bits cada uno. A cada grupo se le denomina octeto y está separado de los demás por un punto. Normalmente las direcciones IP se escriben en formato decimal, y tienen el siguiente aspecto:

192.168.15.1

No obstante, en muchas ocasiones resulta útil representar las direcciones IP en formato binario, considerando los bits de *orden superior* los situados a la izquierda y los de *orden inferior* los situados a la derecha:

11000000.10101000.00001111.00000001

Cada dirección IP contiene una parte de host y una parte de red. El espacio de direcciones disponible se divide en redes de clase A, clase B y clase C. Cada red contiene diferentes cantidades de redes y host por clase. Esto se consigue reservando para cada clase un número de bits de orden superior diferente para la dirección de red, tal como muestra la siguiente tabla:

Clase	Bits de orden superior en el primer octeto	Número de bits adicionales en el campo de red	Número de bits en el campo host	Rango
A	0	7	24	1.0.0.0 – 126.255.255.255
B	10	14	16	128.0.0.0 – 191.255.255.255
C	110	21	8	192.0.0.0 – 223.255.255.255
D	1110	N/D	N/D	224.0.0.0 – 239.255.255.255
E	11110	N/D	N/D	240.0.0.0 – 255.255.255.255

Este esquema está pensado para asignar a las empresas y organizaciones direcciones IP de acuerdo con el número de host para los que necesitan una dirección. Las direcciones de clase A se asignan únicamente a grandes organizaciones y empresas, mientras que las de clase C se asignan a organizaciones y empresas pequeñas. Las redes 0 y 127 están reservadas para propósitos especiales.

Es importante destacar la diferencia entre una dirección IP pública y una dirección IP privada. En el ámbito de la red internet, una IP pública es una dirección de red accesible y reconocible desde cualquier punto de la red. Los organismos encargados de asignar direcciones IP públicas lo hacen siguiendo el criterio de clases expuesto en la tabla anterior según las necesidades de cada compañía.

Una dirección IP privada es aquella que únicamente es visible desde un ámbito de red local, pero no desde el ámbito de internet. En este sentido, las compañías tienen total libertad para escoger direcciones IP de la clase que les resulte más conveniente, ya que no interfieren con el mundo exterior.

Segmentación de Redes

A menudo puede resultar conveniente subdividir una red local en varias subredes. Los motivos pueden ser diversos:

- Si el número de terminales y host es muy elevado.
- Para reducir el número de colisiones en la red.
- Para facilitar la administración.
- Permite crear dominios o subredes lógicas en las que un conjunto relativamente homogéneo de usuarios accede con frecuencia al mismo conjunto de recursos compartidos.

Para subdividir el espacio de direcciones de una red IP, se ha provisto a la dirección IP de dos componentes: la dirección IP en sí misma, y la máscara de subred.

Dirección IP: 192.168.15.1

Máscara de subred: 255.255.255.0

La máscara de subred, en su representación binaria, está formada por una serie de bits a 1 en la parte de orden superior, y una serie de bits a 0 en la parte de orden inferior.

Sin dividirla en subredes, una red de clase A tiene una máscara de subred de 8 bits (255.0.0.0); una red de clase B tiene una máscara de 16 bits (255.255.0.0); y una de clase C tendrá una máscara de subred de 24 bits (255.255.255.0). Es decir, en cualquiera de los

casos para una red sin subdivisiones la máscara de red está formada por el número de bits necesarios para determinar la clase de la red más el número de bits disponibles para identificar la red.

Para subdividir una red se agregan bits a 1 por la derecha a la parte de red de la máscara de subred. Al hacerlo, es necesario tener en cuenta las siguientes consideraciones:

- La primera subred (todos los 0) de cualquier red está reservada.
- Las direcciones primera y última de los host de cualquier subred están reservadas como direcciones de difusión.

En el siguiente ejemplo se muestra una red de clase C dividida en 8 subredes:

Dirección IP: 192.168.1.1

Máscara de subred: 255.255.255.224

Es decir, se han añadido 3 bits a la máscara de red. Esto también suele expresarse de la siguiente forma:

192.168.1.1/27

que indica que la máscara de red tiene en la parte de orden superior 27 bits a 1. El número de subredes en los que se divide la red se calcula como 2 elevado al número de bits añadidos. En este ejemplo, 8 subredes. A esta técnica para subdividir redes también se le conoce con el nombre de subnetting.

Sistema DNS

Cada host dentro de una red local está únicamente identificado por una dirección IP. No obstante, resultaría bastante complicado memorizar las direcciones IP de todos los host. Por este motivo, se ha creado un sistema que establece una asociación entre la dirección IP y un nombre simbólico que resulta fácilmente memorizable. Además de esto, el establecimiento de mnemónicos para identificar las máquinas permite modificar fácilmente el direccionamiento de forma transparente al usuario: se modifica la dirección IP pero no se cambia el mnemónico.

Inicialmente, todos los registros de nombres de host se administraban en un repositorio oficial gestionado por el NIC (Network Information Center). Cada sitio Internet tenía que actualizar su copia local de este repositorio cuando cambiaba. La dificultad de mantener este esquema trajo consigo la creación del servicio DNS (Domain Name Service).

El servicio DNS utiliza una BD distribuida de los registros de nombres que se envía automáticamente a los host que lo necesitan. El DNS utiliza una jerarquía de nombres similar a la usada por los nombres de archivos de un sistema UNIX. Hay un dominio raíz en lo más alto de la jerarquía al que sirven un grupo de servidores de dominios denominados servidores raíz. La siguiente figura representa esta situación:

Estructura de domimios

Los servidores de dominio conforman una red jerárquica a nivel mundial:



Justo debajo del dominio raíz se encuentran los dominios de primer nivel, divididos en dos categorías: territoriales y genéricos. Cada país tiene asociado un dominio territorial representado por un código de dos letras (es, para España; uk, para Reino Unido; etc).

Los dominios genéricos son administrados por una empresa llamada **Internic**. Para solicitar un dominio de segundo nivel es necesario solicitarlo a Internic y abonar una cuota anual. Por otra parte, los dominios territoriales son administrados por organismos nacionales.

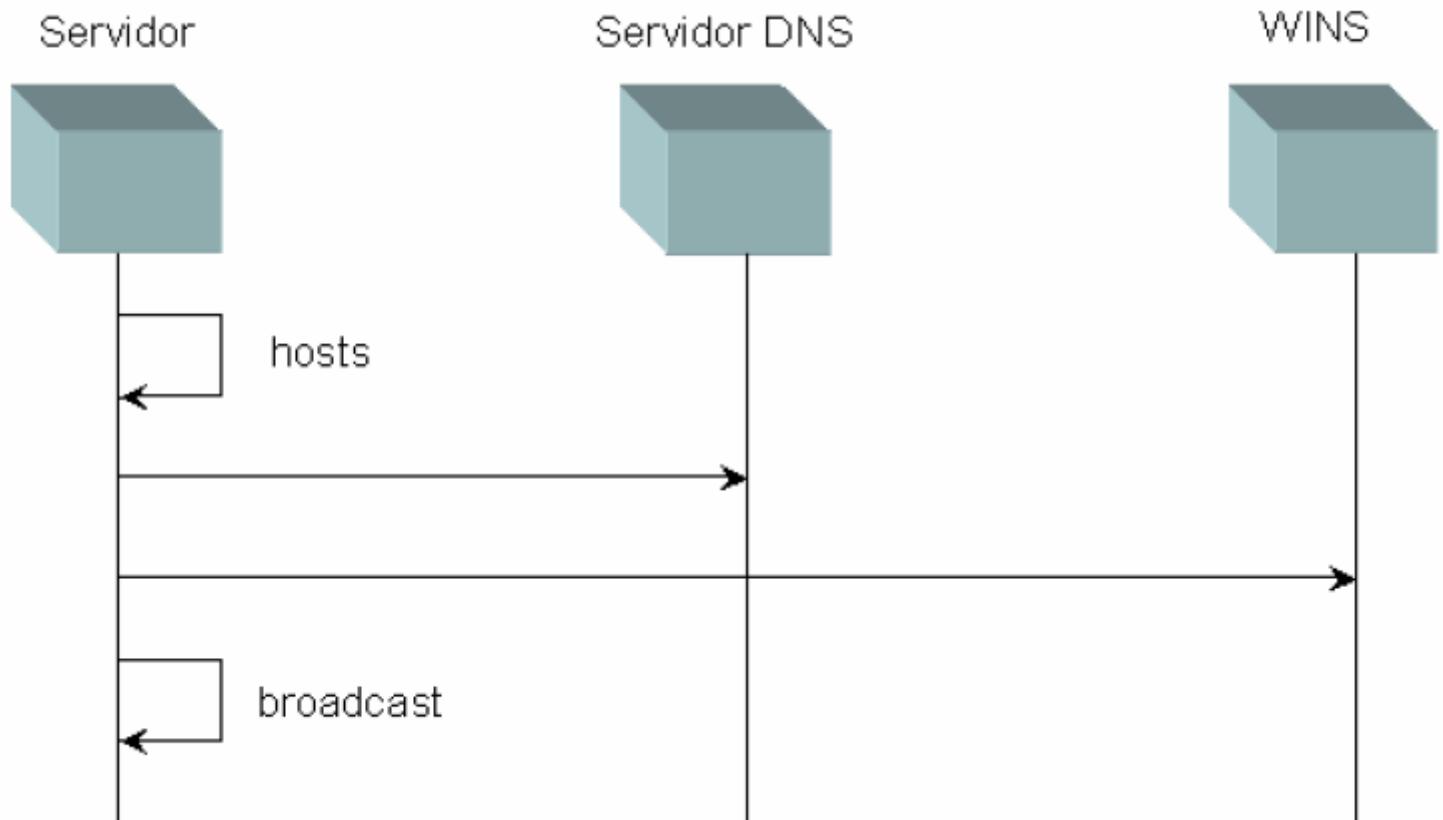
La siguiente tabla muestra una relación de dominios genéricos:

Dominio	Descripción
com	Organizaciones comerciales
edu	Instituciones educativas
gov	Agencias gubernamentales
mil	Organizaciones militares
net	Organizaciones de soporte de red
int	Organizaciones internacionales
org	Otros tipos de organizaciones

Un servidor local de nombres DNS envía una consulta a un servidor raíz de nombres cuando un cliente envía una consulta de búsqueda directa solicitando una dirección IP desde un dominio para el que el servidor local de nombres DNS no tiene autoridad.

Windows

Microsoft Windows utiliza DNS como método principal para la resolución de nombres, pero sigue siendo compatible con el servicio de nombres Internet de Windows (WINS, Windows Internet Naming Services). WINS es el método de resolución de nombres utilizado en Microsoft Windows NT versión 4.0 y anteriores. WINS también suele ser necesario por compatibilidad para resolución de nombres con máquinas Windows 95 y 98.



Es decir, en primer lugar el servidor que necesita obtener la IP busca en el fichero hosts la correspondencia; si no la encuentra busca a través del servicio DNS; como último recurso, busca en el servicio de WINS y si tampoco lo encuentra lanza un broadcast a la red.

El servicio Servidor DNS de Windows proporciona la capacidad de utilizar nombres de dominios completo (FQDN, Fully Qualified Domain Names) jerárquicos en lugar de las convenciones de nomenclatura de NetBIOS que admite WINS. Los clientes utilizan el servicio Servidor DNS para la resolución de nombres y la ubicación de servicios, incluida la ubicación de controladores de dominio que proporcionan autenticación de usuarios.

Red Hat Linux

El software DNS más utilizado para sistemas UNIX y LINUX es BIND (Berkeley Internet Net Daemon).

Los servidores de nombres se pueden dividir en tres categorías dependiendo de cómo estén configurados:

- Maestro: servidor a partir del cual se derivan todos los datos relativos a un dominio. Los datos son suministrados directamente por el administrador.
- Esclavo: al igual que el maestro tiene toda la información sobre un dominio, pero los obtiene del servidor primario.
- De sólo caché: guarda en la caché los resultados de las consultas, siendo éste el único medio que tiene para obtener los datos.

Todos los sistemas UNIX y LINUX tienen una biblioteca denominada *resolver*. Esta biblioteca está asociada a un archivo de configuración */etc/resolv.conf* que permite especificar los servidores DNS a los que se va a consultar.

Para instalar un servidor DNS en un sistema Red Hat se deben instalar los paquetes *bind* y *caching-nameserver*. La configuración de BIND normalmente se almacena en el archivo */etc/named.conf*.

En este archivo se especifican los siguientes parámetros:

- Opciones globales: nombre del directorio de trabajo del servidor; direcciones IP de servidores DNS alternativos a los que se pueden redirigir las peticiones; prevalencia de los servidores alternativos frente al propio host.
- Instrucciones de zona: las zonas BIND son de cuatro clases: maestras (primario), esclavas (secundario), reducidas y de ayuda (caché).

Por otra parte, existen una serie de archivos denominados *archivos de zona* que almacenan la información de las BD de los dominios. Cada archivo de zona tiene una serie de registro de recursos (RR), compuesto por los siguientes campos:

- Nombre del host al que se hace referencia.
- TTL (Time to Live): tiempo, en segundos, que debe almacenarse la información del registro en una caché remota.
- Clase: IN, para Internet.
- Tipo de registro (NS, Name Server; A, address; PTR, pointer; etc).
- Datos relativos al registro.

Gestión de usuarios

La atención a los usuarios y el establecimiento de normas que regulen el uso de la red constituyen una de las razones de ser de los administradores de sistemas. Esta atención va mucho más allá del mero hecho de administrar las cuentas de acceso al sistema, o instalar un software en la estación de trabajo del usuario.

Para que la interacción de los usuarios y los administradores de red sea óptima es necesario un apoyo por parte de la dirección de la compañía u organización en la aplicación de estas normas, de forma que quede claro cómo y cuando deben notificar los usuarios los problemas del sistema al administrador, cómo solicitar que los usuarios lleven a cabo parte del trabajo y cómo hacer un seguimiento de estos problemas y peticiones.

Algunas de las normas a tener en cuenta son las siguientes:

- El usuario no es ignorante: simplemente carece del nivel de conocimientos técnicos sobre la red que tienen los administradores de sistemas.
- El usuario debe recibir el soporte que solicita, siempre que lo haga por las vías establecidas.
- Todos los usuarios deben ser tratados por el mismo patrón, sin favoritismos.
- Las normas deben ser debidamente documentadas y difundidas. Una buena posibilidad es crear documentos en formatos HTML y publicarlos en un servidor web al que tengan acceso todos los usuarios de la red.
- Es necesario revisar periódicamente las normas, para mejorarlas e incorporar nuevos aspectos.
- Las interrupciones de los usuarios deben ser canalizadas convenientemente. Probablemente la mejor opción es una situación de equilibrio entre una atención directa al usuario para incidencias sencillas y una gestión de avisos por correo electrónico para incidencias más complicadas.

- La formación a los usuarios es uno de los elementos clave para reducir las incidencias. Una forma sencilla de ofrecer la formación básica es elaborar una guía de orientación de sistemas que describa donde encontrar determinados elementos en la red, como solicitar nuevo software y cómo utilizar determinadas herramientas.
- Es recomendable la utilización de herramientas online para plantear soluciones y obtener respuestas. Estas herramientas se suelen conocer como *escritorios de ayuda*. Las características de un buen escritorio de ayuda son:
 - Deben ofrecer una interfaz simple a los usuarios.
 - Deben ofrecer un mecanismo que informe a los usuarios sobre el estado de su petición.
 - Deben permitir elaborar automáticamente estadísticas e informes de las incidencias recibidas.
 - Debe tener un sistema para establecer prioridades.

Entre los escritorios de ayuda que podemos encontrar se pueden citar WREQ, herramienta de código abierto utilizada habitualmente en sistemas Linux. En sistemas con Windows pueden utilizarse las *directivas de grupo* ofrecidas por el sistema para reducir los costes de operación de la red. Una directiva de grupo permite establecer plantillas administrativas para configurar las aplicaciones, establecer la apariencia de los escritorios y el comportamiento de los servicios del sistema, manejar opciones para la configuración de los equipos locales y la red, administración central de instalación de software, etc.

Una vez realizada esta introducción, vamos a ver en detalle el procedimiento para gestionar usuarios y sus privilegios de acceso a los recursos del sistema en los SO de red que estamos tratando.

Windows

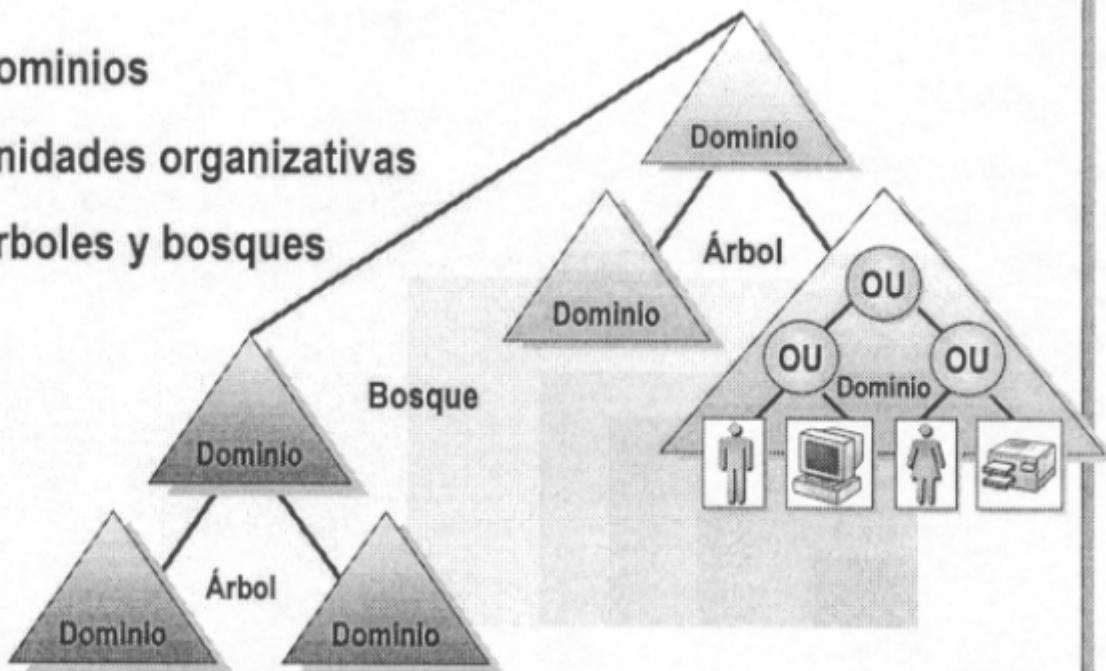
Introducción al Active Directory

La administración de usuarios, grupos y privilegios en Windows se realiza a través del *Active Directory*. Active Directory es el servicio de directorio de una red Windows. Un servicio de directorio es un servicio de red que almacena información sobre los recursos de la red, incluyendo su nombre, descripción, claves de búsqueda, etc.

El Active Directory ofrece un medio para organizar y administrar de forma centralizada el acceso a los recursos de la red, haciendo transparente la topología de la red y sus protocolos. Un usuario puede tener acceso a cualquier recurso sin saber cómo está conectado físicamente.

El Active Directory organiza la información que almacena en secciones que permiten el almacenamiento de una gran cantidad de objetos. Esta estructura proporciona un método para diseñar una jerarquía de directorios comprensible para los usuarios y los administradores. La siguiente figura muestra los componentes lógicos de la estructura del Active Directory:

- Dominios
- Unidades organizativas
- Árboles y bosques



- **Dominios** : conjunto de equipos definidos por el administrador que comparten una BD de directorio común. En una red Windows, el dominio sirve como límite de seguridad, de forma que el administrador de un dominio tienen los permisos necesarios para llevar a cabo la administración únicamente dentro de ese dominio. Una red de múltiples dominios es útil para organizaciones que utilizan un modelo administrativo descentralizado. Cada dominio es gestionado por un servidor denominado *controlador de dominio* . Los controladores de dominio mantienen la información completa del directorio asociado a su dominio.
- **Unidades organizativas (OU)** : objeto contenedor que se utiliza para organizar objetos dentro de un dominio. Una unidad organizativa puede contener objetos como cuentas de usuarios, grupos, equipos, impresora, etc. Las unidades organizativas permiten delegar el control administrativo sobre los objetos que contienen, asignando a usuarios o grupos de usuarios permisos específicos. Las unidades organizativas suelen formar una estructura jerárquica que representa la propia estructura organizativa de la organización o un modelo administrativo de red.
- **Árboles** : disposición jerárquica de dominios que comparten un espacio de nombres contiguo. Cuando se agrega un dominio a un árbol existente, el nombre del dominio secundario (el recién añadido) se combina con el nombre del dominio principal para formar su nombre DNS. Algunas de las razones para tener más de un dominio son:
 - Administración de red descentralizada.
 - Mejor control para la duplicación de los directorios.
 - Necesidad de gestionar una gran cantidad de objetos en la red.
 - Nombres de dominio de Internet diferentes.
 - Requisitos de contraseña diferentes entre organizaciones.
- **Bosques** : grupo de árboles que no comparten un espacio de nombres contiguo. Los árboles de un bosque comparten una configuración, un esquema y un catálogo global comunes.

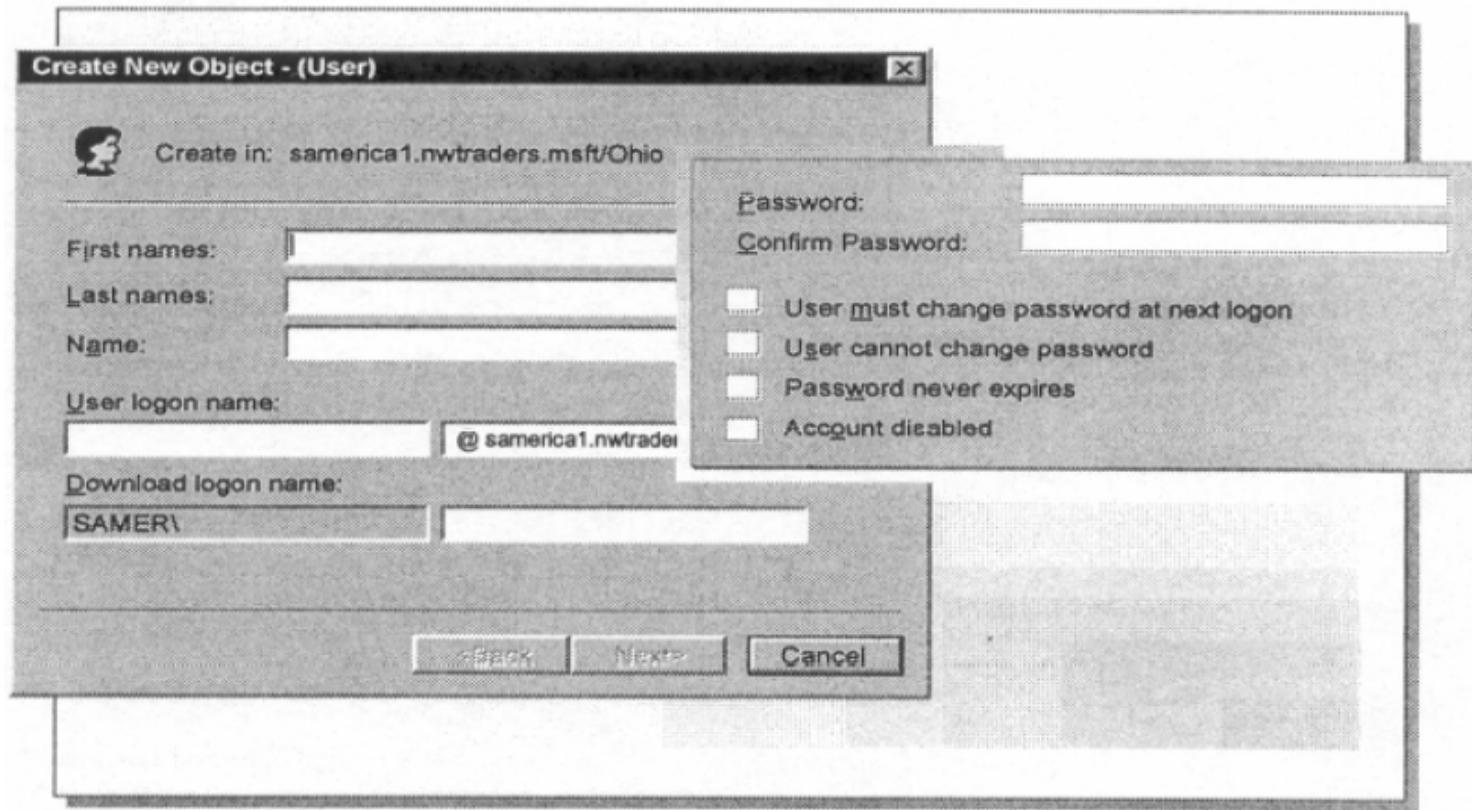
El procedimiento para crear la estructura del Active Directory es crear en primer lugar la estructura de dominios, para completar luego la estructura lógica de la red con una jerarquía de unidades organizativas que contengan usuarios, equipos y otros recursos. Por último, se controlará el acceso a los objetos del Active Directory mediante el establecimiento de permisos.

Cuentas de Usuario

Las cuentas de usuario se emplean para autenticar al usuario y para conceder permisos de acceso a los recursos de red. Una cuenta de usuario se crea dentro de una unidad organizativa, realizando las siguientes operaciones:

1. En el menú *Herramientas Administrativas*, abrir *Usuarios y equipos de Active Directory*.
2. Pulsar el botón derecho del ratón en la unidad organizativa deseada, seleccionar *Nuevo* y pulsar *Usuario*.

Al hacerlo aparecerá el cuadro de diálogo de la siguiente figura (todo depende del sistema Windows utilizado):



Cuentas de Equipo

Las cuentas de equipo son similares a las cuentas de usuario en el sentido de que pueden utilizarse para autenticar y auditar el equipo, y para conceder permisos a los recursos de red. Los equipos cliente deben tener una cuenta de equipo válida para poder unirse al dominio.

Para crear una cuenta de equipo se realizan las siguientes operaciones:

1. En el menú *Herramientas Administrativas*, abrir *Usuarios y equipos de Active Directory*.
2. Pulsar el botón derecho del ratón en la unidad organizativa deseada, seleccionar *Nuevo* y pulsar *Equipo*.
3. Escribir un nombre de equipo que sea único dentro del bosque.
4. Designar al usuario o grupo de usuarios con permisos para hacer que el equipo se una al dominio. Esto hará que durante el proceso de unión del equipo al dominio el sistema muestra un diálogo solicitando una cuenta válida para realizar la operación.

Grupos de Usuarios

Los grupos de usuarios son un mecanismo para simplificar la administración de permisos. Asignar una sola vez un permiso a un grupo de usuarios es más sencillo que hacerlo tantas veces como usuarios tenga el grupo de forma aislada. Existen dos tipos de grupos de Active Directory:

- Grupos de seguridad: se utilizan para conceder o denegar permisos.
- Grupos de distribución: se utilizan para enviar mensajes de correo electrónico. No pueden utilizarse con fines de seguridad.

Los grupos de distribución y seguridad tienen además un atributo de ámbito, que determina quién puede ser miembro del grupo y dónde puede usar ese grupo en la red:

- Grupos locales de dominio: pueden contener usuarios, grupos globales y grupos universales pertenecientes a cualquier dominio del bosque, así como grupos locales del mismo dominio.
- Grupos globales: pueden contener cuentas de usuario y grupos globales del dominio en el que existe el grupo.
- Grupos universales: pueden contener cuentas de usuario, grupos globales y otros grupos universales de cualquier dominio de Windows en el bosque.

Para crear un grupo se realizan las siguientes operaciones:

1. En el menú *Herramientas Administrativas*, abrir *Usuarios y equipos de Active Directory*.
2. Pulsar el botón derecho del ratón en la unidad organizativa deseada, seleccionar *Nuevo* y pulsar *Grupo*.

Al hacerlo se obtiene un cuadro de diálogo desde el que se pueden establecer las opciones para el grupo creado y agregar miembros (otros grupos o cuentas de usuario).

Red Hat Linux

Consideraciones generales

En Linux los usuarios y grupos se utilizan para determinar el propietario y permisos de los archivos y dispositivos, y para establecer permisos de acceso a casi todas las funciones del sistema. El programa *Linuxconf* permite acceder a la información de usuarios y grupos.

Para administrar convenientemente los permisos en Linux hay que evitar en la medida de lo posible el uso de cuentas compartidas, por varios motivos:

- Es importante controlar las acciones de los usuarios para saber quiénes actúan correctamente y quiénes no lo hacen. El uso del mismo identificador de usuario por varias personas hace que esta tarea sea inviable. Asimismo, por motivos de seguridad resulta completamente desaconsejable que un grupo de usuarios comparta la misma contraseña para acceder al sistema.
- Si se utiliza un inicio de sesión compartido los archivos de configuración no pueden ser específicos para cada usuario. Por lo tanto, todos deberán utilizar la misma estructura de escritorio y los mismos alias de la *shell*. Esto significa que todos los cambios realizados pueden afectar al entorno de trabajo de los usuarios que comparten la cuenta, y que cualquiera de ellos puede leer un e-mail enviado a la cuenta compartida.
- Por último, como varias personas pueden acceder a un archivo al mismo tiempo, la última modificación sobrescribirá los cambios que puedan haber realizado otros usuarios.

En vez de utilizar cuentas compartidas, resulta más recomendable utilizar grupos para proporcionar a los usuarios el acceso a las mismas partes del sistema de archivos

manteniendo distintos identificadores de inicio de sesión, archivos de configuración y directorios principales.

Archivos del Sistema

Existen una serie de archivos importantes para configurar y modificar los usuarios y grupos del sistema, que se describen a continuación:

- */etc/passwd* : Contiene información sobre los identificadores de los usuarios, los directorios principales y los comandos que se ejecutan cuando se inicia una sesión.
- */etc/group* : Contiene la información sobre los grupos del sistema.
- */etc/shadow* : Archivo utilizado para ocultar las contraseñas del archivo */etc/passwd*, cuya lectura es universal. Contiene la cadena de la contraseña codificada e información sobre el vencimiento de las cuentas.
- */etc/gshadow* : Realiza para los grupos la misma función que hace */etc/shadow* para los usuarios. Contiene las contraseñas codificadas para los grupos e información de las cuentas.
- */etc/login.defs* : Contiene información predeterminada sobre la creación y mantenimiento de las cuentas.
- */etc/skel* : Contiene archivos de configuración predeterminados que se utilizan al crear cuentas nuevas. En este directorio se guardan los archivos de configuración de la shell, los archivos de configuración del administrador de ventanas o los archivos y directorio que un usuario deba tener en su directorio principal.

Configuración de Cuentas de Usuario

Para crear nuevos usuarios en Linux es posible seguir varios procedimientos. Se pueden modificar directamente los archivos de sistema expuestos en el punto anterior, pero esto supone una labor bastante tediosa. El comando *useradd* simplifica estas operaciones, permitiéndonos crear nuevos usuarios en una sola línea.

Sin embargo, la forma más cómoda de hacerlo es utilizar una herramienta GUI como Linuxconf. Este programa puede ejecutarse en X-Windows o en línea de comandos.

Configuración de Grupos de Usuarios

Al igual que ocurre con las cuentas de usuario, en Linux existen diversas formas de configurar los grupos de usuarios. Esta tarea puede realizarse mediante comandos del sistema o utilizando una herramienta GUI como Linuxconf.

A continuación se presentan algunos de los comandos y scripts de shell para gestión de grupos:

- *groupadd* : añade un grupo al archivo */etc/group*
- *groupdel* : elimina un grupo del archivo */etc/group*
- *groupmod* : modifica las entradas en */etc/group*
- *grpck* : realiza una comprobación de integridad del archivo */etc/group*

Gestión de privilegios

Existen varias alternativas para establecer los permisos de acceso de los usuarios a archivos y directorios. El comando *chown* permite otorgar a un usuario la propiedad de un archivo o directorio. No obstante, éste no es el mejor método si necesitamos que varios usuarios accedan a un archivo, ya que sólo puede existir un propietario en un momento dado.

Otra opción mejor es utilizar el comando *chmod* para conceder al usuario los permisos necesarios. Si el usuario no es propietario del archivo, será necesario cambiar la

configuración de permisos para ‘otros’, de forma que los cambios se apliquen a todos los usuarios que no sean propietarios del archivo.

Por último, se puede conceder permisos a los usuarios utilizando la configuración del grupo propietario de un archivo y estableciendo los permisos para grupos. Manipulando los permisos de los grupos en un archivo o dispositivo es posible proporcionar o anular el acceso a ese archivo a nivel de grupo.

Otros comandos para configuración de Usuarios y Grupos

En este apartado se detallan otros comandos de *shell* útiles para la configuración de usuarios y grupos de usuarios:

- *newusers* : añade usuarios nuevos al sistema en modo de procesamiento por lotes.
- *chpasswd* : actualiza el archivo de contraseñas en modo *batch* . Lee parejas nombre_usuario-contraseña de la entrada estándar y actualiza la contraseña para el usuario dado.
- *usermod* : permite modificar o desactivar cuentas de usuario. Afecta a los archivos / *etc/passwd* y /*etc/shadow* .
- *userdel* : elimina una cuenta de usuario.
- *newgrp* : permite cambiar temporalmente el grupo predeterminado a otro grupo del que sea miembro o del que conozca la contraseña.
- *chage* : permite configurar el contenido de /*etc/shadow* .
- *chgrp* : permite cambiar el grupo propietario de un archivo.
- *chown* : permite cambiar el usuario propietario de un archivo.
- *chsh* : cambia la *shell* de inicio de sesión.
- *gpasswd* : permite a los administradores añadir y eliminar usuarios de los grupos y cambiar las contraseñas de los mismos.
- *groups* : imprime los grupos de los cuales es miembro un usuario.
- *passwd* : cambia las contraseñas.
- *su* : permite a un usuario hacerse pasar temporalmente por otro.

Administración de Recursos de Impresión y Archivos

Windows

Compartición de impresoras

Compartir una impresora de red en el entorno Windows es bastante sencillo. Desde el panel de control del equipo donde está instalada localmente la impresora, seleccionaremos *Impresoras y Faxes* . Esto nos mostrará una vista con todas las impresoras instaladas. Seleccionamos una de ellas y pulsamos con el botón derecho del ratón, seleccionando a continuación la opción *Compartir* . Después únicamente queda llenar las opciones solicitadas por el cuadro de diálogo.

Compartición de Archivos

Administración de los recursos de archivos:

- Publicación de recursos de archivo en el Active Directory: permite a los usuarios encontrar los recursos fácilmente en la red.
- Sistema de archivos *Dfs (Distributed file system)* : permite la transparencia de las estructuras de red y el sistema de archivos de cara a los usuarios, posibilitando el acceso a los recursos de archivo de forma centralizada. Proporciona un árbol lógico único para los recursos del sistema de archivo que pueden estar en cualquier lugar de la red.
- Permisos especiales del sistema de archivos NTFS: Proporcionan un mayor grado de control a la hora de conceder acceso a los recursos. Por ejemplo, permite asignar a

un usuario permiso para cambiar permisos de archivos o carpetas, o la posibilidad de tomar posesión de archivos o carpetas.

- Sistema de archivos de cifrado (*EFS, Encrypting File System*) : Proporciona la tecnología de cifrado de archivos para el almacenamiento de archivos NTFS en disco.
- Asignación de cuotas de disco en volúmenes NTFS: Permiten asignar espacio de disco a los usuarios en función de los archivos y carpetas que poseen. También se puede supervisar la cantidad de espacio de disco duro que los usuarios han utilizado.

Para compartir recursos en una red es necesario compartir carpetas. Esta tarea se lleva a cabo de la siguiente forma:

1. Abrir *Administración de equipos* desde el menú *Herramientas Administrativas*.
2. En el árbol de la consola de *Administración de equipos*, seleccionar *Carpetas compartidas* bajo Herramientas del Sistema.
3. Bajo *Carpetas compartidas*, seleccionar *Recursos compartidos*. Aparecerán en el panel de detalles todas las carpetas compartidas desde el equipo local.
4. Pulsar el botón derecho del ratón en el panel de detalles y hacer clic en *Nuevo recurso compartido de archivo*.
5. Seguir las instrucciones del *Asistente* para la creación de carpetas compartidas.

Para publicar una carpeta en el Active Directory se procederá de la siguiente manera:

1. Abrir *Usuarios y equipos de Active Directory* en el menú *Herramientas Administrativas*.
2. En el árbol de la consola de *Usuarios y equipos de Active Directory* pulsar con el botón derecho del ratón en el dominio en el que se desea publicar la carpeta compartida y seleccionar *Nuevo / Carpeta compartida*.
3. Escribir el nombre con el que se desea que aparezca la carpeta compartida en el Active Directory.
4. Especificar la ruta de la carpeta compartida.

Red Hat Linux

Compartición de Impresoras

Red Hat Linux da soporte a las impresoras de red mediante el demonio *lpd*, o con el paquete LPRng. Para configurar una impresora en red se realiza la instalación de una impresora local y posteriormente se permite el acceso desde otras máquinas de la red.

Para poder instalar una impresora local es necesario que ésta tenga un controlador para *GhostScript*. *GhostScript* es un paquete de software que implementa el lenguaje *PostScript*, y que incluye controladores para muchas impresoras conocidas.

Linux almacena la configuración de las impresoras en el archivo */etc/printcap*. Puede utilizarse la utilidad *PrintTool* como herramienta GUI para crear las entradas de este archivo.

El acceso a otras máquinas a una impresora local a través de la red se realiza configurando entradas en el archivo */etc/hosts.lpd*, en el que se especifican los nombres de las máquinas que pueden acceder al demonio de la impresora en línea en el *host* local.

Por último, es necesario definir la impresora remota en las máquinas a las que se les ha dado acceso. Esto puede hacerse también mediante PrintTool, estableciendo de forma gráfica la entrada correspondiente para la impresora remota en el archivo *printcap*.

Compartición de Archivos

El método estándar para compartir archivos en Linus es *NFS (Network File Services)*. La exportación de sistemas de archivos se lleva a cabo mediante la configuración en el

servidor NFS del fichero `/etc(exports`. Las entradas de este fichero definen qué clientes pueden acceder a los sistemas de archivo compartidos y con qué permisos.

Existen numerosas opciones para controlar los permisos de acceso al sistema de archivos, que se muestran en la tabla siguiente:

Opción	Descripción
ro	Acceso de sólo lectura
rw	Acceso lectura-escritura
noaccess	Deniega el acceso a una rama particular de un árbol exportado
secure	Requiere que las peticiones se hagan desde un rango de puertos seguro
map_daemon	Mapea los identificadores que difieren entre los sistemas clientes y el servidor mediante el demonio <i>ugidd</i>
map_static=file	Mapea los identificadores que difieren entre los sistemas clientes y el servidor mediante un archivo de texto
map_nis=nisdomain	Mapea los identificadores que difieren entre los sistemas clientes y servidor mediante consultas de NIS
no_root_squash	Permite al usuario <i>root</i> en el sistema cliente acceder al sistema de archivos con privilegios de <i>root</i> .
all_squash	Mapea todos los identificadores al usuario anónimo.
anonuid=uid	El identificador al que son mapeadas las cuentas de usuario si se especifican las opciones <i>root_squash</i> o <i>all_squash</i> .
anonguid=guid	El identificador de grupo al que son mapeadas las cuentas de usuario si se activan las opciones <i>root_squash</i> , o <i>all_squash</i> .

En el cliente se ejecutan una serie de operaciones para acceder a un sistema de archivos compartidos mediante NFS:

- Asegurarse de que se tiene acceso al recurso. Para ello se ejecutará en la máquina cliente el comando *showmount*, pasándole como parámetro el nombre del servidor NFS.
- Modificar el archivo */etc/fstab* para incluir información sobre el sistema de archivos que se quiere montar.
- Crear los puntos de montaje adecuados, y montar el sistema de archivos. Esto se lleva a cabo utilizando el comando *mount*, que da instrucciones para montar todos los sistemas de archivos descritos en el archivo */etc/fstab*.

Existen otras alternativas además de NFS para compartir sistemas de archivos en Linux. Entre ellas se encuentran AFS, un sistema de archivos comercial, y Coda, un sistema de archivos gratuito con bastantes similitudes con AFS.

Administración de Dispositivos de Almacenamiento

Configuraciones RAID

La idea básica del RAID (Redundant Array of Independent Disk) es combinar varios discos relativamente poco costosos en un array de discos para obtener eficiencia, capacidad y fiabilidad que supere a la ofrecida por un disco de alta capacidad. El array de discos aparece en la máquina como un disco lógico simple.

Existen dos conceptos básicos necesarios para comprender las distintas configuraciones RAID:

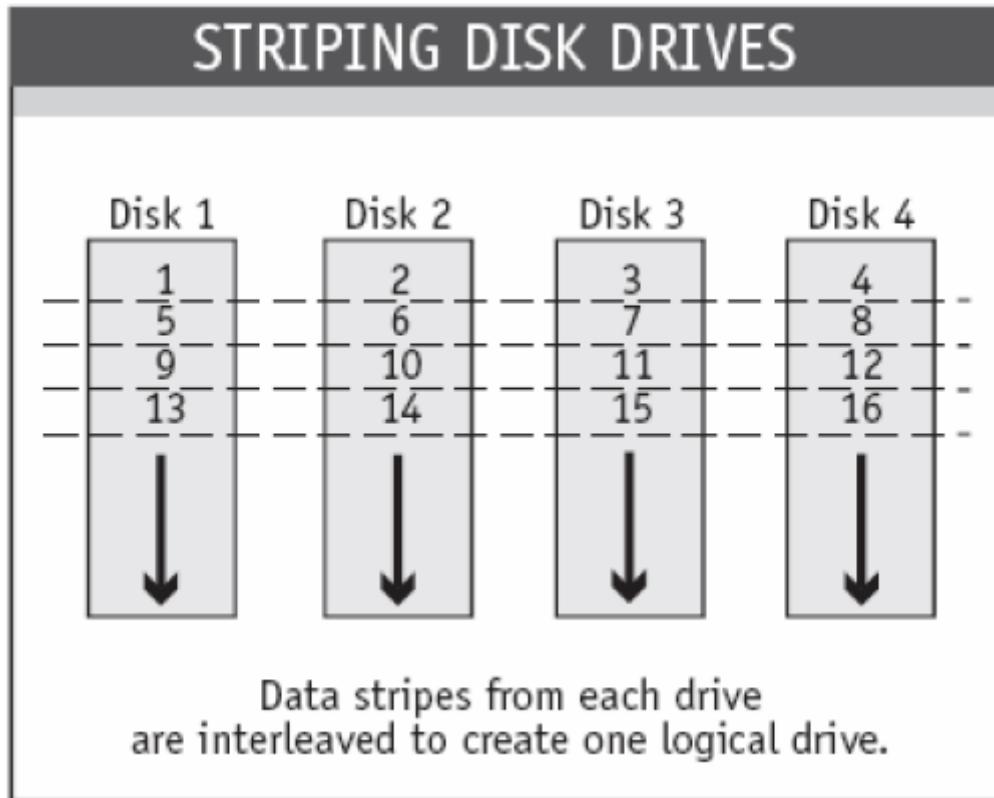
- Striping
- Mirroring (Espejo)

Striping

Constituye la tecnología fundamental del RAID. Es el método para combinar múltiples discos en una unidad de almacenamiento lógica simple. Esta técnica particiona el espacio

de almacenamiento de cada disco en franjas o *stripes*, que pueden ser tan pequeñas como un sector (normalmente 512 bytes), o del orden de megabytes. Estas franjas se intercalan en una secuencia rotativa, de forma que el espacio combinado se compone de franjas alternativas de cada uno de los discos.

La mayoría de los SO soporta actualmente operaciones concurrentes de E/S sobre múltiples discos. Sin embargo, para optimizar el rendimiento, la carga de E/S se debe balancear entre todos los discos de forma que cada disco se mantenga ocupado el máximo tiempo posible. En un sistema de discos múltiples sin striping, la carga de E/S nunca está perfectamente balanceada. Algunos discos contendrán ficheros de datos frecuentemente accedidos, y otros discos serán raramente accedidos.



Creando los discos del array con stripes lo suficientemente grandes para que un registro entre enteramente en un solo stripe, la mayoría de los registros se pueden distribuir entre todos los disco. Esto hace que los discos del array se mantengan ocupados en situaciones de fuerte carga. Con ello se consigue que los discos trabajen de forma concurrente en operaciones de E/S, maximizando así el número de operaciones simultáneas de E/S que se pueden llevar a cabo en el array.

Mirroring (Espejo)

La técnica de mirroring se basa en replicar en dos o más discos los datos para garantizar su disponibilidad. Todas las escrituras se deben ejecutar en todos los discos del array en espejo, de forma que mantengan idéntica información. No obstante, con ello también se multiplica la eficiencia en lectura mientras que la eficiencia en escritura permanece invariable respecto a un solo disco ya que se ejecuta en paralelo.

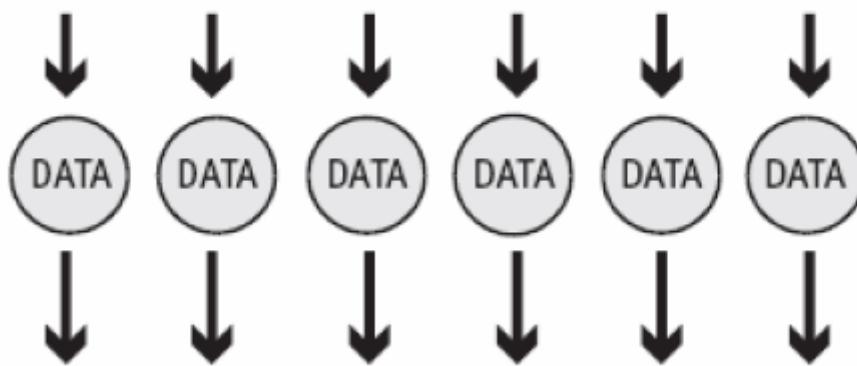
RAID 0

Está constituido por un grupo de discos a los que se le aplica la técnica de striping, sin paridad o información de redundancia. Los arrays RAID 0 ofrecen la mejor eficiencia en cuanto a espacio de almacenamiento y eficiencia de acceso. La desventaja que tienen es que si uno de los discos del array falla, falla el array completo.

RAID 0

Non-Redundant Striped Array

Writes can occur simultaneously on every drive.



Reads can occur simultaneously on every drive.

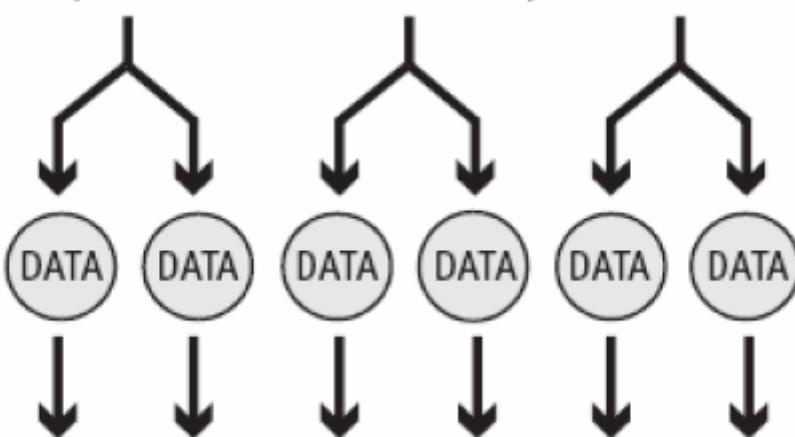
RAID 1

También se le conoce como configuración en espejo. Consiste en una serie de discos (normalmente dos) que almacenan información duplicada, pero que se muestra a la máquina como un disco simple. Aunque la técnica de striping no se usa en un par de discos en espejo, es posible utilizar esta técnica en arrays RAID 1 múltiples, creando un array simple y grande de parejas de discos en espejo.

RAID 1

Mirrored Arrays

Duplicate data is written to pairs of drives.

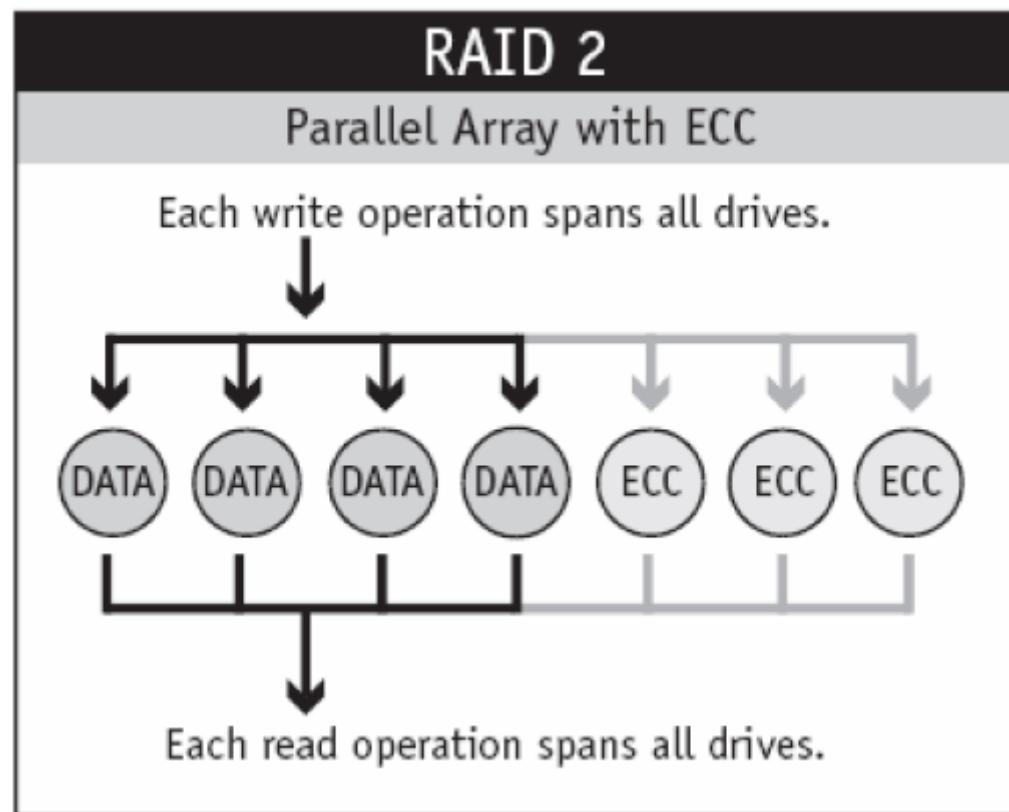


Reads can occur simultaneously on every drive.

RAID 2

Estos arrays utilizan la técnica de striping en sectores de datos a lo largo de grupos de discos, dedicando algunos discos a almacenar información de redundancia. Puesto que

todos los discos contienen actualmente información de redundancia en cada sector, RAID 2 no ofrece ventajas significativas respecto a otras configuraciones RAID.



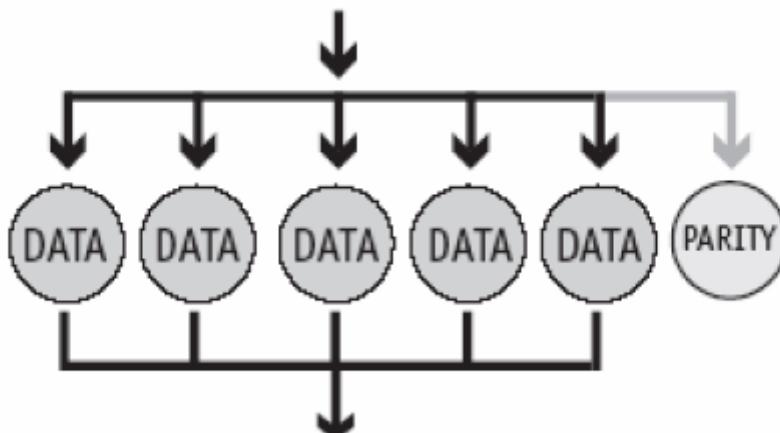
RAID 3

Como en RAID 2, se utiliza la técnica de striping con los sectores de datos en grupos de discos, pero un disco del grupo se dedica a almacenar información de paridad. RAID 3 se apoya en la información de redundancia que tiene cada sector para la detección de errores. Cuando se produce el fallo de un disco, la recuperación de datos conlleva el cálculo XOR de la información almacenada en los discos restantes. Los registros de datos habitualmente se reparten entre todos los discos, lo que optimiza la velocidad de transferencia. Debido a que cada petición de E/S accede a cada disco en el array, RAID 3 sólo puede ejecutar una petición en cada momento. Por ello proporciona la mejor eficiencia para aplicaciones monousuario o monotarea con registros grandes. La configuración RAID 3 requiere discos y controladoras especiales y costosos para mantener los discos sincronizados de forma que se evite la degradación de eficiencia en el acceso a registros cortos.

RAID 3

Parallel Array with Parity

Read and write operations span all drives.



Parallel access decreases data transfer time
for long sequential records.

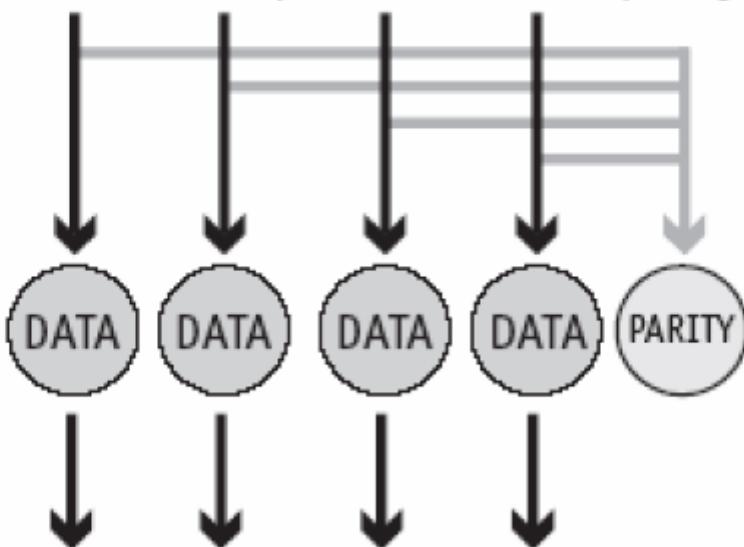
RAID 4

Esta configuración es idéntica a RAID 3 con la excepción de que se utilizan stripes de tamaño largo, de forma que los registros se pueden leer desde cualquier disco individual del array. Esto permite que las operaciones de lectura se puedan solapar. Sin embargo, puesto que todas las operaciones de escritura deben actualizar el disco de paridad, no se pueden solapar.

RAID 4

Parallel Array with Parity

Every write must update dedicated parity drive.



Reads can occur simultaneously
on every data drive.

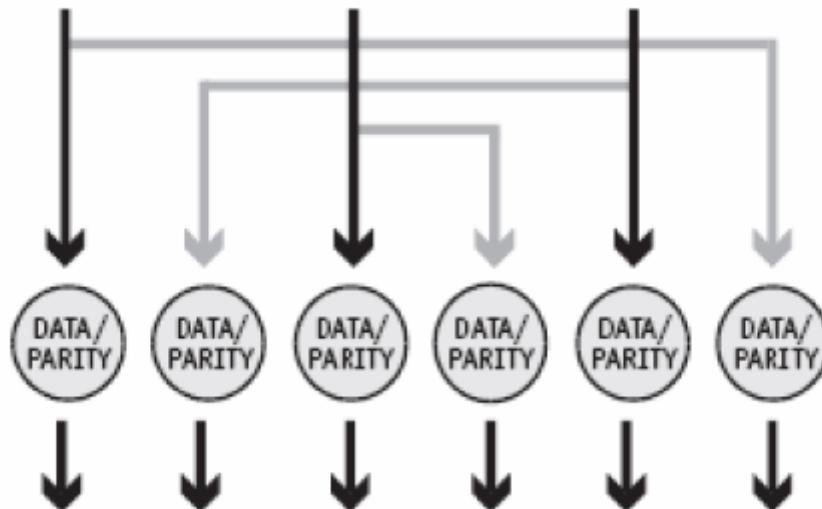
RAID 5

Esta configuración evita los cuellos de botella en escritura que provoca la utilización de un único disco para la paridad en RAID 4. En RAID 5 la información de paridad se distribuye entre todos los discos. Al no haber un disco de paridad dedicado, todos los discos contienen datos y las operaciones de lectura se pueden solapar sobre cualquier disco del array. Las operaciones de escritura accederán habitualmente a un disco de datos y a un disco de paridad. No obstante, ya que los diferentes registros almacenan su paridad en diferentes discos, las operaciones de escritura se pueden solapar con frecuencia.

RAID 5

Striped Array with Rotating Parity

Writes require parity to be updated.



Reads can occur simultaneously on every drive.

Arrays de Doble Nivel

Además de los niveles RAID estándar, se pueden combinar varios arrays RAID en un grupo de arrays simple. Los arrays de doble nivel aportan un equilibrio entre la alta disponibilidad de datos de RAID 1 y RAID 5 y la mayor eficiencia de lectura de RAID 0. Estos arrays se conocen como RAID 0+1 o RAID 10, y RAID 0+5 o RAID 50.

Windows

Windows ofrece dos tipos de almacenamiento en disco: almacenamiento básico y almacenamiento dinámico:

- Almacenamiento básico: Contiene hasta cuatro particiones primarias o tres particiones primarias y una extendida. Este es el sistema de particiones ya utilizado en MS-DOS.
- Almacenamiento dinámico: Ofrece los siguientes tipos de volúmenes:
 - Volumen simple: contiene el espacio de un único disco.
 - Volumen distribuido: contiene el espacio de dos o más discos (hasta 32). Es equivalente a una configuración RAID 0.
 - Volúmenes con espejo: Son dos copias idénticas de un volumen simple, cada una de ellas en un disco duro independiente. Es equivalente a la configuración RAID 1.

- Volumen seccionado: combina en un único volumen áreas de espacio libre de 2 a 32 discos duros.
- Volúmenes RAID 5.

Windows permite realizar la administración de discos a través del *MMC (Microsoft Management Console)* . Con esta herramienta se puede administrar el espacio de almacenamiento, ver las propiedades del disco, ver las propiedades de particiones y volúmenes y actualizar la información de administración de discos.

Red Hat Linux

El soporte para RAID en Red Hat Linux se suministra con el *Kernel* a través del controlador *md* . Este controlador se implementa en dos modos:

- Modo lineal: permite la concatenación de múltiples dispositivos físicos en una única unidad lógica. Sin embargo, no ofrece mejoras en el rendimiento ya que no se accede a los discos en paralelo. La única función del modo lineal es ligar múltiples dispositivos físicos en un único dispositivo lógico, permitiendo sistemas de archivos más grandes.
- Soporte RAID: en este modo se soportan los niveles RAID 0, 1, 4 y 5. La versión 6.0 de Red Hat incluye la posibilidad de reconstrucción de los arrays dañados, la detección automática de RAID en tiempo de arranque y la posibilidad de añadir y extraer dispositivos físicos de un array en ejecución.

Existen una serie de comandos para controlar las configuraciones RAID, que se muestran a continuación:

<i>Comando</i>	<i>Función</i>
<i>mkraid</i>	Inicializa los arrays de dispositivos RAID
<i>raidstart</i>	Configura los dispositivos RAID en el <i>kernel</i> y los activa
<i>raidhotadd</i>	Añade dispositivos de recambio a un array RAID en ejecución
<i>raidhotremove</i>	Extrae dispositivos de un array RAID en ejecución
<i>raidstop</i>	Elimina la configuración de un array de dispositivos RAID

La información sobre configuraciones RAID se almacena en el archivo */etc/raidtab* . Cada entrada de este archivo contiene el nivel RAID, el número y nombres de las particiones que constituyen el array e información complementaria.

Monitorización y Control de Tráfico

La supervisión de la red es una tarea esencial para los administradores de sistemas. Mediante esta actividad, se puede ver la actividad de la red, de manera que los problemas se pueden diagnosticar con rapidez y exactitud. También resulta muy útil para resolver huecos de seguridad y cuellos de botella en el tráfico de red.

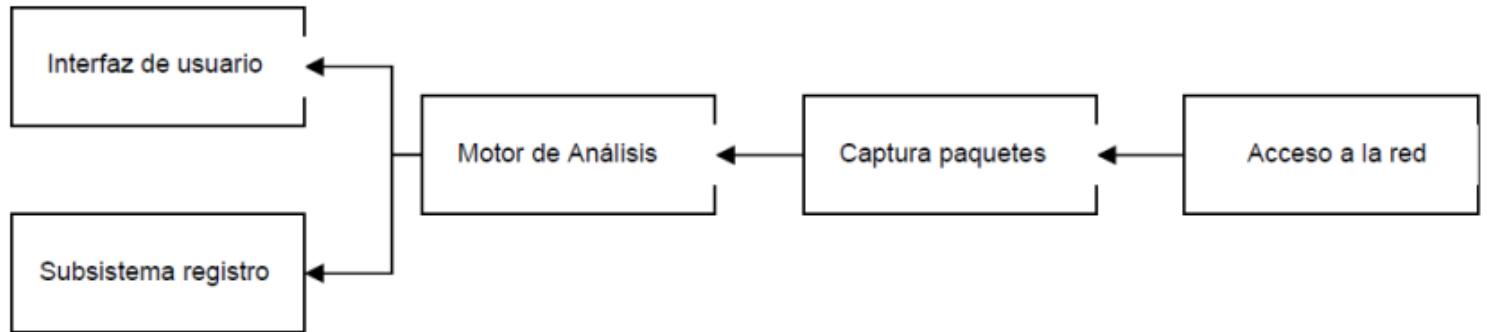
Normalmente la supervisión de red se basa en la captura y examen de paquetes de datos desde una tarjeta de red. Es decir, los paquetes de datos se inspeccionan antes de que se realice ningún tratamiento de los mismos en la máquina. Esto permite realizar numerosas actividades, como la detección de paquetes corruptos o el diagnóstico de problemas de mala configuración.

Por otra parte, con la supervisión de red es posible examinar el tráfico entre los diferentes host de la red. Los cuellos de botella y los conflictos de direcciones IP son algunos de los problemas que se pueden resolver de esta forma.

En cuanto a temas de seguridad, la supervisión de red permite la detección de contraseñas fáciles de adivinar, el uso ilegal de ancho de banda por parte de determinados usuarios o ataques de denegación de servicio.

Esquema de un sistema de supervisión

La siguiente figura muestra el esquema de componentes de un sistema de supervisión de red:



- Punto de acceso a la red: proporciona la conexión física entre el host supervisor y la red supervisada. La ubicación del punto de acceso de red depende mucho de la naturaleza conmutada o no conmutada de la red. En las redes no conmutadas los host comparten un único canal de comunicaciones, por lo que el ancho de banda es compartido entre todos ellos. En una red conmutada cada host tiene su propio canal de comunicación, disponiendo de todo el ancho de banda. También influye en la colocación del punto de acceso la forma en la que esté segmentada la red, siendo una tarea más sencilla en las redes de un solo segmento.
- Sistema de captura de paquetes: captura el tráfico en bruto procedente del punto de acceso a la red. La mayoría de los sistemas de supervisión de red utilizan un API para realizar el acceso a los paquetes capturados.
- Motor de análisis: efectúa la decodificación del protocolo de la red y, si es posible, la reconstrucción de la sesión de red. Además agrupa las estadísticas y examina las tendencias de tráfico. Suelen utilizar plantillas de protocolo para separar el encabezamiento de los paquetes de la carga útil de datos.
- Subsistema de registro: módulo utilizado por el motor de análisis para almacenar los datos conseguidos, incluyendo estadísticas, tendencias y paquetes capturados. El soporte de almacenamiento puede ser variado, desde ficheros de texto a BD.
- Interfaz de usuario: conjunto de vistas a través de las cuales el usuario puede visualizar de forma coherente la información capturada y analizada. El interfaz de usuario mostrará estadísticas de utilización, listados de host activos, estadísticas de protocolos, etc.

Las características de un buen sistema de supervisión de red son las siguientes:

- Portabilidad del hardware del sistema de supervisión.
- Compatibilidad con una o más topologías de red.
- Capaz de supervisar a velocidades cercanas a las del cable.
- Capaz de almacenar grandes cantidades de datos y estadísticas.
- Decodificación del tráfico de red en múltiples niveles de la pila de protocolos.

Existen básicamente dos tipos de sistema de supervisión de red que se estudiarán en los siguientes apartados: los rastreadores y los analizadores de tráfico.

Rastreadores

Un rastreador o *sniffer* captura el contenido real de las sesiones de red y de la transmisión de datagramas. Permiten almacenar información de los protocolos y las actividades del

nivel de aplicación. Esta faceta los hace enormemente potentes porque permite detectar problemas derivados de malas configuraciones de aplicaciones o de red, ataques de denegación de servicio, etc.

El rastreador funciona capturando los datos de los paquetes de acuerdo a criterios de filtrado específico y extrayendo el contenido de carga útil de los paquetes recibidos. Los rastreadores avanzados dan soporte a una técnica denominada *reconstrucción de sesión de red*, que permite que las transmisiones de datos que emplean múltiples paquetes simultáneos sean reconstruidas como un único flujo de comunicaciones coherente.

Dada su capacidad de analizar el tráfico a nivel de aplicación, los rastreadores son extremadamente útiles para proporcionar una visión profunda de la red, analizando el contenido de las sesiones de red en lugar de sólo la información de cabecera del protocolo.

Algunos de los ataques de denegación de servicio más comunes que puede detectar un rastreador son:

- Avalanchas SYN: consiste en inundar un host de destino con peticiones de conexión TCP no válidas, de forma que la cola de conexiones del servidor atacado se colapse.
- Ping de la muerte: consiste en enviar un paquete fragmentado mayor de lo normal para provocar el desbordamiento del identificador de longitud de paquetes en su reconstrucción, de forma que acabe tomando un valor negativo. En las máquinas más vulnerables, esto suele provocar el desbordamiento de la pila.
- Smurf: consiste en el envío de una petición de ping ICMP a la dirección de difusión de una red. Cada host que reciba la petición generará una respuesta, pudiendo provocar una situación de sobrecarga.

Los rastreadores también son útiles para detectar problemas habituales de la red, como los siguientes:

- Direcciones IP duplicadas: dos o más host comparten la misma dirección IP en la red.
- Errores ARP.
- Problemas de encaminamiento.

Analizadores de Tráfico

Los analizadores de tráfico son herramientas para el diagnóstico y solución de problemas, y para la mejora del rendimiento. Muestran exactamente lo que está sucediendo en una red de computadoras.

Funciona capturando paquetes de datos que viajan por la red, decodificando el protocolo y elaborando estadísticas. Algunas de las utilidades que los analizadores de tráfico tienen son:

- Detección de cuellos de botella: un cuello de botella es un punto de la ruta entre dos host en el que el ancho de banda disponible es muy limitado. Esto puede repercutir en una degradación grave del rendimiento de la red. Los cuellos de botella pueden estar producidos por múltiples causas: errores de configuración, problemas hardware, servidores saturados...
- Análisis del tipo de tráfico: permiten catalogar el tráfico que circula por la red mediante el establecimiento de criterios de filtrado.

Vulnerabilidades, Riesgo y Protección

Los Centros de tratamiento de la información, más comúnmente conocidos como Centros de Proceso de Datos (CPD), son los recintos que albergan los equipamientos informáticos principales que dan soporte al conjunto de información de la organización. A grandes rasgos se puede considerar que los Centros de Proceso de Datos (CPD) son los depositarios, los “guardianes”, de la información utilizada por todas las áreas de la organización.

Esta responsabilidad implica que la información confiada al CPD está convenientemente salvaguardada (funcionalidad y disponibilidad), que se procesa de acuerdo con las instrucciones (fiabilidad e integridad) y que se devuelve intacta a quien la solicite y esté autorizado a obtenerla (accesibilidad y confidencialidad).

El objetivo de la seguridad es garantizar la continuidad de la explotación o, lo que es lo mismo, evitar los riesgos potenciales de ataque, robo o daño a los Sistemas de Información de la empresa, tanto accidentales como intencionados, que puedan ocasionar la interrupción total o parcial de las actividades de negocio o bien causar una pérdida a la organización que hubiera podido evitarse.

Un aspecto importante a tener en cuenta es el factor coste-beneficio. Por ejemplo, la costosa instalación de sistemas de supresión de incendios puede ser fundamental para proteger un gran ordenador que procese datos corporativos críticos, pero puede no ser justificable para proteger un simple PC.

Los riesgos potenciales a los que está sometido un sistema informático se pueden clasificar de acuerdo con su origen en accidentales e intencionados. Entre los primeros podemos destacar:

- Desastres naturales: vendavales, sismos, rayos, etc.
- Incendios.
- Inundaciones.
- Averías en los equipos.
- Averías en las instalaciones eléctricas o de suministro.
- Averías en la climatización.
- Perturbaciones electromagnéticas.
- Errores en la introducción, transmisión o utilización de los datos.
- Errores de explotación.
- Errores de diseño y desarrollo de las aplicaciones.

Entre los intencionados:

- Fraudes.
- Robo de elementos del equipo.
- Robo de información.
- Modificación de los datos con la finalidad de causar perjuicio a la organización en beneficio propio.
- Sabotajes y atentados.
- Huelgas.
- Abandono de la empresa de personal estratégico.
- Difusión o salida controlada de información al exterior.

Frente a estos riesgos podemos adoptar unas posturas determinadas:

- Aceptar el riesgo, confiando en su baja probabilidad de incidencia.

- Transferir el riesgo, contratando seguros, aunque ello no repone la información perdida ni compensa los posibles efectos adversos.
- Evitar el riesgo con la elaboración y puesta en marcha de un Plan de Seguridad Informática, cuyas medidas de carácter preventivo minimicen la probabilidad de ocurrencia de un siniestro.

Seguridad Física y Lógica del CPD

La **seguridad física** consiste en el conjunto de mecanismos y normas encaminados a proteger las personas, instalaciones, equipos centrales y periféricos y los elementos de comunicaciones contra daños eventuales. Está relacionada con los controles que protegen de los desastres naturales como incendios, inundaciones o terremotos, de los intrusos o vándalos, de los peligros medioambientales y de los accidentes.

Los controles de seguridad física regulan además de la sala donde se alberga el equipo del ordenador, la entrada de datos, el entorno (bibliotecas, registros cronológicos, medios magnéticos, áreas de almacenamiento de copias de seguridad y salas de instalaciones de servicios) y todos los detalles o requerimientos tanto arquitectónicos como de preinstalación y mantenimiento de todos los servicios e infraestructuras, incluso la previsión de disponer de una seguridad física integral del entorno.

En el pasado, seguridad física significaba mantener un ordenador y su información alejados de cualquier daño físico, rodeando la instalación del ordenador con cierres de seguridad, vallas y guardias (se basaba casi exclusivamente en la seguridad perimetral).

El concepto de seguridad física ha cambiado para acomodarse a las realidades del entorno de los ordenadores de hoy en día, un entorno que con frecuencia es la típica oficina repleta de PCs. Por ello se deben adaptar los conceptos de seguridad perimetral y de controles de acceso físico a la situación concreta de cada CPD.

La **seguridad lógica** consiste en el conjunto de operaciones y técnicas orientadas a la protección de la información contra la destrucción, modificación indebida, divulgación no autorizada o retraso en su gestación. Se centra, sobre todo, en los controles de acceso a la información manteniendo una política de password, utilizando encriptación en el almacenamiento de la información y estableciendo unos niveles de acceso apropiados en las comunicaciones.

La seguridad física y la lógica deben estar completamente coordinadas, ya que ambas están estrechamente relacionadas y comparten objetivos y presupuestos.

Análisis de riesgos y planes de contingencia

Debemos establecer un compromiso entre la necesaria operatividad del sistema frente a los diversos riesgos potenciales, los mecanismos y técnicas que permiten minimizar sus efectos y costes directos e indirectos del empleo de dichas técnicas.

Se considera, pues, la seguridad informática como un problema de gestión, en el que se trata de alcanzar unos objetivos determinados mediante la asignación óptima de unos recursos (humanos, técnicos, tiempo, económicos). La seguridad debe ser, por tanto, cuidadosamente presupuestada y planificada determinándose el nivel aceptable de seguridad para la organización y los medios más idóneos para conseguirlo.

Del estudio pormenorizado de los riesgos y de la criticidad se determina el nivel aceptable de seguridad y se eligen las medidas a adoptar. Estas medidas se traducen en la Seguridad Preventiva y el Plan de Contingencia.

Análisis de Riesgos

Deberemos determinar cuantitativa y cualitativamente los riesgos a que esté sometida la organización. Una vez tipificados procederemos a estimar la probabilidad de ocurrencia de cada uno. Esta probabilidad no depende tanto del riesgo en sí como de las características concretas de cada CPD. Debido a la dificultad de determinar la probabilidad de ocurrencia de un riesgo se establece una escala con unos niveles subjetivos (evaluación de la ocurrencia de los riesgos):

Prácticamente nunca	Nivel 0
Una vez cada 300 años	Nivel 1
Una vez cada 30 años	Nivel 2
Una vez cada 3 años	Nivel 3
Una vez cada 100 días	Nivel 4
Una vez cada 10 días	Nivel 5
Una vez al día	Nivel 6
10 veces al día	Nivel 7

También se puede recurrir a las estadísticas propias de la instalación (caso de existir) o a las editadas por empresas de consultoría o seguros.

Análisis de criticidad

Se establece un listado priorizado de elementos críticos (aplicaciones, bases de datos, software de base, equipos centrales, periféricos, comunicaciones) según el impacto que su carencia o malfuncionamiento causaría en la operatividad del sistema.

Para ello podemos utilizar una escala que marque el tiempo que se podría tolerar un fallo de funcionamiento de cada elemento: 24 horas, 2-3 días, 1 semana, 15 días, más de un mes. También podemos tener en cuenta la diferente criticidad según la época del año, del mes o de la semana.

Determinación del nivel aceptable de seguridad

Debemos encontrar un equilibrio entre los beneficios de las técnicas a emplear y su coste. Para cada riesgo deberemos cuantificar los daños que puede ocasionar y una estimación de los costes de dichos daños junto con los costes de implantación y mantenimiento de las técnicas apropiadas para evitarlo o reducir su impacto.

Elección de medidas a adoptar

Consiste en seleccionar las medidas de seguridad que permitan prevenir los daños en lo posible y corregirlos o minimizarlos una vez acaecidos, determinando los recursos necesarios para su implantación.

Plan de contingencia

Las medidas de corrección se plasman en un plan con unos objetivos concretos:

- Minimizar las interrupciones en la operación normal.
- Limitar la extensión de las interrupciones y de los daños.
- Posibilitar una vuelta rápida y sencilla al servicio.
- Ofrecer a los empleados unas normas de actuación frente a contingencias.
- Proveer los medios alternativos de proceso en caso de catástrofe.

Para garantizar su validez y que no quede obsoleto con el tiempo, deberá estar en continua revisión. Además el personal debe estar entrenado mediante pruebas simuladas periódicas. En la elaboración del plan debe intervenir la dirección, los técnicos de explotación, los técnicos de desarrollo, el personal de mantenimiento, los usuarios y los proveedores.

El plan debe recoger, en forma de planes unitarios, las respuestas a los diferentes problemas que puedan surgir, y se desglosa en:

- **Plan de Emergencia** : Guía de actuación “paso a paso” en cada fallo o daño. Determina una serie de acciones inmediatas (parada de equipos, aviso de responsables, activar o desactivar alarmas, uso de extintores u otros elementos auxiliares, llamada a mantenimiento lanzar salvaguardas o listados, etc), una serie de acciones posteriores como salvamento, valoración de daños, elaboración de informes, relanzar procesos, relanzar el SO, recuperar copias de seguridad, saltar procesos, etc, así como una asignación de responsabilidades, tanto para las acciones inmediatas como para las posteriores. Para mayor eficacia se procederá en cadenas secuenciales de actuaciones y a introducir una duplicitad humana para asegurar su realización.
- **Plan de Recuperación** : Desarrolla las normas de actuación para reiniciar todas las actividades normales de la organización, bien en el propio CPD, bien en otro centro de respaldo. Si se recupera en el propio centro, se deberán activar los equipos duplicados o auxiliares (si no es automático), se utilizarán los soportes de procesamiento alternativos, se iniciarán las actuaciones de mantenimiento o sustitución de equipos dañados y se utilizarán si es preciso las copias de seguridad. Si se utiliza un centro de respaldo, se deben definir los procedimientos y emplear según la causa que originó el problema, se debe realizar una política de traslados (y vuelta posterior al centro original), se debe recuperar el SO, el software de base y las aplicaciones), se deben relanzar las operaciones (recuperando desde la última salvaguarda en caso de necesidad) y se debe revisar la operación mediante la introducción de pruebas que aseguren el correcto funcionamiento.
- **Plan de Respaldo** : Especifica todos los elementos y procedimientos necesarios para operar en el centro de respaldo (si existe) y mantener en el mismo información sobre la configuración del equipo y de las comunicaciones, del SO, software de base, de las aplicaciones, soporte humano y técnico, suministros de documentación y formularios, modo de regenerar el software para su operativa normal, reglas de explotación y operación, política de accesos y confidencialidad, identificación de usuarios, terminales, etc.

Medios de Detección y Protección

Las áreas controladas deben contar con medios de detección de situaciones anómalas, tales como:

- Puertas abiertas.
- Acceso de intrusos.
- Inundación, humos, control de temperatura, fuegos, etc.

Su objetivo es permitir un conocimiento inmediato y preciso del hecho y su localización, por lo que su actuación debe ser absolutamente fiable dentro de unos parámetros previamente establecidos. Ello exige unas revisiones de funcionamiento y un riguroso mantenimiento preventivo cuya periodicidad dependerá del sistema de detección y del tipo de área controlada al que se aplique.

La detección de un hecho anómalo requiere la información necesaria para una reacción proporcionada. Dependiendo de la información suministrada por el medio de detección y los parámetros previamente establecidos, antes de llegar a un estado de alarma se puede pasar por un estado de alerta.

Así los medios de reacción se van organizando en previsión de su posible actuación. Todos los medios de detección deben integrarse en el Sistema de Gestión de la Seguridad para que los gestione y:

- Avise de la anomalía y su gravedad.
- Inicie acciones de corrección automáticas o proponga acciones manuales a realizar por el personal entrenado para ello.
- Controle las actuaciones (qué, quién, cómo, dónde y cuándo).

Este sistema debe estar bajo vigilancia permanente y combinado con los servicios de mantenimiento, para los casos de mal funcionamiento de cualquier medio de detección. Hay que subrayar que los sistemas de detección deben funcionar incluso con el suministro eléctrico de emergencia.

En caso de incendio, su extinción puede realizarse con medios manuales o automáticos. Los medios manuales se basan en extintores portátiles, mangueras, etc. Es importante resaltar que:

- Existen diferentes tipos de fuego (de sólidos, líquidos, gases eléctricos) y hay extintores apropiados para cada tipo.
- El elemento extintor localizado en un área debe ser el apropiado para el previsible tipo de incendio a declararse en ella. Cualquier medio de extinción puede ser excelente utilizado en un área o más dañino que el propio fuego, si es usado en otra.
- Nunca debe emplearse un medio de extinción manual basado en agua donde pueda haber fuego eléctrico por peligro de electrocución.
- No es aconsejable la intervención de personal no entrenado para ello.
- Siempre que se disponga de tiempo, hay que avisar a la brigada interior de incendios (si la hubiera) o al Servicio de Bomberos.

Los medios automáticos se basan en la inundación del área mediante agua, CO² u otros agentes extintores. El más recomendable es el basado en agua, por su bajo coste y su nulo impacto en el entorno. Los sistemas automáticos basados en el agua deben tener un mecanismo de preacción que, en caso de llegar a un estado de alerta o alarma, sustituye el aire de la conducción por agua.

La actuación de estos sistemas de extinción debe ser combinada con la previa desconexión del suministro de energía eléctrica del área afectada.

Los medios basados en el gas halón (basado en clorofluorocarbonos: CFCs) aunque son efectivos, entrañan peligro para las personas y para el medio ambiente (capa de ozono, efecto invernadero), estando totalmente prohibido por una u otra causa en la mayoría de los países firmantes del Protocolo de Montreal (control de uso de los CFCs). Las áreas controladas deben contar con medios automáticos y manuales de extinción de incendios.

Seguridad perimetral

Se refiere a las medidas que podemos establecer para evitar un acceso indebido al conjunto del CPD. Han sido ya referenciadas en la seguridad física. El establecer un área segura es importante para el buen funcionamiento del centro, puesto que la información almacenada y los procedimientos que se realizan en el CPD son vitales para la organización.

Por ello se deben adoptar todas las medidas cuyo coste esté justificado. Entre ellas:

- Servicio de seguridad: que no sólo controle los accesos al recinto, sino que también realice inspecciones periódicas de las dependencias, sobre todo de las que no tengan personal en cada momento. Su importancia se hace evidente en horas nocturnas o días festivos.
- Barreras, puertas de seguridad, ausencia de ventanas. Son medidas que tienden a dificultar el acceso de personal no autorizado.
- Video vigilancia y alarmas volumétricas: controladas por una centralita en la cabina de seguridad.

Control de Acceso Físico

Se basan en medidas de identificación única de las personas que acceden al CPD. El servicio de seguridad debe llevar un registro de las entradas y salidas al centro. Las visitas autorizadas deben llevar obligatoriamente una tarjeta identificativa o etiqueta en lugar visible que indique claramente que es una visita a las áreas a las que puede acceder y el tiempo de validez (suele ser diaria).

El personal propio debe portar una tarjeta identificativa con fotografía. Se pueden utilizar colores para identificar las áreas a las que puede acceder.

En centros de alta seguridad pueden requerirse medidas auxiliares de identificación:

- Huellas dactilares.
- Fondo de ojo (retina).
- Introducción de códigos de acceso para abrir las puertas.

Niveles de seguridad de acceso

Las instalaciones de la empresa deben clasificarse en varias áreas o zonas que, dependiendo de su utilización y los bienes contenidos, estarán sujetas a unos u otros controles de acceso. Las instalaciones pueden clasificarse de acuerdo con los criterios y denominaciones siguientes:

- Áreas Públicas: espacios en los que no hay ningún tipo de restricción de acceso a empleados o personas ajena a la empresa.
- Áreas Privadas: espacios reservados habitualmente a los empleados y personas ajena a la empresa con autorización por motivos de negocio. En ellos puede haber recursos informáticos con un valor bajo.
- Áreas de Acceso Limitado (AAL): espacios cuyo acceso está reservado a un grupo reducido de empleados y personas ajena a la empresa autorizadas por un acuerdo escrito. Pueden concentrarse en ellos recursos informáticos que, en su conjunto, tiene un valor medio.
- Áreas de Acceso Restringido (AAR): espacios cuyo acceso está reservado a un grupo muy reducido de empleados y personas ajena de la empresa autorizadas por un acuerdo escrito, que tengan necesidad de acceder por razones de negocio. En ellos se encuentran recursos informáticos que, en conjunto, tienen un alto valor o contienen activos de información críticos para las actividades del negocio.

A las dos últimas se les denomina Áreas Controladas. Tienen que permanecer cerradas, incluso cuando estén atendidas, y sus accesos controlados. En las áreas controladas, todos los empleados y las personas ajenas a la empresa con autorización para acceder por razones de negocio tienen que llevar permanentemente y en lugar visible un identificador:

- Los empleados, al menos, con fotografía y nombre.
- Las restantes personas, al menos el nombre (legible) y distintivo de la función que cumplen (ej: visita, contratado, suministrador, etc).
- Los identificadores de los empleados con acceso a áreas controladas pueden tener la posibilidad de lectura por banda magnética o por cualquier otro medio, para facilitar el control de accesos y su registro.
- Todo identificador, especialmente los que permiten el acceso a áreas controladas, es personal y debe ser considerado como una contraseña de acceso físico y no compartirlo con nadie, para evitar verse envuelto en algún incidente de seguridad no deseado.

En las áreas controladas tienen que estar prohibido comer, fumar, consumir bebidas alcohólicas y cualquier tipo de drogas. Las dos últimas están consideradas de alto riesgo potencial para la instalación, por lo que adicionalmente debe impedirse la entrada a cualquier área controlada a las personas de quién se sospeche el consumo.

Los suministros informáticos que sean peligrosos o combustibles tienen que ser almacenados a una distancia prudencial y no trasladarlos al área donde se encuentran el resto de recursos informáticos hasta el momento de su utilización. De igual forma, hay que retirarlos de la zona inmediatamente después de finalizar su uso.

Control de Accesos

Los responsables de las áreas controladas deben mantener unos controles de acceso efectivos y proporcionales al valor de los activos a proteger para que puedan cumplir con unos requisitos de auditabilidad mínimos. Los objetivos son:

- Permitir el acceso únicamente a las personas autorizadas por el responsable del área.
- Registrar las entradas y/o salidas (quién, por dónde y cuándo).

Para facilitar el control de los accesos a estas áreas, es recomendable la existencia de un único punto o puerta de acceso habitual para entrada y salida, sin perjuicio de que existan otras salidas para emergencias que se puedan abrir desde el interior mediante el empuje de una barra.

Las entradas en las Áreas de Acceso Limitado (AAL) tiene que efectuarse desde un área interna, nunca desde un área pública. Cada área de acceso limitado debe tener identificado formalmente un responsable o propietario cuyas responsabilidades son:

- Aprobar y mantener actualizada la relación de personas con autorización de acceso permanente. Las personas que tengan su autorización cancelada, por petición de su dirección o por haber causado baja en la empresa, deben ser eliminadas de la relación de acceso en un tiempo razonable.
- Aprobar accesos temporales a estas áreas. En este caso la persona autorizada debe saber que la autorización es para una sola vez.

Las Áreas de Acceso Restringido (AAR) no deben tener ventanas al exterior y la entrada en las mismas tiene que efectuarse desde un área interna o un Área de acceso limitado, nunca desde un Área pública. Tienen que tener barreras de aislamiento de suelo y techo, incluyendo el falso suelo y el falso techo, o bien detectores volumétricos de intrusos.

Cada área de acceso restringido debe tener identificado formalmente un responsable o propietario cuyas responsabilidades son:

- Aprobar y mantener actualizada la relación de las personas con autorización de acceso permanente, generalmente, porque el trabajo a realizar requiere su presencia dentro del área. La lista de acceso debe ser actualizada siempre que haya cambios que así lo aconsejen y revisada formalmente, al menos, cada seis meses. Las personas que tengan su autorización cancelada por petición de su dirección o por haber causado baja en la empresa tienen que ser eliminados de la lista de acceso inmediatamente.
- Aprobar los accesos temporales a estas áreas, incluyendo los accesos del personal que, estando destinado en el área, accede fuera de su jornada laboral. En este caso, la persona autorizada debe saber que la autorización es para una sola vez. Las autorizaciones temporales deben contener:
 - Nombre de quien autoriza si no es el propietario.
 - El nombre de la persona autorizada.
 - Razón social (si corresponde) o motivo.
 - Fecha y hora de acceso y firma.
 - Fecha hora de salida y la firma.

Tienen que ser guardadas como documentos auditables, al menos un año. El propósito de este registro es tener un archivo histórico de accesos, a utilizar en caso de investigación de incidente de seguridad, pero en ningún caso debe ser una herramienta de control de los empleados. El propietario del área debe revisar, al menos mensualmente, que estos registros de acceso contienen la información descrita.

Es preciso revisar y documentar que las salidas de emergencia tengan alarmas y sean audibles y/o visibles, en la propia sala y en el Centro de Control de Seguridad. Esta revisión debe realizarse al menos, anualmente e incluir la verificación de su correcto funcionamiento incluso con alumbrado de emergencia, si hay pérdida de suministro eléctrico. La documentación de revisiones de funcionamiento de alarmas se guarda como documento auditável.

Valoración de las áreas

Los requisitos de control de acceso físico deben basarse en el valor de los sistemas de información contenidos en cada área controlada y en la importancia de las actividades de negocio suministradas por ellos.

El valor de un sistema de información puede obtenerse de acuerdo con los criterios siguientes:

- Alto valor: sistemas corporativos grandes y medios.
- Valor medio: pequeños sistemas corporativos y redes de área local (LAN).
- Bajo valor: pequeños sistemas (PC's) y terminales.

Para cada caso deben considerarse otros aspectos, como el coste y la necesidad de sustitución de los equipos acumulados en un área o el impacto que podría ocasionar a la empresa la carencia prolongada de una actividad y la no disponibilidad de la información que suministra. Sistemas esenciales son aquellos que contengan actividades críticas para el negocio de la empresa.

La valoración final se realiza teniendo en cuenta todos los aspectos descritos para definir los requisitos de control de acceso y seleccionar el tipo de área controlada y deberían estar en consonancia con la tabla siguiente:

SERVICIO	VALOR	REQUISITOS
Sistemas Esenciales para los Procesos de Negocio	Alto	Controles AAR
	Medio	Controles AAR
	Bajo	Controles AAR
Sistemas No-Esenciales para el Negocio de la Empresa	Alto	Controles AAR
	Medio	Controles AAL
	Bajo	Controles AAL
Unidades de control de Teleproceso, independientemente del sistema que las soporte, Servidores, "Bridges", "gateways" y "Routers" de las LANs así como las herramientas que permitan visualizar el tráfico de las líneas (ej.: "Sniffers", "Trace tools")	Controles AAR	

Para determinar la correcta implantación y la efectividad del control de acceso físico, los propietarios de las áreas controladas deben mantener, al menos, la documentación siguiente:

- La identificación del área, el uso a que se destina, el nivel de información clasificada soportada, el valor de los equipos, la valoración del servicio y los requisitos de control requeridos.
- La forma de comunicar a los usuarios de los servicios localizados en el área el nivel de información clasificada soportada, las medidas de seguridad adoptadas y los requisitos para su cumplimiento.

La valoración final, junto con los aspectos considerados, tiene que ser documentada y guardada por el responsable del área controlada como documentos auditables.

Control de Periféricos

Medios de Almacenamiento

No debemos olvidar que la información vital para la organización no sólo está en los discos magnéticos de los equipos, sino que también se puede encontrar en otros soportes como papel y dispositivos de almacenamiento como disquetes, cintas y otros soportes magneto-ópticos.

Ello obliga a introducir sistemas de limitación de acceso a los mismos y sistemas de destrucción o borrado seguro tras su utilización. Como regla general deben almacenarse, mientras sean útiles, en armarios especiales o zonas restringidas.

El custodio ("librarian") es el responsable de almacenar y controlar los medios de almacenamiento desmontables. El custodio tienen que poder controlar todos los movimientos de los medios de almacenamiento desmontables, incluidos los trasladados, a través de una aplicación o producto de uso exclusivo que le facilite la realización del inventario y la reconciliación.

En los casos en que la empresa tenga más de un Centro de Proceso de Datos, tiene que haber un custodio por cada uno de los centros.

En las LAN y sistemas distribuidos, la información suele ser creada, accedida y almacenada en los discos magnéticos no desmontables de estaciones de trabajo y servidores. Siempre que en este tipo de sistema exista información en medios de almacenamiento desmontables, tiene que ser nombrado un custodio por cada LAN o sistema.

En algunas instalaciones, con gran movilidad de personal o temporalidad del mismo, se llega a bloquear o limitar el uso de las disqueteras de los equipos terminales (PC's), de manera que se dificulte la copia masiva de datos sensibles. También se puede recurrir a la encriptación del almacenamiento para dificultar su acceso mediante herramientas de SO o externas a la aplicación.

Todos los medios de almacenamiento bajo el control del custodio deben estar situados en una AAL (área de acceso limitado) o AAR (área de acceso restringido), dependiendo de la ubicación del sistema donde se procesen, y en una zona aislada cerrada a la que pueda acceder exclusivamente el custodio.

Los medios de almacenamiento dedicados a salvaguardas, la recuperación del sistema y los servicios soportados, deben estar situados en otra zona aislada del centro. Lo mismo se aplicaría en el centro alternativo, si lo hubiera. Así si hubiera, por ejemplo, un incendio en el centro podríamos disponer de estos medios de almacenamiento y no se quemarían junto con nuestro CPD.

Tiene que haber un control para evitar que los medios de almacenamiento desmontables sean montados o accedidos sin autorización.

Los movimientos de medios de almacenamiento entre distintos centros, incluidos los dedicados a salvaguarda, tienen que ser registrados y guardados por el custodio de cada centro.

Los medios de almacenamiento en tránsito tienen que ser protegidos contra su pérdida, deterioro o uso indebido, desde que el custodio del centro origen los ponen en manos del transportista hasta que son recibidos por el custodio del centro de destino.

Durante el traslado, el soporte de almacenamiento y la información contenida deben asegurarse teniendo en cuenta:

- Protección física: para que no sean robados, sustituidos o dañados.
- Protección lógica: para que no sean leídos, copiados o modificados.

Para salvaguardar la confidencialidad, integridad y disponibilidad de la información transportada, tienen que usarse medios de transporte fiables, propios o de empresas solventes responsables.

Para la información sensible o con el más alto nivel de clasificación tienen que utilizarse contenedores cerrados y que sólo puedan ser abiertos por los custodios de los centros origen y destino. En casos excepcionales, habrá que fraccionar el envío en más de una entrega y realizarlo por rutas diferentes.

Inventario y Reconciliación

El custodio de medios de almacenamiento es responsable de:

- Implantar el procedimiento de control de inventario.
- Realizar, al menos anualmente, el inventario de medios de almacenamiento.
- Efectuar la correspondiente reconciliación, en caso de haber discrepancias.

El proceso de inventario y reconciliación debe:

- Contemplar todos los medios de almacenamiento removibles incluidos los volúmenes manejados por robots y los que estén sin grabar.
- Ser realizado por personas, al menos una, directamente relacionadas con la responsabilidad de medios de almacenamiento y con la participación del custodio.
- Iniciarse partiendo de las cifras finales del anterior inventario y completarse incluyendo nuevos volúmenes y los recibidos de otros centros, eliminando los volúmenes retirados o destruidos y los enviados a otros centros y obteniendo la cifra final que será utilizada por el próximo inventario. Las discrepancias deben ser documentadas e iniciar el proceso de reconciliación.
- Incluir, en la documentación relativa al proceso, cualquier informe de discrepancias o incidentes. Esta documentación tiene que ser firmada por el custodio y por su línea de dirección o el propietario de la librería de medios de almacenamiento.

La documentación de soporte de los últimos inventarios y reconciliaciones tiene que ser guardada como documentos auditables.

Impresoras y Listados

Según su ubicación, se consideran dos tipos de impresoras:

- Las locales del sistema, situadas en las áreas de acceso limitado/restringido.
- Las impresoras remotas que no están situadas en las áreas dedicadas a sistemas.

Se deben aplicar los siguientes criterios:

- La responsabilidad de especificar las reglas de utilización de cada impresora y de cada sistema de almacenamiento magnético es del propietario o responsable del mismo.
- El control de las salidas impresas es responsabilidad del usuario final que las envía a las impresoras. El control de los soportes de almacenamiento es responsabilidad del que los genera y del que los utiliza.
- El propietario del sistema o servicio no es responsable de que el propietario o el usuario de las impresoras remotas cumplan con los procedimientos de seguridad descritos.
- Las salidas impresas de información clasificada tienen que ser protegidas contra accesos no autorizados.

Las impresoras remotas situadas en áreas internas (no situadas en AAL o AAR) tiene que tener alguno de los controles siguientes:

- Tener designado un responsable de dirigir las salidas impresas al usuario final que las envió.
- Estar directamente atendida por el usuario final.
- Recoger los listados personalmente e inmediatamente después de terminar la impresión.
- Tener la posibilidad de borrado de listados pendientes de impresión.

Las impresoras locales o remotas situadas en AAL o AAR no requieren ningún control adicional para imprimir información clasificada. No puede haber impresoras remotas situadas en áreas públicas.

Sistema Integral de Gestión de la Seguridad

Se contemplará y analizará la seguridad física independientemente de los sistemas de gestión y control implementados en el CPD.

Se instalará un sistema informatizado para la gestión y el control integral de todas las alarmas procedentes del equipamiento informático, de las infraestructuras y de las instalaciones específicas de seguridad del CPD.

Dicho sistema, recibirá las señales de alarma, dispondrá de la gestión de las mismas y de la posibilidad de realizar desde el mismo la modificación de ciertos parámetros u operaciones de parada, arranque o maniobra del equipamiento de las salas de informática o del recinto del CPD:

- Red de incendios (Sala del CPD, áreas de servicios y despachos, zonas del SAI y del grupo electrógeno).
- Alarmas en general.
- Arranque, paro o maniobra del entorno industrial del CPD.
- Control de accesos y movimientos.
- Control de ahorro de energía.
- Control en el bloque de multicasilleros de reparto.
- Control de la expedición de la producción.
- Control de los stocks de almacenes.
- Estado de las baterías de los SAIs y control de los grupos electrógenos.
- Control de climatización, sobrepresión y renovación ambiental.
- Rede de sondas ambientales en falso suelo, techo y sala de ordenador.
- Red de detección de humedad.

Instalaciones

Entorno

Los edificios o instalaciones de los CPDs requieren unas características adicionales de protección física que deben ser consideradas antes de seleccionar su ubicación, teniendo en cuenta aspectos tales como la posibilidad de daños por fuego, inundación, explosión, disturbios civiles, cercanía de instalaciones peligrosas (depósitos de combustible, aeropuertos, acuartelamientos, etc) o cualquier otra forma de desastre natural o provocado.

Se deberán analizar de forma integral las características dominantes de los distintos entornos, evaluando las ventajas y los riesgos potenciales que pudieran afectar al buen funcionamiento del CPD y planteando las respuestas adecuadas en relación con:

Entorno Natural

Tendremos en cuenta:

- Climatología: tormentas, precipitaciones de agua y nieve, temperaturas extremas, huracanes, ventisca y vientos dominantes.
- Geotecnia: mecánica de los suelos (corrimientos de tierra, hundimientos, estructura físico-química, humedad, existencia de minerales magnéticos, sismicidad, etc)
- Hidrología: proximidad de ríos, proximidad del mar, embalses cercanos y posibles avenidas, etc.

Entorno Artificial

Podemos considerar:

- Acceso a medios de emergencia: bomberos, policía, servicios sanitarios, etc.
- Centrales de gas o depósitos de gas, centrales nucleares.

- Redes de telecomunicaciones.
- Suministro eléctrico, redes de suministro accesibles.
- Plantas petroquímicas, fábricas de cemento, betunes, derivados del vidrio, etc.
- Conducciones o depósitos de líquidos (agua potable, aguas residuales, combustibles, etc).
- Contaminación atmosférica: polvo, vapores corrosivos o tóxicos, etc.
- Perturbaciones locales: ruidos, vibraciones, radiaciones parásitas (radares, balizas de navegación, emisoras de radio y televisión, torres de telecomunicaciones, líneas de alta tensión cercanas, grandes transformadores o motores, centrales eléctricas, repetidores, centrales nucleares, aeropuertos).

Entorno Urbanístico

Tenemos entre otras:

- Dotaciones urbanas: metro, autobuses, intercambiadores ferroviarios, autopistas, aeropuertos, puertos marítimos, hospitales, universidades, supermercados, etc.
- Zonas urbanas: parcelas abiertas, edificaciones colindantes, zonas de oficinas y negocios, parques empresariales, recintos feriales.
- Ambiente de trabajo y salud laboral (microclima de trabajo, contaminación ambiental, sobrecargas físicas y psíquicas influyentes, etc).

Actualmente se ha acuñado el término “AMENITIES” para abarcar todos los servicios complementarios que no son estrictamente necesarios para el desempeño de la actividad de proceso de datos, pero que pueden ofrecerse en el conjunto de la oferta inmobiliaria sobre todo en los parques empresariales o zonas singulares.

Entre ellos están:

- Áreas de descanso, ocio y servicios terciarios.
- Guardería.
- Aparcamiento.
- Clubes, gimnasios e instalaciones deportivas.
- Cajeros automáticos.
- Restaurantes y cafeterías.
- Hoteles.
- Centros comerciales.

Características de Construcción

Una vez seleccionada la ubicación física del edificio que albergará el CPD, habrá que analizar las características específicas de las instalaciones, haciendo hincapié en algunos aspectos:

- Deben estar diseñadas de forma que no se faciliten indicaciones de su propósito ni se pueda identificar la localización de los recursos informáticos.
- Incluir zonas destinadas a carga y descarga de suministros y su inspección de seguridad.
- Cumplir, en los elementos constructivos internos (puertas, paredes, suelos, etc), el máximo nivel de protección exigido por la Norma Básica de Edificación (NBE/CPI-91).
- Disponer de canalizaciones protegidas de cableado de comunicaciones y de electricidad, para evitar ataques (sabotajes, fuego, roedores, insectos), intercepción o perturbaciones por fuentes de emisión próximas (radio, electricidad magnetismo, calor).

Habitabilidad

La mayoría de construcciones de edificios públicos, de oficinas o de negocios no empezaron a cubrir las necesidades de preinstalaciones e instalaciones informáticas hasta bien entrados los años ochenta.

Actualmente, el diseño arquitectónico de un CPD debe estar lo más cercano posible a la arquitectura inteligente. Este hecho ha dado cabida a la domótica.

La domótica comprende todos aquellos desarrollos tecnológicos enfocados al diseño de soluciones rentables que pueda tener el inmueble en el marco de la propia génesis del proyecto arquitectónico. Es la automatización del edificio más la disponibilidad de los recursos de las telecomunicaciones y la ofimática.

Los requerimientos de habitabilidad tienen en cuenta la arquitectura informática del momento. Prevén no sólo el crecimiento del equipamiento informático, sino también el cambio total a otro entorno informático y mantienen rentable las infraestructuras y las dotaciones inteligentes o servicios avanzados del inmueble:

- Habitabilidad en horizontal, es el edificio informático óptimo, el de pocas plantas.
- Habitabilidad en torre, las torres pierden en diafanidad, dificultan la extensión horizontal de la sala de ordenadores y complican la evacuación de emergencia, etc. En las edificaciones informáticas se contemplarán y analizarán al menos los siguientes aspectos, aportando las soluciones más adecuadas:
 - Diseño y ergonomía del inmueble: tienen en cuenta las exigencias de explotación, producción y los requerimientos físicos de la arquitectura informática residente. Serán estas exigencias funcionales las que determinarán el diseño exterior e interior así como la estructura del edificio para las salas de informática.
 - Flexibilidad: informa sobre la capacidad del edificio para satisfacer las necesidades futuras, entre las que destaca la posibilidad de modificar distribuciones tanto de la arquitectura informática residente o de nueva implantación como de sus preinstalaciones y las áreas del personal de producción o explotación.
 - Integración de servicios: permite saber cuándo un edificio entra dentro del concepto de "arquitectura inteligente", que al menos debe cumplir:
 - Automatización de la actividad de mantenimiento.
 - Automatización de los servicios comunes del edificio.
 - Mejora de la calidad de vida en el trabajo.
 - Ofimática.
 - Planificación del espacio.
 - Telecomunicaciones.

Requerimientos de las Edificaciones e Instalaciones

Se aplicarán las normas generales de obligado cumplimiento:

- Norma Básica de la Edificación.
- Normas tecnológicas de la Edificación.
- Ordenanzas Municipales.
- Reglamentos electrotécnicos.
- Verificación de los Productos y Suministros Industriales en el marco de la construcción.
- Normas de Preinstalación de las Firmas Informáticas.

Además de la aplicación de las normas generales, el estudio para la elección del edificio deberá comprender todo lo que compete a la arquitectura tradicional y muy especialmente a:

- La estructura y sobrecargas de uso.
- Las fachadas del inmueble.
- Accesos a los almacenes.
- Instalaciones para las salas de informática.
- Muelles de carga y descarga, elevadores, montacargas, etc.
- Acceso al edificio de mercancías pesadas (montacargas industrial).
- Acceso a la Sala de Informática (siempre puertas doble hoja).
- Existencias de salidas de Seguridad al CPD.
- Falso suelo y techo tecnológicos.
- Protección contra las infiltraciones de agua y humedad.
- Suministros de energía eléctrica y agua.
- Iluminación de día y de emergencia.
- Resistencia al Fuego en minutos de la estructura, forjados y muros de carga.
- Muros cortafuegos.
- Puertas contra incendios.
- Situación de las puertas de acceso y evacuación.
- Túneles de seguridad y escaleras de emergencia.
- Particiones interiores o mamparas dobles.
- Que no crucen las salas de informática conducciones de aguas tanto pluviales como de desagües excepto las propias de la climatización.
- Tratamientos referentes a resistencias eléctricas, acústicas y mecánicas.
- Protección contra la energía eléctrica de reacción: Toma de tierra del edificio, pararrayos.

Características de las Infraestructuras

Instalaciones Eléctricas

Los cuadros de mandos se instalarán en lugares fácilmente accesibles, con espacio holgado (previendo las posibles ampliaciones), correcta y claramente etiquetados y por supuesto con el más estricto rigor en materia de calidad de aparatos y montaje (deberán cumplir con las normas habituales de protección y seccionamiento). Se evitarán las perturbaciones electromagnéticas, aislando adecuadamente aquellas máquinas generadoras de campos inductivos y armónicos.

Se evitará la electricidad estática empleando los revestimientos más adecuados, instalando las tomas de tierra convenientes y manteniendo la humedad en el rango adecuado (al menos del 55%).

Los recursos informáticos son sensibles a las variaciones de tensión y de frecuencia de la corriente eléctrica. Los requerimientos básicos para el suministro de energía eléctrica son dos: Calidad y Continuidad.

Relacionado con la Calidad se puede destacar que:

- Las variaciones de frecuencia deben corregirse con equipos estabilizadores que la mantengan dentro de los rangos establecidos por los fabricantes de los recursos informáticos a alimentar, aunque algunos recursos informáticos de nueva tecnología los llevan incluidos.
- Las variaciones de tensión deben ser manejadas por un Sistema de Alimentación Ininterrumpida (SAI en inglés UPS), de modo que se puedan prevenir los efectos de posibles microcortes.

En relación con la continuidad del suministro eléctrico debe tenerse en cuenta que las caídas de tensión pueden ser manejadas por un SAI (UPS), pero sólo por tiempo limitado, ya que el desgaste de sus acumuladores es muy rápido y su recarga muy lenta para utilizarlo en cortes sucesivos y nunca como única alternativa.

Las soluciones habituales se basan en una de las siguientes o en la combinación de varias de ellas:

- Conexión conmutada a dos compañías suministradoras.
- Conexión conmutada a dos estaciones transformadoras de la misma compañía pero situadas en rutas de suministro diferentes.
- Capacidad de transformación de corriente asegurada mediante equipos redundantes.
- Equipos electrógenos de combustión.

Siempre que el volumen de las instalaciones informáticas así lo aconseje, el suministro eléctrico y las tomas de tierra deben ser independientes de las generales del edificio y a suficiente distancia de ellas, correctamente aisladas y rigurosamente mantenidas.

Recinto de Protección Combinada

Son recintos de protección combinada aquellas compartimentaciones dentro de los CPD capaces de garantizar una custodia segura de los soportes magnéticos de respaldo ante los agentes más peligrosos que puedan atacarlos. Estarán dotados al menos con:

- Apantallamiento electromagnético y jaulas de Faraday.
- Protección contra reacciones químicas que produzcan HCL y gases de combustión corrosivos dentro de la cámara.
- Protección contra intrusión y el robo (puerta de seguridad).
- Protección contra el vandalismo y explosiones.
- Protección contra incendios y sus efectos derivados (humos y vapores).
- Protección contra las inundaciones interiores del CPD.
- Protección contra el impulso electromagnético nuclear (NEMP).
- Sellado de las instalaciones y de la cámara contra altas frecuencias e incendio.

Instalaciones de Agua

Se evitará, en lo posible, las canalizaciones de agua en la sala de ordenadores (sobre todo por falso techo, falso suelo o visibles). En todo caso se preverán los mecanismos de detección de fugas y la instalación de válvulas que puedan cerrar las conducciones afectadas. Los detectores de agua se basarán en sensores puntuales o de banda que cubran áreas completas.

El cableado debe estar impermeabilizado cuando discorra por zonas con riesgo de humedad o inundación. Si no es posible separar los conductos de agua del resto de instalaciones, se preverá dotar al techo o solera del forjado, por donde discurren las tuberías, de la inclinación oportuna para evacuar el agua hacia los puntos de drenaje establecidos, evitando su acumulación.

Si existen en el edificio o adosados a él depósitos de agua u otro tipo de líquido, se asegurará la estanqueidad de los mismos y se instalarán de forma que su rotura no afecte a los servicios esenciales ni por supuesto a las personas.

En el caso de salas de informática situadas en sótanos se reforzará la estanqueidad de paredes, pisos, techos, puertas y ventanas. Se preverá la instalación de bombas automáticas para evacuar eventuales inundaciones, que deben alimentarse con un sistema eléctrico independiente del resto de la sala para permitir su funcionamiento independiente.

Si las CPU precisan agua fría para la refrigeración, se preverá la red de tuberías con sus válvulas de corte y retención, sondas detectoras y sistema auxiliar de emergencia desde el contador del canal con filtrado del líquido.

Medidas de las Instalaciones contra Incendios

El fuego causa el mayor número de accidentes en los CPDs. Por ello es imprescindible controlar puntos zonales y además realizar un estudio en función de los agentes extintores (tener en cuenta la prohibición del uso del HALÖN, Protocolo de Montreal sobre CFCs).

Se procederá a estudiar como medidas:

- El acceso de los bomberos a cualquier zona del edificio previendo las tomas de agua a presión convenientes.
- La resistencia al fuego de los materiales de construcción, carpintería, revestimiento, etc. Se evitarán aquellos materiales que generen productos tóxicos o gran cantidad de humo al ser sometidos al fuego (NBE-CPI-91). También hay que evitar que se acumulen listados de control y otros papeles en el CPD.
- El mecanismo más adecuado para cortar la alimentación eléctrica en caso de incendio.
- Los mecanismos idóneos para evitar que los conductores de refrigeración y ventilación actúen como chimeneas y contribuyan a propagar el incendio, parándose automáticamente el aire acondicionado en caso de incendio.
- La compartimentación del edificio, aislando aquellas zonas que contengan materiales fácilmente combustibles, que se limitarán al máximo.
- Tabicados de hormigón con mamparas y puertas ignífugas.
- La Instalación de puertas contra fuegos dotadas de los mecanismos que aseguren su cierre de forma automática.
- La prohibición de fumar, colocando carteles claramente visibles, en las zonas de mayor riesgo.
- El mobiliario, fabricado con materiales resistentes al fuego.
- Los contenedores de papel, materiales plásticos, etc, deberán tener una tapa metálica, que permanecerá cerrada de forma automática.
- La construcción de recintos de protección combinada o la disposición de armarios ignífugos.
- La instalación de un sistema de alarmas cruzadas y centralizadas en el Sistema Integral de Gestión de la Seguridad, para la detección o extinción de incendios en el CPD.

La mayoría de los armarios que se utilizan en las salas de informática no son ignífugos sino refractarios o simples cajas fuertes. No corresponden al grado de vulnerabilidad exigido en la CEE.

Acondicionamiento de Aire

Con la evolución tecnológica ya existen en el mercado recursos informáticos que reducen (prácticamente eliminan) los tradicionales requerimientos de aire acondicionado. Sin embargo, debido al parque existente en España y a su antigüedad media, se deben tener en cuenta las siguientes consideraciones:

- Para mantener el ambiente con la temperatura y la humedad adecuadas, especialmente los de las grandes instalaciones, hay que disipar el calor que generan a través del aire acondicionado.
- La suficiente potencia y redundancia de estos equipos permitirá que trabajen desahogadamente y que las operaciones de mantenimiento sean sencillas y frecuentes.
- Un elemento fundamental del sistema acondicionador de aire es el mecanismo de corte automático tras producirse una detección de incendio.

Planes de Emergencia en Instalaciones y Evacuación

Tiene que haber implantado, de acuerdo con las leyes y reglamentos en vigor (especialmente con la norma NBE/CPI-91), un Plan de Emergencia y Evacuación de las instalaciones de la empresa. Este plan sólo afecta a la protección de las personas que trabajan o se hallan circunstancialmente en las instalaciones de la empresa y por tanto también afecta al CPD.

No tiene relación directa con el plan de seguridad del CPD o de emergencia, aunque deben estar completamente coordinados.

Los objetivos de este Plan deben ser:

- Conocer los edificios y sus instalaciones, las áreas de posibles riesgos y los medios de protección disponibles.
- Evitar, o al menos minimizar, las causas de las emergencias.
- Garantizar la fiabilidad de los medios de protección.
- Informar de las medidas de protección a todos los ocupantes de las instalaciones.
- Disponer de personal organizado y adiestrado para las situaciones de emergencia.
- Hacer cumplir la vigente normativa de seguridad.
- Preparar la posible intervención de recursos externos (Policía, Bomberos, Ambulancias, etc).

El plan de evacuación debe ser estudiado y comprobado. Se colocarán señalizaciones hacia las salidas de emergencia en todas las salas y pasillos de forma que sean fácilmente visibles desde cualquier ubicación y situación (elementos luminosos con baterías propias, elementos fosforescentes).

Se instalarán alarmas acústicas y luminosas para alertar de las emergencias. Se realizarán periódicamente simulaciones de evacuación.

Dimensionamiento de Equipos

Evaluación del Rendimiento de un Sistema Informático

Se define evaluación del rendimiento de un sistema informático como la medida de cómo un software determinado está utilizando el hardware con una determinada carga del sistema. Por ejemplo, para un computador se entiende por carga del sistema a una determinada combinación de programas.

La mayor dificultad que tiene la evaluación de las prestaciones de un sistema informático se atribuye al hecho de que la carga real de un sistema informático cambia continuamente, lo que impide poder repetir la medida a no ser que se trabaje en un entorno controlado de carga. Todas las actividades que forman parte del estudio del comportamiento de un sistema se denominan actividades de evaluación de sus prestaciones.

La necesidad de evaluar las prestaciones de un sistema informático ha surgido como una consecuencia natural del aumento de la potencia y de la complejidad de los sistemas. Esta evaluación no es una tarea sencilla, ya que ha de tener en cuenta muchos y variados aspectos del hardware, del software y de las aplicaciones que se han de llevar a cabo en el sistema informático.

La evaluación de un sistema informático sirve para:

- Comprobar que el funcionamiento del sistema es el correcto.
- Detectar y eliminar los denominados “cuellos de botella”.
- Influir en las decisiones de diseño, implantación, compra y modificación de los sistemas informáticos, es decir, en todas las etapas de su ciclo de vida.
- Comparar un cierto número de diseños alternativos del sistema (diseñador de sistemas).
- Analizar el sistema más adecuado para ejecutar un determinado número de aplicaciones (administrador de sistemas).
- Planificar la capacidad, es decir, predecir el comportamiento del sistema con nuevas cargas.

Por lo tanto, es necesario evaluar un sistema informático cuando se quiere:

- Diseñar una máquina.
- Diseñar un sistema informático.
- Seleccionar y configurar un sistema informático.
- Planificar la capacidad de un sistema informático.
- Sintonizar o ajustar un sistema informático.
- Caracterizar y predecir la carga.

El comportamiento de un sistema es muy dependiente de la carga aplicada al mismo. Debido al crecimiento vegetativo de la carga de un sistema informático se produce una disminución de las prestaciones del mismo.

Para evitar esta disminución es necesario ajustar o cambiar algunos de los parámetros del SO. En ciertos casos, si el sistema no se puede cambiar, hay que intentar mejorar el comportamiento mediante la modificación de la carga (programas).⁷

Sistemas de Referencia

Se distinguen tres tipos de sistemas de referencia o tipos de funcionamiento de un sistema informático a la hora del estudio de las prestaciones y su evaluación:

- *Sistema por lotes (batch)* : Básicamente consiste en que el computador ejecuta una serie de programas que previamente el responsable del sistema deja almacenados en memoria. Es dicha persona quién decide los trabajos que deben estar en ejecución en cada instante. Por lo tanto, la planificación interna del SO está ayudada por la externa humana. Estos trabajos realizan ciclos de uso de la CPU y de los discos de forma continua hasta que finalizan. Algunos índices de las prestaciones de estos sistemas son los siguientes:
 - *Tiempo de respuesta (Turnaround time)* , es el tiempo que transcurre desde que se lanza la ejecución de un trabajo hasta que se termina.
 - *Diseño y evaluación* de configuraciones.
 - *Productividad* medida en trabajos por unidad de tiempo.
- *Sistema transaccional* : es aquél en que un conjunto de terminales remotos conectados al sistema interaccionan con un conjunto de programas. Cada interacción constituye una *transacción* . Ejemplos: el sistema informático de un banco o el de reserva de billetes o el que recibe medidas de un satélite. La planificación interna del SO debe gestionar las peticiones recibidas sin intervención humana. Un entorno de este tipo queda definido por el flujo de transacciones que le llega, siendo su índice de prestaciones característico el *tiempo de respuesta* $tr = tra + ten + tr$, donde:
 - *tra* es el *tiempo de reacción* , que se define como el tiempo que transcurre desde que la transacción llega al sistema hasta que comienza la ejecución.

- *ten* es el *tiempo de ejecución*, que se define como el tiempo que transcurre desde que el sistema comienza la ejecución de la transacción hasta que termina.
- *tro* es el *tiempo de retorno*, que se define como el tiempo que transcurre desde que finaliza la ejecución hasta que, eventualmente, se completa la respuesta hacia el usuario.
- *Sistema interactivo o por demanda* : Un *sistema interactivo* es aquél en que los usuarios acceden a él desde terminales remotos teniendo acceso a la totalidad del SO. En estos sistemas, un usuario da una orden al terminal que pasa a procesarse por el conjunto CPU+discos y, transcurrido un cierto tiempo, produce una respuesta en el terminal. En estos sistemas no existe planificación humana que ayude a la planificación del SO. Además, queda definido por los siguientes índices:
 - *Número de usuarios* que tiene conectados.
 - *Tiempo de reflexión* de los usuarios, que es el tiempo que transcurre desde que el usuario recibe la respuesta y envía otra nueva orden.

Los índices de prestaciones característicos son: el *tiempo de respuesta* y la *productividad*, medida esta última en peticiones por unidad de tiempo.

Técnicas de Evaluación de un Sistema Informático

Se denominan *técnicas de evaluación* a los métodos y herramientas que permiten obtener los índices de prestaciones de un sistema que está ejecutando una carga dada con unos valores determinados de parámetros del sistema. Se distinguen tres tipos de técnicas:

- *Monitorización* . Los monitores son unas herramientas de medición que permiten seguir el comportamiento de los principales elementos de un sistema informático, cuando éste se haya sometido a una carga de trabajo determinada. Estas herramientas hacen un seguimiento de lo que sucede en el sistema, que es lo que se denomina *monitorización* .
- *Modelado* . Es la herramienta que hay que utilizar cuando se trata de evaluar el comportamiento de un sistema en el que hay algún elemento que no está instalado. El modelado se puede realizar de dos formas:
 - *Métodos analíticos* que proporcionan las teorías de colas. Se basan en la resolución de las ecuaciones matemáticas que representan el equilibrio existente en los eventos que se producen en el sistema, mediante algoritmos aproximados. Su principal inconveniente es la limitación para tratar determinadas estructuras de colas que existen en los sistemas informáticos.
 - *Simulación* . Consiste en la construcción de un programa que reproduce el comportamiento temporal del sistema, basándose en sus estados y sus transiciones. Los resultados se obtienen por extracción de estadísticas del comportamiento simulado del sistema. Requieren de más tiempo de cálculo y esfuerzo de puesta a punto que los métodos analíticos. La principal dificultad del modelado reside en la obtención de datos lo suficientemente precisos para ejecutar el modelo y obtener resultados con un grado de aproximación adecuado.
- *Benchmarking* . Se trata de un método bastante frecuente de comparar sistemas informáticos frente a la carga característica de una instalación concreta. La comparación se realiza básicamente a partir del tiempo de ejecución. Las principales dificultades que plantea este método están relacionadas con la utilización de una carga que sea lo suficientemente reducida para ser manejable y lo suficientemente extensa para ser representativa.

Monitores

Un *monitor* es una herramienta utilizada para observar la actividad de un sistema informático mientras es utilizado por los usuarios y para cuantificar los resultados de dicha observación.

En general, los monitores observan el comportamiento del sistema, recogen datos estadísticos de la ejecución de los programas, analizan los datos recogidos y presentan los resultados.

Se define *monitorización* como el seguimiento de la actividad realizada por un sistema informático. Se ha de tener en cuenta que en informática, puesto que no es posible repetir las mismas condiciones de carga en los mismos instantes, el resultado de una medición será distinto unas veces de otras, es decir, no se da la repetibilidad de la medida. La información aportada por el monitor puede ser útil para:

- El usuario y el administrador, ya que les permite conocer toda una serie de características del sistema (capacidad, posibilidad de ampliación, planificación, etc.)
- El propio sistema, para la realización de la adaptación dinámica de la carga.

La calidad de un monitor viene determinada por las siguientes características:

- *Sobrecarga o interferencia* . La energía del sistema consumida por el instrumento de medida debe ser tan poca como sea posible, de forma que la perturbación introducida por el instrumento no altere los resultados de la observación. Los monitores hardware presentan este peligro en sus puntos de conexión y para evitarlo utilizan sondas electrónicas de muy alta impedancia. Por otro lado, los monitores software aumentan la carga del sistema y alteran por consiguiente su comportamiento, por lo que se debe tratar de minimizar al máximo este efecto.
- *Precisión* . Es el error que puede afectar al valor de los datos recogidos. Estos errores son debidos a diferentes causas: la interferencia del propio monitor, una incorrecta instalación o utilización, el número de dígitos para representar la medición, etc.
- *Resolución* . Es la capacidad de la herramienta de separar dos acontecimientos consecutivos en el tiempo. También se define como la máxima frecuencia a la que se pueden detectar y registrar correctamente los datos.
- *Ámbito o dominio de medida* . Hace referencia al tipo de acontecimientos que puede detectar, es decir, a las características que puede observar y medir.
- *Anchura de entrada* . Es el número máximo de bits que el monitor puede extraer en paralelo y procesar cuando se produce un acontecimiento.
- *Capacidad de reducción de datos* . Es la capacidad que puede tener el monitor de analizar, procesar y empaquetar datos durante la monitorización para un mejor tratamiento y compresión de los mismos y para reducir el espacio necesario para almacenar los resultados.
- *Compatibilidad* . El hardware y el software de monitorización debe ser fácilmente adaptable a cualquier entorno y requerimiento de la aplicación.
- *Coste* (adquisición, instalación, mantenimiento, formación y operación).
- *Facilidad de instalación y de utilización* .

Los monitores se pueden clasificar atendiendo a tres aspectos:

- Forma de implantación.
- Mecanismo de activación.
- Forma de mostrar los resultados.

Según su implantación se clasifican en:

- *Monitores software* : Son programas o ampliaciones del SO que acceden al estado del sistema, informando al usuario sobre dicho estado. Son los más adecuados para monitorizar los SO, las redes y las BD, así como las aplicaciones que las utilizan. Cada activación del monitor implica la ejecución de varias instrucciones por parte de la CPU del sistema que está analizando, lo que puede provocar una gran sobrecarga en el sistema si la causa de la activación se produce con gran frecuencia.

- *Monitores hardware* : Son dispositivos para medir las prestaciones de sistemas informáticos y se conectan al hardware del sistema que se va a monitorizar por medio de sondas electrónicas, que son elementos capaces de detectar eventos de tipo eléctrico. Un monitor hardware podrá reconocer todos aquellos acontecimientos que se reflejen en puntos fijos del sistema. Su principal característica es que son externos al sistema que van a medir, lo que implica:
 - No utilizan recursos del sistema que van a monitorizar.
 - No producen interferencias con éste.
 - Son muy rápidos.
 - Sus principales desventajas son:
 - Son más difíciles de instalar.
 - Existen magnitudes a las que no se puede acceder.
 - Requieren para su operación y análisis de resultados de personal especializado.
 - Pueden interactuar a nivel eléctrico con el sistema que se va a monitorizar, provocando perturbaciones que resulten en un funcionamiento anómalo del sistema monitorizado.
- *Monitores híbridos* : Son una combinación de las dos técnicas anteriores, intentando combinar las ventajas de una y otra.

Según su mecanismo de activación se clasifican en:

- *Monitores de eventos o acontecimientos* . Son aquellos que se activan por la aparición de ciertos eventos. Si el evento se da con frecuencia, la sobrecarga que se produce es elevada.
- *Monitores de muestreo* . Son aquellos que se activan a intervalos de tiempo fijos o aleatorios mediante interrupciones de reloj. La frecuencia de muestreo viene determinada por la frecuencia del estado que se desea analizar y por la resolución que se deseé conseguir.

Según su forma de mostrar los resultados se clasifican en:

- *Monitores de tiempo real* , que constan de un módulo analizador que procesa los datos a medida que los recibe.
- *Monitores batch* , que primero recogen la totalidad de la información para posteriormente analizarla.

Técnicas Analíticas: Teoría de Colas

La teoría de *colas* permite determinar el tiempo que un trabajo pasa esperando en las distintas colas de un sistema. Una red de colas es un conjunto de estaciones de servicio y de clientes. Las estaciones de servicio representan los recursos del sistemas y los clientes usualmente representan a los usuarios.

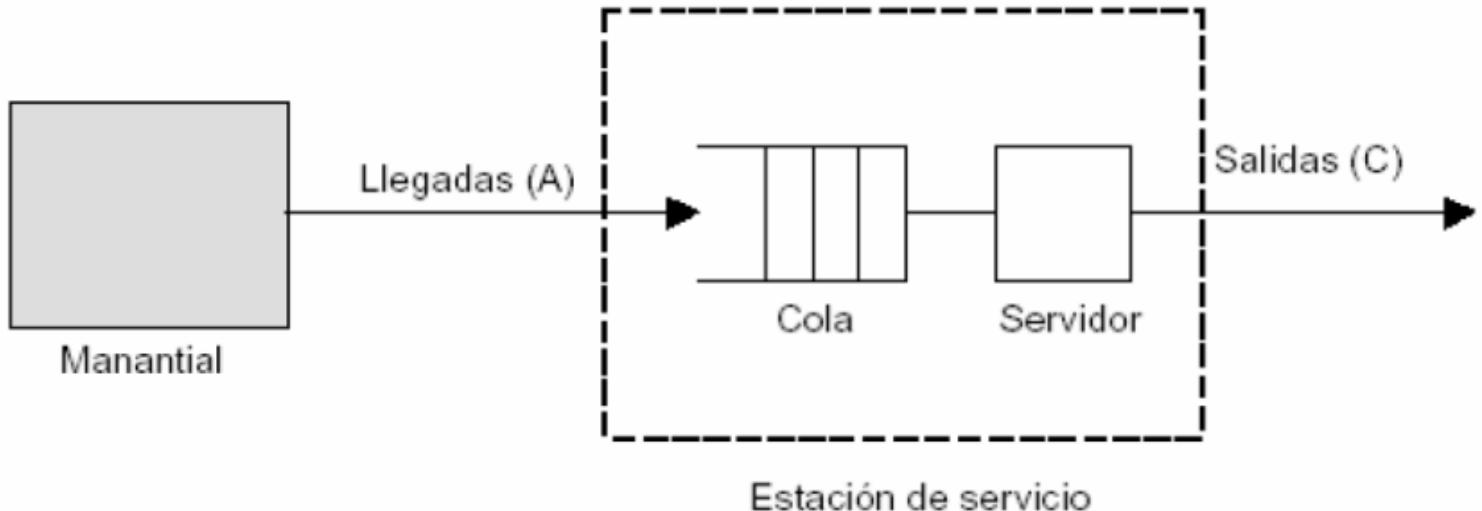
Una *estación de servicio* consta de un servidor más una cola de espera asociada. Se establece por tanto la siguiente relación entre los modelos y los sistemas reales:

Servidor (modelo) <=> Recurso del sistema (hardware)
 Cola (modelo) <=> Cola (software) asociada al recurso

Se denomina *resolución o evaluación analítica* a la obtención de los índices de prestaciones del sistema a partir de un modelo de colas. Básicamente, consisten en la resolución de un conjunto de ecuaciones que se deducen a partir del modelo y de sus parámetros.

El objetivo del es llegar a establecer relaciones entre las variables que caracterizan la carga y las que miden el comportamiento. El término operacional implica que el sistema es directamente medible. Por tanto, una suposición o hipótesis operacional será aquella que puede ser comprobada o verificada mediante la medida.

Esquema de una estación de servicio:



Una petición tendrá que esperar en la cola hasta que el servidor quede libre. Una vez que un cliente recibe el servicio abandona la estación. El conjunto (Estación de servicio + clientes) constituye la versión más simple de un modelo de red de colas. Básicamente este modelo tendrá dos parámetros:

- La intensidad de carga o tasa de llegada de los clientes, que se mide en peticiones/segundo.
- La demanda de servicio, que es el tiempo medio de servicio de un cliente y se mide en segundos.

Los clientes en una estación de cola compiten por el servidor. Por ello, el tiempo de residencia estará compuesto por un posible tiempo de espera y un tiempo de servicio.

Las variables operacionales básicas son las que se pueden medir directamente sobre el sistema durante un intervalo de observación finito:

- T (seg), intervalo de observación o de medida del sistema.
- A (peticiones o clientes), número de peticiones llegadas o clientes durante el intervalo T.
- C (peticiones o clientes), número de peticiones completadas o servidas durante el intervalo T.
- B (seg), tiempo durante el cual el recurso observado (la estación de servicio) ha estado ocupado.

Benchmarks

En general, se puede decir que los benchmarks son programas utilizados para medir el rendimiento de un sistema informático o de alguna de sus partes. La finalidad de sus estudios puede ser muy variada: comparación de sistemas, su sintonización, la planificación de su capacidad, la comparación de compiladores en la generación de código eficiente, el diseño de sistemas o procesadores, etc.

Como programa benchmark se puede usar prácticamente cualquier programa ejecutable, escrito en cualquier lenguaje de programación, incluso en lenguaje máquina. Se denomina benchmarking al proceso de comparar dos o más sistemas mediante la obtención de medidas. Para conseguir un buen paquete de programas benchmark se deben seguir una serie de pasos, siendo los principales:

- Determinar los objetivos de utilización de benchmark.
- Escoger los mejores programas de benchmark según los objetivos determinados en el paso anterior. Por ejemplo, si se desea estudiar el rendimiento de E/S de un sistema,

se elegirán programas con un consumo importante de E/S y no programas con consumo intensivo de CPU.

- Se deben comprobar los aspectos del sistema bajo estudio que influyen en el rendimiento, como pueden ser el SO, los compiladores y su nivel de optimización, el uso de memorias caché, etc. Además se debe comprobar que los programas, versiones y datos usados como benchmark sean los mismos en todas las pruebas.
- Finalmente, obtenidos los resultados y entendiendo perfectamente qué hace cada programa benchmark, se debe intentar estudiar la causa de las diferencias obtenidas en los distintos sistemas evaluados.

La palabra *benchmark* se puede definir de dos formas:

- Definición 1: Los benchmarks son una forma de evaluar las prestaciones de un sistema informático, bien en su conjunto o de alguna de sus partes. Además, si el benchmark está estandarizado, se puede utilizar para comparar diferentes sistemas.
- Definición 2: Los benchmarks se pueden definir como conjuntos de programas completos escritos en lenguaje de alto nivel y que se consideran representativos de la carga real.

Una vez dadas las diferentes definiciones de benchmark, conviene conocer las principales aplicaciones de este tipo de programas:

- En la *comparación del rendimiento de diferentes sistemas informáticos* con vistas a la adquisición de equipos. Se debe remarcar que estas comparaciones serán tanto más importantes cuanto mayores y más complejos sean el sistema y las aplicaciones que debe soportar. De esta forma será más importante en máquinas de tipo UNIX (multitarea y multiusuario) trabajando sobre arquitecturas diversas que en máquinas MSDOS (monopuesto y monotarea) trabajando sobre la base de una misma familia de chips. Con vistas a la consecución de este objetivo, interesa que los benchmarks estén formados por programas extraídos de la carga real, programas estándar o una mezcla de ambos y servirán para comparar el rendimiento de sistemas informáticos.
- En la *sintonización de sistemas*, es decir, cuando se quiere mejorar el rendimiento de un sistema informático que ya está en funcionamiento. En este sentido, interesa que los benchmark permitan detectar qué partes del sistema deben mejorarse o, una vez introducidas las modificaciones, comprobar que efectivamente se ha aumentado el rendimiento del sistema. Suelen ser programas extraídos de la carga real.
- En la *planificación de la capacidad* de un sistema informático, es decir, conocer la capacidad que le queda disponible en previsión de posibles ampliaciones. Interesa que los benchmarks usados lleven el rendimiento del sistema hasta el límite, para así poder prever las carencias que presentará el sistema en el futuro. Para este fin, se utilizan programas artificiales que disponen de parámetros regulables que permiten modificar la cantidad de consumo de recursos en el sistema.
- En la *comparación de compiladores* desde el punto de vista de la generación de código. Los programas elegidos para la comparación de compiladores pueden ser estándar o extraídos de la carga real. Además conviene indicar sobre qué arquitectura y SO se realizan las pruebas.
- En el *diseño de sistemas informáticos o de procesadores*. En este caso, se parte de un sistema general inicial y, mediante simulaciones, tomando como entrada los benchmarks elegidos, se obtiene unos resultados a partir de los cuales se intentará ir mejorando paulatinamente el diseño del sistema. En el diseño se debe tener en cuenta el tipo de compilador que se emplea, ya que deben ser tan independientes como sea posible de la arquitectura. También se pueden estudiar las interrelaciones existentes entre las distintas arquitecturas, sus implantaciones, los lenguajes de programación y los algoritmos que ejecutan.

Cuellos de Botella

Un *cuello de botella* es una limitación de las prestaciones del sistema provocada por diversas causas: un componente hardware, un componente software o la organización del sistema. Un cuello de botella produce una ralentización considerable del tráfico de los procesos en un área del sistema. Así, cuando la demanda de servicio de un determinado componente excede en frecuencia e intensidad a la capacidad de servicio de ese componente, se dan las condiciones para la aparición de un cuello de botella.

El término cuello de botella sólo es apropiado cuando el problema en las prestaciones puede ser atribuido a uno o dos recursos del sistema, ya que en un sistema donde todos o casi todos los componentes están sobrecargados no se pueden encontrar cuellos de botella concretos y se habla de un sistema *sobrecargado* o *saturado*.

En ocasiones, la eliminación de un cuello de botella hace que aparezca otro distinto, diciendo en ese caso que un cuello de botella esconde otro. Se deben eliminar todos los cuellos de botella hasta conseguir que el sistema se encuentre *equilibrado* (*balanced*).

Por otra parte, los cuellos de botella no están directamente ligados a una configuración dada sino que son función, en gran medida, de la carga. Además como la carga en los sistemas suele variar con el tiempo, pueden aparecer *cuellos de botella temporales*, que son aquellos que aparecen por un espacio de medida relativamente corto respecto a la sesión de medida. Por ello para su detección se suele utilizar un método de interpretación de las medidas en tiempo real (on line).

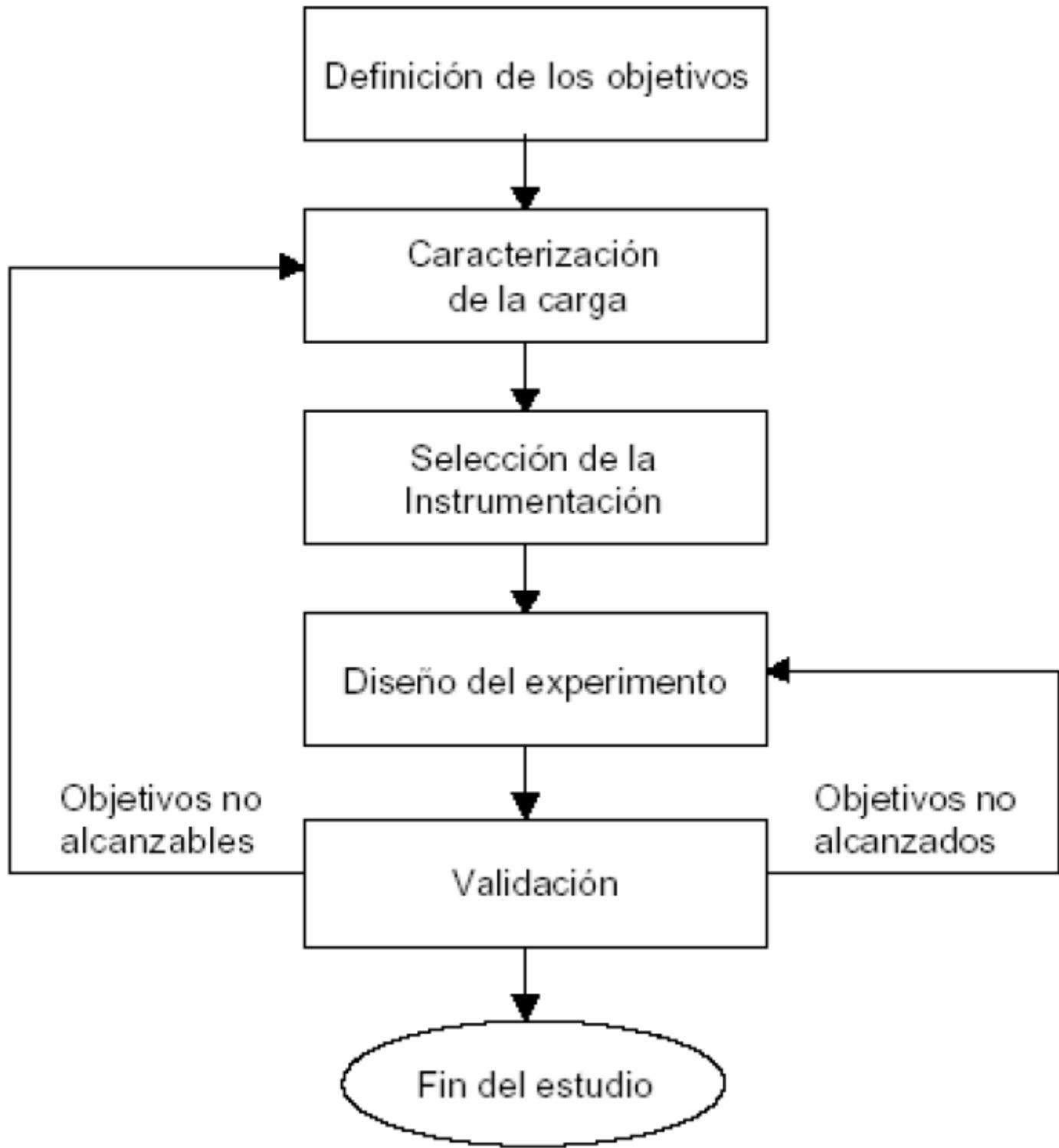
Existen diversas aproximaciones para la detección de los cuellos de botella. Conceptualmente todas ellas son bastante similares, aunque están basadas en técnicas diferentes (simulación, modelos analíticos, medidas sobre el sistema, ...). El método más común es el basado en la interpretación a posteriori (*off-line*) de las medidas realizadas sobre el sistema. El procedimiento que sigue cualquiera de estos métodos es el siguiente:

- Debido a síntomas de ineficiencia o algunos de los estudios de evaluación que se realizan al sistema se sospecha de la existencia de un cuello de botella. Se debe formular una hipótesis acerca de la localización y causa de dicho cuello de botella.
- Se procede a la validación de la hipótesis acerca de la causa del cuello de botella. Para ello se recurre a los datos medidos sobre el sistema o, si éstos son insuficientes, a la recogida de más datos y a su análisis.
- Una vez confirmada la hipótesis se plantea el problema de eliminar el cuello de botella o de reducir sus efectos al menos. Para ello hay que determinar el método más eficaz. En general se pueden distinguir dos tipos de modificaciones del sistema para la eliminación de cuellos de botella:
 - *Terapias de reposición (upgrading)*, hacen referencia a modificaciones del hardware como añadir, reemplazar o incluso eliminar uno o más componentes hardware.
 - *Terapias de sintonización (tunning)*, se trata de modificaciones que no alteran la configuración pero de alguna manera tienen efecto sobre la organización del sistema como por ejemplo, cambiar ficheros de un disco a otro, cambiar un disco de canal, etc. Este tipo de terapias son, en general, más económicas y menos radicales que las terapias de reposición.
- Se procede a la *modificación del sistema* de acuerdo con el método seleccionado.

Sintonización

Para mejorar las prestaciones y la eficiencia de un sistema informático es necesario realizar un estudio de evaluación de las prestaciones de dicho sistema. Las operaciones

que hay que llevar a cabo en un método de mejora de prestaciones se pueden agrupar en las siguientes fases:



La definición de los objetivos del estudio es una fase fundamental, ya que en función de los mismos se determinará el método que se utilizará para analizar las prestaciones del sistema, la cantidad de recursos que es preciso emplear y la forma de justificar la inversión necesaria.

Inicialmente los objetivos deben ser modestos y estar basados en el análisis de los datos proporcionados por las rutinas de contabilidad, que proporcionan los datos de partida del estudio y pueden revelar la existencia de problemas no detectados.

Estos problemas se pueden agrupar en las siguientes clases:

- Análisis de los dispositivos de hardware.
- Eficiencia de los programas.
- Problemas de la carga.

- Localización de los cuellos de botella.

Cada una de estas clases o áreas de estudio requerirá unas herramientas y unas técnicas específicas. Por otra parte, también es conveniente fijar el ámbito de estudio, es decir, si se va a tratar un problema concreto o un problema global.

Ejemplos de objetivos de estudio a *nivel global del sistema* son:

- Verificar la posibilidad de evitar, o de al menos, posponer por algún tiempo, la adquisición de nuevo hardware (memoria, periféricos, CPU, ...).
- Reducir el overhead del sistema y en general todas las actividades que no sean productivas.
- Reducir la carga actual.

La consecución del segundo y tercer objetivo aumentará la capacidad residual del sistema y resultará de interés si lo que se pretende es encontrar espacio para nuevas aplicaciones sin tener que expandir la configuración del sistema.

Por otro lado, el primero y especialmente el segundo de los objetivos suelen requerir la localización de posibles cuellos de botella del sistema.

Ejemplos de objetivos específicos son:

- Reducir el tiempo medio de respuesta en un porcentaje dado.
- Determinar si la tasa de paginación es tan elevada que pueda ser necesario algún tipo de intervención (ej: la ampliación de la memoria principal).
- Determinar la mejor distribución de los archivos en los discos conectados a los diversos canales.
- Determinar la relación existente entre las utilizaciones de memoria y de CPU y el número de usuarios conectados.
- Equilibrar y optimizar la actividad de los canales.

En ocasiones para lograr un determinado objetivo es necesario analizar y resolver previamente otros problemas.

Factores a considerar

Para evaluar el comportamiento de un sistema es necesario disponer de una serie de medidas cuantitativas o parámetros que:

- Caracterizan el comportamiento tanto del hardware como del software del computador.
- Hacen referencia a cómo el usuario (visión externa) y el responsable del sistema (visión interna) ven su comportamiento.

Estas magnitudes o parámetros están relacionadas con tres tipos de medidas correspondientes a:

- Consumo de tiempos.
- Utilización de recursos o dispositivos.
- Trabajo realizado por el sistema o componentes del mismo.

Variables externas o perceptibles por el usuario

- *Productividad (Throughput)* , es la cantidad de trabajo útil ejecutado por unidad de tiempo (u.t.) en un entorno de carga determinado. Normalmente se mide en (trabajos/hora) o en (transacciones/segundo).
- *Capacidad* , es la máxima cantidad de trabajo útil que se puede realizar por u.t. en un entorno de carga determinado.
- *Tiempo de respuesta* , es el tiempo transcurrido entre la entrega de un trabajo o una transacción al sistema y la recepción del resultado o la respuesta.

Variables Internas o del Sistema

- *Factor de utilización de un componente* , es el porcentaje de tiempo durante el cual un componente del sistema informático (CPU, dispositivo de E/S, canal, etc.) ha sido realmente usado.
- *Solapamiento de componentes* , es el porcentaje de tiempo durante el cual dos o más componentes del sistema han sido utilizados simultáneamente.
- *Overhead* , es el porcentaje de tiempo que los distintos dispositivos del sistema (CPU, disco, memoria, etc) han sido utilizados en tareas del sistema no directamente imputables a ninguno de los trabajos en curso.
- *Factor de carga de multiprogramación* , es la relación entre el tiempo de respuesta de un trabajo en un determinado entorno de multiprogramación y su tiempo de respuesta en monoprogramación.
- *Factor de ganancia o multiprogramación* , es la relación entre el tiempo total necesario para ejecutar un conjunto de programas secuencialmente en monoprogramación y en multiprogramación.
- *Frecuencia de fallo de página* , es el número de fallos de página que se producen por unidad de tiempo en un sistema de memoria virtual paginada.
- *Frecuencia de swapping* , es el número de programas expulsados de memoria por unidad de tiempo a causa de falta de espacio o con el fin de permitir su reorganización para recuperar espacio en ella o para disminuir la paginación.

Otras Magnitudes Relativas al comportamiento

- *Fiabilidad* , es una función del tiempo definida como la probabilidad que el sistema trabaje correctamente a lo largo de un intervalo de tiempo dado. Se mide por la probabilidad de fallos por unidad de tiempo o por el tiempo medio entre fallos.
- *Disponibilidad* , es una función del tiempo definida como la probabilidad de que el sistema esté trabajando correctamente y por lo tanto se encuentre disponible para realizar sus funciones en el instante considerado t .
- *Seguridad* , es la probabilidad que el sistema esté realizando correctamente sus funciones o parado de forma tal que no perturbe el funcionamiento de otros sistemas ni comprometa la seguridad de las personas relacionadas con él.
- *Rendimiento* , es una función del tiempo definida como la probabilidad que las prestaciones del sistema estén por encima de un cierto nivel en un instante determinado.
- *Mantenibilidad* , es la probabilidad que un sistema averiado pueda ser reparado y devuelto al estado operacional dentro de un periodo de tiempo determinado.

Magnitudes que caracterizan la carga

Se denomina *carga de prueba* la carga usada en el estudio de las prestaciones de un sistema informático. Se distinguen dos tipos de carga:

- *Carga real*, se observa en un sistema durante su funcionamiento normal. Su principal inconveniente es que no permite repeticiones para eliminar los errores de medición, por lo que es difícilmente utilizable como carga de prueba.
- *Carga sintética*, está constituida por un conjunto de programas extraídos o no de la carga real del sistema informático que la reproduce de forma compacta. Puede utilizarse repetidamente y puede modificarse sin afectar a la operatividad del sistema. En muchos sistemas la evaluación se suele realizar en un sistema distinto pero equivalente al real, es decir, creándose dos sistemas paralelos.

Magnitudes que caracterizan cada componente de la carga

- *Tiempo de CPU de trabajo*, es el tiempo total de CPU necesario para ejecutar un trabajo (programa, transacción, etc) en un sistema determinado. Es una función directa del número de instrucciones que se ejecutan para realizar ese trabajo, del volumen de datos procesados y de la velocidad del procesador.
- *Número de operaciones de E/S por trabajo*, es el número total de operaciones de E/S que requiere la ejecución de un trabajo.
- *Características de las operaciones de E/S por trabajo*, hacen referencia al soporte (cinta, disco, etc) y, en el caso de discos, a la posición que ocupa el archivo sobre el que se efectúan.
- *Prioridad*, es la que el usuario asigna a cada uno de los trabajos que procesa el sistema.
- *Memoria*, es la que requiere ocupar, para su ejecución, un trabajo determinado. Puede ser constante o variable.
- *Localidad de las referencias*, es el tiempo en que todas las referencias hechas por un trabajo permanecen dentro de una página (segmento) o conjunto de páginas (segmentos).

Magnitudes que caracterizan el conjunto de carga

- *Tiempo entre llegadas*, es el tiempo entre dos requerimientos sucesivos para un servicio (ejecución de un trabajo o transacción) del sistema.
- *Frecuencia de llegada*, es el número medio de llegadas de nuevas peticiones de ejecución que se producen por unidad de tiempo. Es la inversa del tiempo entre llegadas.
- *Distribución de trabajos*, define la proporción existente entre las ejecuciones de los distintos trabajos que constituyen la carga.

Magnitudes que caracterizan las cargas convencionales

- *Tiempo de reflexión del usuario*, es el tiempo que el usuario de un terminal de un sistema interactivo necesita para generar una nueva petición al sistema (tiempo de leer la respuesta previa, de pensar en la nueva acción que se vaya a formar y de teclearla).
- *Número de usuarios simultáneos*, es el número de usuarios interactivos que trabajan simultáneamente sobre el sistema en un instante dado.

- *Intensidad del usuario*, es la relación entre el tiempo de respuesta de una petición y el tiempo de reflexión del usuario.

Magnitudes para controlar el comportamiento

Algunas de las modificaciones que se pueden introducir en un sistema para mejorar su comportamiento son:

- Ajuste de los parámetros del SO.
 - *Tamaño del quantum*, es la cantidad de tiempo de uso ininterrumpido de la CPU que un sistema de tiempo compartido asigna a los diferentes trabajos. Si el quantum es demasiado grande se favorece a los trabajos con mucho uso de la CPU, mientras que si es demasiado pequeño se puede introducir un overhead importante debido a los continuos cambios de contexto de un programa a otro cada vez que se agota el quantum.
 - *Prioridad interna*, es el nivel inicial de prioridad interna que recibe un programa en función de la prioridad externa asignada.
 - *Factor de multiprogramación*, es el número máximo de trabajos que están simultáneamente en memoria principal y, por lo tanto, que tienen opción a utilizar la CPU y los demás recursos activos del sistema. Cuanto mayor sea este valor tanto mejor será el aprovechamiento de todos los recursos del sistema, aunque también aumentará el overhead.
 - *Tamaño de la partición de memoria*, es la cantidad fija de memoria principal asignada a una cola de trabajos.
 - *Máxima frecuencia de fallo de página*, es el valor de la frecuencia de fallo de página por encima del cual se produce un excesivo overhead. A partir de este valor de frecuencia se efectúa la suspensión o swapping de alguno de los trabajos en curso.
 - Número máximo de usuarios simultáneos.
- *Modificación de las políticas de gestión del SO*, como por ejemplo cambiar las prioridades de los diferentes tipos de tareas.
- *Equilibrado de la distribución de cargas*, se pretenden utilizar de la forma más uniforme posible todos los dispositivos del sistema informático. Cuando el uso de los mismos está desequilibrado se deben disponer los cambios necesarios para lograr el equilibrio deseado. Este tipo de corrección acostumbra, en muchos casos, a proporcionar mejoras espectaculares en el comportamiento del sistema.
- *Sustitución o ampliación de los componentes del sistema*, cuando los métodos anteriores no funcionan se debe modificar la configuración del sistema, bien sea sustituyendo determinados elementos por otros de mayor capacidad o rapidez, o bien sea por aumento del número de dispositivos que constituyen la configuración del sistema. Es importante darse cuenta de que la ampliación de la configuración debe hacerse de tal forma que se despeje el posible cuello de botella que se pueda haber detectado, ya que de lo contrario el comportamiento conjunto del sistema ampliado no variará de forma significativa.
- *Modificación de los programas*, de tal forma que su ejecución promedio requiera de menos recursos. Esto se puede conseguir bien mediante recodificación de los caminos del programa recorridos con mayor asiduidad, o bien por un montaje que agrupe en la misma página o segmento aquellos módulos del programa que deben coexistir en memoria para la ejecución del programa, etc. Hay que destacar que este método provoca la modificación de la carga del sistema y normalmente se considera la carga como un dato del problema que no se puede modificar.

Conmutación

La conmutación surge en las redes de larga distancia para reducir el número de enlaces, simplificar el mantenimiento y ahorrar costes. La conexión de N elementos de una red todos con todos, con enlaces punto a punto requeriría $N(N-1)/2$ enlaces, que lo hace totalmente inviable.

La transmisión a larga distancia se realiza a través de una red de nodos de conmutación intermedios. Los conmutadores forman un mosaico de NxM puntos de conmutación que conectan todas las líneas. Estos nodos de conmutación no están interesados por el contenido de la información, simplemente los encaminan desde la fuente hacia el destino mediante su conmutación de nodo en nodo.

En una red de conmutación suele haber más de un camino entre cada par de estaciones. Surge así la necesidad de determinar cual es la ruta de encaminamiento óptima.

Existen dos tecnologías de conmutación diferentes: conmutación de circuitos y conmutación de paquetes.

Conmutación de Circuitos

El funcionamiento de las redes de conmutación de circuitos se basa en los dos principios siguientes:

- Se establece un circuito para la comunicación entre dos usuarios que piden el intercambio de información.
- El circuito se asigna durante todo el tiempo que dure la comunicación.

En la conmutación de circuitos quien establece la llamada determina el destino enviando un mensaje especial a la red con la dirección del destinatario de la llamada. Se establece entonces una comunicación directa entre las dos estaciones mediante la conmutación adecuada de todos los nodos intermedios. La red asigna recursos para que tenga lugar la comunicación y se enviará el mensaje cuando quien ha efectuado la llamada tenga conocimiento de que ésta ha sido establecida.

Las redes de conmutación de circuitos pueden ser analógicas, como la Red Telefónica Pública Conmutada, o digitales como la Red Digital de Servicios Integrados.

Aunque la conmutación de circuitos puede utilizarse para la transmisión de datos, resulta ineficiente para este propósito en términos de los recursos de línea puesto que la línea se mantiene ocupada durante toda la sesión, incluso cuando no hay información circulando por ella. Además la necesidad del establecimiento de un enlace de conexión antes de enviar los datos, puede generar un retraso de tiempo significativo respecto al tiempo de transferencia de los datos.

La conmutación de circuitos se caracteriza por:

- Manejar un ancho de banda fijo.
- La información sigue una ruta preestablecida que no puede modificarse. Si existiera un bloqueo, no se podría enviar la información por otro camino.
- Los nodos intermedios no almacenan información. Si existiera un bloqueo, la información se perdería.
- No proporciona control de errores.
- Es muy rápida, tanto como permita el ancho de banda, y segura.

Las redes de conmutación de circuitos trabajan casi exclusivamente con elementos de tipo síncrono, planteándose entonces el problema de sincronización a través de toda la red. Nos encontramos con tres tipos de sincronización de los centros de conmutación:

- Sincronización por valor medio de fase: el reloj de cada nodo de la red se ajusta con el valor obtenido calculando la media de los valores de los relojes de los nodos vecinos y del propio nodo.
- Sincronización por director-subordinado: se ajustan los relojes de los distintos nodos con el valor de un reloj maestro o director.
- Sincronización de forma anárquica: los relojes de los distintos nodos son de gran precisión y cada uno de ellos actúa de forma independiente.

Conmutación de Paquetes

La conmutación de paquetes es la técnica utilizada más comúnmente para la comunicación de datos. Los mensajes que se quieren comunicar se dividen en varias partes denominadas paquetes. Cada paquete se transmite de forma individual a través de la red y pueden incluso seguir rutas diferentes. En el destino se reensamblan los paquetes en el mensaje original.

Al igual que en la conmutación de circuitos las líneas se conectan a centros de conmutación, pero en este caso los paquetes se transmiten al conmutador y son almacenados en una cola en espera de su envío (store and forward) y cuando hay un tiempo muerto en la comunicación entre la fuente y el destino, las líneas pueden ser utilizadas para otras comunicaciones.

En esta técnica se pone un límite al tamaño de los paquetes, que depende, entre otros factores, de:

- El número de enlaces que se pueden conectar al centro de conmutación.
- La ocupación de la línea de transmisión.
- El tiempo de respuesta requerido.

La conmutación de paquetes se caracteriza por:

- Permite realizar varias transmisiones simultáneamente.
- Existe un retardo mínimo entre usuarios.
- El conmutador de paquetes proporciona normalmente tratamiento de errores y conversiones de código entre los distintos terminales y ordenadores.
- Se permite trabajar a distintas cadencias de línea.
- El ancho de banda no es rígido para toda la comunicación.
- Cada paquete lleva una cabecera que contiene la información sobre el camino que debe seguir el paquete.
- Permite la multidifusión. Un mismo mensaje se puede enviar a varios usuarios.
- Se pueden establecer niveles de prioridad en los mensajes.

Existe otra técnica de conmutación, la conmutación de mensajes, que se basa en los mismos principios que la conmutación de paquetes, pero en la que los mensajes no son fragmentados en paquetes sino que se envían como un bloque completo. Comparada con la conmutación de paquetes presenta dos desventajas importantes: el espacio necesario para el almacenamiento de mensajes puede llegar a ser muy grande y lo mismo sucede con los retrasos de tiempo, lo que la hace inapropiada para el trabajo en tiempo real.

Existen dos esquemas en la conmutación de paquetes:

- **Datagramas** : cada paquete se trata como una entidad independiente y es encaminado individualmente a través de la red. La cabecera de cada paquete contiene información completa acerca de su destino.
- **Círculo virtual** : se utiliza una fase inicial para establecer una ruta (un circuito virtual) entre los nodos intermedios para todos los paquetes transmitidos durante la sesión de comunicación entre dos nodos finales.

En el método datagrama los paquetes no siguen una ruta preestablecida, los nodos intermedios examinan la cabecera y deciden a que nodo enviar el paquete para que éste alcance su destino. La entrega no está garantizada y los paquetes pueden llegar al destino en un orden distinto al que fueron enviados, así que deben ser ordenados en el destino para formar el mensaje original. La principal implementación de red de conmutación de paquetes en modo datagrama es Internet con el protocolo IP.

En el método circuito virtual en cada nodo intermedio se registra una entrada en una tabla indicando la ruta que ha sido establecida para la conexión, de esta forma los paquetes pueden tener cabeceras más cortas al contener sólo un identificador del circuito virtual y no información completa sobre su destino. Los paquetes llegan al destino en el orden correcto y hay una cierta garantía de que llegan libres de errores.

Las formas más comunes de redes de conmutación de paquetes de circuito virtual son X.25, Frame Relay y ATM.

Encaminamiento

Como ya hemos comentado, cuando hay mas de un camino entre cada par de estaciones hay que determinar cual es la ruta de encaminamiento óptima para el intercambio de información.

Los protocolos de encaminamiento son aquellos protocolos que proporcionan técnicas para encaminar la información y que además proporcionan técnicas o mecanismos para compartir la información de encaminamiento.

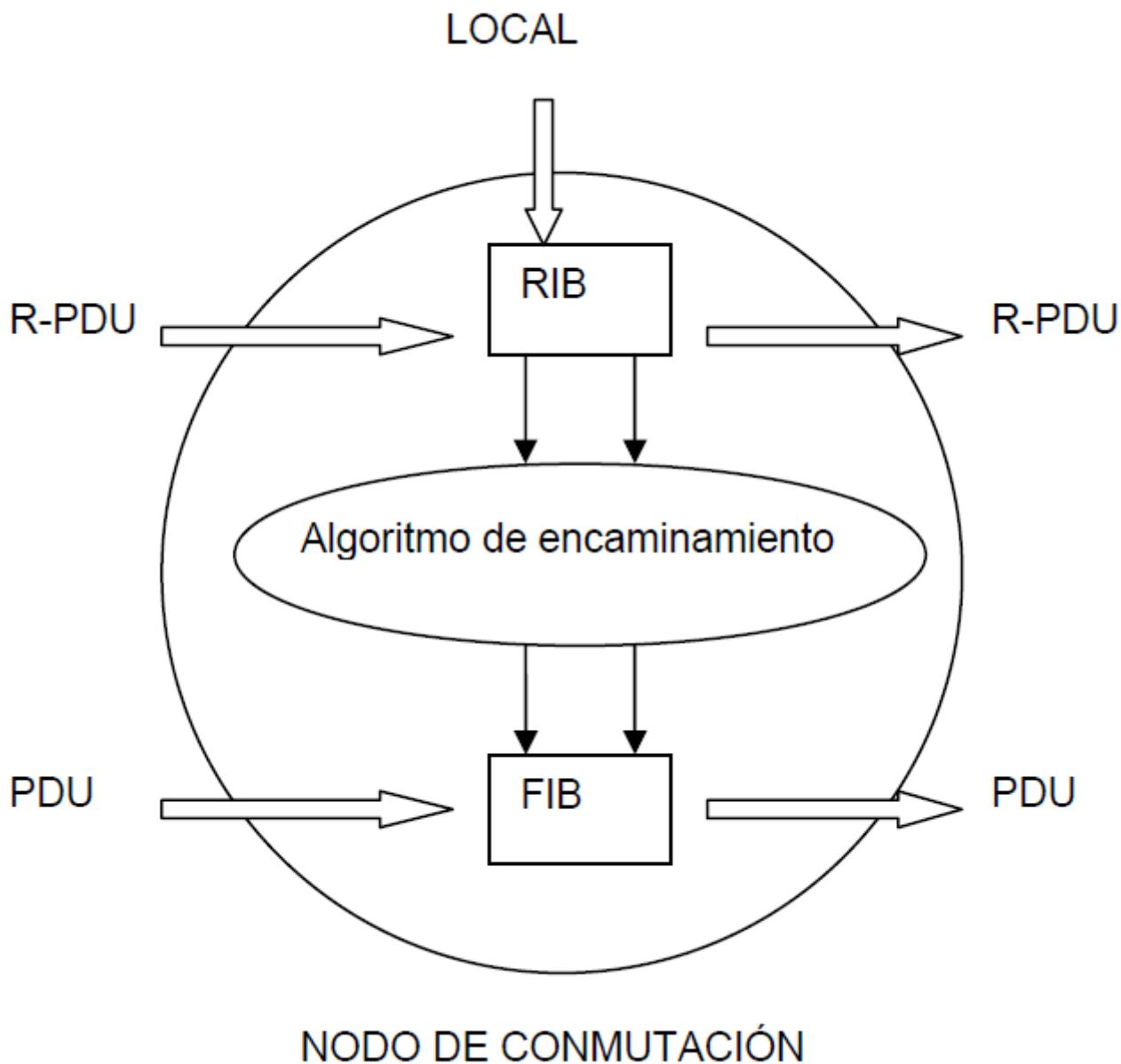
Los distintos protocolos de encaminamiento, que actúan en la capa de red del modelo de referencia OSI, utilizan diferentes algoritmos de encaminamiento y diferentes métricas para evaluar qué camino es el mejor para transportar el paquete, pero entre sus objetivos de diseño todos tienen una o más de las características siguientes:

- Optimalidad: se refiere a la capacidad del algoritmo de encaminamiento empleado de seleccionar la mejor ruta, que va a depender de las métricas utilizadas y de como se ponderen al realizar los cálculos.
- Simplicidad y baja sobrecarga: deberían ofrecer su funcionalidad de una forma eficiente con mínima utilización de información de sobrecarga.
- Ser robustos y estables: se espera que actúen de una forma correcta cuando se enfrenten a situaciones poco usuales o imprevistas, tales como fallos del hardware, implementaciones incorrectas o condiciones de sobrecarga de tráfico.
- Convergencia rápida: la convergencia es el proceso de acuerdo sobre las rutas óptimas por parte de todos los routers.
- Flexibilidad: deberían ser capaces de adaptarse de una manera rápida y apropiada a las condiciones cambiantes de la red.

Las métricas se pueden calcular atendiendo a diferentes factores, entre ellos:

- Ancho de banda: se refiere a la capacidad de transporte de datos de un enlace.
- Número de saltos: el número de routers que se atraviesan en el camino entre el origen y el destino.
- Carga: la cantidad de datos que pasa por una interfaz determinada.
- Fiabilidad: la tasa de error de un enlace determinado.
- Retardo: lo que tarda el paquete desde el origen al destino.
- Coste de la comunicación: término genérico que engloba los gastos de operación de los enlaces.

Antes de ver la clasificación de los métodos de encaminamiento conviene detallar la estructura general de un nodo de conmutación.



- **LOCAL** : información sobre el entorno local del nodo (memoria disponible, enlaces locales, etc.).
- **PDU** (Protocol Data Unit): unidad básica de información.
- **R-PDU** (Routing PDU): información de control entre nodos. Son paquetes de control enviados por otros nodos con información sobre la red, no contienen datos.
- **FIB** (Forward Information Base): es la tabla de encaminamiento que es consultada para el reenvío de los paquetes (representados en la figura por la PDU).
- **RIB** (Routing Information Base): es la tabla que almacena las distancias o costes a los nodos. Es la base de información de encaminamiento que se utiliza para formar la

FIB. La información de la RIB se consigue mediante la recepción de R-PDUs procedentes de otros nodos vecinos y por la interacción con el entorno local de cada nodo.

Los métodos de encaminamiento se pueden clasificar, atendiendo al lugar dónde se decide el encaminamiento, en:

- **Encaminamiento fijado en origen** (Source Routing): la ruta que debe seguir el paquete se fija en el sistema terminal origen y son estos los que contienen las tablas de encaminamiento. Cada paquete lleva un campo que especifica su ruta y los nodos de conmutación simplemente reenvían los paquetes por esas rutas ya especificadas.
- **Salto a Salto** (Hop by Hop): cada nodo, sabiendo donde está el destino, conoce sólo el siguiente salto a realizar.

En función de la adaptabilidad a los cambios se clasifican en:

- **No adaptativos (estáticos)** : las tablas de encaminamiento de los nodos se configuran manualmente y permanecen fijas hasta que se vuelve a actuar sobre ellas. Una variante de los algoritmos estáticos son los **Q-estáticos** que poseen una cierta adaptabilidad: en lugar de dar una sola ruta fija, se proporcionan además varias alternativas para el caso en que falle la principal.
- **Adaptativos (dinámicos)** : Se tienen tres tipos de encaminamiento dinámico:
 - **Adaptativo Centralizado** : todos los nodos se consideran iguales excepto uno, el nodo central. Este nodo cuenta con información de todos los nodos y centraliza el control. Cada nodo envía al nodo central las R-PDUs con información de control. Es el nodo central el encargado de formar la tabla de enrutamiento de cada nodo.
 - **Adaptativo Aislado** : en cada nodo sólo se cuenta con información local, pero el control es distribuido: Cada vez que un nodo recibe un paquete que no es para él lo reenvía por todos los enlaces excepto por el que llegó. Los principales métodos de encaminamiento adaptativo aislado son los **algoritmos de inundación y de estado de enlaces** .
 - **Adaptativo Distribuido** : son los más utilizados. Todos los nodos envían y reciben información de control a sus vecinos y calculan su tabla de encaminamiento en función de su RIB. El control de encaminamiento es distribuido. La adaptación a los cambios es óptima (siempre que estos cambios sean notificados). Entre los métodos de encaminamiento adaptativo distribuido se encuentran los **algoritmos de vector de distancias** .

En los **algoritmos de inundación** los nodos no intercambian información de control. Cuando llega un paquete a un nodo lo que hace es conmutarlo por todos los puertos de salida sin mirar ninguna tabla de enrutamiento. Tienen el problema potencial de que si la topología de la red tiene bucles el paquete puede estar dando vueltas de manera indefinida. Una solución pasa por limitar la vida del paquete en la red, incluyendo un campo en el paquete que contenga el número de saltos que puede dar. Cada vez que se commuta el paquete se decrementa en una unidad el valor de dicho campo y cuando llegue a cero en lugar de conmutarlo se descarta.

En los **algoritmos de estado del enlace** cada nodo difunde a todos los nodos de la red sus distancias con los nodos vecinos, es decir, cada nodo comunica su información local a todos los nodos.

En los algoritmos de vector de distancias cada nodo informa a todos sus nodos vecinos de las distancias o costes conocidos por él mediante los vectores de distancias, que son vectores de longitud variable que contienen un par (nodo, distancia de nodo) por cada nodo conocido por el nodo en cuestión. Estos vectores se envían periódicamente y cada vez que varían las distancias. Con todos los vectores recibidos cada nodo monta su tabla de enrutamiento.

X.25

El escenario de las redes de comunicación en los primeros años 70 estaba caracterizado por la existencia de muchas redes públicas de datos, muy diferentes internamente, en manos de organizaciones gubernamentales y compañías privadas. Con el crecimiento de la interconexión de redes se hizo patente la necesidad de un protocolo de red común capaz de proporcionar la conectividad entre estas redes públicas de datos.

En 1976, la CCITT (ITU-T desde 1993) recomendó X.25 como el protocolo común. Este protocolo ha sido revisado posteriormente en varias ocasiones.

X.25 es un conjunto de protocolos que define una recomendación internacional tanto para el intercambio de datos como para el control de la información entre un dispositivo de usuario, el Equipo Terminal de Datos (ETD), y un nodo de red, el Equipo de Terminación del Circuito de Datos (ETCD), de una red de conmutación de paquetes (PSN, Packet Switched Network). La velocidad típica de una red X.25 está entre 9,6-64 Kbps.

Para el establecimiento de la comunicación entre dispositivos en diferentes emplazamientos, una organización puede construir su propia PSN privada o puede abonarse al servicio de una PSN pública. Las facilidades ofrecidas al abonado y la estructura de tarifas dependerán del proveedor del servicio.

X.25 utiliza un servicio orientado a la conexión. Un nodo extremo indica a la red que desea iniciar una conversación con otro nodo extremo. La red envía la petición al destinatario, que puede aceptarla o rechazarla. Si la acepta, se establece la conexión.

Según la terminología usada en X.25, los dispositivos de red se clasifican en tres categorías:

- **Equipo Terminal de Datos (ETD)** : también conocido por DTE (Data Terminal Equipment). Son los equipos finales que se comunican a través de la red X.25. Generalmente son terminales, PCs o hosts localizados en el local del abonado del servicio.
- **Equipo Terminal del Circuito de Datos (ETCD)** : también conocido por DCE (Data Communication Equipment). Es un dispositivo en el punto de entrada a la red. Son DCEs por ejemplo los módems o los conmutadores de paquetes. Cada ETD debe estar asociado a un ETCD.
- **Equipo de conmutación de datos (ECD)** : también conocido por DSE (Data Switching Equipment). Es un nodo de conmutación en la red de conmutación de paquetes. Transfiere los datos de un DTE a otro DTE.

Cada sistema en una red X.25 tiene una dirección que lo identifica y que es proporcionada por el proveedor. Para asegurar la unicidad de las direcciones, la norma X.121 define un esquema de numeración internacional. Esta dirección se llama dirección de usuario de red (NUA, Network User Address) y está compuesta de dos partes:

- **Código de identificación de la red de datos, DNIC** . Consta de 4 dígitos divididos en:
 - **Código del país, DCC** . Formado por tres dígitos, el primero identifica una zona geográfica a escala mundial y los dos siguientes un país específico.
 - **Código de la red de datos pública** : un único dígito que identifica una PDN específica.
- **NTN, Número de terminal nacional** : son 10 dígitos asignados por el proveedor y que no tienen una regla determinada para su formación.

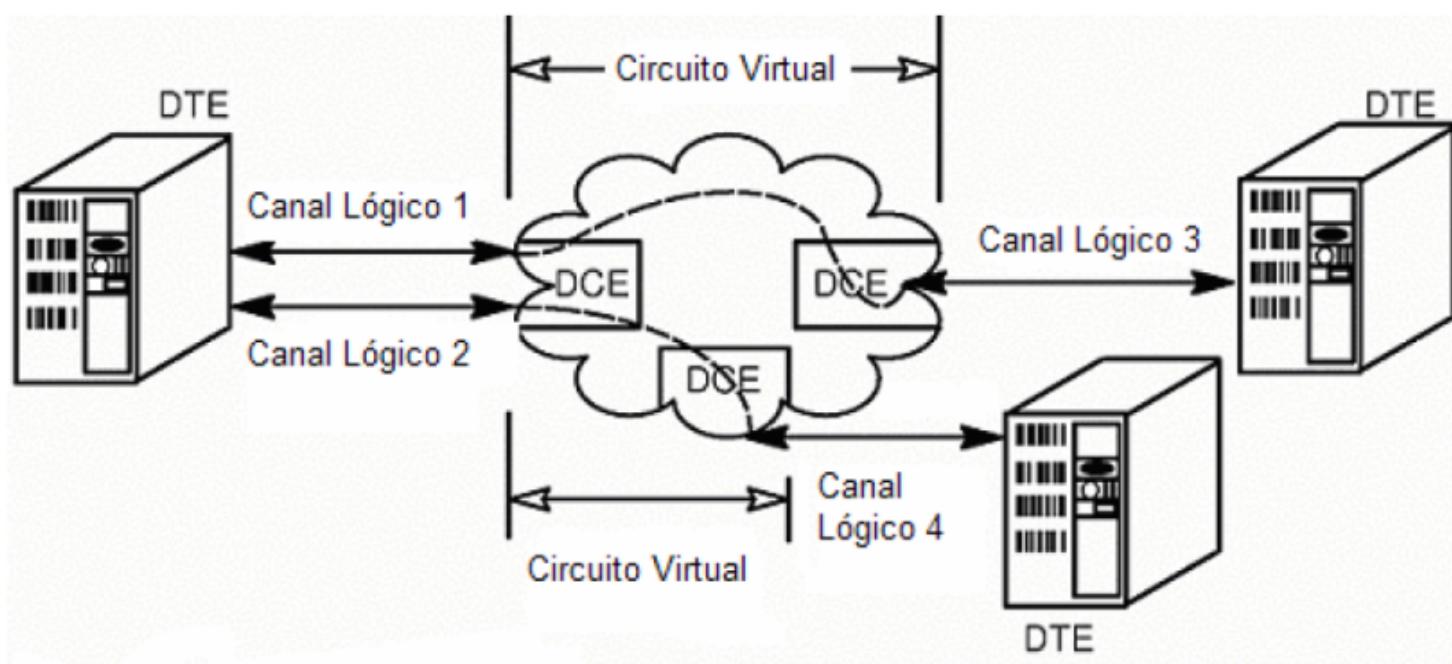
Circuitos Virtuales. Canales Lógicos

La capacidad de transferencia de datos de la línea X.25 puede estar compartida entre un número de sesiones diferentes. El número de sesiones está en función del tipo de suscripción y de las capacidades software y hardware del DTE. Cada sesión constituye lo que se llama un **círculo virtual**.

Un círculo virtual es un camino lógico bidireccional entre los sistemas local y remoto que puede tener su ruta comutada en la red, es decir, físicamente la conexión puede pasar a través de cualquier número de nodos intermedios, tales como DCEs o DSEs.

En X.25 existen dos tipos de circuitos virtuales: **circuitos virtuales conmutados** (SVC, Switched Virtual Circuit) y **circuitos virtuales permanentes** (PVC, Permanent Virtual Circuit). Los circuitos virtuales conmutados son los más habituales, también se les denomina **llamadas virtuales**. Este tipo de circuitos son conexiones temporales utilizadas para transferencias de datos esporádicas. Requieren, cada vez que dos dispositivos DTE necesitan comunicarse, que se establezca, se mantenga y se termine una conexión. Los circuitos virtuales permanentes son conexiones establecidas de forma permanente utilizadas para transferencias de datos frecuentes y no necesitan que se establezca o termine la conexión porque esta está siempre activa.

Para permitir circuitos virtuales simultáneos, ya sean permanentes o conmutados, se utilizan los llamados **canales lógicos**. Los canales lógicos son los caminos de comunicación entre un DTE y su DCE asociado. El proveedor asigna números de canales lógicos específicos, y cada número debe hacer corresponder el DTE con su DCE. El rango de números de canales lógicos válidos está entre 0 y 4095.



Cuando se utilizan SVCs no es necesario conocer el número de canal lógico en uso, el software X.25 asigna el número de canal lógico durante la fase de establecimiento de la comunicación, eligiendo el número dentro de la gama establecida mediante acuerdo con el proveedor en el momento del abono. Para PVCs si es necesario conocer el número, los PVCs están configurados de manera permanente.

Hay tres tipos de canales lógicos para SVC:

- **Unidireccionales Entrantes** : el DTE sólo puede recibir llamadas en ese canal.
- **Unidireccionales Salientes** : el DTE sólo puede iniciar llamadas en ese canal.
- **Bidireccionales** : el DTE puede iniciar y recibir llamadas en ese canal.

Si se utiliza más de un tipo de canal lógico, los números deben asignarse de acuerdo a la siguiente jerarquía de menor a mayor:

- PVCs.
- SVCs unidireccionales entrantes.
- SVCs bidireccionales.
- SVCs unidireccionales salientes.

Niveles en X.25

X.25 tiene tres niveles que se corresponden con las tres primeras capas de la arquitectura de siete niveles del modelo de referencia OSI de ISO. Estos niveles son:

- Nivel físico: describe la interfaz con el medio físico. Es similar a la capa física del modelo de referencia OSI.
- Nivel de enlace: es el responsable de la comunicación fiable entre el DTE y el DCE. Es similar a la capa de enlace del modelo de referencia OSI.
- Nivel de paquete: describe el protocolo de transferencia de datos en la red de paquetes conmutados. Se corresponde con la capa de red del modelo de referencia OSI.

Nivel Físico

El nivel físico, como en cualquier otra pila de protocolos, especifica las características mecánicas, eléctricas, funcionales y de procedimiento que son necesarias para activar, mantener y terminar una conexión física entre un DTE y un DCE.

Se implementa como un controlador de corrientes y realiza las funciones siguientes:

- Activa y desactiva los circuitos físicos mediante el uso de señales eléctricas.
- Mantiene las características de línea de la interfaz seleccionada.
- Indica fallos en las tramas de entrada (por ejemplo trama con longitud incorrecta).

La recomendación X.25 no especifica en sí misma como debe funcionar el nivel físico sino que referencia otras recomendaciones de la ITU-T, especificando cuales de ellas se pueden utilizar. Las recomendaciones especificadas son: X.21, X.21bis, X.31 y las interfaces de la serie V como V.24.

La recomendación X.21 es una interfaz de señalización digital que opera sobre 8 circuitos de intercambio. Sus características funcionales están definidas en la recomendación X.24 y sus características eléctricas en la recomendación X.27.

Los 8 circuitos de intercambio son:

Círcuito	Nombre	Dirección
G	Señal tierra	
Ga	Vuelta común del DTE	DTE -> DCE
T	Transmisión	DTE -> DCE
R	Recepción	DCE -> DTE
C	Control	DTE -> DCE
I	Indicación	DCE -> DTE
S	Temporización del elemento de señal	DCE -> DTE
B	Temporización de byte	DCE -> DTE

X.21bis define la interfaz analógica para permitir el acceso a la red digital de circuitos conmutados utilizando un circuito analógico. Proporciona procedimientos para el envío y recepción de información direccional, lo que permite al DTE establecer circuitos conmutados con otros DTE que tengan acceso a la red.

Nivel de Enlace

En el nivel de enlace se especifica el procedimiento de acceso al enlace para el intercambio de datos a través del enlace físico. El nivel de enlace garantiza una transferencia fiable de los datos entre el DTE y el DCE, transmitiendo los datos como una secuencia de tramas. Esto significa que además de proporcionar los mecanismos para la transmisión debe de proporcionar medios para informar de si los datos han alcanzado el destino correctamente, y si no es así retransmitirlo.

Las funciones realizadas en el nivel de enlace incluyen:

- Transferencia de datos de manera eficiente y ajustada en tiempo.
- Sincronismo de enlace para asegurar que el receptor está en fase con el transmisor.
- Detección y recuperación de errores de transmisión.
- Identificación de errores de procedimiento e informe a las capas superiores para su recuperación.

Hay varios protocolos que se pueden utilizar en el nivel de enlace:

- **LAPB (Link Access Protocol, Balanced)** : es el utilizado generalmente. Se deriva del protocolo HDCL y además de todas las características de éste permite formar una conexión de enlace lógico.
- **LAP (Link Access Protocol)** : es una versión temprana de LAPB y hoy día prácticamente no se utiliza.
- **LLC (Logical Link Protocol)** : es un protocolo IEEE 802 para Redes de Área Local que permite transmitir paquetes X.25 a través de un canal RAL.

En **LAPB** se describen los procedimientos para el intercambio de información entre un DTE y un DCE. Este intercambio de información puede ser a través de un único circuito físico o de varios. El funcionamiento por múltiples circuitos físicos, que es opcional, se denomina **multienlace** y es necesario si se quiere evitar que las averías del circuito interrumpan el funcionamiento de la capa de paquete.

LAPB maneja tres tipos de tramas:

- **Tramas de Información (I)** : contienen la información real que se quiere transmitir.
- **Tramas Supervisoras (S)** : se utilizan para realizar funciones de control de supervisión del enlace de datos, tales como el acuse de recibo, la petición de retransmisión y la petición de una supresión temporal de la transmisión de tramas I.
- **Tramas no Numeradas (U)** : se utilizan para proporcionar funciones adicionales de control del enlace de datos.

Cada trama, excepto las tramas U, llevan una numeración secuencial, pudiendo ser esta numeración en módulo 8, módulo 128 o módulo 32768. Los números secuenciales adoptan cíclicamente todos los valores de la gama, entre 0 y el módulo menos 1. Se tienen así tres modos de funcionamiento de X.25:

- Funcionamiento básico: en módulo 8.
- Funcionamiento ampliado: en módulo 128.
- Funcionamiento superampliado: en módulo 32768.

Ejemplos de tramas S son:

- RECEIVE READY (RR): trama de reconocimiento (ACK) indicando el número de la siguiente trama de información esperada.
- RECEIVE NOT READY (RNR): trama que indica al transmisor que detenga la transmisión debido a problemas temporales.

- REJECT (REJ): es una trama de reconocimiento negativo (NACK) que indica la petición del inicio de un proceso de recuperación por la pérdida de tramas de información.
- SELECTIVE REJECT (SREJ): es como REJ pero además permite la recuperación de errores de secuencia de tramas.

En el modo de funcionamiento básico no se pueden utilizar tramas SREJ, son obligatorias las tramas REJ. En el modo de funcionamiento superampliado es al contrario. En el modo ampliado se elige en el momento del abono el tipo de trama que se va a utilizar, REJ o SREJ.

Ejemplos de tramas U son:

- DISCONNECT (DISC): permiten anunciar a la máquina una desconexión.
- SET NORMAL RESPONSE TIME (SNRT): permite anunciar a la máquina que ha vuelto.
- UNNUMBERED ACKNOWLEDGEMENT (UA): se utiliza para el acuse de recibo y para la aceptación de instrucciones de fijación de modo.
- FRAME REJECT (FRMR): se utiliza para indicar que se ha recibido una trama de semántica imposible.

La estructura general de una trama LAPB es la siguiente:

BANDERA	DIRECCIÓN	CONTROL	DATOS	FCS	BANDERA
---------	-----------	---------	-------	-----	---------

- **Campo de bandera (8 bits)** : indica el comienzo y el final de cada trama. Está formado por la secuencia 01111110.
- **Campo de dirección (8 bits)** : contiene la dirección del DTE o DCE, identifica al receptor previsto en una trama de instrucción o al transmisor en una trama de respuesta.
- **Campo de control (8/16/64 bits)** : contiene los números de secuencia o las instrucciones y respuestas para controlar el flujo de datos entre el DTE y el DCE. En el funcionamiento en módulo 8 siempre tiene una longitud de 8 bits. En el funcionamiento ampliado y superampliado tiene una longitud de 16 y 64 bits respectivamente para tramas que contengan números secuenciales; para tramas que no contengan números secuenciales su longitud es siempre de 8 bits.
- **Campo de información (longitud variable)** : no siempre existe, pero si existe sigue al campo de control y precede al campo de verificación de trama (FCS). Cuando se transmite del DCE al DTE, si el número de bits de información que ha de insertarse en el campo de información no es múltiplo de 8, el DCE rellenará este campo de información con ceros de modo que los octetos del campo de información queden alineados. Cuando se transmita del DTE al DCE, el DTE transmitirá únicamente información alineada por octetos.
- **Campo de secuencia de verificación de trama (FCS) (16 bits)** : sirve para la comprobación de errores de transmisión. Es una variante del CRC.

Nivel de Paquete

El nivel de paquete gobierna la comunicación extremo a extremo entre los diferentes DTEs. Crea unidades de datos de red, denominados paquetes, que contienen información de control y datos de usuario. Se tienen entonces dos tipos de paquetes: paquetes de control y paquetes de datos. Cada paquete que deba transferirse a través de la interfaz DTE/DCE estará contenido en el campo de información de la capa de enlace de datos.

La capa de paquete proporciona procedimientos para manejar los siguientes servicios:

- **Círculo Virtual Conmutado (SVC)** : como ya hemos visto es una asociación temporal entre dos DTEs. Este servicio se inicia por un DTE que envía una señal de petición de llamada (CALL REQUEST) a la red y garantiza una entrega ordenada de paquetes entre dos DTEs en cualquier dirección.
- **Círculo Virtual Permanente (PVC)** : es una asociación permanente entre dos DTEs y que por tanto no requiere una acción de establecimiento o liberación de llamada.
- **Selección rápida** : es un servicio que permite a los paquetes que establecen el SVC llevar también datos.
- **Establecimiento y liberación de llamada** : son servicios requeridos por el SVC.
- **Control de flujo y manejo de errores de cada canal lógico** .

En la tabla siguiente se muestran los distintos tipos de paquetes y su utilización en los servicios SVC y PVC.

TIPO DE PAQUETE		SERVICIO	
Del DCE al DTE	Del DTE al DCE	SVC	PVC
Establecimiento y liberación de la comunicación			
Llamada entrante	Petición de llamada	X	
Comunicación establecida	Llamada aceptada	X	
Indicación de liberación	Petición de liberación	X	
Confirmación de liberación por el DCE	Confirmación de liberación por el DTE	X	
Datos e interrupción			
Datos del DCE	Datos del DTE	X	X
Interrupción por el DCE	Interrupción por el DTE	X	X
Confirmación de interrupción por el DCE	Confirmación de interrupción por el DTE	X	X
Control de flujo y reiniciación			
RR del DCE	RR del DTE	X	X
RNR del DCE	RNR del DTE	X	X
	REJ del DTE	X	X
Indicación de reiniciación	Petición de reiniciación	X	X
Confirmación de reiniciación por el DCE	Confirmación de reiniciación por el DTE	X	X
Rearranque			
Indicación de rearanque	Petición de rearanque	X	X
Confirmación de rearanque por el DCE	Confirmación de rearanque por el DTE	X	X
Diagnóstico			
Diagnóstico		X	X

Todos los tipos de paquetes contienen un encabezamiento que contiene los siguientes campos:

- **Identificador de protocolo (1 byte)** : para funcionamiento en módulo 8 y 128 no aparece en ningún tipo de dato. Para funcionamiento en módulo 32768 aparece en el primer byte de cada paquete de datos.
- **Identificador general de formato (4 bits)** : es un campo codificado que indica el formato general del resto del encabezamiento. El primer bit se utiliza para el bit calificador en paquetes de datos; para el bit de dirección en los paquetes de establecimiento y liberación de la comunicación; y se pone a 0 en el resto de paquetes. El siguiente bit se utiliza para el procedimiento de establecimiento de la comunicación y de confirmación de entrega en paquetes de datos; se pone a 0 para los demás paquetes. Los dos siguientes bits indican el modo de funcionamiento: 01 para módulo 8, 10 para módulo 128 y 11 para 32768.

- **Identificador de canal lógico (12 bits)** : está formado por el número de grupo de canales lógicos (4 bits) y el número de canal lógico (8 bits). Como en número 0 se reserva para futuros usos, los 12 bits dan un máximo de 4095 posibles números.
- **Identificador del tipo de paquete (8 bits)** : es una codificación dependiente del tipo de paquete y modo de funcionamiento.

El formato del resto del paquete depende del tipo.

Establecimiento y Liberación de la Comunicación

Cuando un DTE A quiere establecer una comunicación con un DTE B, A indica una petición de llamada transfiriendo un paquete PETICIÓN DE LLAMADA (CALL REQUEST) a su DCE por la interfaz DTE/DCE. El canal lógico seleccionado por el DTE está en estado "*DTE en espera*". El paquete incluye la dirección del DTE llamado y puede incluir también la dirección del DTE llamante.

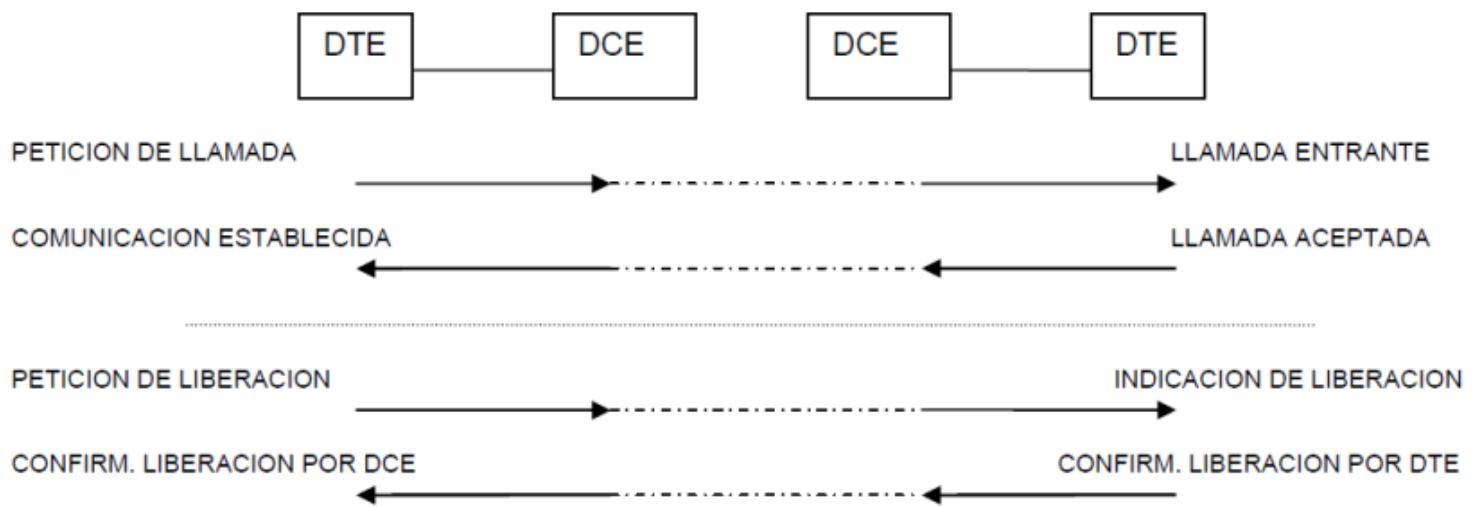
Cuando el paquete llega al DCE asociado a B, el DCE indica que hay una llamada entrante transfiriendo por la interfaz DTE/DCE, un paquete de LLAMADA ENTRANTE (INCOMING CALL). Esto hace pasar al canal lógico al estado "*DCE en espera*". El paquete incluye la dirección del DTE llamante.

B indicará su aceptación de la llamada transfiriendo un paquete de LLAMADA ACEPTADA (CALL ACCEPTED) que especifique el mismo canal lógico que el paquete LLAMADA ENTRANTE. Esto hace pasar al canal lógico especificado al estado "*transferencia de datos*". Cuando el paquete llega al DCE asociado a A transfiere por la interfaz DTE/DCE un paquete de COMUNICACIÓN ESTABLECIDA (CALL CONNECTED). La recepción de este paquete en A con el mismo canal lógico que el establecido en el paquete PETICIÓN DE LLAMADA, le indica que la llamada a sido aceptada por el DTE llamado B y el canal lógico pasa al estado "*transferencia de datos*". Cuando el DTE A recibe el paquete de COMUNICACIÓN ESTABLECIDA se establece el Circuito Virtual.

El DTE determina el número de canal lógico para la petición de llamadas y el DCE para las llamadas entrantes. Cuando un DTE y un DCE transfieren simultáneamente un paquete PETICIÓN DE LLAMADA y un paquete de LLAMADA ENTRANTE con el mismo número de canal lógico se produce una **colisión de llamadas**. En este caso en la norma X.25 determina que se dar curso a la petición de llamada y se cancela la llamada entrante.

En cualquier momento cualquiera de los DTEs puede indicar la liberación de la llamada transfiriendo por la interfaz DTE/DCE un paquete de PETICIÓN DE LIBERACIÓN (CLEAR REQUEST). El canal lógico está en este caso en estado "*petición de liberación por el DTE*". El DCE del otro extremo indica la liberación transfiriendo un paquete de INDICACIÓN DE LIBERACIÓN (CLEAR INDICATION). El canal lógico está entonces en el estado de "*indicación de liberación por el DCE*". El DTE responde transfiriendo un paquete de CONFIRMACIÓN DE LIBERACIÓN POR EL DTE, quedando el canal lógico en estado "*preparado*".

Cuando un DTE y un DCE transfieren simultáneamente un paquete PETICIÓN DE LIBERACIÓN y un paquete de INDICACIÓN DE LIBERACIÓN con el mismo número de canal lógico se produce una **colisión de liberaciones**. En este caso el DCE considera completada la liberación.



Frame Relay

Frame Relay (FR) es una tecnología WAN de alto rendimiento que opera en los niveles 1 y 2 (nivel físico y nivel de enlace de datos) del modelo de referencia OSI. Diseñado originalmente para operar con las interfaces de RDSI, en la actualidad se usa también sobre otras interfaces de red. FR es un ejemplo de tecnología de conmutación de paquetes que se considera como una evolución de X.25 y un paso de transición hacia ATM.

Técnicamente es una tecnología diseñada para transmitir y distribuir datos a alta velocidad en unidades de longitud variable denominadas tramas.

FR consigue rendimientos muy superiores a los de X.25 al eliminar la mayoría de los controles de errores, con lo que se disminuye el trasiego de información, pero esto hace que se necesiten líneas que garanticen un mínimo de fiabilidad. En FR se soportan velocidades de transmisión de hasta 45 Mbps aunque las implementaciones típicas no pasan de 1.5/2 Mbps.

Se adapta muy bien a la interconexión de LANs y al tráfico en ráfagas que éstas presentan. Utiliza técnicas de multiplexación estadística que permiten que la conexión virtual que tiene tráfico en un momento dado utilice parte del ancho de banda que no está siendo utilizado en las otras conexiones con las que comparte el mismo enlace físico. La multiplexación estadística proporciona a la red ancho de banda bajo demanda: la red es capaz de obtener el ancho de banda que necesita cuando lo necesita sin tener que reservarlo por adelantado y mantenerlo sin usar hasta que sea requerido.

FR nació como un estándar de facto en 1990 resultado del acuerdo entre un grupo de fabricantes de equipos de telecomunicaciones. Surgió como una solución transitoria para cubrir necesidades del mercado no satisfechas hasta ese momento, pero ha logrado una gran aceptación y en la actualidad juega un papel importante en la interconexión de redes.

FR se desarrolló básicamente por las siguientes motivaciones:

- Demanda de mayores velocidades: esta necesidad venía provocada por el aumento creciente del tráfico de datos a ráfagas, la proliferación de las LAN, la arquitectura cliente/servidor y la integración LAN-WAN.
- Disponibilidad de mejores medios de transmisión: protocolos como X.25 o SNA se desarrollaron con una cierta complejidad para introducir procedimientos de corrección de errores, uno de los factores principales en la limitación del ancho de banda. Con la aparición de las líneas digitales, sobre todo la fibra óptica, se puede

eliminar la mayor parte de estos procedimientos al ser medios con una tasa muy baja de errores.

- Mayor capacidad de procesamiento de los equipos conectados a la red: equipos como PCs o estaciones de trabajo demandan mayores velocidades a la vez que tienen capacidades de procesamiento que facilitan la gestión de tramas. Además existen ya familias de protocolos como TCP/IP capaces de gestionar el control de errores y secuenciación de tramas.

Existen tres estándares FR: el estándar del FR Forum (asociación de fabricantes entre los que están Cisco, DEC y Nortel), el estándar ANSI y el estándar de la ITU-T. Estas tres fuentes de normas no son siempre coincidentes, lo que no ocurría en X.25.

Para la transmisión de datos entre estaciones finales se utiliza el protocolo Q.992, que es una versión mejorada del protocolo LAP-D utilizado en RDSI. En FR sólo se utilizan las funciones básicas de Q.922:

- Delimitación, alineamiento y transparencia de tramas por medio de las banderas típicas de la familia de protocolos HDCL.
- Alineamiento de los límites de la trama.
- Detección de errores de transmisión mediante un campo FCS incluido en la trama.
- Multiplexación y desmultiplexación de tramas mediante el campo de direcciones incluido en la trama.
- Control de congestión.

Circuitos Virtuales e Interfaces Frame Relay

Frame Relay es un sistema orientado a conexión en el sentido de que antes de establecer la comunicación entre dos o más puntos se requiere previamente haber definido un camino o ruta por la cual tenga lugar la comunicación, es decir, existe una conexión definida entre cada par de dispositivos y estas conexiones están asociadas con un identificador de conexión.

De manera similar a X.25 este servicio se implementa por medio de un circuito virtual Frame Relay. Los circuitos virtuales ofrecen una trayectoria de comunicación bidireccional de un dispositivo DTE a otro y se identifican de manera única por medio del **Identificador de Conexiones de Enlace de Datos (DLCI, Data-Link Connection Identifier)**. Además se pueden multiplexar muchos circuitos virtuales en un único circuito físico.

Los DLCI generalmente son asignados por el proveedor del servicio y su valor tiene significación local, es decir, los DLCI son únicos en la LAN pero no necesariamente en la WAN Frame Relay.

Al igual que en X.25 se tienen circuitos virtuales permanentes y conmutados. En FR lo más común son los PVC, que una vez programados permanecen en funcionamiento (se les use o no) hasta que se les desconecta.

Tanto para los PVCs como para los SVCs se distinguen dos interfaces:

- **UNI (User-to-Network Interface)** : interfaz de usuario red. La interfaz UNI se establece entre el dispositivo de acceso a la red del usuario y un conmutador de la red.
- **NNI (Network-to-Network Interface)** : interfaz red a red. La interfaz NNI se establece entre dos conmutadores que pueden ser de la misma o de diferentes redes Frame Relay.

En Frame Relay Forum (FRF) recomienda las siguientes interfaces físicas en la UNI:

- ANSI T1.403: interfaz a 1.5 Mbps.

- UIT-T V35: interfaz full-duplex a 56 o 64 Mbps.
- UIT-T G.703: interfaz a 2 Mbps.
- UIT-T X.21: interfaz síncrona.

Una sesión de comunicación a través de un SVC consta de los siguientes estados operacionales:

- **Establecimiento de llamada** : se establece el circuito virtual entre dos dispositivos DTE.
- **Transferencia de datos** : los datos se transmiten entre los dispositivos DTE a través del circuito virtual.
- **Desocupado (Idle)** : la conexión entre los dispositivos DTE permanece activa pero no hay transferencia de datos. Si un SVC permanece durante un determinado periodo de tiempo en estado desocupado, la llamada puede darse por terminada.
- **Finalización de llamada** : se da por finalizado el circuito virtual entre los dispositivos DTE.

En los PVC, al ser conexiones establecidas de forma permanente, no se requiere ni el establecimiento de llamada ni la finalización. Operan siempre en los estados transferencia de datos o desocupado.

Control de Congestión y Control de Tráfico

Frame Relay, más que un control de flujo explícito por cada circuito virtual, implementa sencillos mecanismos de notificación de congestión. En general Frame Relay se implementa sobre medios de transmisión de red fiables para no sacrificar la integridad de los datos, el control de flujo se puede realizar por medio de los protocolos de las capas superiores.

En FR se tienen dos mecanismos de notificación de congestión: **FECN** (Forward-Explicit Congestion Notification) y **BECN** (Backward-Explicit Congestion Notification). Tanto FECN como BECN se implementan mediante los bits del subcampo control de congestión del campo direcciones de la trama FR. Este campo de control de congestión está formado por tres bits: el bit FECN, el bit BECN y el bit DE (Discard Eligibility).

El **mecanismo FECN** se inicia en el momento en que un DTE envía tramas a la red. Si la red está saturada, los DCE fijan el valor del bit FECN a 1. Cuando las tramas llegan al DTE destino, el bit FECN activado indica que en su trayectoria del origen al destino hubo problemas de congestión. El DTE puede enviar esta información a los protocolos de las capas superiores para su procesamiento.

En el **mecanismo BECN** los DCEs establecen el valor del bit BECN a 1 en aquellas tramas que viajan en sentido opuesto a las tramas con bit FECN igual a 1. Esto permite al DTE receptor saber que una trayectoria específica en la red está saturada. Posteriormente el dispositivo DTE envía información a los protocolos de las capas superiores para su procesamiento.

Dependiendo de la implementación, la indicación de congestión puede ser ignorada o puede que se inicien los procedimientos de control de flujo.

Los DTEs pueden fijar el valor del **bit DE** de una trama a 1 para indicar que esta trama tiene una importancia menor respecto a otras tramas. En caso de congestión de la red DCE descartaran las tramas con el bit DE a 1 antes de descartar aquellas que no la tienen. Con esto se disminuye la probabilidad de que se eliminen datos críticos durante el blindaje de saturación.

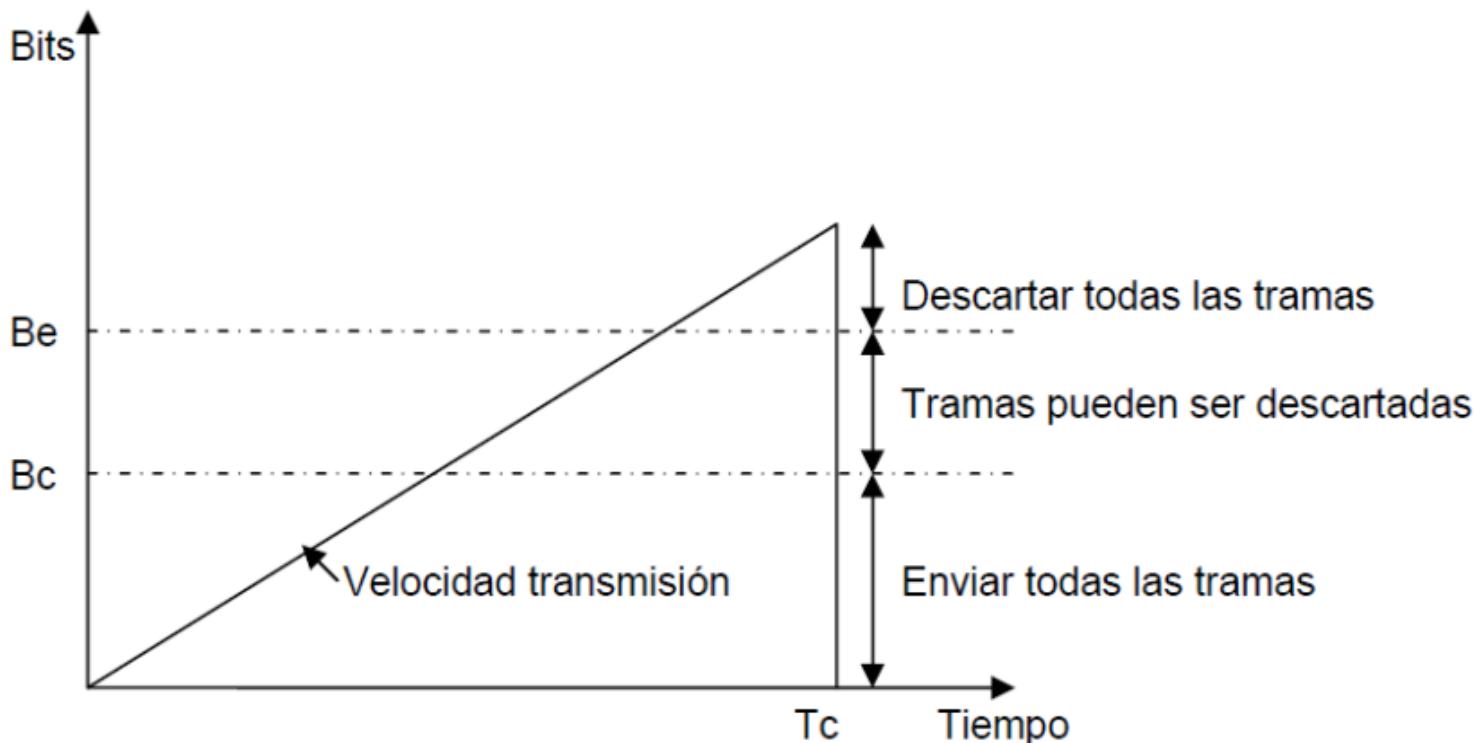
El control de tráfico en Frame Relay se basa en la especificación de varios parámetros de usuario:

- **CIR** (Committed Information Rate): es el caudal medio garantizado que la red se compromete a dar en una conexión.
- **Bc** (Committed Burst Size): máxima cantidad de bits que la red se compromete a enviar, en condiciones normales, en un intervalo de tiempo definido **Tc**. Este periodo Tc es acordado entre la red y el usuario y es por tanto conocido por ambos.
- **EIR** (Excess Information Rate): especifica un caudal adicional que la red no debería sobrepasar nunca ya que las tramas recibidas por encima de este valor serán descartadas directamente por el conmutador.
- **Be** (Excess Burst Size): máxima cantidad de bits que se le permite al usuario sobrepasar Bc durante el periodo Tc.

Cada PVC y SVC tiene asignado un valor CIR que indica la capacidad de transmisión que ha sido acordada para el servicio. La relación entre estos parámetros es $CIR = Bc/Tc$ y $EIR = Be/Tc$.

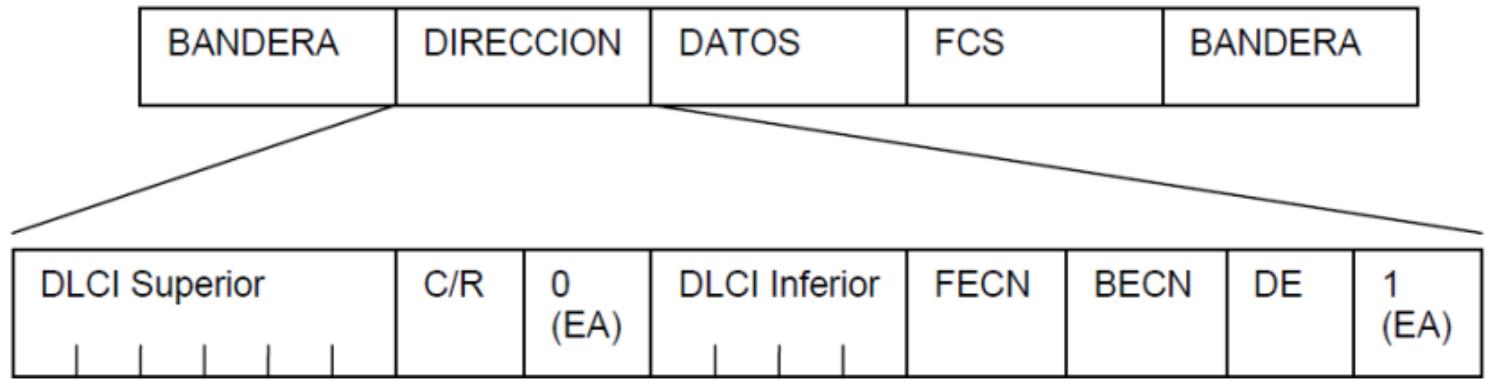
Aunque la red no puede impedir que un usuario exceda el valor CIR sí incluye funciones de aviso de situaciones problemáticas y de recomendación de reducción de la velocidad de transmisión. Si la velocidad de transmisión es inferior al CIR todo el tráfico es cursado con garantías. Si la velocidad supera este valor las tramas se marcan como descartables activando el bit DE y serán transportadas por la red con una política "best effort", llegarán al destino si durante el camino no se encuentran con congestiones importantes. Si se sobrepasa el valor EIR, todas las tramas recibidas por encima de ese valor serán descartadas.

En la siguiente figura se esquematiza este mecanismo de control de tráfico.



Formato de Trama Frame Relay

El formato de una trama FR es el siguiente:



Formato de trama FR

- **Bandera** (8 bits): delimitan el comienzo y el final de cada trama. Su valor es siempre el mismo 01111110 (7E en hexadecimal).
- **Dirección** (16 bits): contiene la información siguiente:
 - **DLCI** (10 bits): como ya se ha comentado identifica de manera única al circuito virtual. No todos los valores de DLCIs pueden ser asignados a conexiones virtuales para datos de usuarios, sino los valores del 16 al 1007. Otros DLCIs están reservados para señalización (como lo son los valores 0: reservado para la señalización de control de llamada y 1023: reservado para la Interfaz de Gestión Local) y otros valores están reservados para definiciones futuras como los valores del 1 al 15 y del 108 al 1022.
 - **EA** (2 bits): bit de dirección extendida. Se utiliza para indicar el último byte del campo de dirección mediante un valor 1. Aunque todas las implementaciones actuales de Frame Relay utilizan un DLCI de dos bytes, esta característica permite que en el futuro se utilicen DLCIs más largos. El octavo bit de cada byte del campo de direcciones se utiliza para indicar el EA.
 - **C/R** (1 bit): este bit comando/respueta está sin definir.
 - **Control de congestión** (3 bits): consta de los bits FECN, BECN y DE.
- **Datos** (longitud variable): contiene información encapsulada de las capas superiores. Contiene los datos de usuario o carga útil, que tiene una longitud variable. Pudiendo llegar hasta los 8.188 bytes.
- **FCS** (16 bits): el campo de secuencia de verificación de trama sirve para la comprobación de errores de transmisión.

LMI, Interfaz de Administración Local

LMI (Local Management Interface) es una interfaz que fue definida por los primeros fabricantes de sistemas FR (Cisco Systems, StrataCom y DEC entre otros) para ofrecer un conjunto de extensiones a la especificación FR básica. LMI ha servido de modelo para los estándares de señalización de Frame Relay de otros organismos de estandarización.

Las principales extensiones de LMI son direccionamiento global, mensajes de status de circuito virtual y multidifusión.

El **direccionamiento global LMI** da a los valores DLCI un significado global más que local. Los convierten en direcciones DTE únicas en la WAN Frame Relay. Esto agrega funcionalidad y facilita la administración de las redes Frame Relay al permitir, por ejemplo, que las interfaces de red individuales y los nodos terminales conectados a ellos se puedan identificar por medio de técnicas estándares de descubrimiento y resolución de direcciones. Además, toda la red Frame Relay aparece como una típica LAN.

Los **mensajes de status de circuitos virtuales LMI** permiten la sincronización entre los DTEs y DCEs y la comunicación de su estado. Los mensajes de status se utilizan para

informar de manera periódica del estado de los PVCs con lo que se previene el envío de datos a agujeros negros (esto es, a través de PVCs inexistentes).

La **multidifusión LMI** permite que se asignen grupos de multidifusión. Con la multidifusión se ahorra ancho de banda pues permite que los mensajes sobre la resolución de direcciones y de actualizaciones de encaminamiento sean enviados solamente a grupos específicos de routers.

ATM

La tecnología de Modo de Transferencia Asincrónica (ATM), está basada en los estudios del Grupo de Estudio XVIII de la ITU-T para el desarrollo de la Red Digital de Servicios Integrados de Banda Ancha (B-ISDN) para la transferencia de voz, vídeo y datos a altas velocidades a través de la red pública.

En 1991 se crea el Forum ATM (fundado por Cisco Systems, NET/ADAPTIVE, Nortern Telecom y Sprint) que ha jugado un papel importante en el desarrollo de la tecnología ATM. En la actualidad con ATM es posible transferir voz, vídeo y datos a través de redes privadas y a través de redes públicas. ATM continúa evolucionando con varios grupos estándares que tratan de finalizar las especificaciones que permitan la interoperabilidad entre los equipos de diferentes vendedores en las industrias de redes públicas y privadas.

ATM es una tecnología de conmutación de paquetes orientada a conexión que como tal crea circuitos virtuales entre los sistemas que desean intercambiar información. Estos circuitos se denominan **canales virtuales (VC, Virtual Channel)** en el estándar. Los nodos terminales de las redes ATM se denominan hosts y los nodos intermedios de encaminamiento se denominan conmutadores (de forma análoga a X.25 o Frame Relay). Los conmutadores ATM son siempre equipos de comunicaciones especializados y de elevadas prestaciones, nunca ordenadores de propósito general.

En ATM existen tanto PVCs como SVCs. Los PVCs se configuran de manera estática en los conmutadores. Un PVC está establecido siempre que estén operativos los conmutadores por los que pasa y los enlaces que los unen, es decir siempre que la red está operativa. Los SVCs se crean y destruyen dinámicamente, según se necesita. El protocolo utilizado para establecer SVCs en ATM es el Q.2931, y está basado en el Q.931 utilizado en la señalización de RDSI.

Una red (o subred) ATM está formada por un conjunto de conmutadores unidos entre sí por líneas punto a punto de alta velocidad, normalmente SONET/SDH de 155,52 Mbps, aunque también existen interfaces de velocidades inferiores. La interfaz que conecta los hosts con la subred es la **UNI** (User-Network Interface), y la que comunica los conmutadores entre sí es la **NNI** (Network-Network Interface).

En ATM se agrupan los VCs entre dos nodos terminales en los denominados **caminos o trayectos virtuales (VP, Virtual Path)**. Tanto los VCs como los VPs se numeran para su identificación. Para establecer una comunicación entre dos nodos es preciso especificar el número de VP y de VC.

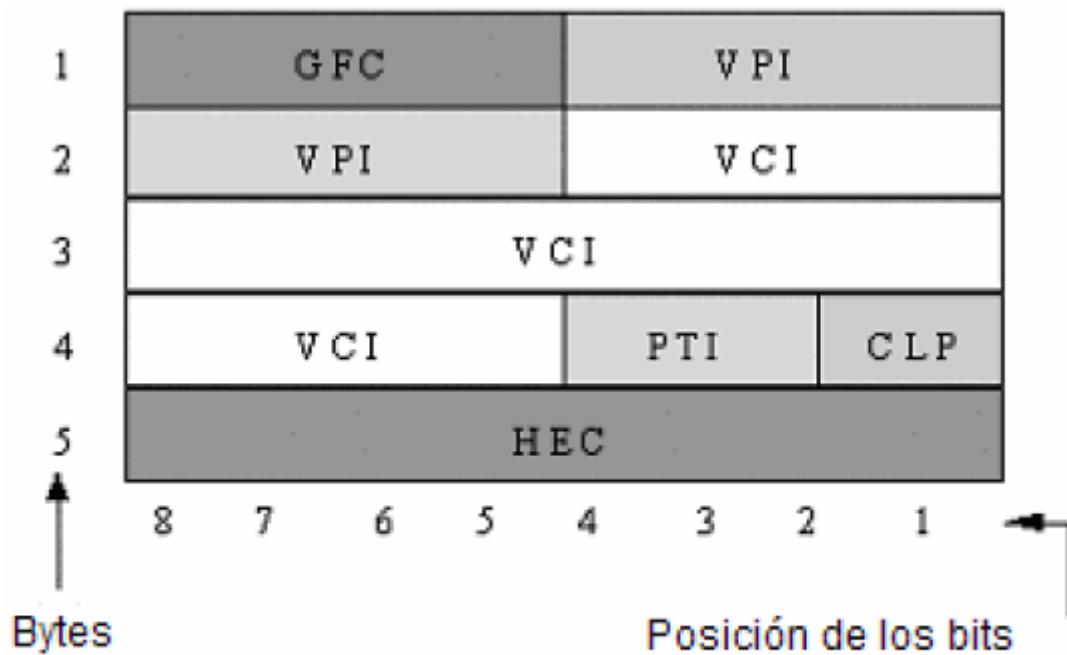
Cuando un nodo desea establecer un VC con otro ha de enviar un mensaje solicitando la conexión por el VC reservado para señalización, que por convenio es el VP 0 VC 5.

Los estándares de ATM han definido paquetes de tamaño fijo llamados **celdas con una longitud de 53 bytes**. Una celda ATM consta de dos partes: la carga útil (payload) de 48 bytes que transporta la información generada por el emisor o transmisor, y la cabecera de 5 bytes que contiene la información necesaria para la transferencia de la celda.

Formato de Cabecera de la Celda ATM

En los estándares ATM se definen dos formatos de cabecera: el formato de la cabecera de la UNI y el formato de la cabecera de la NNI. La recomendación I.361 de la ITU-T es la base de estas definiciones, con clarificaciones más amplias dadas en las especificaciones ANSI T1.627 y el Forum ATM UNI o B-ICI.

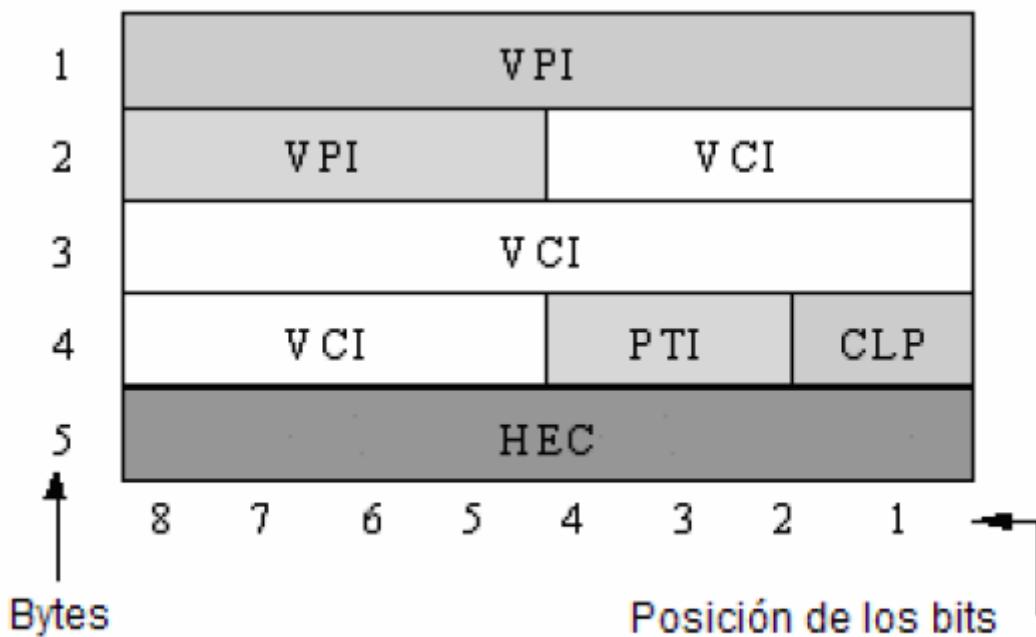
El formato de la cabecera UNI es el siguiente:



Formato Cabecera UNI

- **GFC, control de flujo genérico (4 bits)** : este campo generalmente no se utiliza y se le asigna un valor por defecto, pero puede ser utilizado para proporcionar estaciones locales como por ejemplo la identificación de estaciones que comparten una única interfaz ATM.
 - **VPI, identificador de camino virtual (8 bits)** : identifica el camino virtual por el que debe circular una celda.
 - **VCI, identificador de canal virtual (16 bits)** : identifica el canal virtual por el que debe circular la celda dentro del VP especificado por el VPI. El VCI junto con el VCI identifica el próximo destino de la celda que pasa a través de una serie de commutadores ATM en su trayecto hasta el destino final.
 - **PTI, tipo de carga (3 bits)** : el primer bit indica si la celda contiene datos de usuario o datos de control. Si la celda contiene datos de usuario, el segundo bit indica si se detecta o no congestión; y el tercer bit indica si la celda es la última en una serie de celdas que representan a una única trama AAL5.
 - **CLP, prioridad de pérdida de celda (1 bit)** : indica si la celda debería o no descartarse en el caso de que se encuentre en congestión extrema cuando se mueve a través de la red.
 - **HEC, control de error de cabecera (8 bits)** : es un CRC que suministra la información de verificación de error de la cabecera.

La cabecera NNI se diferencia de la cabecera UNI en que no existe el campo GFC y el campo VPI ocupa 12 bits.



Formato de cabecera NNI

Categorías de Servicio

Para poder satisfacer una amplia gama de necesidades se han definido en ATM las llamadas categorías de servicio. Cada una de ellas da al usuario un nivel de garantía diferente respecto a la disponibilidad de los recursos de red solicitados.

Se han definido cuatro categorías de servicio: CBR (Constant Bit Rate), VBR (Variable Bit Rate), ABR (Available Bit Rate) y UBR (Unspecified Bit Rate).

La categoría de servicio **CBR** garantiza una capacidad determinada y constante, que está continuamente disponible durante el tiempo de vida de la conexión, con independencia de la utilización que hagan de la red otros usuarios. Este servicio es el más sencillo de implementar y el más seguro de todos, ya que la red reserva la capacidad solicitada en todo el trayecto de forma estática. No se realiza ningún tipo de control de congestión, ya que se supone que ésta no puede ocurrir. El servicio CBR es equivalente a una línea dedicada punto a punto.

La categoría de servicio **VBR** está pensada para un tráfico a ráfagas. El usuario especifica un caudal medio pero puede utilizar ocasionalmente caudales superiores en función de sus necesidades y del estado de la red. Esto da mayor flexibilidad al permitir ajustar el caudal a las necesidades medias. En algunos servicios VBR el tráfico excedente sale marcado con el bit CLP. Desde el punto de vista de la red, VBR tiene una complejidad superior a CBR.

VBR tiene dos modalidades definidas por el Forum ATM: **RT-VBR** (Real-Time Variable Bit Rate), con requerimientos de bajo retardo y jitter para aplicaciones en tiempo real (videoconferencia, vídeo bajo demanda, etc.); y **NRT-VBR** (Non-REAL-TIME Variable Bit Rate) para aplicaciones en las que el control del retardo no es tan importante, como por ejemplo la transferencia de ficheros.

La categoría **ABR**, también pensada para tráfico a ráfagas, es de todas las categorías de servicio que ofrece ATM la que más se parece a Frame Relay. Permite establecer un ancho de banda mínimo garantizado y fijar un valor máximo orientativo. ABR es la única categoría de servicio ATM en la que se prevé que la red suministre control de flujo al emisor para que reduzca el ritmo de transmisión en caso de congestión, lo que hace de

ABR apropiada para tráfico de datos pero menos apropiada para aplicaciones isócronas. Debido a su funcionalidad ABR es la categoría de servicio más compleja de implementar.

La categoría de servicio UBR puede considerarse la de más baja calidad. No existe ningún tipo de garantía en cuanto al retardo o ancho de banda, y tampoco se informa al emisor en caso de congestión. UBR utiliza la capacidad sobrante en la red de las demás categorías de servicio. Puede utilizarse para enviar tráfico IP cuando el costo sea el factor principal y la calidad de servicio no sea importante.

El ATM Forum ha definido una variante del servicio UBR denominada UBR+. Añade al servicio UBR la posibilidad de especificar una capacidad mínima requerida, lo que la hace similar a ABR pero sin el control de congestión.

Calidad de Servicio y Descriptores de Tráfico

Una de las grandes virtudes de ATM es la posibilidad de establecer una Calidad de Servicio (QoS, Quality of Service) garantizada. En las redes ATM se pueden establecer una larga serie de parámetros que definen los niveles mínimos de calidad que el operador debe ofrecer al usuario para cada una de las categorías de servicio mencionadas en el punto anterior. Estos parámetros se pueden clasificar en dos grupos: parámetros de tráfico y parámetros de QoS. No todos los parámetros tienen sentido en todas las categorías de tráfico.

Los parámetros de tráfico son los siguientes:

- **MCR** (Minimum Cell Rate): velocidad mínima que se considera aceptable para establecer el circuito ATM.
- **PCR** (Peak Cell Rate) y **CDVT** (Cell Delay Variation Tolerance): máximo caudal que permite el VC y tolerancia (pequeña) respecto a este caudal.
- **SCR** (Sustainable Cell Rate) y **BT** (Burst Tolerance): caudal medio máximo permitido y tolerancia a ráfagas (grande) respecto a este caudal.

Los parámetros de Calidad de Servicio son los siguientes:

- **MCTD** (Maximum Cell Transfer Delay): es el retardo máximo permitido, es decir, el tiempo máximo que puede tardar la red en transmitir una celda de un extremo a otro del circuito.
- **Peak-to-Peak CDV** (Peak To Peak Cell Delay Variation): es el jitter o fluctuación máxima que se podrá producir en el retardo de las celdas.
- **CLR** (Cell Loss Ratio): es el porcentaje máximo aceptable de celdas que la red puede descartar debido a congestión. Cuando una celda es entregada en el destino con un retardo superior a MCTD se considera una celda perdida.

No todos los parámetros son aplicables a todas las categorías de servicio. Por ejemplo en un servicio CBR se especifica PCR, pero no SCR ni MCR. En un servicio UBR no se especifica ningún parámetro.

En la siguiente tabla se muestran los parámetros que se especifican normalmente en cada categoría de servicio:

	CBR	rt-VBR	nrt-VBR	ABR	UBR+	UBR
MCR	NO	NO	NO	SI	SI	NO
PCR/CDVT	SI	SI	SI	NO	NO	NO
SCR/BT	NO	SI	SI	NO	NO	NO
MCTD	SI	SI	NO	SI	NO	NO
Pk-t-Pk CDV	SI	SI	NO	NO	NO	NO
CLR	SI	SI	SI	SI	NO	NO

Modelo de Capas de ATM

El modelo de capas de ATM está formado por tres capas: capa física, capa ATM y capa de adaptación ATM.

Capa Física

La Capa Física controla la transmisión y recepción de bits en el medio físico y mantiene el rastro de los límites de las celdas y de los paquetes de celdas dentro del tipo de trama apropiado al medio físico utilizado.

Está dividida en dos partes: subcapa dependiente del medio físico y subcapa de convergencia de transmisión.

La **Subcapa Dependiente del Medio Físico (PMD)** es responsable de enviar y recibir un flujo constante de bits, junto con información de temporización con el fin de sincronizar la transmisión y la recepción.

La **Subcapa de Convergencia de Transmisión (TC)** es la responsable de:

- Delimitación de Celdas: mantiene los límites de las celdas ATM.
- Generación y verificación del HEC: genera y chequea el código de control de error de cabecera para garantizar datos válidos.
- Desacoplamiento de velocidad de la celda: inserta o suprime celdas ATM no asignadas para adaptar a la velocidad válida de celdas ATM la capacidad de carga útil del sistema de transmisión.
- Adaptación a la trama de transmisión: empaqueta celdas ATM en tramas aceptables para su implementación en un medio físico en particular.
- Generación y recuperación de tramas de transmisión: genera y mantiene la estructura apropiada de la trama de la capa física.

Capa ATM

La Capa ATM es responsable del establecimiento de las conexiones y del paso de las celdas a través de la red ATM, para lo que utiliza la información contenida en la cabecera de cada celda ATM. Aquí es donde aparecen los conceptos de VCs y VPs vistos anteriormente.

Entre las funciones de esta capa se incluyen:

- Control Genérico de Flujo.
- Generación/Extracción del encabezado de la celda.
- Enrutamiento de las celdas basado en los VPI/VCI de la celda.
- Detección de errores basado en el campo HEC.
- Multiplexación y Demultiplexación de celdas.

Capa de Adaptación ATM (AAL)

Dentro del modelo ATM la capa que se ocupa de la comunicación host-host, y que por tanto podemos considerar de transporte, es la denominada Capa de Adaptación ATM (AAL, ATM Adaptation Layer).

La ITU-T define la capa AAL en la recomendación I.363. Esta recomendación ha sido fruto de diversos compromisos y reajustes sobre la marcha.

Dado que el objetivo de la capa AAL es adaptar diversos tipos de tráfico para su transporte sobre redes ATM, la ITU-T empezó estudiando y clasificando las clases de tráfico que podían tener cierto interés. Desde el punto de vista de la ITU-T los parámetros

relevantes para esa clasificación eran tres: tiempo real o no tiempo real (tráfico isócrono o asíncrono); caudal de tráfico constante o variable; y servicio orientado a conexión o no orientado a conexión. Considerando entonces cuatro clases de tráfico:

- **Clase A** : en tiempo real con caudal de tráfico constante y servicio orientado a la conexión.
- **Clase B** : en tiempo real con caudal de tráfico variable y servicio orientado a la conexión.
- **Clase C** : no en tiempo real con caudal de tráfico variable y servicio orientado a la conexión.
- **Clase D** : no en tiempo real con caudal de tráfico variable y servicio no orientado a la conexión.

Para las cuatro clases descritas se definieron inicialmente cuatro protocolos denominados de AAL1 a AAL4 para las cuatro clases descritas. Posteriormente se observó que los requerimientos de los protocolos AAL3 y AAL4 eran similares y fueron agrupados en un protocolo conjunto AAL3/4. Las empresas fabricantes de equipos informáticos (conmutadores y adaptadores ATM), que se incorporaron tarde al proceso de estandarización de los protocolos AAL decidieron crear un nuevo protocolo que denominaron AAL5, para transportar la misma clase de datos que AAL3/4 pero de forma más eficiente.

Se tienen entonces 5 tipos de protocolos AAL:

- **AAL1** : soporta tráfico de clase A normalmente con una categoría de servicio CBR. El AAL1 es apropiado para soportar tráfico de voz y tráfico de vídeo no comprimido. Garantiza un mínimo retardo, un jitter pequeño y un reducido overhead de proceso y de información de control.
- **AAL2** : es el estándar de protocolo utilizado para soportar el tráfico de clase B, generalmente con una categoría de servicio rt-VBR.
- **AAL3/4** : soporta tráfico de clases C y D y puede utilizar cualquiera de las categorías de servicio, aunque la más utilizada es VBR. Es adecuado para tráfico de datos sensibles a pérdidas de celdas, pero no a retardos.
- **AAL5** : soporta el transporte del tráfico de clase C y al igual que la clase anterior utiliza generalmente VBR, aunque se puede usar cualquiera de las categorías de servicio.

La capa AAL está compuesta de dos subcapas:

- **Subcapa de Segmentación y Reensamblado (SAR, Segmentation And Reassembly)** : es la subcapa inferior y se ocupa, como indica su nombre, de crear en el emisor las celdas a partir de los datos recibidos de la subcapa superior, y de reconstruir en el receptor los datos originales a partir de las celdas recibidas.
- **Subcapa de Convergencia (CS, Convergence Sublayer)** : es la subcapa superior y actúa de interfaz entre la capa AAL y la aplicación. Esta subcapa a su vez se subdivide en otras dos subcapas:
 - **Subcapa de Convergencia de Parte Común (CPCS, Common Part Convergence Sublayer)** : constituye la parte baja de la CS. No depende de la aplicación, de ahí su nombre, pero sí depende del tipo de protocolo AAL utilizado.
 - **Subcapa de Convergencia Específica del Servicio (SSCS, Service Specific Convergence Sublayer)** : representa la parte alta de la CS y es específica de la aplicación. Esta subcapa puede no estar presente, puede ser nula y hasta ahora está definida para Frame Relay y para SMDS. No se requiere para IP, pues IP es soportado directamente por la CPCS.

Integración de Voz y Datos

El concepto básico para la integración de voz y datos es relativamente simple: se trata de transformar la voz en paquetes de información que puedan transmitirse por una red de commutación de paquetes. A lo largo de los últimos años el avance tecnológico ha creado un entorno que posibilita la transmisión de voz sobre este tipo de redes. Entre los factores que han posibilitado este desarrollo se encuentran:

- Técnicas avanzadas de digitalización de voz.
- Protocolos de transmisión en tiempo real.
- Nuevos mecanismos de control y priorización del tráfico.
- Nuevos estándares que permiten la calidad de servicio sobre redes de paquetes.

Uniendo a lo anterior el espectacular desarrollo de Internet, junto al ahorro que el uso de este tipo de tecnologías trae consigo, se ha llegado a la situación de considerar la integración de voz y datos un tema estratégico para muchas Organizaciones.

Las tecnologías para el transporte de voz en forma de paquetes se enfrentan a una serie de retos:

- Retardo. Para proporcionar una calidad de servicio aceptable el retardo inducido en la red debe ser minimizado. El retardo deteriora la calidad de la voz y provoca interrupciones en una conversación normal extremo a extremo.
- Supresión de silencios. En el flujo de una conversación normal existen pausas y periodos de silencio. Esta característica se puede utilizar para el ahorro de ancho de banda mediante la parada de la transmisión de paquetes en estos periodos de silencio.
- Señalización. La señalización se refiere al uso eficiente de los recursos para la transferencia de información de control.

Estos tres aspectos se ven afectados a su vez por las características de la red de transporte subyacente que puede ser FR, IP o ATM. Cuando la red de transporte es una red FR se habla de VoFR (Voice over Frame Relay), si es una red IP se habla de VoIP (Voice over IP) y si es una red ATM de VoATM (Voice over ATM).

VoFR

VoFR representa una tecnología relativamente estable para la integración de voz y datos. El Forum FR ha publicado dos especificaciones para permitir la interoperatividad entre productos VoFR de diferentes fabricantes:

- FRF.11 (VoFR): especifica los tipos de codificación y los formatos de trama para el transporte de tráfico de fax y de voz sobre redes FR.
- FRF.12 (Frame Relay Fragmentation): especifica los medios para la fragmentación y el reensamblado de grandes tramas de datos de forma que se minimice el retardo que se produciría al tener que encolar estas tramas en las tramas VoFR que son más pequeñas.

VoIP

VoIP es una tecnología, más reciente que VoFR, para la integración de voz en redes IP. El objetivo de VoIP es asegurar la interoperabilidad entre equipos de diferentes fabricantes. Se especifican aspectos tales como la supresión de silencios, la codificación de la voz y el direccionamiento, a la vez que se establecen nuevos elementos para permitir la

conectividad con la infraestructura telefónica tradicional. Estos elementos se refieren básicamente a los servicios de directorio y a la transmisión de señalización por tonos multifrecuencia (DTMF).

Está basado en la familia de estándares H.323 de la ITU-T que ya cubría la mayor parte de las necesidades para la integración de voz y gracias a otros protocolos de comunicación, como el RSVP (Resource ReSerVation Protocol), es posible reservar cierto ancho de banda dentro de la red que garantice la calidad de la comunicación.

Dentro de esta familia, el protocolo de señalización H.225 cubre las especificaciones de registro, autorización y status de entidades H.323 con el Gatekeeper H.323. Para la negociación del establecimiento de llamada y de los parámetros de la codificación y sesión de voz se tienen los protocolos Q.931 y H.245 respectivamente. H.245 especifica el protocolo RTPC (Real-Time Control Protocol) para el canal de control de señalización de llamada y finalmente se tiene el protocolo RTP (Real-Time Transport Protocol) con la especificación de los canales bidireccionales entre las entidades llamantes y las entidades llamadas.

VoATM

VoATM es la opción más reciente para el transporte de voz sobre redes de datos. Permite a los conmutadores ATM transportar tráfico de voz sobre una red ATM. Su principal ventaja es el amplio soporte que proporciona para QoS. Cuando se envía tráfico de voz sobre ATM, éste es encapsulado en paquetes AAL1 o AAL2, aunque el servicio AAL1 CBR es menos eficiente para VoATM que el servicio AAL2 VBR.

Control de Accesos

Introducción

¿Qué nos sugiere el título principal de este apartado?, “**Control de accesos**”, un análisis de estas tres palabras nos lleva a hacernos las siguientes preguntas:

- ¿Acceder a **qué** y por parte de **quién** ?
- ¿ **Qué significa controlar** el acceso?
- ¿ **Por qué** es preciso controlar el acceso a algo?

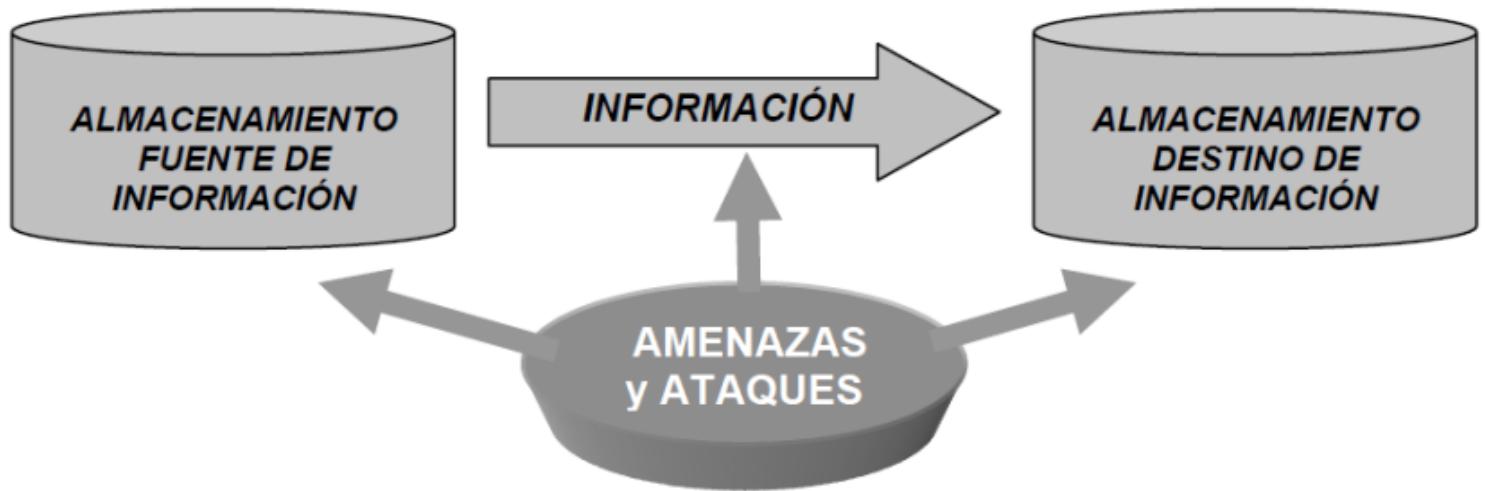
Responder a la primera pregunta resulta sencillo, en el entorno de la seguridad en redes y en la informática en general, **s e accede a información**, a datos. También parece obvio que quién accede a estos datos, en última instancia, será una persona, aunque en pasos intermedios del proceso los datos puedan haber sido obtenidos por mediación de máquinas, que hayan tenido accesos unas a otras; pero en última instancia será una persona la que interprete los datos. Una máquina puede tomar decisiones en función de unos datos, pero siempre habrá sido programada por una persona.

El conocimiento de las cosas supone poder y control sobre ellas, históricamente el tener datos o conocimientos no poseídos por rivales ha supuesto la victoria. Es decir, los datos son un bien preciado y también privado. En definitiva, el poseedor de información trascendental debe controlar el acceso que otros individuos pueden tener a estos datos, pues su publicidad no deseada podría ser peligrosa para él. Con este razonamiento se responde a la segunda y tercera preguntas planteadas. **Controlar el acceso** significa evitar que nuestros datos sean conocidos, modificados o borrados por intrusos, normalmente con malas intenciones. Tener el control de la privacidad de los datos **porque** de no hacerlo podría suponer un grave riesgo en múltiples facetas de nuestra vida. Siempre hay un riesgo de amenaza que atenta contra nuestra información.

Cualquier dato privado puede estar amenazado por individuos que podrían hacer uso fraudulento y perjudicial para el propietario. Este riesgo se multiplica por un factor enorme si además estos datos van a estar en circulación, con esto estamos indicando que son datos que necesitan ser conocidos por distintas personas en distintos lugares. Si ya existe un riesgo en la privacidad y seguridad de los datos estando a “buen recaudo” por su dueño, mucho mayor será ese riesgo si los datos han de transmitirse de una fuente a un destino. La transmisión de datos implica un alto riesgo para la integridad y privacidad de los datos transmitidos.

En nuestros días, el advenimiento de Internet, la red de redes, supone el intercambio de datos de forma inimaginable hace tan sólo unas décadas, el riesgo de amenazas y ataques ante expuesto hace que haya que tomarse muy en serio el control de acceso a la información y todo lo que ello implica. Un esquema básico para centrar las ideas antes expuestas y comenzar afrontando su estudio lo muestra la siguiente figura:

Amenazas y ataques a la posesión y la transmisión de información.



Por tanto, el control de acceso que vamos a estudiar en este apartado trata sobre “proteger” el acceso a los datos privados o esenciales de una entidad, ya sea una persona o una empresa. Protegerlos de cualquier tipo de ataque y/o de amenaza que pueda poner en manos no deseadas la información que se desea proteger.

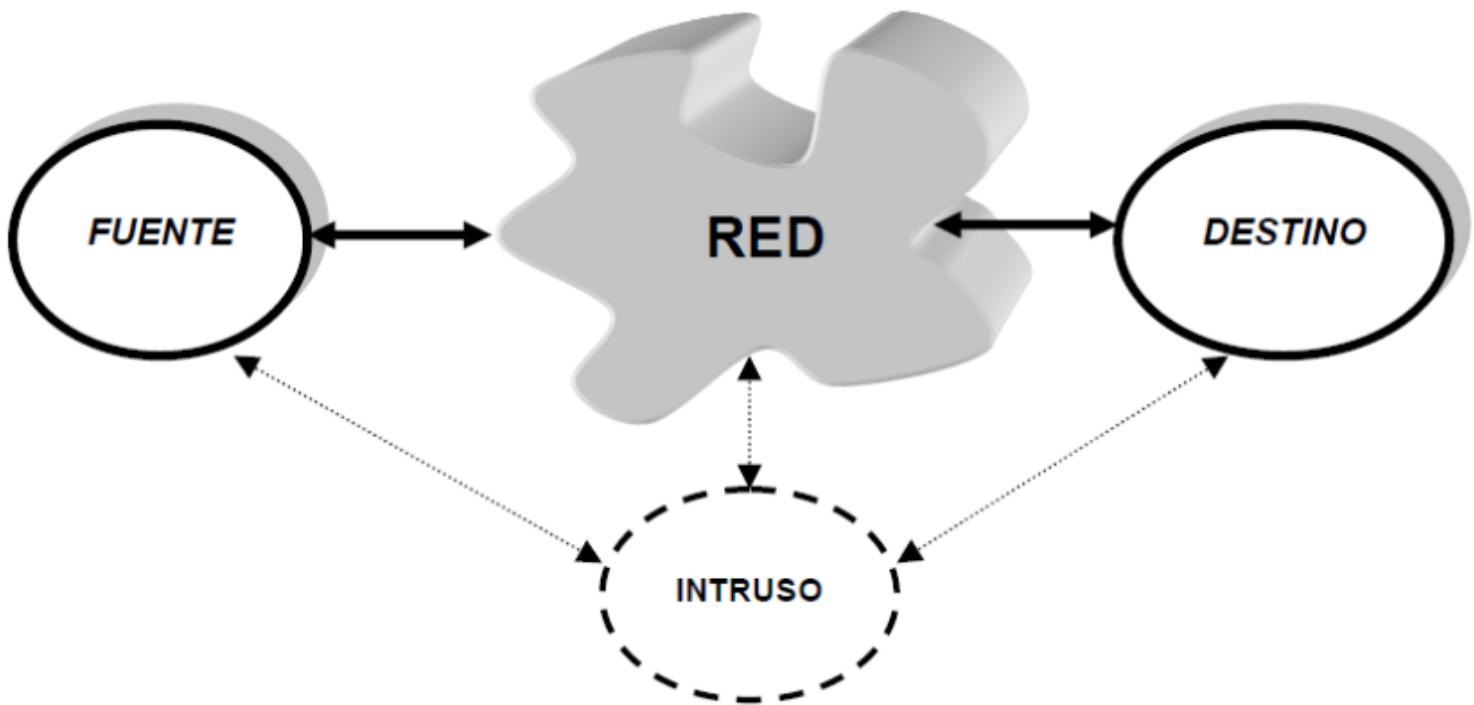
Seguridad en la Red

El concepto de seguridad en la información es más amplio que la simple protección de los datos. Para proporcionar una seguridad real se han de tener en cuenta múltiples factores internos y externos. En primer lugar habría que caracterizar el sistema que va a albergar la información para poder identificar las amenazas, y en este sentido se podrá realizar una división, muy general, entre:

- **Sistemas aislados** : Son los que no están conectados a ningún tipo de red.
- **Sistemas interconectados** : Hoy casi cualquier ordenador pertenece a alguna red, enviando y recogiendo información del exterior. Esto hace que las redes de ordenadores sean cada día más complejas y más peligrosas.

Durante las primeras décadas de su existencia, las redes de ordenadores fueron usadas principalmente por investigadores universitarios para el envío de correo electrónico, y por empleados corporativos para compartir impresoras. En estas condiciones, la seguridad no recibió mucha atención. Pero ahora, cuando millones de ciudadanos comunes usan redes para sus transacciones bancarias, compras y declaraciones de impuestos, la seguridad de las redes es un problema potencial de grandes proporciones.

Amenaza o ataque a una red.



La mayoría de los problemas de seguridad en una red son causados intencionalmente por gente maliciosa que intenta ganar algo o hacerle daño a alguien. Algunos tipos más comunes de intrusos son:

ADVERSARIO	META
Estudiante	Divertirse husmeando el correo de la gente.
Hacker	Probar el sistema de seguridad de alguien. Robar datos.
Hombre de negocios	Descubrir el plan estratégico de un competidor.
Ex empleado	Vengar su despido.
Timador	Robar números de tarjeta de crédito.
Espía	Conocer la fuerza militar de un enemigo.
Terrorista	Robar Secretos de guerra bacteriológica.

Los problemas de seguridad de las redes pueden dividirse en términos generales en cinco áreas interrelacionadas:

- **Confidencialidad** : La confidencialización, o secreto, tiene que ver con mantener la información fuera de las manos de usuarios no autorizados. Esto es lo que normalmente viene a la mente al pensar en la seguridad de las redes.
- **Validación de Identificación** : La validación de identificación determina con quién se está hablando antes de revelar información delicada o hacer un trato de negocios. Debemos tener seguridad en la identificación del remitente y destinatario, con lo cual aseguramos que “en el otro lado” está el usuario deseado y reconocido.
- **No Repudio** : El no repudio se encarga de las firmas digitales (analogía informática de las firmas del mundo real), con las que se pretende solventar el problema planteado por preguntas como la siguiente: ¿cómo comprobar que su cliente realmente solicitó (firmó) una orden electrónica por 2 millones de unidades del producto “X” a un precio de 5 euros cada una si más tarde el cliente alega que el precio estipulado fueron 4 euros?
- **Integridad** : Característica que previene contra la modificación o destrucción no autorizadas de los datos. También supone responder a la pregunta: ¿cómo puede asegurarse de que un mensaje recibido realmente fue el enviado, y no algo que un adversario malicioso modificó en el camino o preparó por su propia cuenta?

- **Disponibilidad** : Los elementos de un sistema deben estar disponibles para las entidades autorizadas. Puede ocurrir que haya elementos del sistema no disponibles por algún tipo de reparación o mejora técnica ya prevista por los propietarios de tal sistema. Lo que no debe tolerarse es que el sistema deje de estar en funcionamiento debido al ataque de un intruso malintencionado, es decir, que se pierda la disponibilidad como consecuencia de una seguridad pobre.

Control de Acceso a Redes

Uno de los problemas esenciales de la seguridad cuando se involucra la conexión en red es la extensión, posiblemente incontrolada, de nuestro perímetro de seguridad. Sin embargo, esta extensión se produce en dos ámbitos bien distintos: uno controlable, pues afecta a dispositivos que se encuentran bajo nuestro dominio y responsabilidad, y otro incontrolado, pues corresponde al "mundo exterior". Echemos un vistazo a las medidas de control físico que debemos aplicar dentro de nuestra organización. No conviene olvidar que frecuentemente los ataques más fáciles y más frecuentes se producen desde dentro. Es decir, lo primero que hay que considerar es la propia red local (LAN).

Una red de área local (LAN - Local Area Network) normalmente conecta equipos físicamente cercanos (por ejemplo, en el mismo edificio o grupo de edificios) que pertenecen a la misma organización, de quien depende su instalación, administración y mantenimiento. Las redes locales tienen como características genéricas relevantes su extendida utilización, su flexibilidad de uso, su facilidad para ampliar la conectividad y su inherente descentralización. Representan la primera capa en el esquema de conectividad de una organización.

Pasemos ahora a analizar los principales elementos físicos a considerar en una red de área local.

Elementos de Conexión

Las LANs multiplican los problemas de seguridad física de una instalación informática. Por un lado, los accidentes, como por ejemplo, los cortes de corriente, pueden tener un efecto mucho más devastador que cuando afectan a un equipo aislado. Por otro lado, las redes incorporan nuevos elementos susceptibles de sufrir ataques a la seguridad física: los cables de una instalación pueden ser saboteados, pinchados o derivados y las tomas de datos pueden además ser reconectadas con fines ilícitos.

Para hacer frente a estas eventualidades se recomienda reforzar todos los mecanismos de seguridad física. Los *servidores de red* deben estar especialmente protegidos, sobre todo en el aspecto eléctrico: tomas de tierra seguras, protección antiestática, e instalación de dispositivos de suministro ininterrumpido de corriente.

Además, el *cableado* de la instalación debe estar siempre documentado, procurando que esté a la vista o que sea fácilmente registrable, así como que sea de apariencia homogénea y ordenada. Es imprescindible su inspección metódica y periódica. En algunos casos debe contemplarse la instalación de cableado redundante para proporcionar rutas alternativas en caso de problemas en las líneas de datos. El uso de cables blindados es recomendable si se desea evitar la colocación de vampiros, y cuando se trate de fibra óptica es importante tener en cuenta los repetidores, en los que la señal se torna eléctrica, y por tanto más vulnerable. Si se utiliza radiofrecuencia o infrarrojos para la transmisión, es imprescindible el cifrado de la señal.

Los *armarios de conexiones y concentradores* deben disponer de cerradura, ésta debe usarse sistemáticamente (cosa que se verifica menos a menudo) y debe ser sólida (cosa aún más infrecuente).

Ubicación y Uso de los Ordenadores

Los servidores de red tienen el peligro de volverse “invisibles”, en el sentido de que, una vez funcionando, no requieren mucha atención. Su acceso debe estar restringido al máximo, o en todo caso debe deshabilitarse su disquetera, para evitar contaminación por virus. En el otro extremo están los equipos y rosetas de conexión poco utilizados, que pueden convertirse en puntos de acceso ilícito privilegiados. Ello no se debe solamente a que carezcan de vigilancia, sino también a que tienden a ser olvidados en las actualizaciones de los planes de seguridad. En este sentido una buena solución es instalar mecanismos de gestión de inventario, que se ocupen de rastrear e informar para que el administrador de seguridad pueda saber con precisión qué equipos están conectados en cada momento, dónde están y qué configuración tienen.

Por último debemos citar el problema de las terminales desatendidas, que pueden provocar graves incidentes de seguridad si son aprovechadas por personal desaprensivo. Aparte del trabajo de concienciación de los usuarios, suele dar buen resultado instalar facilidades que despiden automáticamente la conexión a la red en caso de producirse un período de inactividad preestablecido.

La Interconexión de Redes

Con toda la importancia que tiene el no descuidar los aspectos citados en el capítulo anterior, sin duda el mayor desafío a la seguridad se produce cuando se conectan varias redes para formar una unidad superior. La organización de estas redes de redes es absolutamente distribuida, ya que conectan entre sí equipos de usuarios y organizaciones de todos tipo. Su soporte físico puede ser extraordinariamente variopinto: líneas telefónicas (cable, microondas o fibra óptica), transmisiones vía satélite, cables de TV, etc. Y, lo peor de todo, los problemas de seguridad en estos casos son la suma de los de sus componentes. Y, como suele decirse, un sistema es tan seguro como el más inseguro de sus componentes.

Intrusiones

El análisis de los problemas de seguridad asociados a las redes de ordenadores se puede desdoblar en dos aspectos:

- Por un lado, la red puede ser vista como un punto de acceso adicional desde el cual nuestros bienes informáticos pueden ser atacados, dañados, sustraídos, etc. Es decir, susceptibles de algún tipo de **intrusión** por parte de individuos normalmente malintencionados.
- Por otro lado, la propia comunicación es un bien en sí mismo y la necesidad de su protección se añade a las que teníamos previamente.

En el primer caso nos preocupa preservar los servicios de seguridad tradicionales (confidencialidad, integridad, disponibilidad, ...) y contemplar las amenazas derivadas de la conexión en red y las posibles formas de ataque a nuestros datos. En definitiva, nos preocupará especialmente el problema de la intrusión y, por consiguiente, daremos relevancia al problema de la autenticación remota.

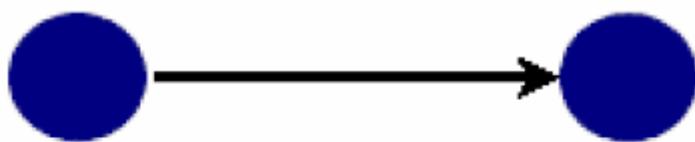
En el segundo, aunque los ataques a la comunicación también se describen en función de los servicios de seguridad amenazados, suele ser frecuente hacer una clasificación algo distinta de las mismas. Si un ordenador (emisor) envía información de cualquier clase y por cualquier medio a otro (receptor).

Formas Genéricas de Amenaza

En el mundo “normal”, fuera de la informática, las personas validan la identificación e otras personas al reconocer sus caras, voces, letra, etc. Las pruebas de firmas se manejan mediante firmas en papel, sellos, etc. Generalmente puede detectarse la alteración de documentos con el auxilio de expertos en escritura, papel y tinta. Ninguna de estas opciones está disponible electrónicamente. Es obvio que se requieren otras soluciones.

En general, no se puede proteger un sistema pensando únicamente en un tipo de amenaza, las amenazas reales suelen ser combinaciones de varios tipos específicos. En términos globales, las formas de amenazas a la seguridad de un sistema informático lo podemos caracterizar teniendo en cuenta como esta información es almacenada, suministrada o transmitida por el sistema. Ya sabemos que, en general, hay un flujo de un almacén fuente de información a un almacén destino. El flujo ideal de información sería el que no tuviese ninguna amenaza, tal como indica la siguiente figura:

Flujo ideal de información.



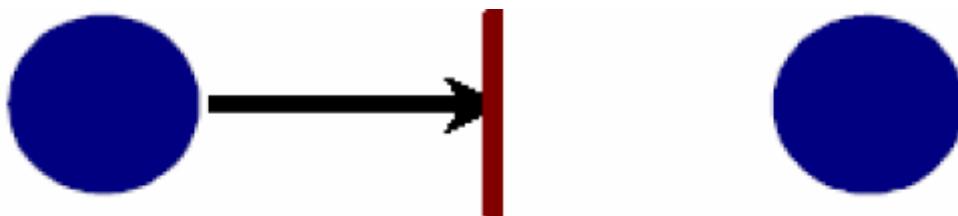
Fuente de Destino de información información

El flujo ideal de información raramente sucede. Lo habitual es que la trasmisión esté supeditada a algún tipo de amenaza o riesgo potencial. Teniendo en cuenta esto, podemos señalar cuatro categorías de forma de amenazas.

Interrupción

Se produce **interrupción** cuando una tercera parte impide que la comunicación se establezca, evitando que los datos del emisor lleguen al receptor. Se puede realizar con conocimiento de los agentes de la comunicación o sin él, aunque este segundo supuesto es más difícil. En esencia es cuando un elemento del sistema es destruido o se hace inservible. Es una amenaza a la **disponibilidad**. Ejemplos son la destrucción de algún elemento hardware (discos, líneas de comunicación, etc.) o a la desactivación del sistema de gestión de ficheros.

Interrupción.

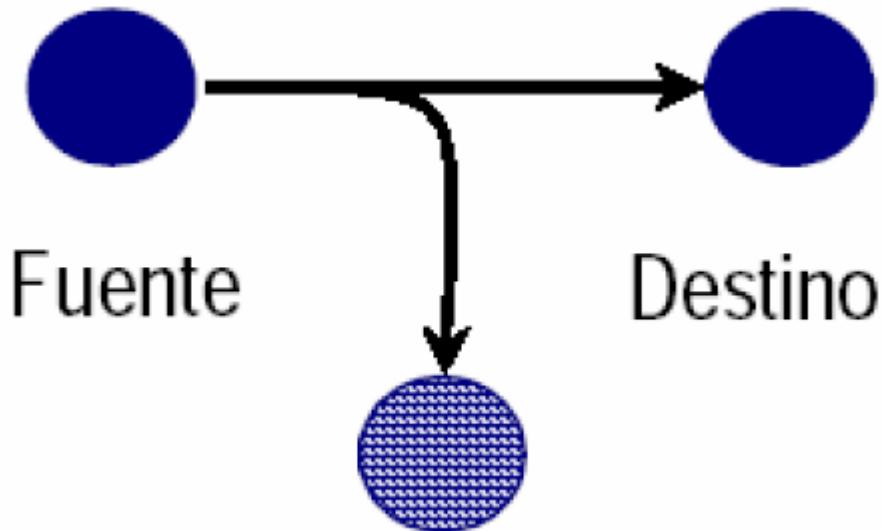


Fuente Destino

Intercepción

Se produce **intercepción** (vulgarmente hablando, una *escucha*) cuando una tercera parte no autorizada accede al contenido de la comunicación mientras esta se está produciendo. Normalmente la escucha se realiza sin necesidad de dejar huella alguna en la comunicación, por lo que ni el emisor y ni el receptor tienen por qué apercibirse de que se ha producido. Se trata de una amenaza contra la **confidencialidad** de los datos transmitidos. La parte no **autorizada** puede ser una persona, un programa o un ordenador. Ejemplos son la copia ilícita de programas y la visualización de ficheros que han de permanecer ocultos.

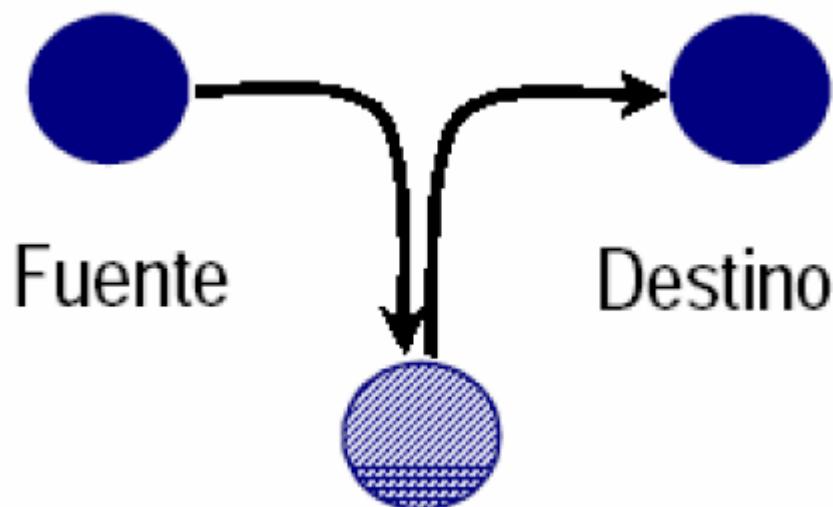
Intercepción.



Modificación

Se produce **modificación**, también llamada *manipulación*, cuando una tercera parte no autorizada accede al contenido de la comunicación y lo modifica de forma que los datos que llegan al receptor difieren en algo de los enviados originalmente por el emisor. Si la manipulación está bien hecha también resulta transparente a los agentes de la comunicación, aunque a medida que transcurre el tiempo van aumentando sus posibilidades de ser descubierta. Se trata de una amenaza contra la **integridad** de los datos transmitidos. Una parte no autorizada no sólo obtiene acceso sino que puede modificar un elemento relacionado con la seguridad. Ejemplos son la alteración del contenido de un fichero y modificar un programa para que funcione de forma diferente.

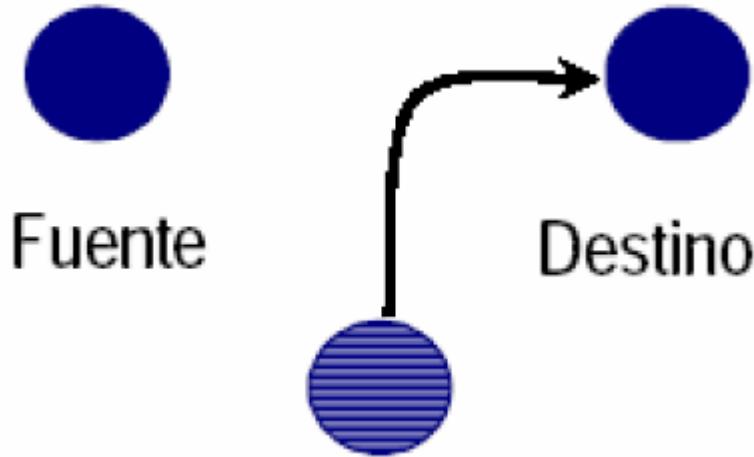
Modificación.



Suplantación

Se produce **suplantación**, también llamada *impostura* o *fabricación*, cuando una tercera parte no autorizada introduce mensajes propios en la comunicación para hacer creer al receptor que proceden del emisor. Como en el caso anterior, el propósito de la suplantación suele ser mantener el engaño durante un lapso de tiempo suficiente para realizar algún tipo de acción maligna. Se trata de una amenaza contra la *integridad* de los datos transmitidos. Una parte no autorizada inserta nuevos elementos en el sistema. Por ejemplo, la adición de registros a un fichero y la inclusión de mensajes espurios en una red.

Suplantación.



Tipos de Ataques

En principio, consideramos que una vez que se cumple la amenaza ya estamos hablando de un ataque, aunque pensando únicamente en la seguridad, ambos conceptos pueden tomarse como equivalentes. Pueden considerarse como la misma cosa, pero una en potencia (amenaza) y otra en acto (ataque).

Vamos a analizar varios tipos de ataques bajo distintas consideraciones. En ocasiones algunas de estas distinciones se solapan, pero lo importante es tener una idea del amplio repertorio de ataques que existen.

Una distinción básica de los ataques es:

- **Ataques Pasivos** : Las agresiones pasivas son del tipo de las escuchas, o monitorizaciones. La meta del oponente es obtener información que está siendo transmitida. Existen dos tipos:
 - *Divulgación del contenido de un mensaje* : Una conversación telefónica, un correo electrónico o un archivo transferido, puede contener información confidencial; por lo tanto sería deseable prevenir que el oponente se entere del contenido de estas transmisiones.
 - *Análisis de Tráfico* : El agresor podría determinar la localización y la identidad de los computadores que se están comunicando y observar la frecuencia y la longitud de los mensajes intercambiados. Esta información puede ser útil para extraer la naturaleza de la comunicación que se está realizando.

Los ataques pasivos son muy difíciles de detectar ya que no implican la alteración de la información. Sin embargo, es factible prevenir el éxito de estas agresiones. Así, el énfasis para tratar estas agresiones es la prevención antes que la detección.

- **Ataques Activos** : Suponen alguna modificación del flujo de datos o la creación de flujos falsos. Se subdividen en cuatro categorías:
 - *Enmascaramiento* : Tiene lugar cuando una entidad pretende ser otra entidad. Por ejemplo, se puede captar una secuencia de autenticación y reemplazarla por otra secuencia de autenticación válida, así se habilita a otra entidad autorizada con pocos privilegios a obtener privilegios extras suplantando a la entidad que los tiene.
 - *Repetición* : Supone la captura pasiva de unidades de datos y su retransmisión posterior para producir un efecto no autorizado.
 - *Modificación de mensajes* : Alguna porción de un mensaje legítimo se altera para producir un efecto no deseado. Por ejemplo, un mensaje con significado “Permitir a X leer el archivo confidencial cuentas” se modifica para tener el significado “Permitir a Y leer el archivo confidencial cuentas”.
 - *Denegación de un servicio* : Previene o inhibe el uso o gestión normal de una comunicación. Por ejemplo, una entidad puede suprimir todos los mensajes dirigidos a un destino particular. Otro ejemplo es la perturbación sobre una red completa, deshabilitándola o sobrecargándola con mensajes de forma que se degrade su rendimiento.

Otra distinción de los tipos de ataque es:

- Ataques **Accidentales** : No son premeditados y en ellos podemos incluir los posibles fallos del hardware y software de nuestra instalación.
- Ataques **Intencionados** : Por medio de algo o de alguien se produce un *ataque* a nuestra información para fines distintos de los que fueron creados.

También tenemos:

- Ataques **Indiscriminados** : Suelen ser los más frecuentes, y también los menos dañinos. Precisamente por su carácter general, existen programas específicos que nos protegen de ellos, como los antivirus.
- Ataques **a Medida** : Menos comunes que los anteriores, y más peligrosos, usualmente ataques que generalmente llevan a cabo los hackers. En estos casos las víctimas son casi siempre grandes corporaciones, y muchas veces la información ni siquiera es destruida o comprometida, puesto que los hackers sólo persiguen enfrentarse al reto que supone para ellos entrar en un sistema grande.

Desde el punto de vista del *acceso a la información*, podemos distinguir dos grandes grupos de ataques:

- Ataques al **Almacenamiento** de la información.
- Ataques a la **Transmisión** de la información.

Los **ataques al almacenamiento** se refieren al acceso no permitido a los lugares donde se guarda la información. Un ejemplo, fuera del mundo informático, sería el robo de una caja fuerte, es decir, un acceso fraudulento al contenido. En el mundo de las redes informáticas estas “cajas fuertes” son los sistemas de almacenamiento de datos de las computadoras, principalmente de servidores de información.

En cuanto al almacenamiento de la información, hoy en día ésta se almacena en unidades de memoria masiva, principalmente en soporte magnético o soporte óptico. Se pueden atentar contra estos soportes de información de dos formas, a saber:

- Ataque **Interno** : Obtener el propio soporte de la información o poder acceder directamente a él, de forma que se pueda conseguir obtener la información que contiene, o modificarla o borrarla. Por ejemplo, sustraer un disco duro y leer su información o tener acceso a la computadora que contiene el disco.
- Ataque **Externo** : Acceder al soporte de información pero sin acceder físicamente a él. Normalmente por intromisión en el medio de transmisión de los datos del soporte fuente al soporte destino (cable eléctrico, onda electromagnética, fibra óptica, etc). Pero no es esta la única forma de ataque externo a un sistema, pueden existir otras maneras tal como veremos más adelante.

Los **ataques a la transmisión**, también llamados ataques a las líneas de transmisiones, hace mención al peligro que existe de que unos datos que se transmiten por algún medio sean leídos, cambiados, eliminados, o cualesquiera otra acción por parte de terceros no autorizados a tener acceso a esos datos durante la mencionada transmisión. Se pueden distinguir los siguientes tipos de amenazas a las líneas de transmisiones:

- Ataques **Pasivos** : Tratan de monitorizar transmisiones de una organización. Son ataques relacionados con la **intercepción** y que afectan a la **confidencialidad**. Puesto que estas amenazas son difíciles de detectar, los esfuerzos deben encaminarse hacia su prevención más que a su detección y solución.
- Ataques **Activos** : Implican alguna modificación del flujo de datos, o la creación de un flujo de datos falso. Podemos subdividirlas en tres categorías:
 - *Modificación del flujo de información* : Para producir un efecto no autorizado; afecta a la integridad.
 - *Denegación de servicio* : Inhibiendo el uso normal de las facilidades de comunicación; afecta a la disponibilidad. Por ejemplo: La supresión de mensajes dirigidos a ciertos destinos, el trastorno del servicio, la deshabilitación de una red o sobrecargándola con mensajes, etc.
 - *Enmascaramiento* : Cuando una entidad pretende ser otra; afecta a la integridad. Normalmente, un ataque de este tipo incluye alguno de los anteriores.

Desde otro ángulo, una vez que se ha conseguido acceder a la información, podemos distinguir otros dos tipos de ataques que suponen una amenaza, estos son:

- Ataques al **Hardware**.
- Ataques **vía Software**.

Los **ataques al hardware**, cuyo objetivo es la destrucción de datos por medio del acceso físico a éstos, puede ser:

- Destrucción del soporte físico que contiene los datos.
- Borrado directo de los datos que contienen el soporte, sin la destrucción de este. Por ejemplo, los datos en soportes magnéticos pueden ser borrados o alterados mediante la aplicación de un campo magnético externo que no sea el del propio cabezal de escritura y/o lectura. Otra forma típica es el borrado directo por medio de órdenes al SO, del tipo *format*, *delete*, etc.

Los **ataques vía software** tienen su origen en programas que explotan las debilidades de los sistemas. Estos programas se dividen en dos grupos: aquellos que necesitan un programa anfitrión y aquellos que son independientes. Los primeros son trozos de programas que no pueden existir de forma autónoma, mientras que los segundos son programas completos que pueden ser planificados y ejecutados por el SO.

También hay que distinguir entre programas que no se replican y los que lo hacen. Estos últimos son programas o trozos de programas que cuando se ejecutan pueden generar una o más copias de ellos mismos, que serán posteriormente activadas, y donde se pueden distinguir los siguientes tipos de ataques de origen software:

- **Bomba Lógica** : Es código incrustado en un programa que comprueba si ciertas condiciones se cumplen, en cuyo caso ejecuta alguna acción no autorizada. Estas condiciones pueden ser la existencia de ciertos ficheros, una fecha particular, la ejecución de una aplicación concreta, etc. Una vez que la bomba explota, puede alterar o eliminar datos, parar el sistema, etc.
- **Puerta Falsa (Trapdoor)** : Es un punto de entrada secreto en un programa, de forma que alguien que conozca la existencia de dicha puerta puede obtener permisos de acceso sin tener que pasar por los mecanismos normales de autentificación. La puerta falsa es un código que reconoce alguna secuencia de entrada especial o se dispara si es ejecutado por cierto usuario o por la ocurrencia de una secuencia determinada de sucesos.
- **Caballo de Troya (Trojan Horse)** : Es una rutina oculta en un programa de utilidad. Cuando el programa se ejecuta, se ejecuta la rutina y ésta realiza acciones no autorizadas y perniciosas. Estos programas permiten realizar de forma indirecta acciones que no puede realizar de forma directa. Por ejemplo, un programa caballo de Troya puede ser un editor que cuando es ejecutado modifica los permisos de los ficheros que edita de forma que éstos puedan ser accedidos por cualquier usuario.
- **Virus** : Es código introducido en un programa que puede infectar otros programas mediante la copia de sí mismo en dichos programas. Además de propagarse, un virus realiza alguna función no permitida.
- **Bacteria** : Programa que consume recursos del sistema replicándose, pero no daña ningún fichero. Se suele reproducir exponencialmente, por lo que puede acaparar recursos como CPU, memoria y disco.
- **Gusano (Worm)** : Es un software que usa las redes de computadores para pasar de unos sistemas a otros. Una vez que llega a un sistema, el gusano se puede comportar como un virus o una bacteria, puede implantar programas caballo de Troya, o puede realizar acciones no autorizadas. Para replicarse, los gusanos emplean algunos programas que proporcionan servicios de red, como correo electrónico, ejecución remota de programas y conexión a sistemas remotos.
- **Señuelos** : Son programas diseñados para hacer caer en una trampa a los usuarios. Señuelos muy comunes consisten en instalar, o de alguna forma hacer que la víctima instale (mediante algún tipo de engaño), un programa que registre las teclas presionadas para después analizar la información en busca de contraseñas, y posteriormente transmitir esta información a otro lugar de la red.

Muchos de estos ataques o amenazas descritas son difíciles de prevenir, por tanto se deben dedicar esfuerzos sobre todo a la prevención, y obviamente a su detección y la recuperación de los trastornos o retardos que puedan causar. Esto también puede tener un efecto disuasorio, que refuerza la propia prevención.

Amenazas Globales

Supóngase una red de ordenadores locales, únicamente se comunican entre sí por medio de conexiones directas por cable eléctrico y están totalmente aislados, en términos de comunicación eléctrica, con cualquier computadora o sistema exterior. En suma, estamos convencidos de que no hay posible transmisión fuera de la red local.

Por otra parte también vamos a suponer que nadie no autorizado puede tener acceso a los ordenadores, de forma que en apariencia la amenaza física no existe. ¿Están pues los datos contenidos en nuestra red local seguros?. La respuesta es NO. Esto es debido a que las computadoras actuales funcionan con circuitos eléctricos, y cualquier circuito eléctrico en funcionamiento emite radiación electromagnética. Esta radiación es una "huella" que

revela la historia de funcionamiento del circuito. Además, si no se hace nada por “apantallarla” esta radiación puede alcanzar centenares de metros de distancia. Con lo que la seguridad física se hace muy difícil, resulta casi imposible controlar un espacio físico de centenares de metros cuadrados (por ejemplo, un edificio de oficinas con múltiples plantas y sus alrededores) de forma permanente en el tiempo.

Existen dispositivos capaces de “rastrear” estas ondas y poder interpretarlas, obteniendo, con un proceso complicado pero factible, los datos que contienen los ordenadores que emiten la radiación. Durante muchos años muchos piratas informáticos se han dedicado a rondar edificios de importantes empresas con dispositivos de rastreo en busca de un precio botín (datos referentes a contraseñas, número secretos de cuenta y cualquier otro tipo de información privada o confidencial).

No queda claro si la forma de amenaza descrita anteriormente es puramente lógica o puramente física, probablemente tenga un poco de ambas. Por una parte, aunque no se accede directamente a los ordenadores sí que hay que estar cerca para poder acceder a los datos. Por otro lado, aunque no estamos accediendo al medio formal de comunicación entre los ordenadores (por el cable de conexión directa), sí que estamos accediendo a un medio de transmisión no formal (residual) fruto de un efecto físico inevitable.

Las dos formas básicas de evitar esta amenaza físico-lógica son:

- **Apantallamiento Electromagnético** : Aunque no se puede evitar la radiación electromagnética de los circuitos eléctricos, si se pueden crear pantallas para reducir la emisión de radiación, tal que ésta apenas alcance unos metros, con lo cual la seguridad física puede garantizarse. Materiales como el hierro o el plomo son los más utilizados para generar estos “apantallamientos”. Normalmente los ordenadores con información crítica son mantenidos en armarios de plomo y los cables de comunicación entre ellos utilizan algún tipo de cubierta que apantalla la radiación y elimina interferencias.
- **Utilizar dispositivos de detección de rastreo** : Al igual que existen sofisticados equipos de rastreo de señales, también existen dispositivos muy avanzados dedicados a localizar posibles rastreadores.
- **Computación Óptica** : En los últimos tiempos se está investigando mucho en computación sin circuitería eléctrica, de forma que todos los cálculos se hacen mediante procedimientos ópticos, que no producen la radiación espuria propia de los circuitos eléctricos.

Las formas de amenazas y ataques rara vez se corresponde de forma pura a alguna de las expuestas hasta ahora, normalmente la amenaza suele ser una combinación de varios tipos, una amenaza de índole general o global. En definitiva, los ataques pueden venir de múltiples lados, en ocasiones de muchos.

Jamás se puede decir que la seguridad es de un cien por cien, ni siquiera apagando todos los ordenadores y no haciendo transmisiones tenemos seguridad total, pues podemos ser víctimas de ataques de almacenamiento de tipo interno; siempre puede producirse un robo de algún soporte de información crítica, y no siempre este robo es realizado por extraños o terceros, también puede ser hecho por personas conocidas. No sería la primera vez que un empleado ha robado datos de su empresa para proporcionárselos a otra empresa rival (el conocido espionaje industrial). Como ya hemos comentado, muchas veces el ataque viene “desde dentro”.

Valoración de las Amenazas

Dependiendo del contexto y la situación, el hecho de que cada una de las amenazas o ataques descritos anteriormente puede ser más o menos peligroso para los emisores y destinatarios de la información. Existen situaciones en las que lo importante es que llegue una información fidedigna, aunque sea interceptada por terceros. En otras circunstancias

es preferible que no llegue ninguna información a que llegue información falsa. Lo que si se tiene claro es que hay que procurar que ninguna de las amenazas antes descritas llegue a concretarse.

Desde el punto de vista conceptual el medio carece de importancia. Pues el hecho en sí, por ejemplo, de interrumpir un dato es independiente del medio. Si que dependerá del medio el cómo interrumpir, interceptar, modificar o suplantar los datos. A título de ejemplo, los métodos técnicos para interceptar información transmitida a través de un cable eléctrico no son los mismos que para interceptar una emisión por ondas electromagnéticas (radio, televisión, microondas, ...).

Otro factor a tener en cuenta cuando se valora la importancia de una amenaza es la relación entre la posibilidad de que esa amenaza se dé y el coste que supondría tener una protección muy elevada, rara vez puede ser absoluta, frente a esa amenaza. ¿Hasta qué punto merece la pena gastarse muchísimo dinero para conseguir una seguridad cercana al cien por cien sobre unos datos cuyo contenido no son de vital importancia?. Se pueden definir medidas de la valoración de un conjunto de amenazas con cocientes parecidos al siguiente:

$$V.A. \text{ (Valoración Amenazas)} = \frac{[(\text{probabilidad AMENAZA 1}) + \dots + (\text{probabilidad AMENAZA N})]}{\text{COSTE TOTAL PROTECCIÓN}}$$

Puede variarse el factor anterior añadiendo algún tipo de “peso” o “sumando” a considerar, pero esencialmente la valoración responde a la idea de probabilidad de amenaza frente al coste de la protección. No tener en cuenta esto puede llevar a que el coste de la protección sea superior al valor de la propia información que se desea proteger, cosa totalmente irrazonable.

Control de Acceso a la Información

En esta sección daremos respuesta a esta pregunta: ¿Cómo protegerse de los ataques contra nuestros datos?. Veremos cuales son los principios de diseño de una buena protección para mantener nuestros datos y transmisiones seguras.

Principios de Diseño para la Seguridad

Algunos principios fundamentales de diseño relacionados con la seguridad, válidos para cualquier tipo de ataque o amenaza a la que haya que enfrentarse, son:

- **Diseño abierto** : La seguridad de un sistema no debe de depender de la suposición de que su diseño es secreto. Asumir que un intruso no lo conoce es engañar a los diseñadores.
- **Negación en caso de duda** : Por defecto, se deben negar los accesos no autorizados.
- **Verificación completa** : Toda operación debe ser contrastada con la información de control de acceso.
- **Principio del menor privilegio** : Todos los procesos deberían tener los mínimos privilegios necesarios para realizar sus tareas.
- **Economía** : El mecanismo de protección debe ser simple, uniforme e integrado hasta las capas más bajas del sistema.
- **Aceptabilidad** : El sistema elegido debe ser psicológicamente aceptable por los nuevos usuarios, ya que en caso contrario éstos no lo usarán.

Protección Física

La **protección física**, también llamada protección o seguridad interna, intenta mantener los datos privados sólo accesibles a los usuarios autorizados por medio de la seguridad física, es decir, por medio de elementos como:

- **Personal de seguridad .**
- **Sala de acceso restringido por identificación personal .**
 - Reconocimiento de alguna **propiedad física** , como por ejemplo:
 - Huella **dactilar** .
 - Reconocimiento **facial** .
 - Análisis de **retina** .
 - Control de **escritura** .
 - Reconocimiento de **voz** .
 - Análisis de **ADN** .
 - Tarjeta **Identificadora** .
 - Supervisión por **personal autorizado** .
- **Recintos vallados** .

En general, la práctica de la seguridad interna se basa en gran medida en la utilización de políticas de contraseñas y control de acceso a los contenedores de información.

Protección Lógica

Puede considerarse como **protección lógica** toda aquella que no es física, es el tipo de protección que no se fundamenta en restricciones impuestas por algo físico, como ocurren en la protección física de la información.

Estas técnicas y mecanismos de seguridad constituyen la lógica que implanta un servicio de seguridad particular y responden a *cómo* implantar los servicios de seguridad. Los principales mecanismos de protección lógica son:

Autenticación del Usuario

Es una protección básica en cualquier sistema de información. Muchos esquemas de protección se basan en que el sistema conoce la identidad de todos los usuarios. El problema de identificar a los usuarios cuando éstos se conectan se denomina autenticación. La mayoría de los métodos de autenticación se basan en identificar algo que el usuario tiene o conoce. El mecanismo más común de autenticación consiste en que todo usuario ha de introducir una contraseña, solicitada por el programa de conexión cuando el usuario introduce su nombre. El inconveniente es que las contraseñas pueden ser averiguadas si el usuario utiliza su nombre, dirección, o datos similares como contraseña. Existen sistemas que procuran evitar esto asignándole a cada usuario un libro con una secuencia de contraseñas, de forma que cada vez que se conecta tiene que introducir la palabra de paso siguiente. El problema está en qué hacer si se pierde el libro de contraseñas. Otra forma obvia de averiguar una contraseña consiste en probar todas las combinaciones de letras, números y símbolos de puntuación hasta adivinar la contraseña.

En instalaciones en las que la seguridad es prioritaria, estas medidas se pueden complementar con las protecciones de tipo físico vistas anteriormente, como son las restricciones de acceso a la habitación en la que se encuentran los terminales, asignar a cada usuario un terminal concreto, una tarjeta de identificación, establecer un horario concreto de trabajo, etc.

Cortafuegos (Firewalls)

Podríamos definir un cortafuegos como aquel sistema de red expresamente encargado de separar redes de comunicación de datos, efectuando un control del tráfico existente entre

ellas. Este control consiste, en última instancia, en permitir o denegar el paso de la comunicación de una red a otra.

En definitiva, son sistemas que controlan el tráfico dentro de las redes utilizando programas de seguridad situados en un servidor u ordenador independiente. Se diseñan para restringir el acceso a las redes de las organizaciones, especialmente desde el exterior. Analizando dónde se originan los paquetes, los dejan pasar o no. Los cortafuegos pueden tener distintas formas: filtrador de paquetes, cortafuegos a nivel de circuitos y a nivel de aplicación.

El concepto que subyace detrás de un sistema cortafuegos es el de Seguridad Perimetral Centralizada, es decir, la creación de perímetros de separación implantados mediante puntos donde se centraliza el control de las comunicaciones. El caso más básico involucra a dos redes, una red a proteger (normalmente una red corporativa) y una red externa (normalmente Internet).

Software de protección

Aplicaciones utilizadas para proteger de los ataques vía software (virus, bacterias, troyanos, ...) o en caso de que se haya producido ya el ataque, reparar el sistema y recuperar las pérdidas en lo posible. Los más conocidos de estas aplicaciones de protección son los antivirus.

Criptografía

Consiste en modificar los datos almacenados o el mensaje transmitido de forma que no sea entendible para un intruso que no conozca el sistema de alteración usado. A este proceso se le denomina **cifrado**. La filosofía subyacente es que si no se puede evitar tener accesos a los datos, por lo menos se puede evitar que se entienda su significado, pero que solamente sea ininteligible para los intrusos, sin embargo, puede ser entendible para la persona a quien queremos enviar el mensaje.

Realmente la criptografía es una rama de otra ciencia más amplia que es la criptología, que contempla tanto el cifrado de datos como el descifrado (operación inversa al cifrado).

Firma digital

Se basa en técnicas criptográficas, y cumple las mismas funciones que la firma manual: el receptor debe ser capaz de validar la firma del emisor, no debe ser falsificable y el emisor de un mensaje no debe poder repudiarlo posteriormente.

Relleno del tráfico

Se basa en introducir tráfico espurio junto con los datos válidos para que no se pueda conocer si se está enviando información o qué cantidad de datos útiles se están enviando.

Etiquetas de seguridad

Permiten que los mensajes sean clasificados para facilitar un correcto control de acceso y la separación de datos según clases de seguridad.

Funciones de dispersión seguras (Hash)

Son funciones matemáticas sin inversa, que aplicadas a un elemento o dato que se transfiere impiden que este sea descifrado. También sirven para verificar la correcta recepción de los mensajes.

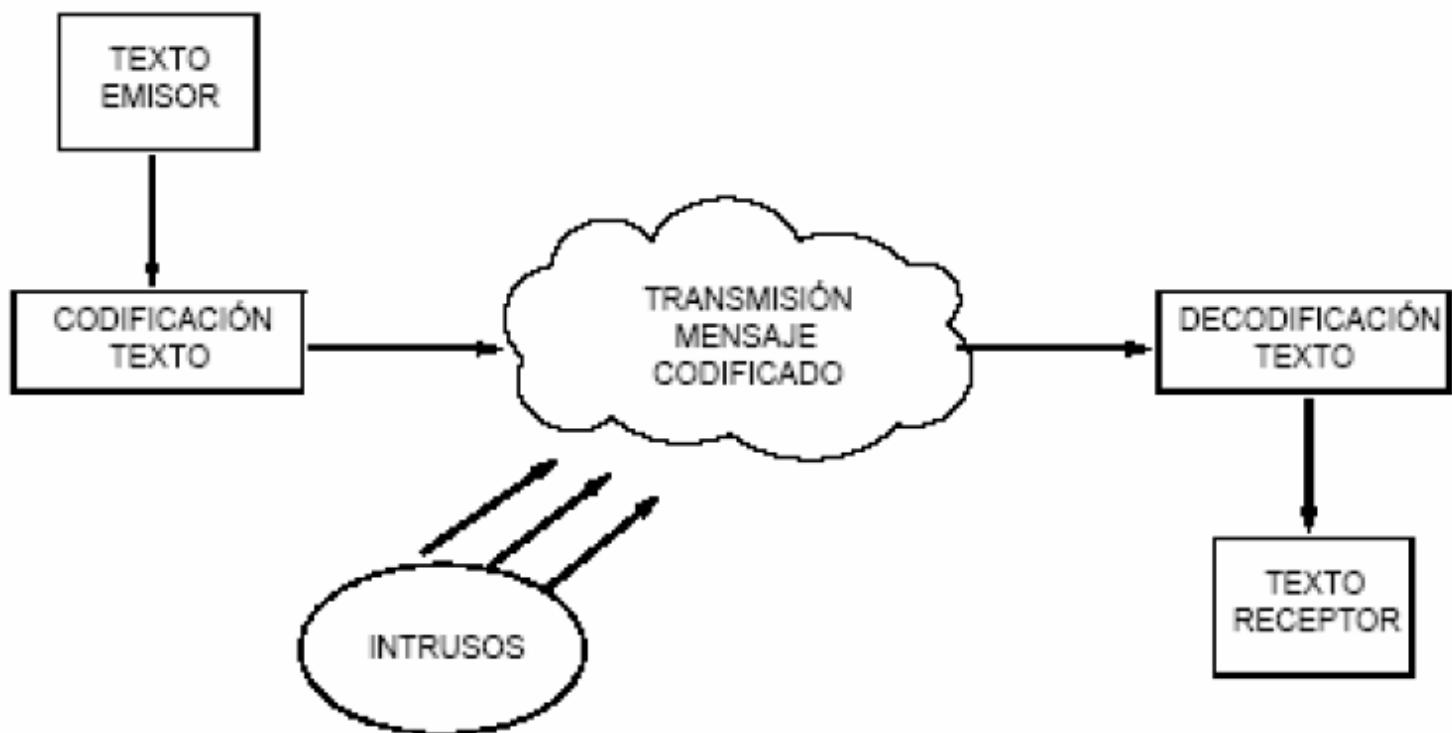
Terceras Partes de Confianza (TTP)

Son entidades cuyos informes se consideran fiables por todos los elementos del dominio de seguridad. Pueden tener registros y firmas digitales y emitir certificados dentro del sistema.

Técnicas Criptográficas

La Criptología (del griego kryptos (oculto) y logos (estudio), estudio de lo oculto, lo escondido) es la ciencia que trata los problemas teóricos relacionados con la seguridad en el intercambio de mensajes en clave entre un emisor y un receptor a través de un canal de comunicaciones. Como hemos indicado en un apartado anterior, la idea es “*si no podemos evitar que capturen nuestro mensaje, por lo menos intentaremos que sea ininteligible para cualquier posible intruso, pero no para el destinatario legal*”.

Uso de la criptografía.



Esta ciencia está dividida en dos grandes ramas:

- **Criptografía** : Se ocupa del cifrado de mensajes.
- **Criptoanálisis** : Es la parte contraria a la criptografía. Trata de descifrar los mensajes en clave y determinar la forma (el algoritmo) bajo el cual se ha obtenido el mensaje en clave.

Es decir, conviene distinguir entre la palabra **Criptografía**, que sólo hace referencia al uso de códigos y la palabra **Criptoanálisis**, que engloba a las técnicas que se usan para romper dichos códigos. En cualquier caso ambas disciplinas están íntimamente ligadas; no olvidemos que cuando se diseña un sistema para cifrar información, hay que tener muy presente su posible criptoanálisis, para evitar sorpresas desagradables.

Finalmente, como ya hemos comentado, el término **Criptología** se emplea habitualmente para agrupar tanto la Criptografía como el Criptoanálisis.

Criptografía

Los sistemas criptográficos están teniendo un gran auge últimamente ante el miedo de que una transmisión en Internet pueda ser interceptada y algún desaprensivo pueda enterarse de alguna información que no debería. Y no estamos hablando de un correo electrónico en el que pensamos quedar con unos amigos, nos referimos a, por ejemplo, una transacción comercial o una información sobre temas empresariales.

Desde la Antigüedad todas las civilizaciones han desarrollado sistemas de criptografía para que las comunicaciones no fueran públicas. Incluso hoy en día muchas personas utilizan lenguajes específicos para que solamente los iniciados en ellos puedan comprender la conversación como, por ejemplo, las jergas utilizadas en ambientes delictivos.

Hay muchos sistemas para “camuflar” lo que escribimos. Quizá el más fácil sea la **transposición** o **sustitución** del texto. Consiste en cambiar cada letra del texto por otra distinta. Por ejemplo, si escribimos “boujwjsvt”, solamente las personas que supieran que hemos puesto la letra siguiente del alfabeto para escribir la palabra “antivirus” podrían entender la palabra. Otro sistema sencillo es el de elegir una frase fácil que recordar por el usuario y eliminar alguna parte también fácil de recordar, por ejemplo, si escribo “qtrlsvcls”, puede observarse que son las consonantes correspondientes a la frase “**q ui t a r l a s v o c a l e s**”.

Evidentemente los sistemas criptográficos actuales van mucho más allá de un sistema como el de transposición, o semejantes; fáciles de descubrir en unos cuantos intentos. Incluso si en lugar de trasponer un determinado número de espacios elegimos aleatoriamente las letras a sustituir, también bastaría con un ordenador que tuviera un simple corrector ortográfico para, en unos cuantos intentos, descubrir el significado de un mensaje. Los sistemas criptográficos de hoy en día se basan en la utilización de matemática avanzada, sobre todo a ciertas cualidades de los números, cuyo estudio pertenece a una rama avanzada de la matemática denominada teoría de números. Un tipo de números cuyas características son muy aprovechadas en la criptografía son los números primos. Una de las tareas que más tiempo ocupa a los grandes sistemas de ordenadores es el cálculo de números primos cada vez mayores. El objetivo sería poder obtener un número que sirva para cifrar mensajes y que luego sea muy complicado descifrarlos.

Criptosistema

Se define un criptosistema como una quíntupla (**M** , **C** , **K** , **E** , **D**), donde:

- **M** : Representa el conjunto de todos los mensajes sin cifrar (lo que se denomina texto claro, o plaintext) que pueden ser enviados.
- **C** : Representa el conjunto de todos los posibles mensajes cifrados, o criptogramas.
- **K** : Representa el conjunto de claves que se pueden emplear en el criptosistema.
- **E** : Es el conjunto de transformaciones de cifrado o familia de funciones que se aplica a cada elemento de M para obtener un elemento de C. Existe una transformación diferente E_k para cada valor posible de la clave k.
- **D** : Es el conjunto de transformaciones de descifrado, análogo a E.

Veamos un ejemplo. Retomemos el ejemplo de la transposición visto anteriormente, pero en este caso vamos a expresarlo con nuestra nueva terminología.

Ejemplo de algoritmo de transposición.

Trasponer cada letra por la siguiente

antivirus

boujwjsvt

Trasponer cada letra por la anterior

Vamos a suponer una transposición cíclica, es decir, la letra siguiente a la “z” es la “a” y la letra anterior a la “a” es la “z”. También suponemos, para mayor sencillez, que sólo usamos las 27 letras minúsculas del abecedario “*abcdefghijklmnñopqrstuvwxyz*” y palabras de 8 letras.

La palabra *antivirus* es una más de nuestro conjunto **M**, puesto que es la palabra sin cifrar; y la “palabra” (por llamarla de alguna forma) *boujwjsvt* pertenece al conjunto **C**. En realidad, ambas palabras pertenecen a ambos conjuntos, pues ambas son susceptibles de ser cifradas y de ser descifradas, aunque no parezca tener mucho sentido cifrar una “palabra” como *boujwjsvt*. Es fácil ver que en nuestro caso **M = C**.

Respecto al número de palabras posibles en M o en C, como tenemos 27 letras, en palabras de 8 letras donde el orden importa y puede darse la repetición, lo que tenemos (recordando la combinatoria) es una variación con repetición de 27 elementos tomados de 8 en 8. Así pues:

$$VR(27,8) = (27)^8 = 282.429.536.481 \text{ palabras posibles.}$$

Como se ha dicho, **E** es el conjunto de posibles transformaciones a aplicar a nuestro mensaje, en nuestro caso se trata de sustituir o trasponer cada letra por la siguiente, pero podría ser, por ejemplo, no sustituir por la siguiente sino por la siguiente de la siguiente (la 2^a siguiente), y así sucesivamente, (3^a siguiente), etc.

Así pues, considerando para el cifrado desplazamientos fijos hacia adelante, hacia la siguiente letra en el abecedario, tendremos:

E = Transformaciones con **desplazamiento fijo hacia delante** sobre palabras de ocho letras en minúsculas (abecedario de 27 letras).

Obsérvese que, al considerar un ciclo de 27 letras, la transposición (28^a siguiente) es igual a (1^a siguiente). Decimos que es fijo porque consideraremos que se aplica el mismo desplazamiento a todas las letras de la palabra, no vamos a considerar desplazamiento del estilo: la primera letra por la siguiente, la segunda por la 2^a siguiente, la tercera por la 3^a siguiente, etc.

Ya tenemos el algoritmo, pero si nos dan un mensaje cifrado (perteneciente a **C**) y nos piden descifrarlo, si no sabemos la cantidad de desplazamientos, tendríamos que probar con los 26 desplazamientos posibles y aún así no sabríamos con qué palabra de las 26 resultantes quedarnos. Es decir, se necesita saber el desplazamiento. Este desplazamiento es la denominada **clave de cifrado** del algoritmo.

En nuestro ejemplo existen 26 posibles desplazamientos hacia delante, pues el desplazamiento 27 equivale al desplazamiento 0 (sin desplazamiento), el desplazamiento 28 equivale al desplazamiento 1, y así sucesivamente.

Por tanto:

$$K = \{1, 2, 3, \dots, 26\}$$

En nuestro ejemplo es fácil ver que k=1. Esto lo indicaremos de la siguiente forma:

E₁ = Cifrado por transposición con desplazamiento fijo hacia delante con k=1

De análoga forma, tendremos que:

D = Transformaciones con desplazamiento fijo hacia atrás sobre palabras de 8 letras en minúscula (abecedario de 27 letras).

D₁ = Descifrado por transposición con desplazamiento fijo hacia atrás con k=1.

Y esto es todo. Vemos pues que únicamente hemos dado una notación matemática a lo que ya teníamos en mente. Sin embargo, esta definición de criptosistema y la notación definida nos va a venir muy bien para analizar ciertas propiedades de la criptografía. Por ejemplo, en nuestro caso tenemos:

E₁(antivirus) = boujwjsvt

D₁(boujwjsvt)=antivirus

Veamos ahora cómo expresar con esta notación una propiedad de todo criptosistema.

Todo criptosistema ha de cumplir la siguiente condición :

D_k(E_k(m)) = m (Siendo m el mensaje a cifrar; m pertenece a M).

Es decir, que si tenemos un mensaje **m** , lo ciframos empleando la clave **k** y luego lo desciframos empleando la misma clave, obtenemos de nuevo el mensaje original **m** . En nuestro ejemplo:

m = antivirus → E₁(m) = boujwjsvt → D₁(E₁ (m)) = antivirus ⇒ [D₁(E₁ (m)) = m]

Vemos claramente como la ventaja de la notación utilizada es que, definiendo el criptosistema con la quíntupla (M,C,K,E,D) podemos expresar con ecuaciones, tan sencillas y cortas como la anterior, propiedades que nos llevaría mucho explicarlas de palabra. Y no sólo eso, sino que con esta notación podremos realizar operaciones algebraicas con las ecuaciones obteniendo resultados nuevos.

Veamos ahora cuales son los dos tipos fundamentales de criptosistemas que existen:

- **Criptosistemas Simétricos o de Clave Privada** . Son aquellos que emplean la misma clave **k** tanto para cifrar como para descifrar. Presentan el inconveniente de que para ser empleados en comunicaciones la clave k debe estar tanto en el emisor como en el receptor, lo cual nos lleva al grave problema de cómo transmitir la clave de forma segura; puesto que en muchos casos el emisor y receptor están a distancia y no es posible una comunicación personal entre ellos para intercambiar la clave de forma privada.
- **Criptosistemas asimétricos o de llave pública, que emplean una doble clave (kp, kP)** . **kp** se conoce como **Clave Privada** y **kP** se conoce como **Clave Pública** . Una de ellas sirve para la transformación E de cifrado y la otra para la transformación D de descifrado. En muchos casos son intercambiables, esto es, si empleamos una para cifrar la otra sirva para descifrar y viceversa. Estos criptosistemas deben cumplir además que el conocimiento de la clave pública kP no permita calcular la clave privada kp. Ofrecen un abanico superior de posibilidades, pudiendo emplearse para establecer comunicaciones seguras por canales inseguros - puesto que únicamente viaja por el canal la clave pública, que sólo sirve para cifrar-, o para llevar a cabo autentificaciones.

En la práctica se emplea una combinación de estos dos tipos de criptosistemas, puesto que los segundos presentan el inconveniente de ser computacionalmente mucho más costosos que los primeros. Cuando utilizamos el término "computacional" nos referimos al tiempo que tarda un ordenador en calcular un algoritmo. Se dice que un algoritmo es

computacionalmente intratable (con los ordenadores y tecnologías de hoy en día) cuando su resolución llevaría una cantidad de tiempo desorbitada, por ejemplo, cientos de años.

En el “mundo real” se codifican los mensajes (largos) mediante algoritmos simétricos, que suelen ser muy eficientes, y luego se hace uso de la criptografía asimétrica para codificar las claves simétricas (cortas).

En la inmensa mayoría de los casos los conjuntos M y C definidos anteriormente son iguales (como hemos visto en nuestro ejemplo). Esto quiere decir que tanto los textos claros como los textos cifrados se representan empleando el mismo alfabeto -por ejemplo, cuando se usa el algoritmo DES, ambos con cadenas de 64 bits-. Por esta razón puede darse la posibilidad de que exista algún k perteneciente a K tal que $E_k(m) = m$, lo cual sería catastrófico para nuestros propósitos, puesto que el empleo de esas claves dejaría todos nuestros mensajes sin codificar. (En nuestro ejemplo anterior, esto pasaría si se utilizase $k = 0, 27, 45$, etc. De ahí que hallamos definido $K = \{1, 2, \dots, 26\}$).

También puede darse el caso de que ciertas claves concretas generen textos cifrados de poca calidad. Una posibilidad bastante común en ciertos algoritmos es que algunas claves tengan la siguiente propiedad: $E_k(E_k(m)) = m$, lo cual quiere decir que basta con volver a codificar el criptograma para recuperar el texto claro original. Estas circunstancias podrían llegar a simplificar enormemente un intento de violar nuestro sistema, por lo que también habría que evitarlas a toda costa.

La existencia de claves con estas características, como es natural, dependen en gran medida de las peculiaridades de cada algoritmo en concreto, y en muchos casos también de los parámetros escogidos a la hora de aplicarlo. Llamaremos en general a las claves que no codifican correctamente los mensajes claves débiles (weak keys).

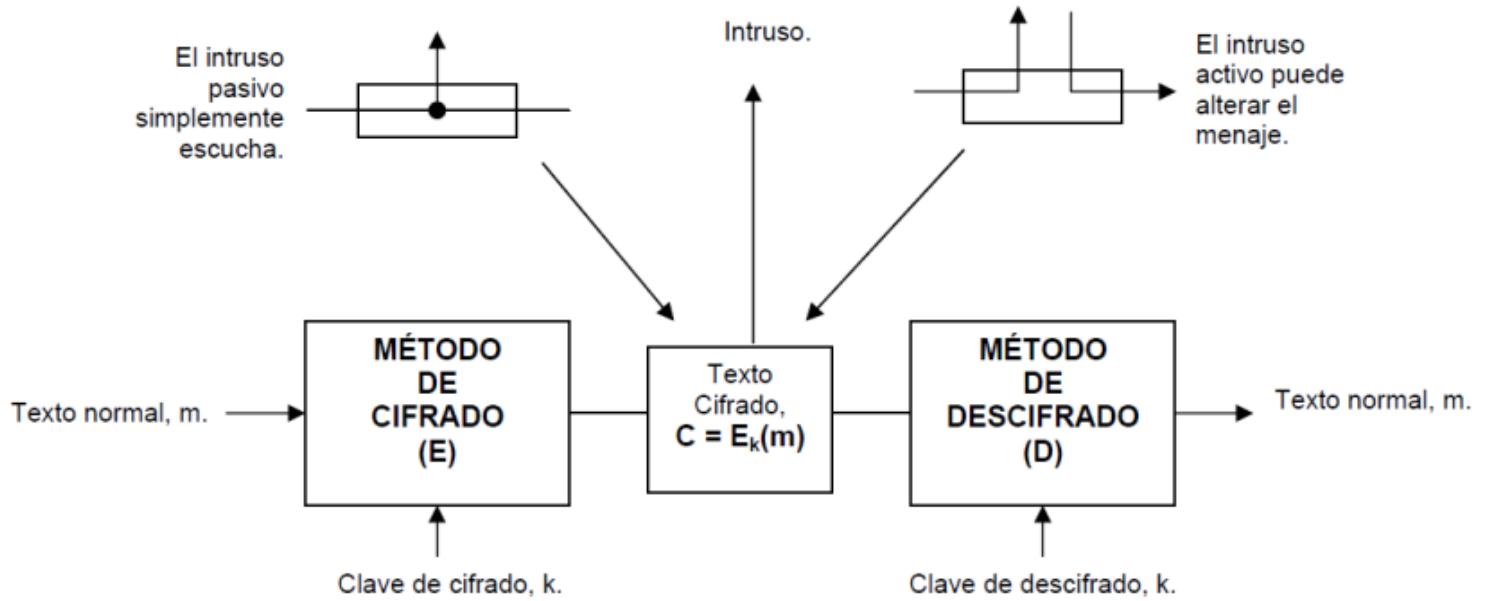
Normalmente en un buen criptosistema la cantidad de claves débiles es cero o pequeña en comparación con el número total de claves. Pero conviene conocer esta circunstancia para evitar en lo posible sus consecuencias.

Algoritmos Simétricos

Desde muy antiguo han existido algoritmos de cifrado, algunos se remontan incluso, como el algoritmo de César, a la Roma Imperial. Todos estos algoritmos son algoritmos de clave privada o simétricos.

Hasta la llegada de las computadoras, una de las restricciones principales del cifrado había sido la capacidad de la persona encargada del codificado para realizar las transformaciones necesarias, frecuentemente en un campo de batalla (la historia de la criptografía está muy unida al ejército) con poco equipo. Una restricción adicional ha sido la dificultad de cambiar rápidamente de un método de cifrado a otro, puesto que esto significa el reentrenamiento de una gran cantidad de gente. Sin embargo, el peligro de que un empleado fuera capturado por el enemigo ha hecho indispensable la capacidad de cambiar el método de cifrado al instante, de ser necesario. De estos requisitos en conflicto se deriva el modelo de la figura siguiente:

El modelo de cifrado tradicional.



Hoy en día, cualquier ordenador doméstico podría descifrarlos rápidamente, pero que fueron empleados con éxito hasta principios del siglo XX. Conviene detenerse someramente en su estudio pues mantienen un interés teórico que nos van a permitir explotar algunas de sus propiedades para entender mejor los algoritmos modernos.

Veamos los más conocidos de estos algoritmos simétricos clásicos.

- **Cifrados Monoalfabéticos.** Se engloban en este apartado todos los algoritmos criptográficos que, sin desordenar los símbolos dentro del mensaje, establecen una correspondencia única para todos ellos en todo el texto. Es decir, si al símbolo A le corresponde el símbolo D, esta correspondencia se mantiene a lo largo de todo el mensaje. Los ejemplos más típicos de cifrados monoalfabéticos son:
 - *Algoritmo de César* : El algoritmo de César, llamado así porque es el que empleaba Julio César para enviar mensajes secretos, es uno de los algoritmos criptográficos más simples. Consisten en sumar 3 al número de orden de cada letra. De esta forma a la A le corresponde al D, a la B la E, y así sucesivamente. Si asignamos a cada letra un número ($A = 0, B = 1, \dots$), y consideramos un alfabeto de 27 letras anterior, la transformación criptográfica sería: $C = (M + 3) \text{ mod } 26$; siendo $A \text{ mod } B = \text{Resto de la división } A/B$. Obviamente, para descifrar basta con restar 3 al número de orden de las letras del criptograma. Al ser siempre la clave fija, en este caso puede decirse que este algoritmo no tiene clave. Obsérvese la forma de expresar el algoritmo de transposición o desplazamiento mediante la operación (mod). Nuestro ejemplo anterior es equivalente a este algoritmo pero con $C = (M+1) \text{ mod } 26$ (obviando el hecho de que habíamos fijado el tamaño de las palabras en ocho caracteres, por cuestiones de comodidad, cosa que evidentemente no ocurre en el algoritmo de César).
 - *Sustitución Afín* : La sustitución afín es el caso más general del algoritmo de César. Su transformación sería: $E(a,b)(M) = (aM + b) \text{ mod } N$. Siendo a y b dos números enteros menores que el cardinal N del alfabeto, y cumpliendo que $\text{mcd}(a,N) = 1$. La clave de cifrado k viene entonces dada por el par (a, b). El algoritmo de César sería pues una transformación afín con $k = (1, 3)$. Nuestro ejemplo anterior es una transformación afín con $k=(1,1)$.
- **Cifrado Monoalfabético General.** Es el caso más general de cifrado monoalfabético. Se considera cualquier algoritmo que realice una asociación biyectiva (1 a 1) de las letras del alfabeto consigo mismo. Normalmente la clave se proporciona a través de una tabla que indica la correspondencia aleatoria entre las diferentes letras, salvo el caso de la existencia de una ecuación matemática que pueda servir para establecer la relación biyectiva, como ocurre en la sustitución afín.

La sustitución ahora es arbitraria, siendo la clave k precisamente la tabla de sustitución de un símbolo por otro.

- **Cifrado Polialfabético** . En los cifrados polialfabéticos la sustitución aplicada a cada carácter varía en función de la posición que ocupe éste dentro del texto claro. En realidad corresponde a la aplicación cíclica de n cifrados monoalfabéticos. El ejemplo más típico de cifrado polialfabético es el *Cifrado de Vigenere* , que debe su nombre a Blaise de Vigenere, su creador, y que data del siglo XVI. La clave está constituida por una secuencia de símbolos $K = \{k_0, k_1, \dots, k_{d-1}\}$, y que emplea la siguiente función de cifrado: $E_k(m_i) = m_i + k(i \bmod d) \pmod n$ siendo m_i el i-ésimo símbolo de texto claro y n el cardinal del alfabeto de entrada.

Algoritmo DES

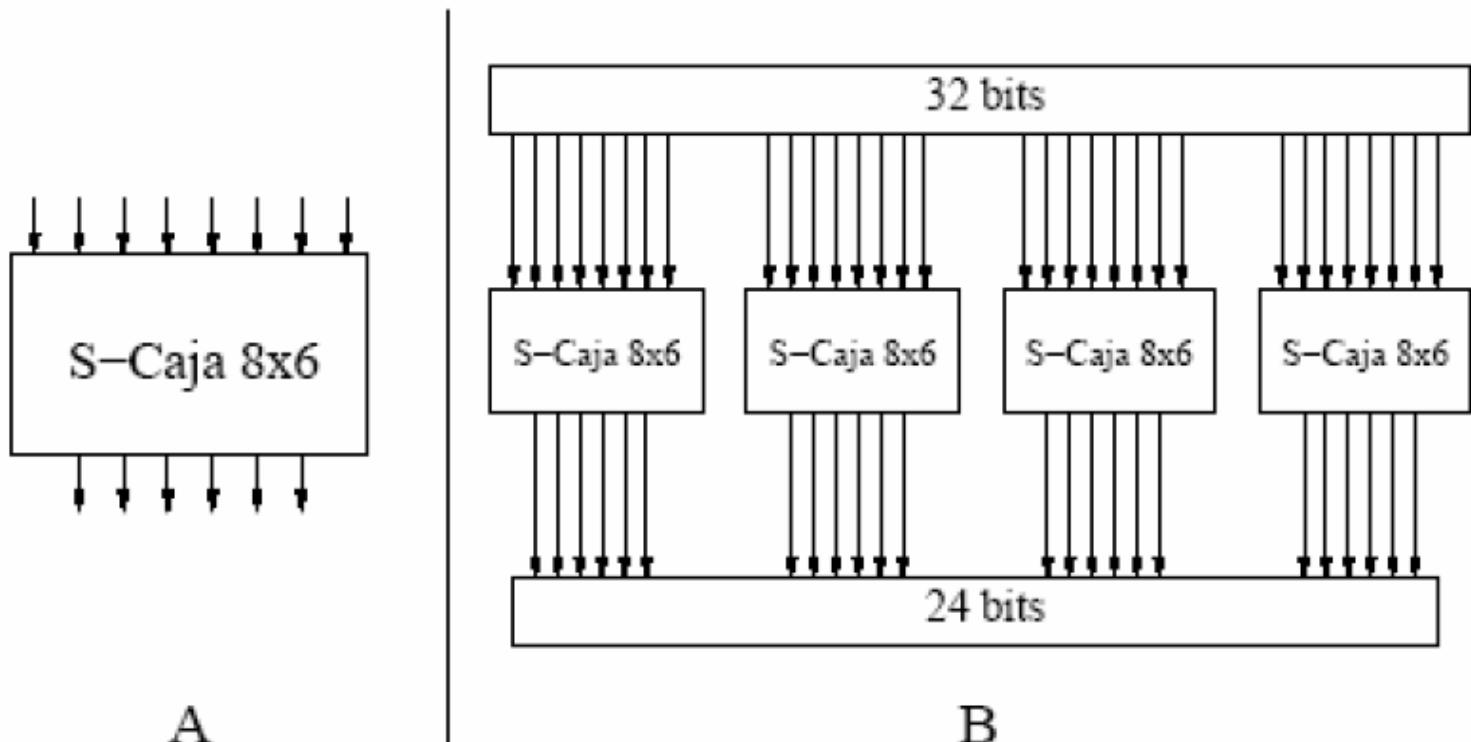
El **algoritmo DES** es el algoritmo simétrico más extendido mundialmente. Se basa en el algoritmo LUCIFER, desarrollado por IBM a principios de los setenta, y adoptado como estándar por el Gobierno de los EE.UU. para comunicaciones no clasificadas en 1976.

El algoritmo DES fue diseñado por la NSA (A gencia N acional de S eguridad de los EE.UU.) para ser implementado por hardware, creyendo que los detalles iban a ser mantenidos en secreto, pero la Oficina Nacional de Estandarización publicó su especificación con suficiente detalle como para que cualquiera pudiera implementarlo por software. No fue casualidad que el siguiente algoritmo adoptado (Skipjack) fuera mantenido en secreto.

A mediados de 1998, se demostró que un ataque por la fuerza bruta a DES era viable, debido a la escasa longitud que emplea en su clave. No obstante, el algoritmo aún no ha demostrado ninguna debilidad grave desde el punto de vista teórico, por lo que su estudio sigue siendo plenamente interesante.

El algoritmo DES se basa en las denominadas S-Cajas. Una S-Caja de $m \times n$ bits es una tabla de sustitución que toma como entrada cadenas de m bits y da como salida cadenas de n bits.

(A) *S-Caja individual.* (B) *Combinación de cuatro S-Cajas.*



DES emplea ocho S-Cajas de 6*4 bits. La utilización de las S-Cajas es sencilla: se divide el bloque original en trozos de m bits y cada uno de ellos se sustituye por otro de n bits,

haciendo uso de la S-Caja correspondiente. Normalmente, cuanto más grandes sean las S-Cajas, más resistente sería el algoritmo resultante, aunque la elección de los valores de salida para que den lugar a un buen algoritmo no es en absoluto trivial.

Las características básicas del algoritmo DES son:

- Codifica secuencialmente **bloques de 64 bits**.
- Emplea **claves de 56 bits**.
- Se usa el **mismo algoritmo tanto para cifrar como para descifrar**.

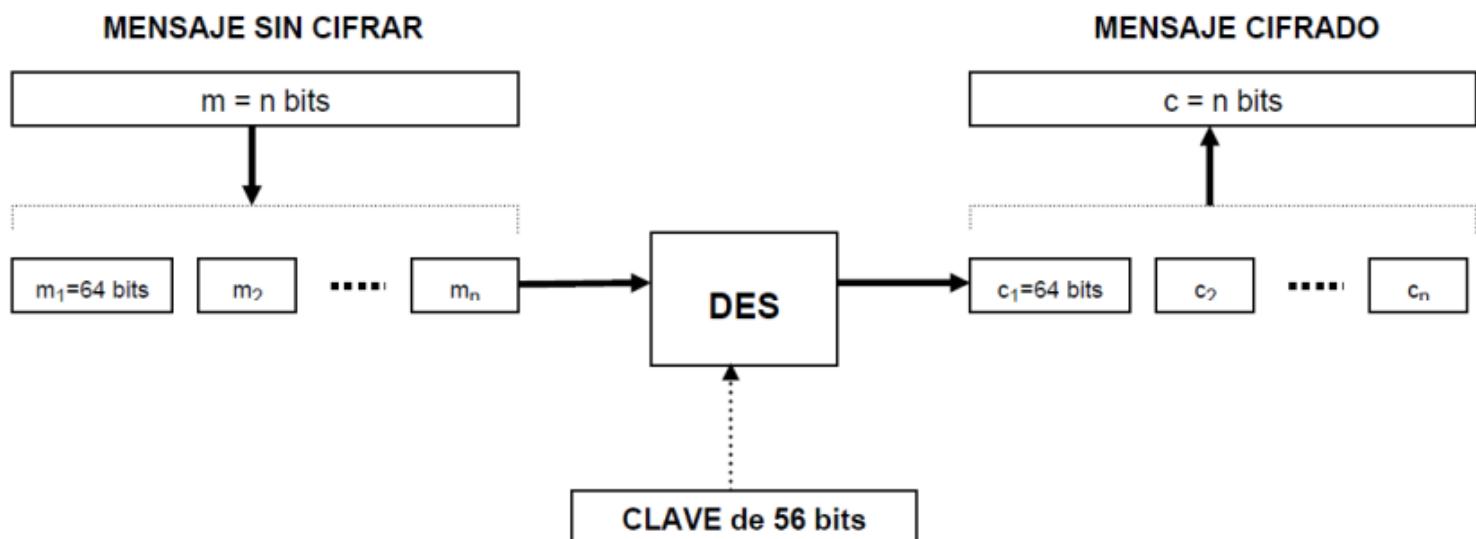
Es decir, el mensaje original se trocea en bloques o submensajes de 64 bits (8 bytes) de tamaño, que son las unidades que procesa el DES, y luego se unen a la salida para formar el mensaje cifrado.

Gráficamente:

m_1, m_2, \dots, m_n : Bloques de 64 bits que conforman el mensaje cifrado.

c_1, c_2, \dots, c_n : Bloques cifrados resultado de aplicar DES a los bloques anteriores.

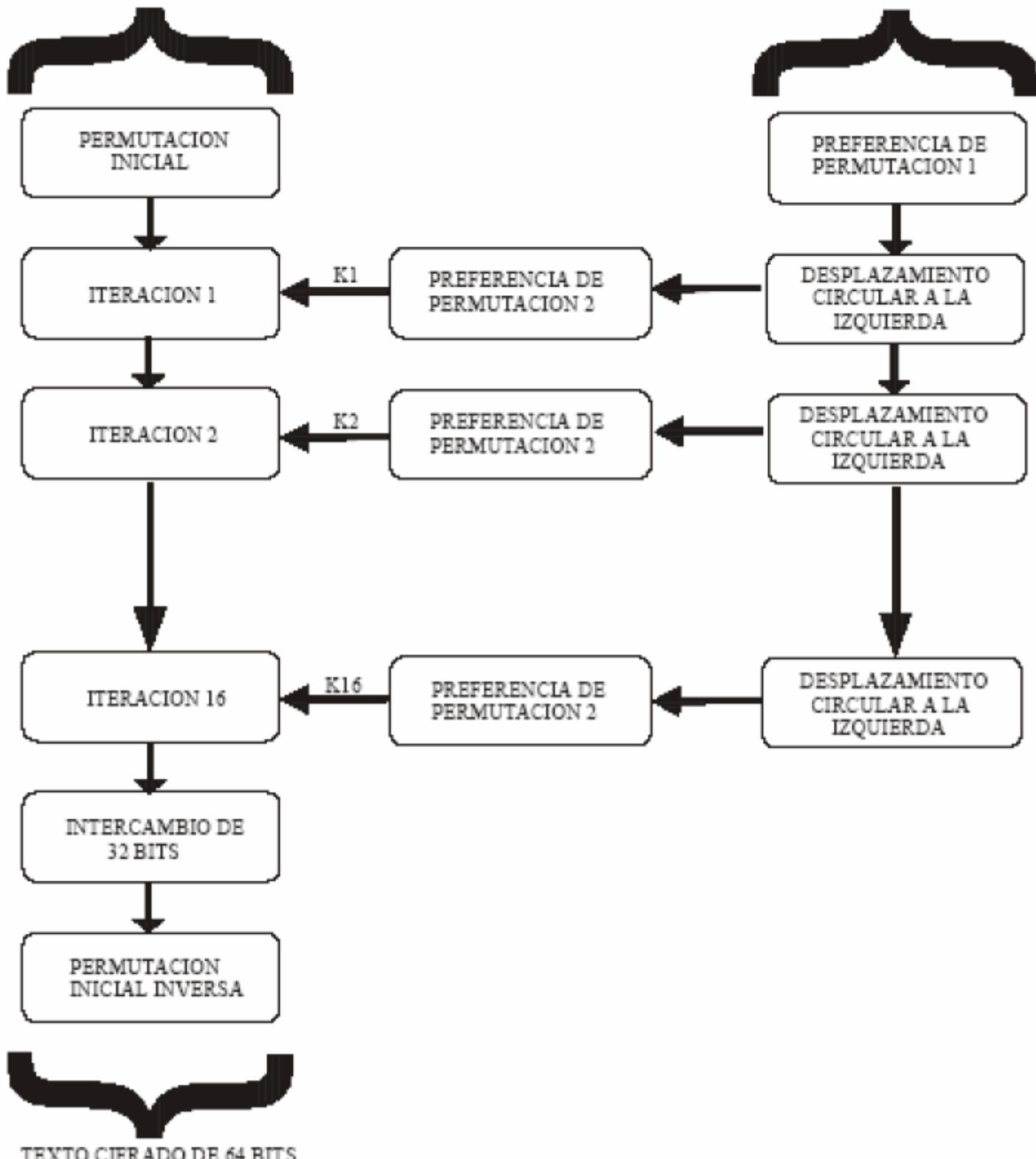
Funcionamiento global del algoritmo DES.



Funcionamiento específico del algoritmo DES.

TEXTO NATIVO DE 64 BITS

CLAVE DE 56 BITS



El texto plano o nativo debe tener una longitud de 64 bits y la clave 56 bits; los textos nativos más grandes se procesan en bloques de 64 bits. La parte izquierda de la figura muestra que el procesamiento del texto nativo se realiza en tres partes:

- Los 64 bits de texto nativo se transforman por medio de una permutación inicial que reordena a los bits para producir la entrada permutada.
- Luego sigue una fase de 16 iteraciones de la misma función. La salida de la última iteración consta de 64 bits que son función de la clave y del texto nativo. A continuación se intercambia la mitad derecha con la mitad izquierda para producir la salida previa.

- La salida previa se permuta con la inversa de la función de permutación inicial para producir los 64 bits de texto cifrado.

La parte derecha de la figura muestra como se usan los 56 bits de la clave. Inicialmente se transforma la clave por una función de permutación. Luego se produce una subclave k_i para cada una de las 16 iteraciones por medio de un desplazamiento circular y una permutación. La función de permutación es la misma para las 16 iteraciones; pero se produce una subclave distinta para cada una debido al desplazamiento circular de los bits de la clave.

El algoritmo DES presenta algunas claves débiles. En general, todos aquellos valores de la llave que conducen a una secuencia inadecuada de k_i serán poco recomendables. Distinguiremos entre claves débiles, que son aquellas que generan un conjunto de diecisésis valores iguales de k_i y que cumplen $E_k(E_k(M)) = M$, y claves semidébiles, que generan dos valores diferentes de k_i , cada uno de los cuales aparece ocho veces. En cualquier caso, el número de llaves de este tipo es tan pequeño en comparación con el número total de posibles claves, que no debe suponer un motivo de preocupación.

Variantes del Algoritmo DES

El algoritmo DES padece de un problema que no radica en su diseño, sino en que emplea una clave demasiado corta (56 bits), lo cual hace que con el avance actual de las computadoras los ataques por la fuerza bruta comiencen a ser opciones realistas. Mucha gente se resiste a abandonar este algoritmo, precisamente porque ha sido capaz de sobrevivir durante veinte años sin mostrar ninguna debilidad en su diseño, y prefieren proponer variantes que, de un lado evitarían el riesgo de tener que confiar en algoritmos nuevos, y de otro permitirían aprovechar gran parte de las implementaciones por hardware existentes de DES. De ahí que se hayan desarrollado múltiples variantes del DES que alivien el problema de la clave demasiado corta pero que permitan aprovechar la gran cantidad de implantación hardware que existe.

- **DES Múltiple** : Consiste en aplicar varias veces el algoritmo DES con diferentes claves al mensaje original. El más común de todos ellos es el Triple-DES, cuya longitud de clave es de 112 bits.
- **DES con Subclaves Independientes** : Consiste en emplear subclaves diferentes para cada una de las 16 rondas de DES. Puesto que estas subclaves son de 48 bits, la clave resultante tendría 768 bits en total. Sin pretender entrar en detalles, puede demostrarse empleando criptoanálisis diferencial que esta variante podría ser rota con 261 textos claros escogidos, por lo que en la práctica no presenta un avance sustancial sobre DES estándar.
- **DES Generalizado** : Esta variante emplea n trozos de 32 bits en cada ronda en lugar de dos, aumentando tanto la longitud de la clave como el tamaño de mensaje que se puede codificar, manteniendo sin embargo el orden de complejidad del algoritmo. No sólo se gana poco en seguridad, sino que en muchos casos se pierde.

Algoritmo IDEA

El **algoritmo IDEA** (International Data Encryption Algorithm) es más joven que DES, pues data de 1992. Para muchos constituye el mejor y más seguro algoritmo simétrico disponible en la actualidad. Sus características son:

- Trabaja con **bloques de 64 bits** de longitud.
- Emplea una **clave de 128 bits**.
- Como en el caso de DES, se usa el mismo **algoritmo tanto para cifrar como para descifrar**.

IDEA es un algoritmo seguro, y hasta ahora resistente a multitud de ataques, entre ellos el criptoanálisis diferencial. No presenta claves débiles, y su longitud de clave hace imposible en la práctica un ataque por la fuerza bruta. Como todos los algoritmos

simétricos de cifrado por bloques, IDEA se basa en los conceptos de confusión difusión, haciendo uso de las siguientes operaciones elementales (todas ellas fáciles de implantar):

- XOR.
- Suma módulo 216.
- Producto módulo 216 + 1.

Como idea general, diremos que el algoritmo IDEA consta de ocho rondas. Dividiendo el bloque X a codificar, de 64 bits, en cuatro partes X1, X2, X3 y X4 de 16 bits. Para la interpretación entera de dichos registros se emplea el criterio *big endian*, lo cual significa que el primer byte es el más significativo.

Las primeras ocho subclaves se calculan dividiendo la clave de entrada en bloques de 16 bits. Las siguientes ocho se calculan rotando la clave de entrada 25 bits a la izquierda y volviendo a dividirla, y así sucesivamente.

Algoritmo Rijndael (AES)

En octubre de 2000 el NIST (National Institute for Standards and Technology) anunciaba oficialmente la adopción del algoritmo Rijndael como nuevo Estándar Avanzado de Cifrado (AES) para su empleo en aplicaciones critográficas no militares, culminando así un proceso demás de tres años, encaminado a proporcionar a la comunidad internacional un nuevo algoritmo de cifrado potente, eficiente, y fácil de implementar.

La palabra Rijndael -en adelante, emplearemos la denominación AES- es un acrónimo formado por los nombres de sus dos autores, los belgas *Joan Daemen* y *Vincent Rijmen*. Su interés radica en que todo el proceso de selección, revisión y estudio tanto de este algoritmo como de los restantes candidatos, se ha efectuado de forma pública y abierta, por lo que, prácticamente por primera vez, toda la comunidad criptográfica mundial ha participado en su análisis, lo cual convierte a Rijndael en un algoritmo digno de la confianza de todos.

AES es un sistema de cifrado por bloques, diseñado para manejar longitudes de clave y de bloque variables, ambas comprendidas entre los 128 y los 256 bits. Realiza varias de sus operaciones internas a nivel de byte, interpretando éstos como elementos de un cuerpo de Galois GF(28) (un cuerpo de Galois -Matemático Francés del siglo XVIII- es un tipo de estructura de números perteneciente a la teoría de Grupos). El resto de operaciones se efectúan en términos de registros de 32 bits. Sin embargo, en algunos casos, una secuencia de 32 bits se toma como un polinomio de grado inferior a 4, cuyos coeficientes son a su vez polinomios en GF(28).

Si bien, este algoritmo soporta diferentes tamaños de bloque y clave, en el estándar adoptado por el Gobierno Estadounidense en noviembre de 2001 (FIPS PUB 197), se especifica:

- Longitud fija de **bloque de 128 bits**.
- Longitud de **clave a escoger entre 128, 192 y 256 bits**.

Algoritmos Asimétricos

Históricamente el problema de distribución de claves siempre ha sido la parte débil de la mayoría de los criptosistemas. Sin importar lo robusto que sea un criptosistema, si un intruso puede robar la clave, el sistema no vale nada. Dado que todos los criptólogos siempre daban por hecho que la clave de cifrado y la clave de descifrado eran la misma (o que se podía derivar fácilmente una de la otra) y que la clave tenía que distribuirse a todos los usuarios del sistema, parecía haber un problema inherente: las claves se tenían que proteger contra robo, pero también se tenían que distribuir, por lo que no podían simplemente guardarse en una caja fuerte.

En 1976, dos investigadores de Stanford, Diffie y Hellman (1976), propusieron una clase nueva de criptosistema, en el que:

- Las claves de cifrado y descifrado son diferentes.
- La clave de descifrado no puede derivarse de la clave de cifrado, y viceversa.

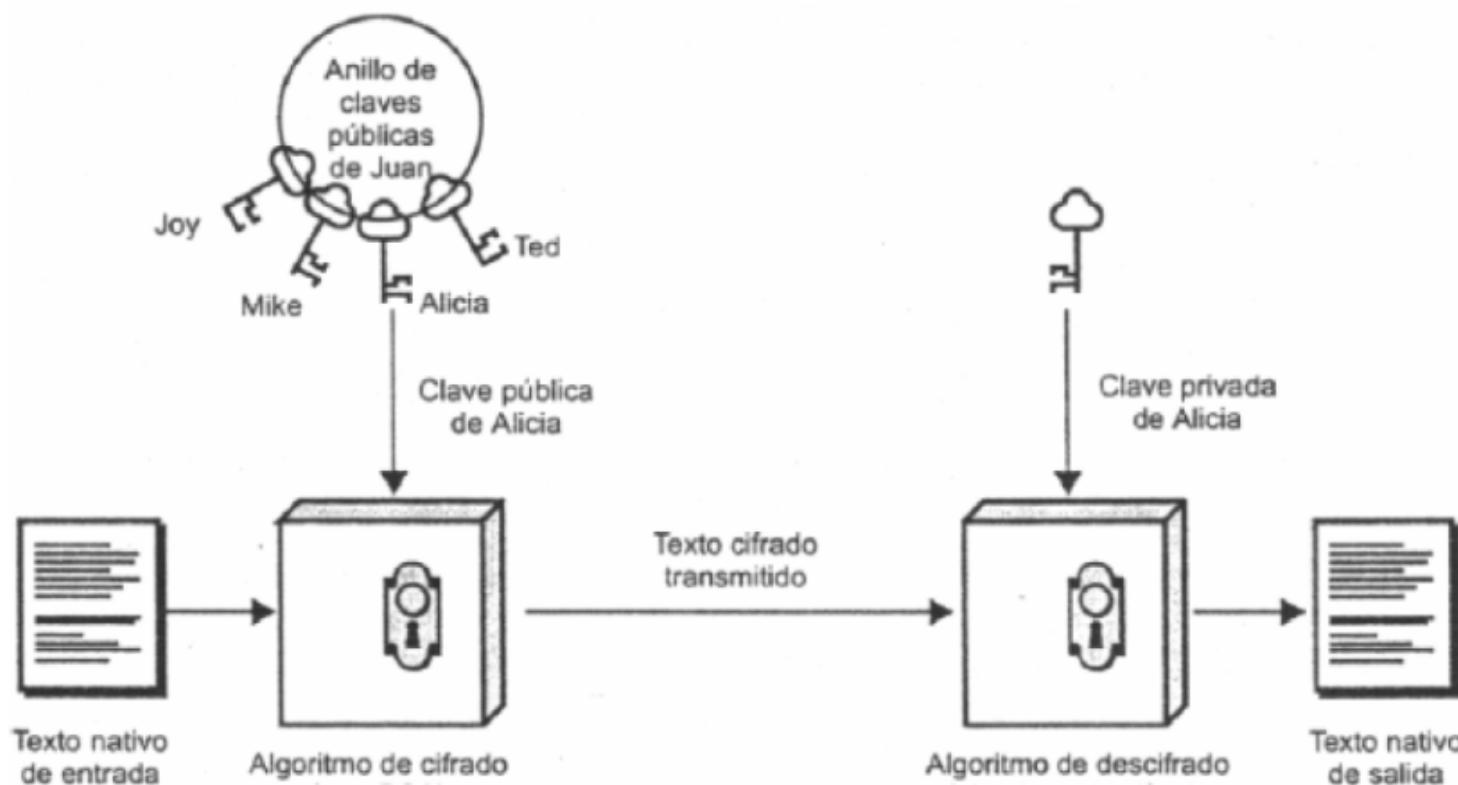
El algoritmo de cifrado (con clave), E, y el algoritmo de descifrado (con clave), D, tenían que cumplir 3 requisitos:

- $D(E(m)) = m$.
- Es extraordinariamente difícil deducir D de E.
- E no puede descifrarse mediante un ataque de texto normal seleccionado.

El primer requisito dice que, si aplicamos D a un mensaje cifrado, $E(m)$, obtenemos nuevamente el mensaje de texto normal original, m . El segundo requisito no requiere explicación. El tercer requisito es necesario porque, como veremos en un momento, los intrusos pueden experimentar a placer con el algoritmo. En estas condiciones, no hay razón para que una clave de cifrado no pueda hacerse pública.

El método funciona como sigue. Una persona, llamémosla Juan, que quiera recibir mensajes secretos, primero diseña dos algoritmos, Ea y Da , que cumplan los requisitos anteriores. El algoritmo de cifrado y la clave, Ea , se hacen públicos, de ahí el nombre de criptografía de clave pública (para contrastar con la criptografía tradicional de clave secreta). Esto podría hacerse poniéndolos en un archivo accesible a cualquiera que quiera leerlo. Alicia publica el algoritmo de descifrado, pero mantiene secreta la clave de descifrado. Por tanto, Ea , es pública, pero Da es secreta. Ahora veamos si podemos resolver el problema de establecer un canal seguro entre Juan y Alicia, que nunca han tenido contacto previo. Se supone que tanto la clave de cifrado de Juan, Ea , como la clave de cifrado de Alicia, Eb , están en un archivo de lectura pública. (Básicamente, se espera que todos los usuarios de la red publiquen sus claves de cifrado tan pronto como se vuelven usuarios de la red).

Idea básica de los conceptos de clave pública y privada.



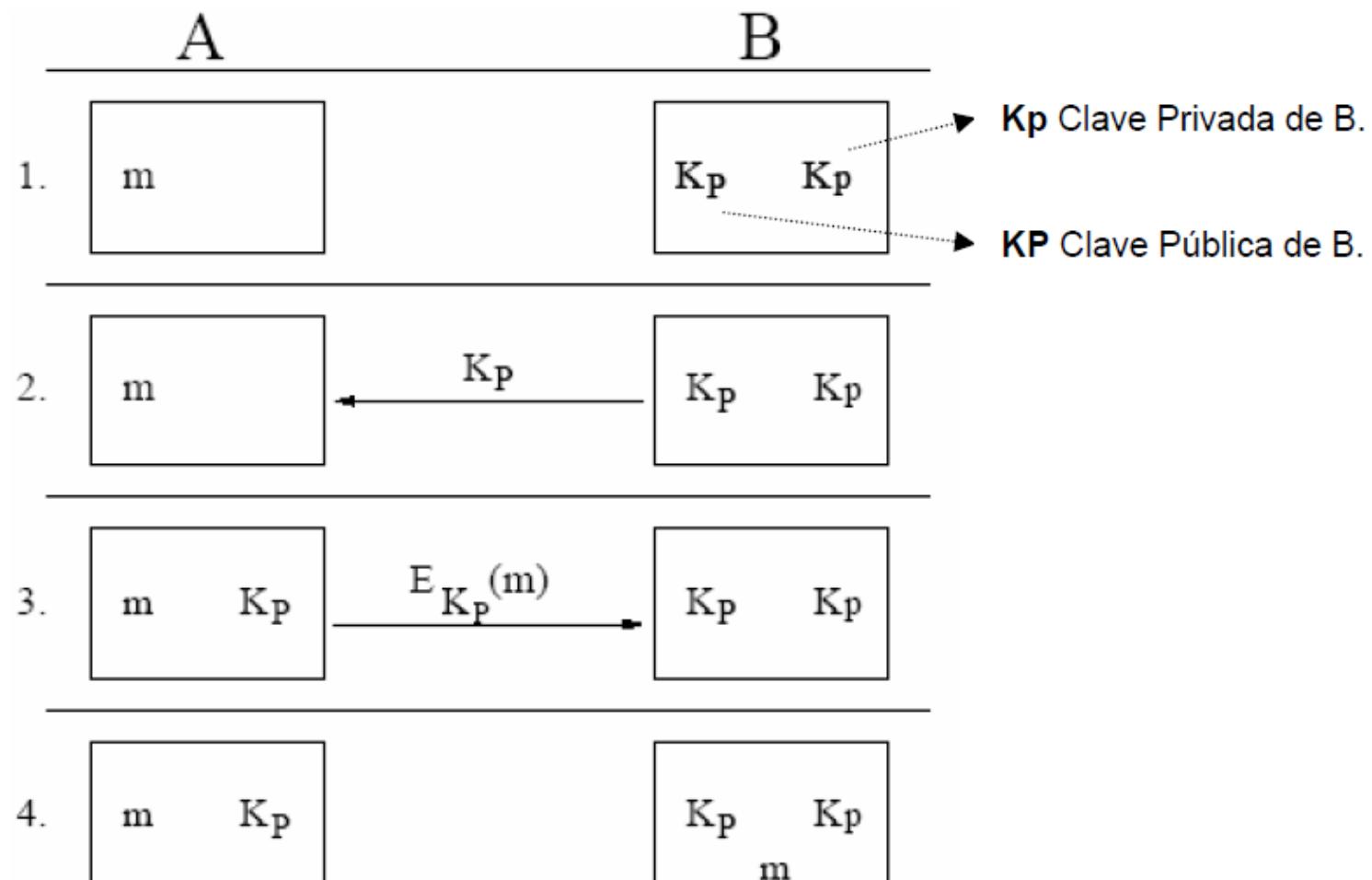
Llegado a este punto, Juan toma su primer mensaje, m , calcula $Eb(m)$ y lo envía a Alicia. Alicia entonces lo descifra aplicando su clave secreta Db [es decir, calcula $Db(Eb(m)=m)$]. Nadie más puede leer el mensaje cifrado, $Eb(m)$, porque se supone que el sistema de cifrado es robusto y porque es demasiado difícil derivar Db de la Eb públicamente conocida. Alicia y Juan ahora se pueden comunicar con seguridad.

Es útil una nota sobre terminología. La criptografía de clave pública requiere que cada usuario tenga dos claves: una clave pública, usada por todo el mundo para cifrar mensajes a enviar a ese usuario, y una clave privada, que necesita el usuario para descifrar los mensajes. Consistentemente nos referiremos a estas claves como *claves públicas y privadas*, respectivamente, y las distinguiremos de las claves *secretas* usadas tanto para cifrado como descifrado en la criptografía convencional de clave simétrica.

Estos algoritmos de llave pública, o algoritmos asimétricos, han demostrado su interés para ser empleados en redes de comunicación inseguras (Internet). Hasta la fecha han aparecido multitud de algoritmos asimétricos, la mayoría de los cuales son inseguros; otros son poco prácticos, bien sea porque el criptograma es considerablemente mayor que el mensaje original, bien sea porque la longitud de la clave es enorme. Se basan en general en plantear al atacante problemas matemáticos difíciles de resolver.

La única dificultad estriba en que necesitamos encontrar algoritmos que realmente satisfagan los tres requisitos indicados anteriormente. Debido a las ventajas potenciales de la criptografía de clave pública, muchos investigadores están trabajando día y noche, y ya se han publicado algunos algoritmos. En la práctica muy pocos algoritmos son realmente útiles. El más popular por su sencillez es RSA, que ha sobrevivido a multitud de ataques, si bien necesita una longitud de clave considerable. Otros algoritmos son los de ElGamal y Rabin.

Funcionamiento genérico de la criptografía asimétrica.



Expresado de manera formal, tal como muestra la figura, el proceso de transmisión de información empleando algoritmos asimétricos es el siguiente:

- **Paso 1** : A tiene el mensaje m y quiere enviárselo a B .
- **Paso 2** : B envía a A su clave pública, KP .
- **Paso 3** : A codifica el mensaje m y envía a B el criptograma $EKP(m)$.
- **Paso 4** : B decodifica el criptograma empleando la clave privada Kp .

Los algoritmos asimétricos emplean generalmente longitudes de clave mucho mayores que los simétricos. Por ejemplo, mientras que para algoritmos simétricos se considera segura una clave de 128 bits, para algoritmos asimétricos -si exceptuamos aquellos basados en curvas elípticas- se recomiendan claves de al menos 1024 bits. Además, la complejidad de cálculo que comportan estos últimos lo hacen considerablemente más lentos que los algoritmos de cifrado simétricos. En la práctica los métodos asimétricos se emplean únicamente para codificar la clave simétrica de cada mensaje o transacción particular.

A modo de resumen de lo dicho anteriormente, es importante recordar que:

- Los algoritmos asimétricos poseen 2 claves diferentes: **Kp (Clave privada) y KP (Clave pública)** .
- Se emplea **una de ellas para codificar** , mientras que **la otra se usa para decodificar** .
- **Dependiendo de la aplicación** que demos al algoritmo, **la clave pública será la de cifrado o viceversa** .
- Para que estos criptosistemas sean seguros ha de cumplirse que **a partir de una de las claves resulte extremadamente difícil calcular la otra** .

Algoritmo RSA

El algoritmo RSA, cuyo nombre deriva de las iniciales de sus tres descubridores (Rivest, Shamir, Adleman), se basa en ciertos principios de la teoría de los números. Hay cuatro pasos previos para encontrar un algoritmo RSA, estos son:

- **Paso 1** : Seleccionar dos números primos grandes, p y q (generalmente mayores que 10 elevado a 100).
- **Paso 2** : Calcular $[n = p \times q]$ y $[z = (p - 1) \times (q - 1)]$.
- **Paso 3** : Seleccionar un número primo con respecto a z , llamándolo d .
- **Paso 4** : Encontrar e tal que $e \times d = 1 \text{ mod } z$.

Con estos parámetros calculados por adelantado, estamos listos para comenzar el cifrado. Dividimos el texto normal (considerado como una cadena de bits) en bloques, para que cada mensaje de texto normal, P , caiga en el intervalo $0 \leq P < n$. Esto puede hacerse agrupando el texto normal en bloques de k bits, donde k es el entero más grande para el que $2k < n$ es verdad.

Para cifrar un mensaje, m , calculamos $C = m \text{ elevado a } e \text{ (mod } n\text{)}$. Para descifrar C , calculamos $m = C \text{ elevado a } d \text{ (mod } n\text{)}$. Puede demostrarse que, para todos los m del intervalo especificado, las funciones de cifrado y descifrado son inversas. Para ejecutar el cifrado, se necesitan e y n . Para llevar a cabo el descifrado, se requieren d y n . Por tanto, **la clave pública consiste en el par (e, n) , y la clave privada consiste en (d, n)** .

La seguridad del método se basa en la dificultad para factorizar números grandes. Si el criptoanalista pudiera factorizar n (conocido públicamente), podría encontrar p y q y, a partir de éstos, z . Equipado con el conocimiento de z y de e , puede encontrar d usando el algoritmo de Euclides. Afortunadamente, los matemáticos han estado tratando de factorizar números grandes durante los últimos 300 años, y las pruebas acumuladas sugieren que se trata de n problema excesivamente difícil.

De acuerdo con Rivest y colegas, suponiendo el uso del mejor algoritmo conocido y de una computadora con un tiempo de instrucción de 1 microsegundo, la factorización de un número de 200 dígitos requiere 4.000 millones de años de tiempo de cómputo (la edad de la tierra se estima en 4.500 millones de años); la factorización de un número de 500 dígitos requeriría la astronómica cifra de 10e25 años. Aun si las computadoras continúan aumentando su velocidad en un orden de magnitud cada década, pasarán siglos antes de que sea factible la factorización de un número de 500 dígitos, y para entonces nuestros descendientes simplemente pueden escoger un p y un q todavía más grandes.

La aplicación más inmediata de los algoritmos asimétricos, obviamente, es el cifrado de la información sin tener que transmitir la clave de decodificación, lo cual permite su uso en canales inseguros.

Algoritmo DIFFIE-HELLMAN

El **algoritmo Diffie-Hellman** es un algoritmo asimétrico basado, como su nombre indica, en el problema matemático de Diffie-Hellman, que se emplea fundamentalmente para acordar una clave común entre dos interlocutores, a través de un canal de comunicación inseguro. La ventaja de este sistema es que no son necesarias llaves públicas en el sentido estricto, sino una información compartida por los dos comunicantes.

Algoritmo el ELGAMAL

El **algoritmo ELGAMAL** fue ideado en un principio para producir firmas digitales, pero después se extendió también para codificar mensajes. Se basa en el problema de los logaritmos discretos, que está íntimamente relacionado con el de la factorización, y en el de Diffie-Hellman. Para generar un par de llaves, se escoge un número primo n y dos números aleatorios p y x menores que n . Se calcula entonces:

$$y = p \text{ elevado } x \pmod{n}$$

La **llave pública es (p, y, n)**, mientras que **la llave privada es x** . Escogiendo n primo, garantizamos que sea cual sea el valor de p , el conjunto $\{p, p^2, p^3, \dots\}$ es una permutación del conjunto $\{1, 2, \dots, n-1\}$.

Nótese que esto no es necesario para que el algoritmo funcione, por lo que podemos emplear realmente un n no primo, siempre que el conjunto generado por las potencias de p sea lo suficientemente grande.

Algoritmo de RABIN

El sistema de llave asimétrica de **RABIN** se basa en el problema de calcular raíces cuadradas módulo con un número compuesto. Este problema se ha demostrado que es equivalente al de la factorización de dicho número. En primer lugar escogemos dos números primos, p y q , ambos congruentes con 3 módulo 4 (los dos últimos bits a 1). **Estos primos son la clave privada. La clave pública es su producto, $n = pq$** . Para codificar un mensaje m , simplemente se calcula: $c = me^2 \pmod{n}$.

Algoritmo DSA

El **algoritmo DSA** (Digital Signature Algorithm) es una parte del estándar de firma digital **DSS** (Digital Signature Standard). Este algoritmo, propuesto por el NIST, data de 1991, es una variante del método asimétrico de ElGamal.

Dispersión de Claves

La potencia de cualquier sistema de cifrado se apoya en una técnica de distribución de claves. En concreto, de nada serviría el sistema asimétrico de claves si las propias claves privadas no tuviesen una distribución segura.

La distribución de claves se puede efectuar de varias formas. Para dos partes A y B:

- **Opción 1** : A puede seleccionar una clave y entregársela físicamente a B .
- **Opción 2** : Una tercera parte selecciona la clave y la entrega físicamente a A y a B .
- **Opción 3** : Si A y B han utilizado previamente y recientemente una clave, una de las partes podría transmitir la nueva clave a la otra cifrada utilizando la clave previa.
- **Opción 4** : Si A y B tienen cada uno una conexión cifrada a una tercera parte C , C podría entregar una clave a través de los enlaces cifrados a A y a B .

La opción 1 y 2 son razonables para cifrado de enlace ya que cada dispositivo de cifrado va a intercambiar datos con su pareja en el otro extremo del enlace. La opción 3 es válida tanto para cifrado de enlace como para cifrado extremo a extremo; pero si un agresor llegara a conseguir una clave, todas las claves siguientes serán reveladas.

La opción 4 es la preferible para proporcionar claves de extremo a extremo. Se identifican dos clases de claves:

- Clave de sesión: Cuando dos sistemas finales desean comunicarse, establecen una conexión lógica. Durante la duración de dicha conexión, todos los datos de usuario se cifran con una clave de sesión de un solo uso. Terminada la conexión, la clave se destruye.
- Clave permanente: Es la clave usada entre entidades con el objetivo de distribución de claves de sesión.

Criptoanálisis

A la hora de atacar un texto cifrado, existen dos formas de hacerlo:

- **Criptoanálisis** : El criptoanálisis, concepto ya presentado anteriormente, se basa en la naturaleza del algoritmo más algún conocimiento de las características generales del texto nativo o incluso de algunos pares texto nativo-texto cifrado. Este tipo de ataque explota las debilidades del algoritmo (si es que las tiene) o sus puntos menos fuertes para intentar deducir un texto nativo o deducir la clave que se está utilizando.
- **Fuerza bruta** : Consiste en probar cada clave posible en un trozo de texto cifrado hasta que se obtenga una traducción inteligible del texto nativo.

En este apartado nos centraremos únicamente en el criptoanálisis, pues es obvio que el ataque por fuerza bruta requiere poca consideración. Basta con programar un ordenador, y tiempo (la mayoría de las veces mucho) para llegar a dar con la clave, o claves de cifrado. Es obvio que cuanta mayor potencia tenga el ordenador, más rápido se producirá el descifrado. Hoy en día se utilizan redes de ordenadores trabajando juntos para probar la fortaleza de los algoritmos de cifrado sometidos a ataques por la fuerza. Aún así, ni con los ordenadores más potentes de hoy en día se conseguiría descifrar en un tiempo razonable ciertos algoritmos de cifrado modernos.

En cuanto al criptoanálisis, éste se comenzó a estudiar seriamente con la aparición de DES. Mucha gente desconfiaba del algoritmo propuesto por la NSA. Se dice que existen estructuras extrañas, que muchos consideran sencillamente puertas traseras colocadas por la Agencia para facilitar la descodificación de los mensajes. Nadie ha podido aún

demostrar ni desmentir este punto. El interés por buscar posibles debilidades en él ha llevado a desarrollar técnicas que posteriormente han tenido éxito con otros algoritmos.

Ni que decir tiene que estos métodos no han conseguido doblegar a DES, pero sí representan mecanismos significativamente más eficientes que la fuerza bruta para criptoanalizar un mensaje. Los dos métodos que vamos a comentar seguidamente parten de que disponemos de grandes cantidades de pares (texto claro-texto cifrado) obtenidos con la clave que queremos descubrir.

- **Criptoanálisis Diferencial.** Descubierto por Biham y Shamir en 1990, permite efectuar un ataque a DES, con texto claro escogido, que resulta más eficiente que la fuerza bruta. Se basa en el estudio de los pares de criptogramas que surgen cuando se codifican dos textos claros con diferencias particulares, analizando la evolución de dichas diferencias a lo largo de las rondas de DES. Para llevar a cabo un criptoanálisis diferencial se toman dos mensajes cualesquiera (incluso aleatorios) idénticos salvo en un número concreto de bits. Usando las diferencias entre los textos cifrados, se asignan probabilidades a las diferentes claves de cifrado. Conforme tenemos más y más pares, una de las claves aparece como la más probable. Esa será la clave buscada.
- **Criptoanálisis Lineal** . El criptoanálisis lineal, descubierto por Mitsuru Matsui, basa su funcionamiento en tomar algunos bits del texto claro y efectuar una operación XOR entre ellos, tomar algunos del texto cifrado y hacerles lo mismo, y finalmente hacer un XOR de los dos resultados anteriores, obteniendo un único bit. Efectuando esa operación a una gran cantidad de pares de texto claro y criptograma diferentes podemos ver si se obtienen más ceros o más unos. Si el algoritmo criptográfico en cuestión es vulnerable a este tipo de ataque, existirían combinaciones de bits que, bien escogidas, den lugar a un sesgo significativo en la medida anteriormente definida, es decir, que el número de ceros (o unos) es apreciablemente superior. Esta propiedad nos va a permitir poder asignar mayor probabilidad a unas claves sobre otras y de esta forma descubrir la clave que buscamos.

Mecanismos de Firma Digital

Métodos de Autenticación

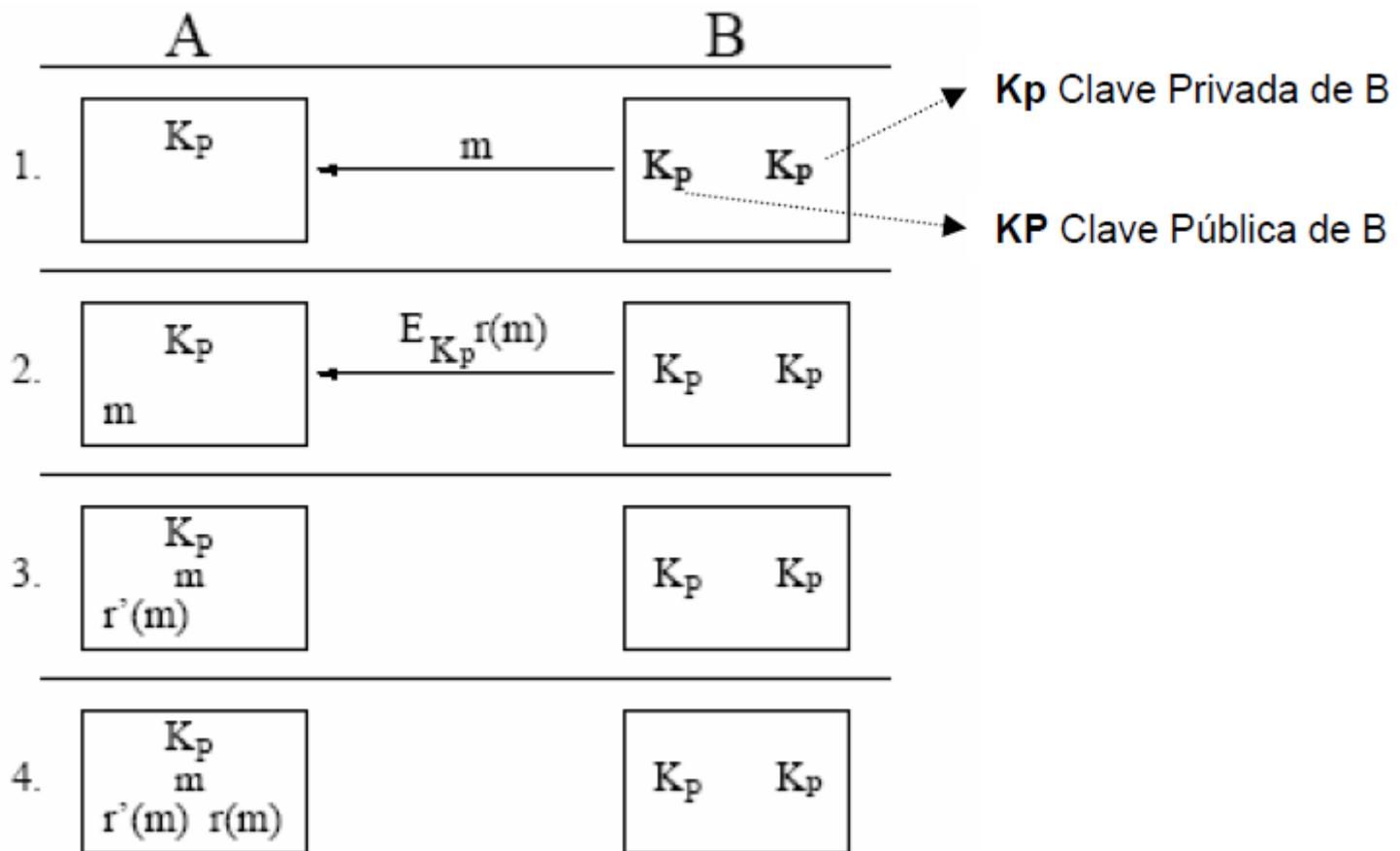
La segunda gran utilidad de los algoritmos asimétricos es la autenticación de mensajes, con ayuda de unas funciones llamadas “funciones resumen” que nos permitirán una **firma digital**, también denominada *signatura*, a partir de un mensaje.

A dicha firma hay que exigirle que cumpla:

- Ser mucho **más pequeña que el mensaje original**.
- Que sea **muy difícil encontrar otro mensaje que dé lugar a la misma**.

Veamos el siguiente ejemplo ilustrado en la siguiente figura. Supóngase que *A* recibe un mensaje *m* de *B* y quiere comprobar su autenticidad. Para ello *B* genera un resumen del mensaje *r(m)* y lo codifica empleando la clave de cifrado, que en este caso será privada.

Proceso de autentificación con algoritmo asimétrico.



La autentificación de información empleando algoritmos asimétricos tal como se indica en la figura precedente es:

- **Paso 1 :** A , que posee la clave pública KP de B , recibe el mensaje m y quiere autenticarlo.
- **Paso 2 :** B genera el resumen de m y envía a A el criptograma asociado $EKp(r(m))$.
- **Paso 3 :** A genera por su cuenta $r'(m)$ y decodifica el criptograma recibido usando la clave KP .
- **Paso 4 :** A compara $r(m)$ y $r'(m)$ para comprobar la autenticidad del mensaje m .

La clave de descifrado se habrá hecho pública previamente, y debe estar en poder de A . B envía entonces a A el criptograma correspondiente a $r(m)$. A puede ahora generar su propia $r'(m)$ y compararla con el valor $r(m)$ obtenido del criptograma enviado por B . Si coinciden, el mensaje será auténtico, puesto que el único que posee la clave para codificar es precisamente B .

Nótese que en este caso la clave que se emplea para cifrar es la clave privada, justo al revés que para la simple codificación de mensajes. En muchos de los algoritmos asimétricos ambas claves sirven tanto para cifrar como para descifrar, de manera que si empleamos una para codificar, la otra permitirá decodificar y viceversa. Esto ocurre con el algoritmo RSA, en el que un único par de claves es suficiente para codificar y autenticar.

Funciones de Dispersión Segura (HASH)

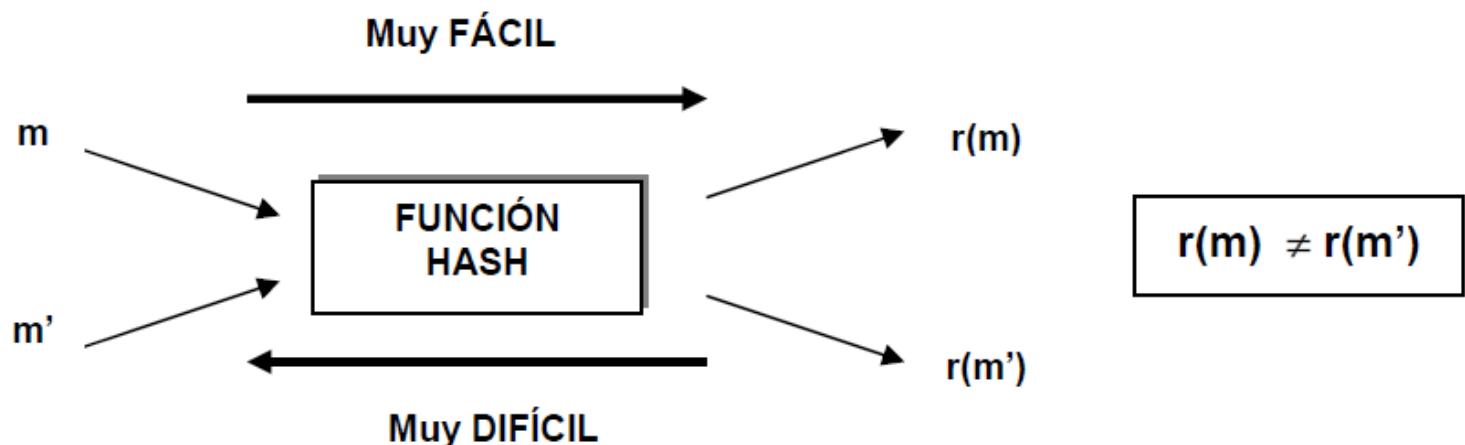
Acabamos de ver que la criptografía asimétrica permite autenticar información. Asimismo también que la autentificación debe hacerse empleando una función resumen y no codificando el mensaje completo. En esta sección vamos a estudiar dichas funciones resumen, también conocidas como **Funciones de Dispersión Segura** ; que nos van a permitir crear firmas digitales estudiadas en el capítulo anterior. Estas funciones también llamadas muy comúnmente **funciones Hash** , son, en esencia, *funciones matemáticas sin inversa* , que aplicadas a un elemento o dato que se transfiere impiden que este sea descifrado. Se utilizan para comprobar la integridad de los datos según un mecanismo por

el cual se cifra una cadena comprimida de los datos a transferir mediante una función Hash; este mensaje se envía al receptor junto con los datos ordinarios; el receptor repite la compresión y el cifrado posterior de los datos mediante la aplicación de la función Hash y compara el resultado obtenido con el que le llega, para verificar que los datos no han sido modificados. Para que sea segura, la función de dispersión segura o Hash $r(m)$, siendo m el mensaje, debe cumplir:

- **$r(m)$ es de longitud fija**, independientemente de la longitud de m .
- Dado m , es fácil calcular $r(m)$.
- Dado $r(m)$, es **computacionalmente intratable recuperar m** .
- Dado m , es **computacionalmente intratable obtener un m' tal que $r(m) = r(m')$** .

Gráficamente:

Características fundamentales de una función Hash.



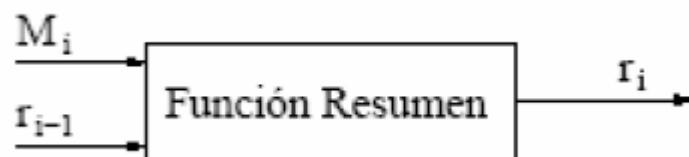
Existen dos tipos de funciones resumen, a saber:

Funciones MDC (M odification D etection C odes)

Como sabemos, un mensaje m puede ser autenticado codificando con la llave privada K_p el resultado de aplicarle una función resumen, $EK_p(r(m))$. Esa información adicional (que denominaremos firma o firma del mensaje m) sólo puede ser generada por el poseedor de la clave privada K_p . Cualquiera que tenga la llave pública correspondiente estará en condiciones de decodificar y verificar la firma.

En general, las funciones resumen se basan en la idea de funciones de compresión, que dan como resultado bloques de longitud n a partir de bloques de longitud mayor m . Estas funciones se encadenan de forma iterativa, haciendo que la entrada en el paso i sea función del i -ésimo bloque del mensaje y de la salida del paso $i-1$ (tal como mostramos en la siguiente figura).

Estructura iterativa de una función resumen.



En general, se suele incluir en alguno de los bloques del mensaje m -al principio o al final- información sobre la longitud total del mensaje. De esta forma se reducen las probabilidades de que dos mensajes con diferentes longitudes den el mismo valor en su resumen.

Conviene no confundir las funciones resumen con las funciones de relleno. Estas son funciones que generan texto cifrado continuamente, incluso en ausencia de texto nativo. Cuando hay disponible texto nativo, este se cifra y se transmite. En ausencia de texto nativo, los datos aleatorios se cifran y se transmiten. Esto hace imposible que un agresor distinga entre flujo de datos verdaderos y ruido, resultando imposible deducir la cantidad de tráfico.

Funciones MAC (Message Authentication Codes)

Frente a las MDC vistas, realmente existe otra clase de funciones resumen, llamada genéricamente **MAC** (Message Authentication Codes). Los MAC se caracterizan fundamentalmente por el empleo de una clave secreta para poder calcular la integridad del mensaje. Puesto que dicha clave sólo es conocida por el emisor y el receptor, el efecto conseguido es que el receptor puede, mediante el cálculo de dicha función, comprobar tanto la integridad como la procedencia del mensaje.

Algoritmos de Generación de Firma Digital

En este apartado vamos a estudiar dos algoritmos de generación de firmas muy utilizados: **MD5** y **SHA-1**.

Algoritmo MD5

El **algoritmo MD5** es el resultado de una serie de mejoras sobre el algoritmo MD4, diseñado por Ron Rivest. Las características básicas de este algoritmo son:

- Procesa los mensajes de entrada en bloques de 512 bits.
- Produce una salida de 128 bits.

Siendo m un mensaje de b bits de longitud, en primer lugar se alarga m hasta que su longitud sea exactamente 64 bits inferior a un múltiplo de 512. El alargamiento se lleva a cabo añadiendo un 1 seguido de tantos ceros como sea necesario. En segundo lugar, se añaden 64 bits con el valor de b , empezando por el byte menos significativo. De esta forma tenemos el mensaje como un número entero de bloques de 512 bits, y además le hemos añadido información sobre la longitud.

Algoritmo SHA-1

El **algoritmo SHA-1** fue desarrollado por la NSA, para ser incluido en el estándar **DSS** (Digital Signature Standard). Al contrario que los algoritmos de cifrado propuestos por esta organización, SHA-1 se considera seguro y libre de puertas traseras, ya que el hecho de que el algoritmo sea realmente seguro favorece a los propios intereses de la NSA. Produce firmas de 160 bits, a partir de bloques de 512 bits del mensaje original.

El algoritmo es similar a MD5, con la diferencia de que usa la ordenación *big endian*. Se inicializa de igual manera, es decir, añadiendo al final del mensaje un uno seguido de tantos ceros como sea necesario hasta completar 448 bits en el último bloque, para luego juxtaponer la longitud en bits del propio mensaje -en este caso, el primer byte de la secuencia será el más significativo-. A diferencia de MD5, SHA-1 emplea cinco registros de 32 bits en lugar de cuatro.

Cortafuegos

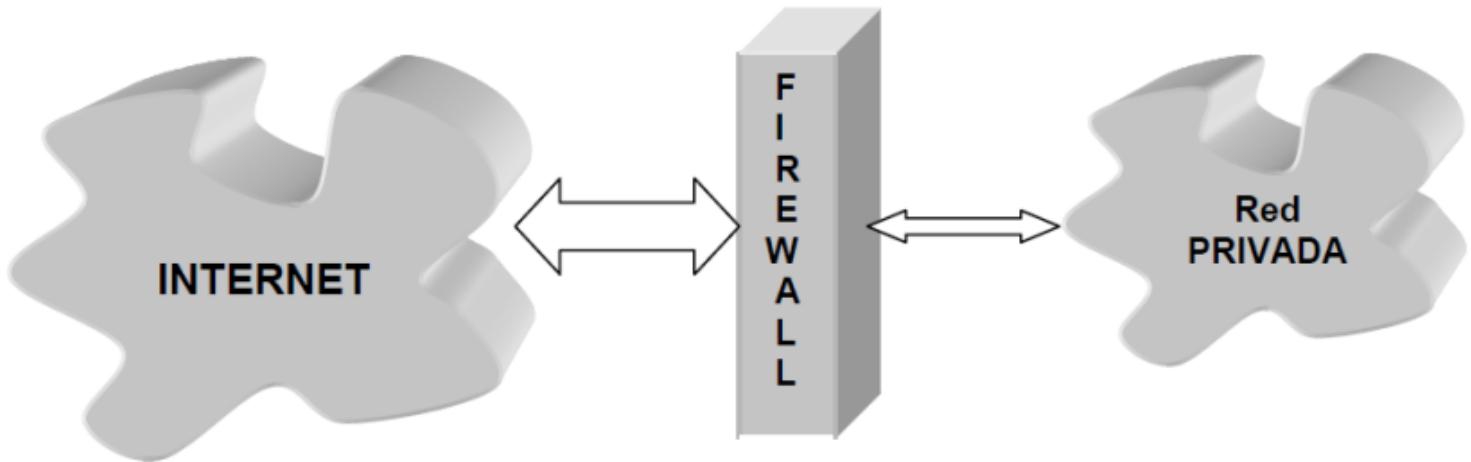
La tecnología de **Cortafuegos**, o **Firewalls** (muros de fuego) es relativamente nueva y se ha potenciado al comprobar que una red abierta como es Internet ha incorporado un

nuevo tipo de usuario no corporativo, y por tanto más difícil de controlar por las medidas y reglas implantadas en los propios "host" (potentes ordenadores corporativos que suelen tener las grandes empresas).

Estos cortafuegos fueron diseñados para impedir a los *Hackers* o intrusos que utilizan Internet el acceso a redes internas de las empresas. Algunos cortafuegos incluso controlan la información que se mueve por dichas redes.

Se define como **Tecnología de Cortafuegos** al sistema que controla todo el tráfico hacia o desde Internet utilizando software de seguridad o programas desarrollados para este fin, que están ubicados en un servidor u ordenador independiente .

Control de acceso a Internet por Cortafuegos (Firewall).



Este sistema comprueba que cada paquete de datos se encamine a donde debe, desde la red Internet a nuestra red privada y viceversa, al mismo tiempo que contiene la política de seguridad especificada por el Administrador del Sistema. Pueden ayudar asimismo a prevenir la entrada de virus encapsulados en los paquetes transmitidos con destino a la red privada. Se utiliza la expresión cortafuegos para designar pasarelas u otras estructuras más complejas, existentes entre la red propia e Internet, con la finalidad de restringir y filtrar el flujo de información entre ambas. Para prevenir o permitir el tráfico de red, comprueba el *host*, la red y la puerta desde la cual el paquete es originado o destinado. Para conectar directamente ordenadores de un sistema corporativo en red Internet, existe una aplicación que reside en el servidor para permitir un buen acceso a los servicios Internet facilitando al mismo tiempo la mayor seguridad que sea posible.

Este servidor comprueba que:

- El *host* desde el cual se **origina la conexión** .
- El *host* al cual la **conexión es solicitada** .
- Los **comandos que se producen** en la conexión.

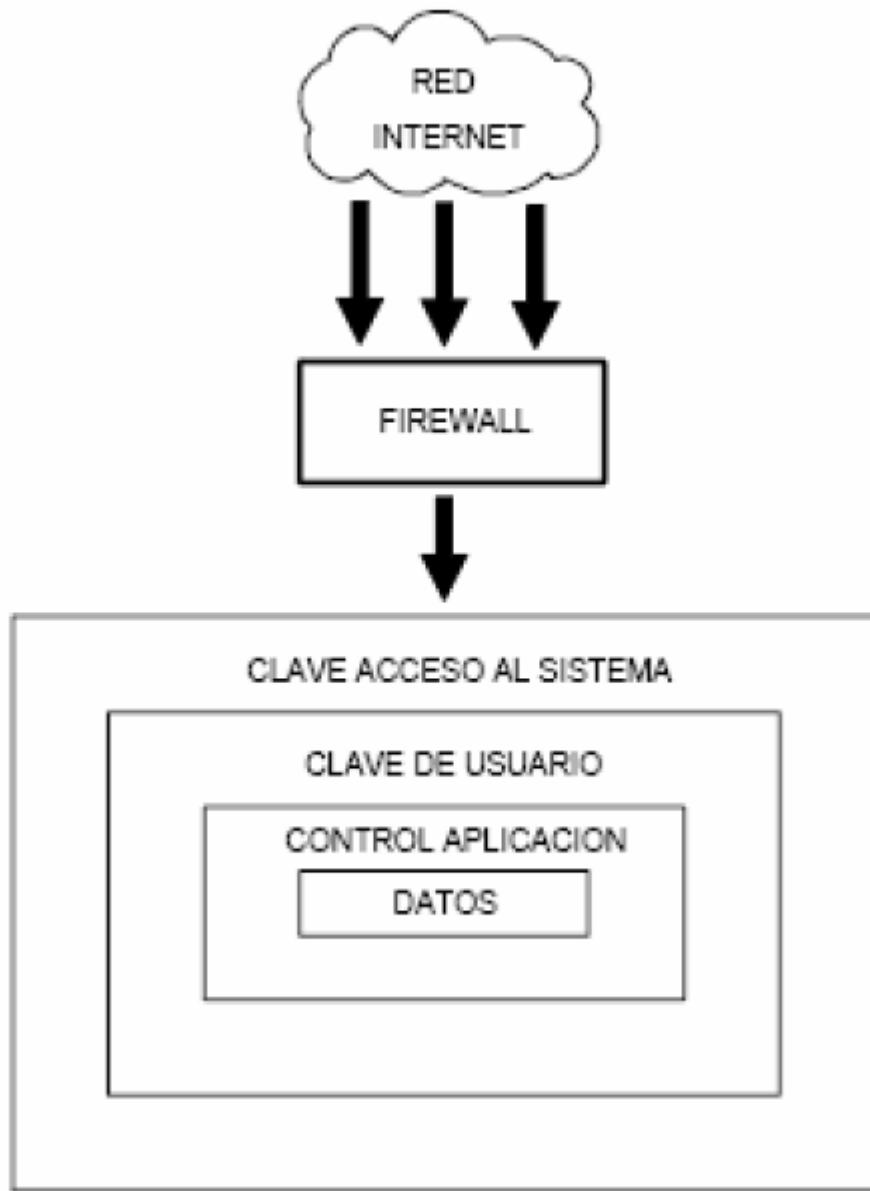
Todo ello puede facilitar al Administrador del Sistema la prevención de todas las conexiones desde *host* especificados o de redes en Internet. Desde esta puerta de entrada, el sistema puede también prevenirse de aquellos usuarios que a través de comandos añaden un factor de riesgo a nuestra seguridad. Se trata de prevenir, por ejemplo, la exportación de información contenida en el cortafuegos o en los servidores al exterior.

Mediante aplicaciones residentes en el servidor o cortafuegos se puede:

- Definir qué usuarios tienen palabra clave de acceso autorizada.
- Configurar las palabras clave de acceso que deben ser aceptadas por los diferentes hosts configurados en nuestra red privada.
- Controlar las cuentas de aplicación autorizadas.
- Evitar que la intrusión pueda cambiar la configuración de la aplicación residente.

- Controlar los accesos entre la red privada y el servidor como punto de entrada.
- Llevar un registro de todas las incidencias que se produzcan.

Barreras de protección de los datos frente a intrusos en Internet.

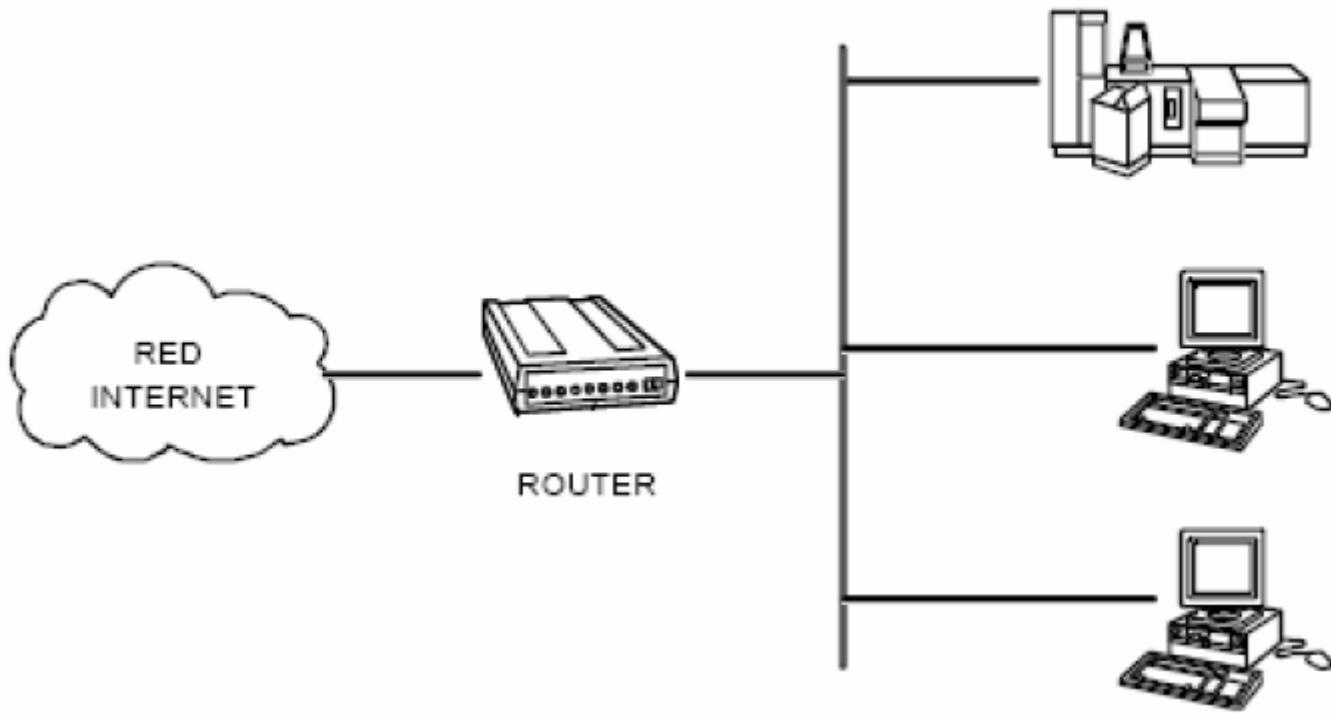


Los esquemas que se describen a continuación o cualquier combinación de ellas responden a distintas formas de implantar un cortafuegos.

Cortafuegos Filtrador de Paquete

El cortafuegos filtrador de paquetes se basa en un dispositivo denominado encaminador, o más comúnmente conocido por su nombre en inglés, router, que puede filtrar los paquetes de datos, tanto los que salen como los entrantes a la red de la empresa, con destino u origen en Internet. Es el sistema más sencillo de establecer conexión con Internet. También se pueden configurar los protocolos de filtrado para que permitan la entrada en la red sólo de determinados tipos de transmisiones o de aquellas que tengan su origen en emisiones redeterminados.

Esquema de un cortafuego por filtrado de paquetes.



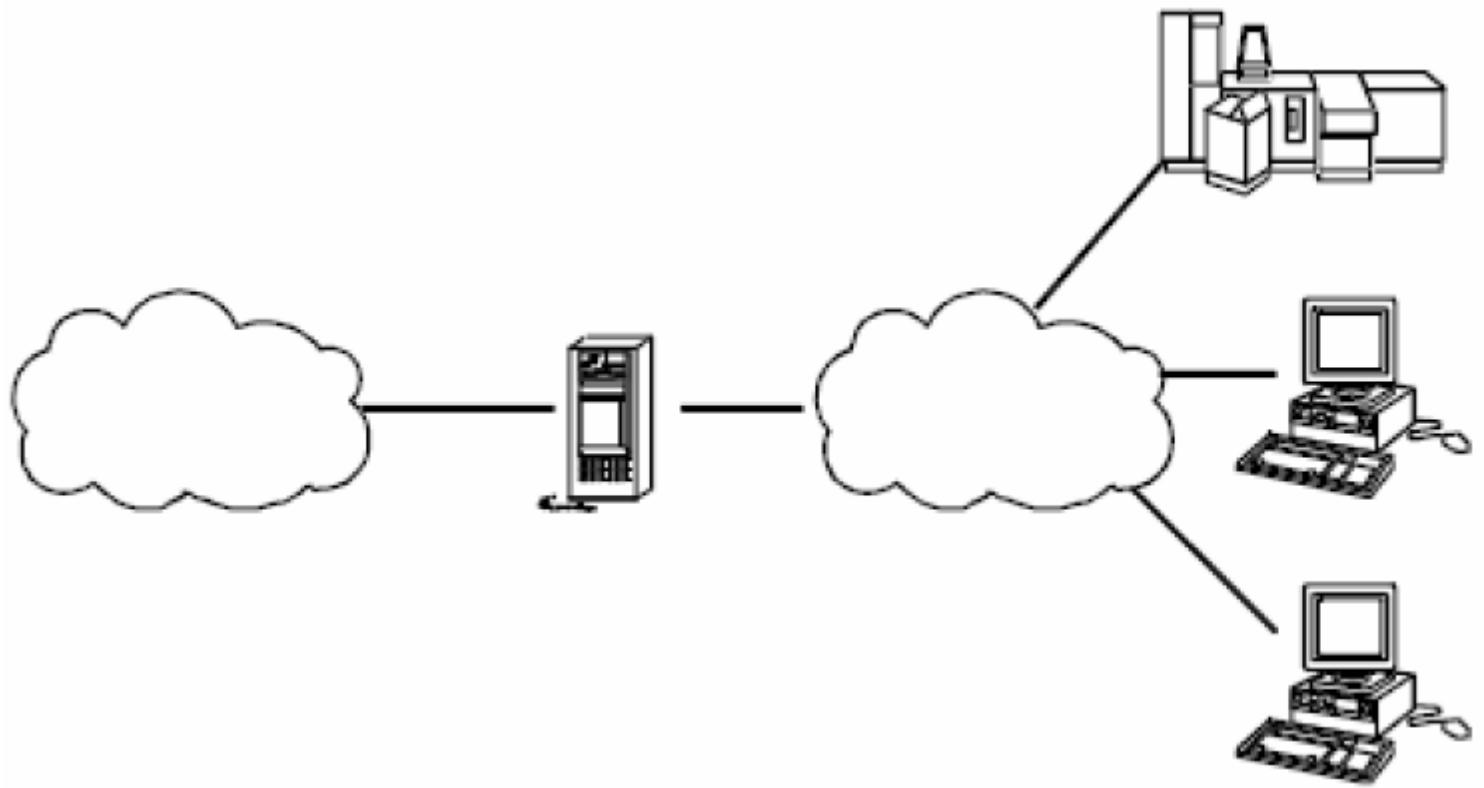
El problema surge cuando además de controlar esa puerta, tiene que filtrar los encaminamientos a todos o algunos de los hosts y a los distintos tipos de acceso. Por ejemplo, una red interna sólo puede recibir correo electrónico, otra no ser accesible desde Internet, una tercera solo puede facilitar información al exterior, etc.

Aunque no pueden facilitar un alto nivel de seguridad, es una implantación bastante sencilla y sólida. Es mucho más asequible tanto en cuanto a coste, como en relación a la experiencia tecnológica necesaria con respecto a otros sistemas. Este sistema se debe apoyar o complementar con un mecanismo de seguridad propio de la aplicación que vaya a tratar la información, con el fin de impedir que lleve otro contenido que no sea el solicitado.

Cortafuegos a Nivel de Circuito

En este caso las medidas de seguridad se establecen a nivel de circuitos, un dispositivo interpuerto que transmite las comunicaciones entre las redes internas y externas. Suele ser un host provisto de dos interfaces operando a modo de pasarela y realiza las tareas de filtrado de paquetes, que en el apartado anterior realizaba el router, pero en este caso pueden añadirse más funciones de seguridad como las de autenticación mediante el uso de contraseñas (passwords) o palabras clave asignadas previamente a los usuarios. De esta forma cualquiera que intente acceder a la red interna mediante la técnica de generar dinámicamente "passwords" aleatorios, se encontrará con un impedimento adicional a las medidas adoptadas, haciendo el acceso mucho más difícil.

Esquema de un cortafuegos a nivel de circuito.



Cortafuegos a Nivel de Aplicación

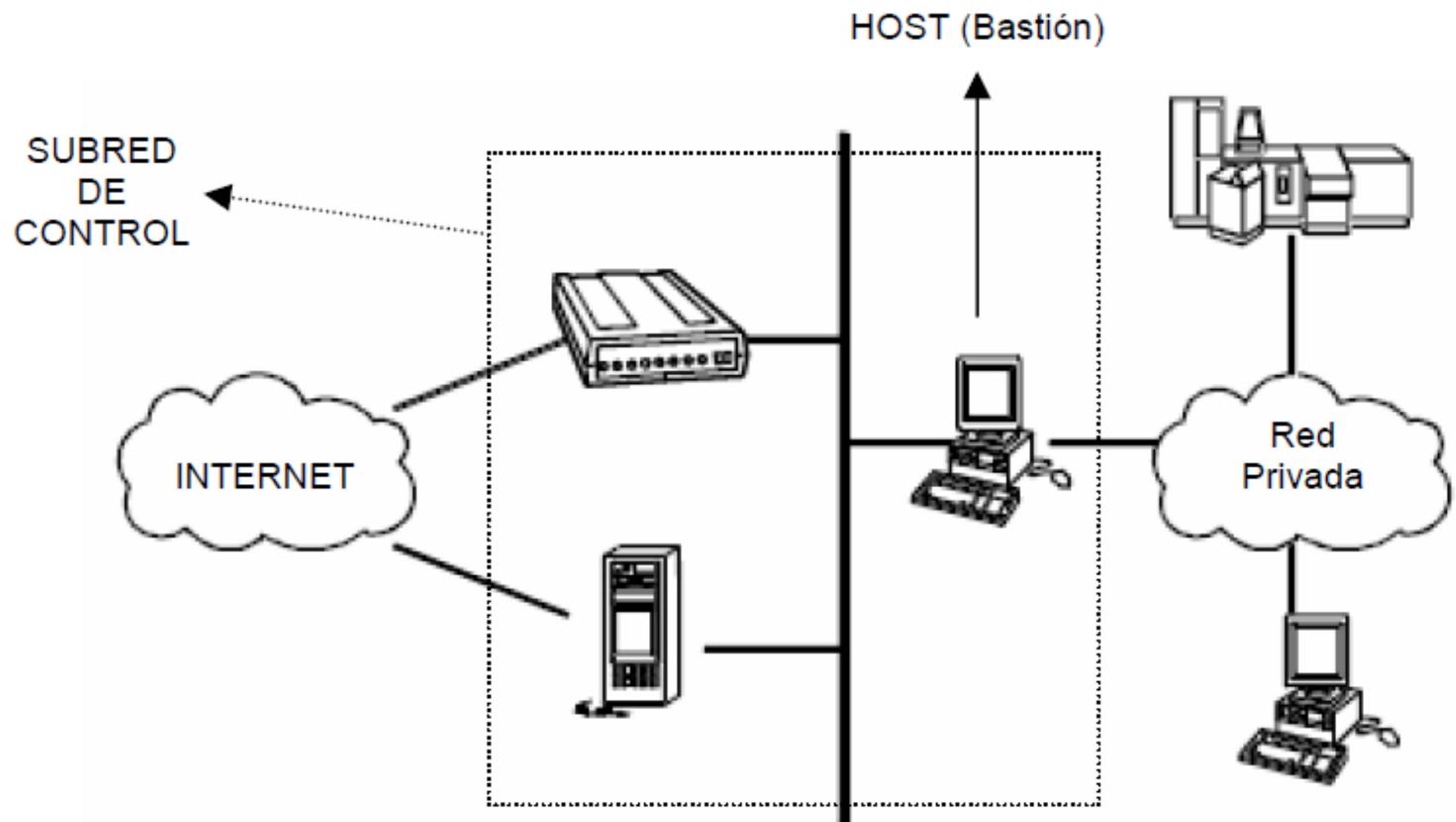
La forma de protección más completa, además de la más conocida y experimentada, es la utilización de cortafuegos a nivel de aplicación. Consiste en crear una subred, que constituya una zona de separación entre las redes internas e Internet. Por ejemplo mediante un router o encaminador, pero también se puede colocar un cortafuegos de acceso a la red interna.

Tiene que haber un segundo dispositivo, casi siempre un *host* (denominado bastión), que se situará delante de la red interna. Los usuarios, tanto de entrada como de salida acceden a este *host* mediante una operación *Telnet* (conexión remota) para trabajar con una determinada aplicación ubicada en el mismo. Este *Host* gestiona las tareas de autenticación de usuarios, limitación de tráfico de entrada y salida, realiza un seguimiento de todas las actividades manteniendo un registro de incidencias.

Este tipo de cortafuegos debe incorporar código escrito, especialmente para especificar todas y cada una de las aplicaciones para las cuales existe autorización de acceso.

La ventaja de este sistema es que el usuario externo nunca tiene acceso a las redes internas de la empresa, por lo tanto nunca podrá realizar ningún tipo de intrusión. El inconveniente es que supone una fuerte inversión en tiempo y dinero para proporcionar un servicio cuyo grado de utilización puede no llegar a ser rentable.

Esquema de un cortafuegos a nivel de aplicación.



Aplicaciones de los Cortafuegos

Las siguientes aplicaciones no son más que diversas formas de usar el concepto de cortafuegos y los tipos de cortafuegos básicos vistos en el apartado anterior. Algunos ejemplos de estos usos son:

Aplicación al correo electrónico en uso corporativo.

Se puede crear un grupo cerrado que interconecte todos los ordenadores situados en puntos geográficos distintos incluidos en la red Internet, pero que pueden conectarse entre sí para aplicaciones de correo electrónico, transferencia de ficheros y poder conectarse usuarios de otro ordenador incluido en este Grupo, controlados por un sistema de seguridad corporativo. Cada uno de ellos, a su vez, puede conectarse con otros sistemas informáticos bien por Internet, o por otros medios de comunicación. Para evitar intrusiones a través de correo electrónico, se puede colocar un servidor o punto único de entrada que controle los accesos y salidas del grupo cerrado. Se le puede dar funciones de sólo entrada de correo electrónico pudiendo actuar como cortafuegos, y a cada ordenador del grupo darle funciones de sólo salida. Así limitaremos los puntos de acceso a uno solo y con un mayor control en todos los sistemas.

Servidor como punto de entrada única a la red interna.

Este tipo de entrada proviene de una conexión directa entre el usuario de Internet y la red interna. El filtro de acceso a nuestras aplicaciones controla aquellos usuarios que pueden acceder al interior desde el exterior y viceversa. Sin embargo una aplicación corriendo en este servidor puede establecer conexiones desde el exterior hacia cualquier punto de la red interna. Un problema se nos presenta al ser superada esta barrera, entonces nuestra red está totalmente desprotegida.

Servidor como punto de entrada única a las aplicaciones.

Como complemento del servidor único como punto de entrada única, se pueden colocar diferentes servidores exclusivos para cada aplicación. Nos garantiza que filtrarán todos

los accesos, dejando sólo los que corresponden a la aplicación permitida. Deben estar colocados entre Internet y el servidor único y pueden contener la lista de usuarios que tienen permitido el acceso al host, y/o aplicaciones.

Servidor como punto de entrada único al correo electrónico.

Una puerta de entrada puede ser la colocación de otro servidor localizado en la red interna. Distribuye los mensajes del correo electrónico entre los diferentes ordenadores que están en la red privada. Contiene a su vez la lista de usuarios que pueden recibir mensajes y los que pueden enviar mensajes al exterior. Se puede usar también para el tráfico interior de la red.

Servidor sólo para facilitar la información.

Cuando nuestro propósito sea permitir únicamente la lectura de información, deberemos adecuar el control de acceso de usuarios a la función de lectura únicamente. En grandes instalaciones, el ideal sería disponer de otro ordenador único que contenga la información solamente, de esta manera evitaremos cualquier posibilidad de que se pueda obtener otra información distinta a nuestro propósito debido a tener cortados los accesos a cualquier otro sistema.

Redes Privadas Virtuales (VPN)

El término **VPN** (**V**irtual **P**rivate **N**etwork) se ha asociado tradicionalmente a los servicios de conectividad remota de datos ofrecidos por las operadoras de telefonía mediante líneas dedicadas, aunque en la actualidad se refiere al uso de los acceso vía Internet para realizar comunicaciones remotas con las mismas características de seguridad que las líneas dedicadas, con un coste económico mucho menor. Es decir, se aprovecha el bajo coste del acceso a Internet, al que se añaden técnicas de encriptación fuerte para conseguir seguridad y se simulan las clásicas conexiones punto a punto.

De esta forma, un usuario o una sede remota que se conecta a través de Internet a su organización y establece un túnel VPN puede estar funcionando *como si estuviera dentro de la misma* a todos los efectos de conectividad.

El factor crucial que convierte a un VPN en una red privada virtual es lo que se denomina un túnel, que no indica una ruta fija y delimita entre ambos extremos que se comunican vía Internet, sino que se refiere a que únicamente ambos extremos son capaces de ver lo que se mueve por el túnel, convenientemente encriptado y protegido del resto de Internet. **La tecnología del túnel encripta y encapsula los protocolos de red que utilizan los extremos sobre el protocolo IP**, lo que le permite funcionar como si se tratara de un enlace dedicado convencional, de modo transparente para el usuario.

El protocolo estándar para el soporte de túneles, encriptación y autenticación en las VPN es IPSec (IPSecurity), que se diseñó para proteger el tráfico de red y que permite gestionar el control de acceso, la integridad de las conexiones, la autenticación del origen de los datos y la confidencialidad del flujo de datos.

Este protocolo permite dos modos operacionales. En el modo transporte se utiliza para proteger conexiones individuales de usuarios remotos y las comunicaciones se cifran entre un ordenador remoto (el cliente VPN) y el servidor de VPN, mientras que en el modo túnel se encriptan las comunicaciones entre dos dispositivos de tipo encaminador (o un encaminador y el servidor de VPN), con lo que se protegen todas las comunicaciones de todas las máquinas situadas detrás de cada encaminador.

A pesar de que IPSec es el protocolo más utilizado para VPNs y ser el único que está pensado para soportar desarrollos futuros de VPNs, existen otros muy utilizados, fundamentalmente por haber sido desarrollados por Microsoft como soluciones temporales mientras se estandarizaba totalmente IPSec, se trata de:

- L2TP (Layer 2 Tunneling Protocol)
- PPTP (Point-to-Point Tunneling Protocol)

En el caso de L2TP, se necesita IPSec para proporcionar las funciones de encriptación, mientras que PPTP tiene capacidades propietarias de encriptación y autenticación, aunque se le considera ya obsoleto.

Bibliografía

- [Scribd \(Ibiza Ales\)](#)

Introducción

Estos últimos años la comunicación inalámbrica ha evolucionado notablemente, el uso de teléfonos móviles ya es masivo y a éstos cada vez se la han ido agregando nuevas funcionalidades, sin embargo esto es sólo el inicio de la comunicación inalámbrica, ya que éstas continúan apostando por convertir al aire en el mejor medio de transporte de datos.

La apuesta anterior se basa en las tecnologías de comunicación inalámbrica WI-FI y Bluetooth, las cuales han permitido que muchas personas utilicen tiempo que antes era de ocio pudiendo contar con Internet en situaciones tan comunes como la espera en el aeropuerto, el tiempo de viaje, entre otras ventajas.

Estas tecnologías aparentemente pueden parecer similares, sin embargo hay grandes diferencias entre ellas, lo que hace que sean complementarias.

Las Redes inalámbricas facilitan la operación en lugares donde la computadora no puede permanecer en un solo lugar, como en almacenes o en oficinas que se encuentren en varios pisos. Pero la realidad es que esta tecnología está todavía en pañales y se deben resolver varios obstáculos técnicos y de regulación antes de que las redes inalámbricas sean utilizadas de una manera general en los sistemas de cómputo de la actualidad.

Hasta ahora solo se ha percibido los aspectos positivos de estas tecnologías, sin embargo no hay que olvidar los aspectos más débiles de ésta. El mayor problema al que se enfrenta hoy este tipo de comunicación es la seguridad, y aunque se ha trabajado en este aspecto aún falta para que sea tanto o más confiable que el cable. Otro aspecto que no hay que dejar de lado es el coste que significa contar con éstas. Con respecto a estos problemas, es de esperar que en unos años más se llegue a una solución para así confirmar que el aire es el mejor medio de transportes de datos.

No se espera que las redes inalámbricas lleguen a reemplazar a las redes cableadas. Estas ofrecen velocidades de transmisión mayores que las logradas con la tecnología inalámbrica. Sin embargo se pueden mezclar las redes cableadas y las inalámbricas, y de esta manera generar una "Red Híbrida" y poder resolver los últimos metros hacia la estación. Se puede considerar que el sistema cableado sea la parte principal y la inalámbrica le proporcione movilidad adicional al equipo, pudiendo el operador desplazarse con facilidad dentro de un almacén o una oficina.

La disponibilidad de conexiones inalámbricas y redes LAN inalámbricas puede ampliar la libertad de los usuarios de la red a la hora de resolver varios problemas asociados a las redes con cableado fijo y, en algunos casos, incluso reducir los gastos de implementación de las redes. Sin embargo, a pesar de esta libertad, las redes LAN inalámbricas traen consigo un nuevo conjunto de desafíos.

El amplio interés del sector para que exista interoperabilidad y compatibilidad entre los SO ha permitido resolver algunas de las cuestiones relacionadas con la implementación de las redes LAN inalámbricas. Con todo, las redes LAN inalámbricas exponen nuevos retos en lo que respecta a la seguridad, la movilidad y la configuración.

Redes Inalámbricas

Situación Actual

En los últimos años el crecimiento en la demanda de soluciones inalámbricas para la empresa, y últimamente también para el hogar, ha crecido de una manera espectacular. Las distintas tecnologías inalámbricas permiten dar una cobertura casi en cualquier rincón del planeta. Cientos de millones de personas en todo el mundo se comunican e intercambian información todos los días usando una u otra tecnología inalámbrica, permitiendo el envío de datos y con una movilidad sin precedentes.

Las tecnologías más usadas y conocidas hoy día son las de los teléfonos móviles, sistemas de navegación, "buscas", servicios de mensajes, y un largo etcétera. Pero el gran exponente hoy de la revolución digital y de como los bits forman parte de nuestro día a día, proviene del intercambio de datos digitales. Estos datos no sólo se quedan limitados al ámbito de las computadoras, sino también en una gran cantidad de aplicaciones, pasando desde los grandes sistemas de tratamiento de datos empresariales y científicos hasta las más pequeñas utilidades personales destinadas a mejorar nuestro día a día.

Ya no estamos atados a redes cableadas sino que podemos acceder y compartir datos llevándoles con nosotros, dondequiera que vayamos.

Desde el principio de los años 70 hemos tenido una red, la red Ethernet, que ha supuesto una estandarización a la hora de transmitir datos, con un gran éxito en todo el mundo. Aunque no es el único, si es el que más ha influenciado el uso habitual de las redes de área local (LAN). Nos hemos acostumbrado a tener redes de computadores de bajo coste, altas velocidades y una relativamente fácil instalación, más que aptas para la mayoría de las aplicaciones. Pero el hecho de tener una infraestructura nos limita a esas líneas preinstaladas, y los costes ante fallos, mantenimiento y reestructuración se disparan. La flexibilidad es muy baja e incluso no resulta factible la instalación de estas líneas en algunos edificios, antiguos de valor histórico o peligrosos con asbestos u otros materiales. Todo esto puede ser solucionado con las nuevas tecnologías inalámbricas, puesto que ya no es necesaria la instalación de cables.

Las redes de área local inalámbricas utilizan el aire como medio de transporte, y su característica principal es complementar, y en algunos casos, reemplazar las redes de área local alámbricas, generalmente basadas en los estándares Ethernet.

Pero no todo lo que reluce es oro. Las WLANs tienen sus propios problemas, característicos del medio de transmisión, que dificultan en gran medida la transmisión de datos.

Se trata de un medio de transmisión difícil. Al tratarse de bandas libres no tienen la protección de una banda con licencia donde se sabe de antemano que es la única o una de

las pocas tecnologías transmisoras en ese medio, teniendo así cierta seguridad. Pero no, las frecuencias empleadas para la transmisión son las ISM, bandas de propósito general que ahora se usan para WLANs.

Estas bandas son las de 900MHz, 2.4Ghz y 5 Ghz. En WLANs se usan las dos últimas, siendo la banda de 2.4Ghz la más utilizada y por tanto el medio más contaminado. Por otra parte la banda de 5Ghz se usa en tecnologías más recientes, como puede ser el IEEE802.11a o el Hyperlan/2, pero tiene otros problemas, las licencias.

En éstos medios, la calidad no se puede asegurar, con dispositivos móviles la cobertura no siempre está asegurada, se tiene una alta tasa de errores de bit, el problema de los nodos ocultos/expuestos, etc, por lo que la calidad no se puede asegurar ni en tiempo ni en espacio.

¿Cómo Trabajan las WLAN?

Se utilizan ondas de radio o infrarrojos para llevar la información de un punto a otro sin necesidad de un medio físico. Las ondas de radio son normalmente referidas como portadoras de radio ya que éstas únicamente realizan la función de llevar la energía a un receptor remoto. Los datos a transmitir se superponen a la portadora de radio y de este modo pueden ser extraídos exactamente en el receptor final. Esto es llamado modulación de la portadora por la información que está siendo transmitida.

De este modo la señal ocupa más ancho de banda que una sola frecuencia. Varias portadoras pueden existir en igual tiempo y espacio sin interferir entre ellas, sin las ondas son transmitidas a distintas frecuencias de radio. Para extraer los datos el receptor se sitúa en una determinada frecuencia ignorando el resto.

En una configuración de LAN sin cable los puntos de acceso (transceiver) conectan la red cableada de un lugar fijo mediante cableado normalizado. El punto de acceso recibe la información, la almacena y transmite entre la WLAN y la LAN cableada. Un único punto de acceso puede soportar un pequeño grupo de usuarios y puede funcionar en un rango de al menos treinta metros y hasta varios cientos.

El punto de acceso (o la antena conectada al punto de acceso) es normalmente colocado en alto pero podría colocarse en cualquier lugar en que se obtenga la cobertura de radio deseada. El usuario final accede a la red WLAN a través de adaptadores. Estos proporcionan una interfaz entre el SO de red del cliente (NOS: Network Operating System) y las ondas, vía una antena.

La naturaleza de la conexión sin cable es transparente al sistema del cliente.

Configuraciones de las WLAN

Básicamente, las WLAN están compuestas por dos elementos:

- Punto de acceso (Access Point - AP por sus siglas en inglés): este elemento es la estación base que generalmente tiene conectividad con el mundo alámbrico (red Ethernet, por ejemplo). El AP crea un anillo a sí un área de cobertura donde los usuarios o dispositivos clientes se pueden conectar. El AP cuenta con una o dos antenas y con una o varias puertas Ethernet.
- Dispositivos clientes: estos elementos son PCs, PDAs u otros que cuentan con tarjeta de red inalámbrica. Estas tarjetas existen en diferentes tipos tales como: PCCard, PCI, Compact Flash, etc.

Las redes pueden ser simples o complejas. La más básica se da entre dos ordenadores equipados con tarjetas adaptadoras para WLAN, de modo que pueden poner en

funcionamiento una red independiente siempre que estén dentro del área que cubre cada uno. Esto es llamado red de igual a igual (peer to peer).

Cada cliente tendría únicamente acceso a los recursos de otro cliente pero no a un servidor central. Este tipo de redes no requiere administración o preconfiguración.

Uno de los últimos componentes a considerar en el equipo de una WLAN es la antena direccional. Por ejemplo: se quiere una LAN sin cable a otro edificio a 1Km de distancia. Una solución puede ser instalar una antena en cada edificio con línea de visión directa. La antena del primer edificio está conectada a la red cableada mediante un punto de acceso. Igualmente en el segundo edificio se conecta un punto de acceso, lo cuál permite una conexión sin cable en esta aplicación.

Estandarización y Compatibilidad

Hoy en día existen varias tecnologías y estándares para las comunicaciones de redes de área local inalámbricas. Estos estándares, definen una red formada por un medio compartido y transmisión encriptada de la información.

IEEE 802.11 es un estándar para redes inalámbricas definido por la organización Institute of Electrical and Electronics Engineers (IEEE), instituto de investigación y desarrollo, de gran reconocimiento y prestigio, cuyos miembros pertenecen a decenas de países entre profesores y profesionales de las nuevas tecnologías.

El estándar IEEE 802.11 es un estándar en continua evolución, debido a que existen cantidad de grupos de investigación, trabajando en paralelo para mejorar el estándar, a partir de las especificaciones originales.

- **802.11b** : Este estándar especifica transmisiones en la banda de frecuencias de los 2.4GHz, con velocidades de hasta 11 Mbps. Es sin lugar a dudas la tecnología más popular y la más económica.
- **802.11a** : Este estándar, posterior al 802.11b, especifica transmisiones en la banda de los 5GHz (una banda menos ruidosa que la de los 2.4GHz) y con una velocidad de hasta 54 Mbps. Posee una menor cobertura que 802.11b.
- **802.11g** : Especifica transmisiones de hasta 54 Mbps en la banda de los 2.4GHz y asegura compatibilidad con dispositivos 802.11b.

	<i>IEEE 802.11a</i>	<i>IEEE 802.11b</i>	<i>IEEE 802.11g</i>
<i>Rango de frecuencia</i>	5,15 a 5,35 GHz (banda UNII)	2,40 GHz (banda ISM)	2,40 GHz (banda ISM)
<i>Velocidad de transferencia de datos</i>	De 6 a 54 Mbps	De 1 a 11 Mbps	De 20 a 54 Mbps
<i>Rango en espacio libre (según la velocidad de transferencia de datos)</i>	30 (54 Mbit) a 300 metros (6 Mbit)	120 (11 Mbit) a 460 metros (1 Mbit)	120 (54 Mbit) a 460 metros (6 Mbit)
<i>Rango en habitaciones (según la velocidad de datos)</i>	12 (54 Mbit) a 90 metros (6 Mbit)	30 (11 Mbit) a 90 metros (1 Mbit)	30 (54 Mbit) a 90 metros (6 Mbit)
<i>Números de canales independientes</i>	8	3	13
<i>Número de usuarios admitido por punto de acceso</i>	512	192	?
<i>Aplicación</i>	Multimedia	Datos	Multimedia
<i>Técnicas de modulación</i>	OFDM	DSSS	OFDM
<i>Protocolo</i>		TCP/IP	

La alianza WI-FI (Wireless Fidelity) es una organización sin fines de lucro formada en 1999 para certificar la interoperabilidad de los productos 802.11 y para promocionarlos con un estándar global de WLAN en todos los segmentos.

Se trata de una especificación en continua evolución con posibilidad de adaptarse a nuevos requerimientos y demandas de usuario en un futuro.

IEEE 802.11. Tecnología

El estándar permite el uso de varios medios y técnicas para establecer conexiones. El estándar original permite usar infrarrojos y espectro expandido, tanto en salto en frecuencias como secuencia directa, con la ventaja de usar una capa de acceso al medio (MAC) común. Ello da mucha flexibilidad a los desarrolladores e investigadores, que pueden olvidarse de ciertos aspectos ya que no existe dependencia directa entre ellos.

Los estándares de IEEE 802.11 son de libre distribución y cualquier persona puede ir a la página Web del IEEE y descargarlos. Estos estándares sólo definen especificaciones para las capas físicas y de acceso al medio y para nada tratan modos o tecnologías a usar para la implementación final.

Esto debe permitir y facilitar la interoperabilidad entre fabricantes de dispositivos IEEE 802.11 y para asegurarse de ello se ha creado una alianza denominada WECA para crear y definir procedimientos para conseguir certificados de interoperabilidad y de cumplir las especificaciones, todo dentro de un estándar llamado WiFi (Wireless Fidelity). El nombre además es un indicativo del enfoque doméstico y muy enfocado hacia el usuario final.

El bloque constructivo básico de una red inalámbrica 802.11 es el denominado **conjunto de servicio básico** (BSS, Basic Service Set), que es un área geográfica en la que las estaciones inalámbricas se pueden comunicar. El tipo más sencillo de BSS consiste en dos o más equipos que han entrado dentro de las áreas de transmisión respectivas. Este proceso por el que los dispositivos entran en un BSS se denomina **asociación**.

Capa Física (PHY)

La capa física en cualquier red define la modulación y características de señalización para la transmisión de datos en ese medio. Para poder transmitir para redes inalámbricas en bandas sin licencia se necesitan usar técnicas de espectro expandido, definidas en los requerimientos de casi todos los países.

En el estándar IEEE 802.11 se definen tres medios de nivel físico. Uno usa señales de infrarrojos y los otros dos utilizan señales de radio frecuencia (RF).

Los medios de RF 802.11 funcionan en la banda de 2.4Ghz, con un ancho de banda de 83Mhz entre 2.400 y 2.483GHz.

Las definiciones para la transmisión por radiofrecuencia en los estándares son espectro expandido por salto en frecuencias (FHSS) y espectro expandido por secuencia directa (DSSS). Ambos están definidos para trabajar en la banda de 2.4Ghz, y DSSS además tienen una variante en la banda de los 5Ghz, que consigue mayores velocidades de transmisión.

Infrarrojos

Las comunicaciones por infrarrojos utilizan frecuencias entre 850 y 950 nanómetros, justo por debajo del espectro de la luz visible. La implementación IEEE 802.11 de infrarrojos, a

diferencia de la mayoría de los medios infrarrojos, no requiere comunicación de visión directa, puede funcionar mediante señales reflejadas.

Sin embargo, debido a su limitado alcance comparado con los medios de RF y a que sólo puede funcionar adecuadamente en un ambiente interior cuando las superficies proporcionan una buena reflexión de las señales, es raro que se implemente en las redes inalámbricas. Además impone más restricciones en la ubicación física del dispositivo inalámbrico que FHSS o DSSS.

FHSS

Salto de frecuencias se refiere a un sistema que periódicamente cambia las frecuencias en las que transmite. Se utiliza la banda entera lo que contribuye a aumentar la seguridad frente a escuchas a la vez que ayuda a suprimir el ruido o las interferencias.

FHSS tiene 22 patrones de saltos predefinidos usando 79 canales de 1Mhz a un mínimo de 2.5 saltos por segundo, y para resolver los problemas de sincronización, para que tanto transmisor como receptor salten a la vez, se definen paquetes de sincronización.

La velocidad de los cambios de frecuencia es independiente de la velocidad de bit de la transmisión de datos. Si la velocidad del salto de frecuencia es menor que la velocidad de bit de la señal, la tecnología se denomina **sistema de salto lento**, y si es mayor se denomina **sistema de salto rápido**.

Para la modulación FHSS usa FSK gaussiano de 2 ó 4 niveles. Las velocidades típicas conseguidas son de 1 y 2 Mbps para FHSS.

DSSS

El sistema de radio usando DSSS trabaja en un canal fijo y preconfigurado, lo que le permite obtener mayores tasas de transferencia, pero con la desventaja de ser más sensible a interferencia y a señales procedentes de otros dispositivos usando la misma frecuencia. Es posible tener tres puntos de acceso con tres canales diferentes, sin solapar en un mismo emplazamiento, si tener en cuenta ningún tipo de planificación. Aunque para más de tres puntos de acceso sí es necesaria cierta planificación, para poder mantener las velocidades, puesto que el solape de celdas y frecuencias tendrá un deterioro sobre el rendimiento.

Las modulaciones usadas para DSSS son BPSK y DQPSK para el estándar original. Para 11b, que permite conseguir 11Mbps, se utiliza CCK.

Además se ha definido una variante de IEEE 802.11, incorporada recientemente a la especificación que permite conseguir 54 Mbps en la banda de 5Ghz, con un ancho de banda de hasta 300MHz y usando una modulación OFDM.

Tramas de la Capa Física

En lugar de tener un esquema de señalización relativamente simple como en Ethernet y Token Ring que utilizan Manchester y Manchester diferencial respectivamente, los medios que funcionan en 802.11 tienen su propio formato de tramas, que encapsulan las tramas generadas en el nivel de enlace de datos.

La trama de FHSS contiene los siguientes campos:

- Preámbulo (10 bytes): contiene 80 bits de 1 y 0 alternos utilizados por el receptor para detectar la señal y sincronizar los tiempos.
- Delimitador de comienzo de trama (SFD) (2 bytes): indica el comienzo de la trama.
- Longitud (12 bits): indica el tamaño del campo de datos.

- Señalización (4 bits): contiene un bit para indicar si se está utilizando la velocidad de 1 o 2 Mbps. Los otros 3 bits se reservan para uso futuro. Sólo el campo de datos se puede transmitir a 2 Mbps.
- CRC (2 bytes): contiene un valor de comprobación de redundancia cíclica.
- Datos (de 0 a 4.095 bytes): contiene la trama del nivel de enlace de datos que se transmite.

La trama DSSS contiene los siguientes campos:

- Preámbulo (16 bytes): contiene 128 bits que el sistema receptor utiliza para ajustarse a la señal entrante.
- Delimitador de comienzo de trama (SFD) (2 bytes): indica el comienzo de la trama.
- Señal (1 byte): especifica la velocidad de transmisión.
- Servicio (1 byte): contiene el valor hexadecimal 00 que indica que el sistema cumple con el estándar 802.11.
- Longitud (2 bytes): indica el tamaño del campo de datos.
- CRC (2 bytes): contiene un valor de comprobación de redundancia cíclica.
- Datos (variable): contiene la trama del nivel de enlace de datos que se transmite.

La trama de infrarrojos contiene los siguientes campos:

- Sincronización (SYNC) (entre 57 y 73 ranuras): utilizadas por el sistema receptor para sincronizar el tiempo y opcionalmente para estimar la relación señal/ruido.
- Delimitador de comienzo de trama (SFD) (2 ranuras): indica el comienzo de la trama.
- Velocidad de datos (3 ranuras): especifica la velocidad de transmisión.
- Ajuste del nivel de DC (CDLA) (32 ranuras): utilizado por el receptor para estabilizar el nivel DC después de transmitir los campos precedentes.
- Longitud (12 bits): indica el tamaño del campo de datos.
- CRC (2 bytes): contiene un valor de comprobación de redundancia cíclica.
- Datos (de 0 a 2.500 bytes): contiene la trama del nivel de enlace de datos que se transmite .

La Capa de Acceso al Medio (MAC)

La especificación de la capa MAC del IEEE 80.11 tiene muchas similitudes con el estándar de Ethernet cableado (IEEE 802.3). El protocolo del 802.11 es un esquema de protocolo conocido como detección de portadora, acceso múltiple, evitando colisiones (CSMA/CA). Este protocolo evita las colisiones, en vez de detectarlas como el algoritmo 802.3. Es extremadamente difícil detectar colisiones en una red de transmisión de radiofrecuencias y de ahí de que se trate de evitar las colisiones.

La capa MAC opera junto con la capa física muestreando la energía del medio transmisor de datos. El protocolo CSMA/CA permite opciones para que se pueda minimizar las colisiones usando tramas de transmisión RTS/CTS (Request-to-send/Clear-to-send), datos y reconocimientos de una manera secuencial. En estas tramas se suelen incorporar datos de duración de los envíos con el objetivo de asegurar que esos envíos no van a ser interrumpidos: los demás nodos saben que deben estar callados durante ese intervalo de tiempo. Todo ello además se asegura y confirma con tramas de reconocimiento (ACK).

Pero un problema común a cualquier WLAN es el problema de los nodos ocultos. Esto puede llegar a reducir las prestaciones en un 40% en una WLAN con alta carga. Ocurre cuando un nodo no puede escuchar transmisiones de un nodo y trata de transmitir a un nodo que si puede escucharlas, allí se puede generar muchas colisiones. Algunas mejoras se han incorporado para evitar el problema con el uso de RTS/CTS de una manera inteligente.

Además se utiliza tiempos entre tramas para evitar colisiones, ello a parte de evitar colisiones, permite además cierto uso de clases de calidad o por lo menos de preferencia

de un tráfico sobre otro, utilizando funciones de coordinación puntual y permitir el acceso al medio de tráfico prioritario antes que a los demás.

Tramas del Nivel MAC

El estándar 802.11 define tres tipos básicos de tramas en este nivel:

- Tramas de datos: se usan para transmitir datos de los niveles superiores entre estaciones.
- Tramas de administración: se usan para el intercambio de información para realizar funciones de red como la autenticación y la asociación.
- Tramas de control: se usan para regular el acceso al medio y para reconocimiento de las tramas de datos transmitidas.

Una trama MAC genérica contiene los siguientes campos:

- Control de la trama (2 bytes): contiene 11 subcampos que habilitan las distintas funciones del protocolo:
 - Versión de protocolo (2 bits): especifica la versión del estándar que se está utilizando.
 - Tipo (2 bits): indica si la trama es de administración (00), control (01) o datos (00).
 - Subtipo (4 bits): identifica la función específica de la trama.
 - A DS (1 bit): si vale 1 indica que la trama se transmite al sistema de distribución a través de un punto de acceso.
 - De DS (1 bit): si vale 1 indica que la trama se ha recibido de un sistema de distribución.
 - Más fragmentos (1 bit): si vale 1 indica que el paquete contiene un fragmento de una trama y que hay más fragmentos para su transmisión.
 - Reintentos (1 bit): un valor de 1 indica que la trama se está retransmitiendo debido a una falta de recepción de un reconocimiento.
 - Administración de energía (1 bit): si vale 0 indica que la estación está funcionando en modo activo; si vale 1 en modo ahorro de energía.
 - Más datos (1 bit): si vale 1 indica que el AP tiene más paquetes almacenados para la estación y en espera de transmisión.
 - WEP (1 bit): si vale 1 indica que el cuerpo de la trama se ha cifrado utilizando WEP.
 - Orden (1 bit): si vale 1 indica que la trama de datos se está transmitiendo utilizando la clase de servicio estrictamente ordenado.
- Duración/AID (2 bytes): en las tramas de control de sondeo de energía contiene la identidad de asociación (AID) de la estación transmisora. En el resto de tramas contiene el tiempo (en microsegundos) necesarios para transmitir una trama más el intervalo entre tramas.
- Dirección 1 (6 bytes): contiene una dirección que identifica al receptor de la trama, utilizando uno de los 5 tipos de direcciones definidas en la capa MAC 802.11, dependiendo de los valores de los subcampos A DS y De DS.
- Dirección 2 (6 bytes): contiene una dirección utilizando uno de los 5 tipos de direcciones utilizadas en el subnivel MAC, dependiendo de los valores de los subcampos A DS y De DS.
- Dirección 3 (6 bytes): contiene una dirección utilizando uno de los 5 tipos de direcciones utilizadas en el subnivel MAC, dependiendo de los valores de los subcampos A DS y De DS.
- Control de secuencia (2 bytes): contiene dos subcampos:
 - Número de fragmento (4 bits): contiene un valor que identifica un fragmento particular en una secuencia.
 - Número de secuencia (12 bits): contiene un valor que identifica los fragmentos de la secuencia que componen el conjunto de datos.

- Dirección 4 (6 bytes): contiene una dirección utilizando uno de los 5 tipos de direcciones utilizadas en el subnivel MAC, dependiendo de los valores de los subcampos A DS y De DS.
- Cuerpo de la trama (0 a 2.312 bytes): contiene la información que se está transmitiendo a la estación receptora.
- Secuencia de verificación de trama (4 bytes): contiene un valor CRC.

Los cinco tipos de dirección del subnivel MAC son:

- Dirección origen (SA): una dirección MAC individual que identifica al sistema que generó la información que va en el cuerpo de la trama.
- Dirección destino (DA): una dirección MAC individual o de grupo que identifica al receptor final de una unidad de datos de servicio.
- Dirección del emisor (TA): una dirección MAC individual que identifica al sistema que transmitió la información que va en el cuerpo de la trama en el medio inalámbrico actual (un AP).
- Dirección del receptor (RA): una dirección MAC individual o de grupo que identifica al receptor inmediato de la información en el cuerpo de la trama en el medio inalámbrico actual (un AP).
- ID del conjunto de servicio básico (BSSID): en una red ad hoc el BSSID es un valor generado aleatoriamente durante la creación del BSS; en una red con infraestructura es la dirección MAC de la estación que funciona como AP del BSS.

Tipos de Redes Inalámbricas

Bluetooth

Es una tecnología inalámbrica de corto alcance diseñada para reemplazar los cables entre dispositivos. Se ha convertido en la solución inalámbrica ideal para conectar teléfonos móviles con portátiles para su conexión a Internet, o para que otros organizadores de mano, como PDAs puedan conectarse al PC para coordinar sus contactos, e incluso para poder imprimir desde un ordenador de forma inalámbrica.

Las características intrínsecas de las tecnologías con bluetooth permiten establecer conexiones seguras, con capacidad de encriptación del canal, autenticación de la red y otros parámetros de seguridad como la localización y dispositivo del usuario.

Historia de Bluetooth

La historia de Bluetooth comienza en el año 1994, cuando la compañía Ericsson comenzó un estudio para investigar si era factible o no una comunicación vía radio, que fuera barata y que no consumiera tanto, con la única finalidad de conectar celulares y así eliminar cables. Fue así como se conoció un enlace de radio de poco alcance, el que podía ser usado en distintos equipos celulares, e incluso en otro tipo de dispositivos, ya que el chip utilizado era de bajo coste. Este enlace se llamó MC link. En 1997, Ericsson despertó la curiosidad y el interés en otras empresas fabricantes de celulares y equipos portátiles, y fue así que llegado el año de 1998 se creó un grupo llamado SIG, formado por cinco compañías: Ericsson, Nokia, IBM, Toshiba e Intel. De esta forma, el grupo estaba compuesto de dos líderes en telefonía celular, dos en la fabricación de computadores y uno en la fabricación de chips. Así, se podía crear una tecnología que fuese compatible con los distintos equipos fabricados por esas empresas.

Comenzando el trabajo, se pensó primero en utilizarlo con productos que usarían los ejecutivos que viajan frecuentemente, y entre los primeros dispositivos en que se podría

ocupar dicha tecnología debían estar los celulares, notebooks o computadores portátiles, PDAs y auriculares.

Teniendo esto como base, se definieron como objetivos del sistema:

- Operación en todo el mundo.
- Poco consumo de energía por parte del emisor de radio (porque los equipos móviles utilizan baterías).
- Poder transmitir voz y datos.
- La banda de frecuencia debería ser libre.

Este último escollo se saltó utilizando la banda ISM, una banda médico-científica internacional, de 2,54Ghz, que varía en rangos de 2400MHz y 2500MHz y quee sutilizada en todo el mundo.

En el mercado existen varias versiones de Bluetooth. Las primeras son BT 1.0 y BT 1.1. La diferencia entre ambas radica en que en la primera sólo se puede conectar un dispositivo con otro al mismo tiempo, mientras que en la segunda se puede conectar un dispositivo con siete simultáneamente, claro que todos deben ser compatibles entre sí. Se puede tener todos los dispositivos que se quiera conectados a través de esta vía, pero sólo se puede establecer comunicación con siete de ellos.

Además de estas dos, también existen la BT 1.2 y la BT 2.0, las que han sido desarrolladas por ingenieros de Ericsson Technology Licensing y estudiadas por el SIG. La diferencia principal entre ambas es que la BT 1.2 es para media velocidad, con una velocidad de transferencia de 2 o 3 Mbits por segundo, mientras que la BT 2.0 es para alta velocidad, con una velocidad de transferencia de 4, 8 o 12 Mbits por segundo.

¿Cómo Funciona Bluetooth?

La frecuencia de radio en la que trabaja está en el rango de 2.4 y 2.48 Ghz, con amplio espectro y saltos de frecuencia, con posibilidad de transmitir en full duplex con un máximo de 1600 saltos/seg. Los saltos de frecuencia se realizan entre un total de 79 frecuencias con intervalos de 1Mhz, lo cual permite brindar seguridad y robustez. La frecuencia en la cual trabaja le permite atravesar paredes y maletines, por lo cual es ideal tanto para el trabajo móvil, como en oficinas.

La potencia de salida para transmitir a una distancia máxima de 10m es de 0dBm (1 mW), mientras que la versión de largo alcance, hasta 100m, transmite entre -30 y 20dBm (100 mW).

Para lograr alcanzar el objetivo de bajo consumo y bajo coste, se ideó una solución en un sólo chip utilizando circuitos CMOS. De esta manera, se logró crear una solución de 9x9mm y que consume aproximadamente 97% menos energía que un teléfono móvil común.

Cada uno de los cuatro canales de voz en la especificación Bluetooth puede soportar una tasa de transferencia de 64 Kb/s en cada sentido, la cual es suficientemente adecuada para la transmisión de voz. Un canal de datos asíncrono puede transmitir 721 Kb/s en una dirección y 56 Kb/s en la dirección opuesta, sin embargo, para una conexión asíncrona es posible soportar 432,6 Kb/s en ambas direcciones si el enlace es simétrico.

Para relacionarse e intercambiar información los dispositivos Bluetooth ofrecen distintos servicios, llamados técnicamente Perfiles, entre estos perfiles se encuentran el Acceso a Redes Locales (LAN), Acceso Telefónico, Fax, Transferencia de Archivos, Sincronización, Intercomunicador o Telefonía inalámbrica, entre otros. De esta manera cuando dos dispositivos se comunican por primera vez intercambian esta información para conocer sus posibilidades de intercomunicación.

Las compañías más destacadas en el desarrollo de esta tecnología Bluetooth han sido Ericsson y Nokia. La primera versión del estándar Bluetooth se lanzó en mayo de 1998. Esta tecnología inalámbrica tiene una velocidad de transferencia de datos de 1Mbps y cuenta con un alcance máximo de 100 metros. Sin embargo, lo más utilizado son los 10 metros, ya que el consumo eléctrico aumenta rápidamente con una mayor potencia de recepción o transmisión.

Ventajas

- Menor Radiación. El SIG de Bluetooth recomienda potencias bajas para sus dispositivos móviles, esto es 1 a 10mW, limitando los 100mW para los puntos de acceso, lo cual puede ser comparado con 100mW que utilizan las tarjetas WiFi o los 125 mW de los teléfonos móviles digitales o los 600 mW de los teléfonos móviles analógicos.
- Bajo Consumo Energético. Bluetooth ha sido concebida como una tecnología de bajo consumo, lo cual permite ser utilizada en equipos móviles como Palms e iPAQ. Junto con sus bajos niveles de consumo se incorpora un administrador de energía el cual disminuye la potencia de transmisión si el dispositivo está cerca del punto de acceso o si está en modo de espera.
- Alta Movilidad. Producto del bajo consumo energético y el concepto de latencia incorporado por la especificación Bluetooth constituye la mejor alternativa al otorgar la mayor autonomía a los distintos dispositivos que lo utilizan, constituyendo la mejor opción frente a otras tecnologías inalámbricas.

WI-FI

Las redes de área local inalámbricas pueden operar en tres modalidades:



- **Ad-Hoc :** En esta modalidad, no existen puntos de acceso, por lo que los dispositivos clientes se comunican entre ellos directamente, y por lo tanto, no existe conectividad con el mundo alámbrico.
- **Infraestructura :** En este modo de operación, el más común en el ambiente corporativo, uno o varios APs generan una red inalámbrica que permite que los dispositivos clientes tengan acceso a los recursos de la empresa.
- **Bridge :** Esta es una aplicación especial que consiste en utilizar dos APs para implementar un enlace inalámbrico entre dos sitios o redes. Generalmente, se utilizan APs que operan en la interperie con antenas especiales para grandes distancias.

¿Cómo Funciona?

Se llega a las dependencias del cliente con un acceso banda ancha, ADSL, donde se instala un gateway WI-FI (inalámbrico con norma 802.11b).

Los computadores del cliente deben estar provistos con tarjetas de red inalámbricas para que puedan comunicarse con el gateway WIFI. Todos los equipos para comunicarse entre

sí deben pasar por el Gateway. Este gateway posee puertas de conexión alámbricas (Ethernet tradicional RJ-45) y puertas virtuales inalámbricas.

A las puertas alámbricas es posible conectar equipos que tengan puertas Ethernet. Mientras que a las puertas inalámbricas lo equipos a conectar deben poseer tarjetas WI-FI norma 802.11b o 802.11g.

Ventajas

- Acceso a Internet Banda Ancha.
- Acceso simultáneo de varios computadores a Internet.
- Permite montar una RED LAN sin necesidad de cables, permitiendo llegar a zonas difíciles de cablear.
- Permite movilidad del usuario dentro del recinto con WI-FI.
- De rápido y fácil montaje.

Restricciones

- Los computadores del cliente deben estar dentro de una misma dirección y casa o departamento.
- El cliente debe poseer tarjeta WI-FI instalada en cada uno de los computadores a conectar.
- La señal WI-FI NO atraviesa paredes de hormigón.
- El alcance máximo de la señal es aproximadamente de 50 mts en planta libre.
- Requiere que la señal tenga buena recepción entre gateway y PC de cliente.

Diferencias entre WI-FI y Bluetooth

	WI-FI	Bluetooth
Alcance	Poco más de 90 metros	Poco más de 9 metros
Usos	Considerado como una LAN (red de área local), para ser utilizada en grandes ámbitos como casas u oficinas.	Considerada como una Pan (personal área network o red de área personal), en la que los dispositivos están separados únicamente por una serie de centímetros.
Comunicación	Se conectan generalmente a Internet o a una red directamente.	Se comunican con otros dispositivos cercanos o se unen a grandes redes a través de, por ejemplo, un teléfono móvil.
Procesamiento	Tiene una mayor capacidad de procesamiento, lo que significa que puede enviar más datos de lo que puede hacer Bluetooth.	Menor capacidad de procesamiento.
Energía	Mayor consumo de energía.	Menor consumo de energía.
Coste	Mayor coste.	Menor coste.

Seguridad

Cuando un medio de red nuevo se introduce en un nuevo entorno siempre surgen nuevos retos. Esto es cierto también en el caso de las redes LAN inalámbricas. Algunos retos surgen de las diferencias entre las redes LAN con cable y las redes LAN inalámbricas. Por ejemplo, existe una medida de seguridad inherente en las redes con cable, ya que la red

de cables contiene los datos. Las redes inalámbricas presentan nuevos desafíos, debido a que los datos viajan por el aire, por ondas de radio.

Otros retos se deben a las posibilidades únicas de las redes inalámbricas. Con la libertad de movimiento que se obtiene al eliminar las ataduras (cables), los usuarios pueden desplazarse de sala en sala, de edificio en edificio, de ciudad en ciudad, etc, con las expectativas de una conectividad ininterrumpida en todo momento.

Las redes siempre han tenido retos, pero éstos aumentan cuando se agrega complejidad, tal como sucede con las redes inalámbricas. Por ejemplo, a medida que la configuración de red continúa simplificándose, las redes inalámbricas incorporan características (en ocasiones para resolver otros retos) y métrica que se agrega a los parámetros de configuración.

Retos de Seguridad

Una red con cable está dotada de una seguridad inherente en cuanto a que un posible ladrón de datos debe obtener acceso a la red a través de una conexión por cable, lo que normalmente significa el acceso físico a la red de cables. Sobre este acceso físico se pueden superponer otros mecanismos de seguridad.

Cuando la red ya no se sustenta con cables, la libertad que obtienen los usuarios también se hace extensiva al posible ladrón de datos. Ahora, la red puede estar disponible en vestíbulos, salas de espera inseguras, e incluso fuera del edificio. En un entorno doméstico, la red podría extenderse hasta los hogares vecinos si el dispositivo de red no adopta o no utiliza correctamente los mecanismos de seguridad.

Desde sus comienzos, 802.11 ha proporcionado algunos mecanismos de seguridad básicos para impedir que esta libertad mejorada sea una posible amenaza. Por ejemplo, los puntos de acceso (o conjunto de puntos de acceso) 802.11 se pueden configurar con un identificador del conjunto de servicios (SSID). La tarjeta NIC también debe conocer este SSID para asociarlo al AP y así proceder a la transmisión y recepción de datos en la red. Esta seguridad, si se llegase a considerar como tal, es muy débil debido a estas razones:

- Todas las tarjetas NIC y todos los AP conocen perfectamente el SSID.
- El SSID se envía por ondas de manera transparente (incluso es señalizado por el AP).
- La tarjeta NIC o el controlador pueden controlar localmente si se permite la asociación en caso de que el SSID no se conozca.
- No se proporciona ningún tipo de cifrado a través de este esquema.

Aunque este esquema puede plantear otros problemas, esto es suficiente para detener al intruso más despreocupado.

Las especificaciones 802.11 proporcionan seguridad adicional mediante el algoritmo WEP (Wired Equivalent Privacy). WEP proporciona a 802.11 servicios de autenticación y cifrado. El algoritmo WEP define el uso de una clave secreta de 40 bits para la autenticación y el cifrado, y muchas implementaciones de IEEE 802.11 también permiten claves secretas de 104 bits. Este algoritmo proporciona la mayor parte de la protección contra la escucha y atributos de seguridad física que son comparables a una red con cable.

Una limitación importante de este mecanismo de seguridad es que el estándar no define un protocolo de administración de claves para la distribución de las mismas. Esto supone que las claves secretas compartidas se entregan a la estación inalámbrica IEEE 802.11 a través de un canal seguro independiente del IEEE 802.11. El reto aumenta cuando están implicadas un gran número de estaciones, como es el caso de un campus corporativo.

Para proporcionar un mecanismo mejor para el control de acceso y la seguridad, es necesario incluir un protocolo de administración de claves en la especificación. Para hacer frente a este problema se creó específicamente el estándar 802.1x.

Retos para los Usuarios Móviles

Cuando un usuario o una estación se desplaza de un punto de acceso a otro punto de acceso, se debe mantener una asociación entre la tarjeta NIC y un punto de acceso para poder mantener la conectividad de la red. Esto puede plantear un problema especialmente complicado si la red es grande y el usuario debe cruzar límites de subredes o dominios de control administrativo.

Si el usuario cruza un límite de subred, la dirección IP asignada originalmente a la estación puede dejar de ser adecuada para la nueva subred. Si la transición supone cruzar dominios administrativos, es posible que la estación ya no tenga permiso de acceso a la red en el nuevo dominio basándose en sus credenciales.

Más allá del simple desplazamiento dentro de un campus corporativo, otros escenarios de usuarios móviles son muy reales. Los aeropuertos y restaurantes agregan conectividad inalámbrica con Internet y las redes inalámbricas se convierten en soluciones de red populares para el hogar.

Ahora es más probable que el usuario pueda abandonar la oficina para reunirse con alguien de otra compañía que también disponga de una red inalámbrica compatible. De camino a esta reunión, el usuario necesita recuperar archivos desde la oficina principal y podrá encontrarse en una estación de tren, un restaurante o un aeropuerto con acceso inalámbrico. Para este usuario sería de mucha utilidad poder autenticarse y utilizar esta conexión para obtener acceso a la red de la empresa. Cuando el usuario llegue a su destino, puede que no tenga permiso de acceso a la red local de la empresa que va a visitar. Sin embargo, sería fortuito que el usuario pudiera obtener acceso a Internet en este entorno extraño. Entonces, dicho acceso podría utilizarse para crear una conexión de red privada virtual con la red de su empresa. Despues, el usuario podría irse a casa y desear conectarse a la red doméstica para descargar o imprimir archivos para trabajar esa tarde. Ahora, el usuario se ha desplazado a una nueva red inalámbrica, que posiblemente incluso puede ser de la modalidad *ad hoc*.

Para este ejemplo, la movilidad es una situación que debe pensarse muy detenidamente. La configuración puede ser un problema para el usuario móvil, ya que las distintas configuraciones de red pueden suponer un reto si la estación inalámbrica del usuario no tiene capacidad para configurarse automáticamente.

Retos de Configuración

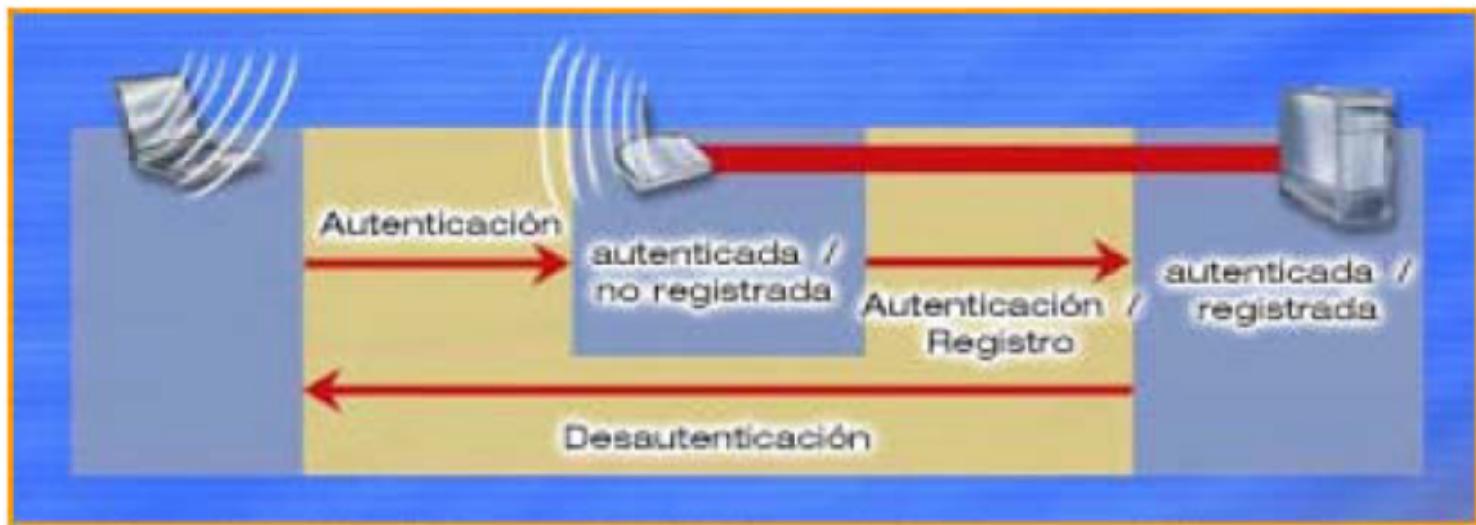
Ahora que tenemos una conexión de red inalámbrica y la complejidad ha aumentado, posiblemente hay muchas más configuraciones que realizar. Por ejemplo, podría ser necesario configurar el SSID de la red a la que se va a realizar la conexión. O bien, podría ser necesario configurar un conjunto de claves WEP de seguridad; posiblemente, varios conjuntos de claves si es necesario conectarse a varias redes. Podría ser necesario tener una configuración para el trabajo, donde la red funciona en modo de infraestructura, y otra configuración para el domicilio, donde funciona en modo *ad hoc*. Entonces, sería necesario elegir qué configuración se va a utilizar en función del lugar donde nos encontremos.

Mecanismos de Seguridad

Control de Acceso: Codificación y Autenticación

Por sí solo, el modo de modulación no garantiza una red local inalámbrica segura. Se necesitan varios requisitos para obtener una seguridad real:

- Asignación de clave SSID: Cada usuario (cliente o punto de acceso) de una red WLAN recibe su propia identificación SSID que ha sido asignada por el administrador de red al configurar la red inalámbrica.
- Direcciones MAC: El fabricante asigna una única dirección MAC global a cada adaptador de una red WLAN. La dirección se debería introducir en las listas de acceso para el punto de acceso. Todos los demás adaptadores de red se rechazan automáticamente.
- Autenticación: Cada estación debe probar que está autorizada para conectarse a la red WLAN correspondiente. Por este motivo, los productos para WLAN actuales utilizan el algoritmo WEP.
- Codificación WEP: El estándar 802.11 implementa WEP como su tecnología de codificación. La versión más segura a 128 bits se debería utilizar (como hace Intel) para disfrutar de una mayor seguridad.
- Utilice la tecnología de red privada virtual (VPN): Las redes privadas virtuales (VPN) llevan ya bastante tiempo operativas y están consideradas como muy seguras. Se obtendrá una red local inalámbrica segura si se sabe sacar provecho de esta tecnología para redes inalámbricas.



Un escenario típico para muchos usuarios es una pequeña oficina (agencias, etc) que se reparte en varias habitaciones de una misma planta.

Cifrado WEP

Resulta evidente a todas luces que las comunicaciones inalámbricas ofrecen un punto de vulnerabilidad en la transmisión de datos, puesto que las emisiones difícilmente pueden acotarse a la zona de cobertura, sino que habitualmente suelen alcanzar puntos fuera del área de transmisión deseada. Para paliar que otros receptores ajenos a la red corporativa y a los intereses de la empresa puedan hacer un uso indebido de la información que viaje por el aire se ha adoptado un sofisticado mecanismo de control de acceso al medio (DSSS), lo cual evita en gran medida las escuchas indiscretas. No obstante, este sistema no es suficiente, por lo que opcionalmente se puede realizar un proceso de cifrado de los datos que se transmiten por la red inalámbrica.

A la hora de proteger la información que viaja por el espacio mediante sistemas de cifrado se puede hacer uso de las técnicas WEP-40 y WEP-128. Estos dos sistemas son funciones

opcionales de la especificación IEEE 802.11 que proporcionan una confidencialidad de datos equivalente a la de una LAN cableada sin cifrar. Es decir, el sistema WP hace que el enlace LAN inalámbrico en una red sea tan seguro como el enlace con cable.

Como se especifica en el estándar, WEP (Wired Equivalent Privacy) utiliza el algoritmo RC4 con una clave de 40 bits para WEP-40 o una clave de 128 bits para WEP-128. Cuando la función WEP está activada, cada estación (cliente o punto de acceso) se le asigna una clave común. Esta clave desordena los datos y se mezcla entre la información antes de ser transmitida, de tal modo que si una estación recibe un paquete que no está mezclado con la clave correcta, la estación descartará el paquete.

A la hora de la verdad, la instalación de esta función es opcional, aunque sumamente sencilla de poner en marcha. Simplemente habrá que seleccionar el tipo de encriptación WEP que se desea implementar, 40 o 128 bits, y a continuación elegir la clave que se utilizará. Obviamente, si la función WEP está activada en uno o más puntos de acceso, todos los dispositivos inalámbricos de la red deberán tener el mismo código WEP, que se establece fácilmente mediante las utilidades de software suministradas.

No obstante, existe la posibilidad de establecer una comunicación con células combinadas. Una célula combinada es una red de radio en la que algunos dispositivos utilizan WEP y otros no. Esta opción es posible mediante la simple activación del parámetro “Allow Association To Mixed Cells”.

OSA (Open System Authentication)

Es otro mecanismo de autenticación definido por el estándar 802.11 para autenticar todas las peticiones que recibe. El principal problema que tiene es que no realiza ninguna comprobación de la estación cliente. Además las tramas de gestión son enviadas sin encriptar, aún activando WEP. Por lo tanto es un mecanismo poco fiable.

ACL (Access Control List)

Este mecanismo de seguridad es soportado por la mayoría de los productos comerciales. Utiliza, como mecanismo de autenticación, la dirección MAC de cada estación cliente, permitiendo el acceso a aquellas MAC que consten en la Lista de Control de Acceso.

CNAC (Closes Network Access Control)

Este mecanismo pretende controlar el acceso a la red inalámbrica y permitirlo solamente a aquellas estaciones cliente que conozcan el nombre de la red (SSID) actuando este como contraseña.

WPA

WPA (Wi-Fi Protected Access) es un nuevo protocolo para reemplazar al desacreditado WEP. Incrementa el nivel de protección de datos y el control de acceso para las WLAN existentes y las futuras.

WPA está diseñado para correr en el hardware existente como una actualización de software. Se deriva de la próxima versión del estándar IEEE 892.11, el 802.11i, y desde luego será compatible con él.

WPA utiliza el protocolo temporal de intercambio de claves, TKIP, una tecnología de cifrado por claves más segura que la RC4 de WEP. Además proporciona autenticación de usuario, que no estaba casi presente en WEP. Para fortalecer la autenticación de usuario WAP implementa los protocolos 802.1x y EAP (Extensible Authentication Protocol).

En un entorno empresarial se utiliza un servidor central de autenticación, como por ejemplo RADIUS, para autenticar a cada usuario antes de que se una a la red y además se

utiliza la “autenticación mutua”, de forma que el usuario no se une accidentalmente a una red maliciosa que pudiera robarle sus credenciales de red.

En entornos domésticos o de pequeña oficina donde no hay servidores centrales de autenticación o EAP, WAP corre en un modo especial denominado PSK (Pre-Shared Key). En este modo el usuario introduce claves manualmente en los puntos de acceso o pasarelas domésticas inalámbricas y en cada PC que está en la red inalámbrica. Desde este punto WPA toma automáticamente el control: por un lado se permite que se unan a la red sólo dispositivos con la misma password y por otro lado arranca automáticamente el proceso de encriptación TKIP.

Beneficios

Entre los beneficios principales de las redes de área local inalámbricas se encuentran:

- Alternativa o extensión de una solución de cableado: Existen múltiples casos en que es más económico implementar una solución inalámbrica, en reemplazo a un cableado o como extensión de una red alámbrica: segmento residencial, edificios sin cableado estructurado, oficinas arrendadas, campamentos, plantas libres de gran tamaño, etc.
- Aumentar la productividad de los empleados: Una implementación de WLAN permite que los empleados puedan tener acceso a la información en cualquier lugar. Esto permite que su tiempo sea más eficiente, y cuenten en todo momento con la información necesaria para la toma efectiva de decisiones. En general, cuando se realiza un análisis de retorno de inversión de una solución de WLAN, este aumento de productividad, por sí solo, justifica la implementación de una WLAN.
- Reutilización de infraestructura: En algunas industrias es necesario reconfigurar la distribución física de una red de datos, o incluso, cambiar de lugar las instalaciones frecuentemente, en estos casos, las WLAN proveen una solución eficiente en coste que permiten la reutilización de la inversión.
- Información en tiempo real: Una solución de rede de área local inalámbrica permite utilizar dispositivos de captura de datos o computadores portátiles para ingresar información en tiempo real.

Glosario de Términos

- **Ad-Hoc** : Es un sistema de red inalámbrica (802.11) que permite que los clientes que están situados en un rango determinado puedan conectarse entre ellos sin necesidad de la presencia de un Punto de Acceso. En este tipo de red, cualquier cliente puede hacer las veces de punto de acceso proporcionando a los demás acceso a Internet o cualquier otro servicio. También se le denomina Peer to Peer, o IBSS (referido al estándar 802.11). El otro sistema de red inalámbrica que emplea Puntos de Acceso se le denomina Infraestructura.
- **Bluetooth** : Tecnología y protocolo de conexión entre dispositivos inalámbricos que integran un chip específico para comunicarse en la banda de frecuencias 2,402-2,480 GHz con un alcance máximo de 10 metros y tasas de transmisión de datos de hasta 721 Kbps (en la segunda generación de Bluetooth). Cada dispositivo posee una dirección única de 48 bits que lo identifica de manera inequívoca, siguiendo el estándar IEEE 802.
- **Bridge (Puente)** : Dispositivo que pasa todos los mensajes de una red a otra sin distinguir a cuál red pertenece el destino del mensaje.

- **Cobertura** : Área geográfica próxima a un nodo o estación base que recibe suficiente señal para mantener una conexión. Depende de diversos factores como tipo de antena, ubicación, topografía del terreno, potencia de la señal, etc.
- **Enlace Punto a Punto** : Enlace en el que las comunicaciones están dirigidas entre dos puntos de conexión concretos. En Redes Wireless suelen ser las conexiones que enlazan dos nodos para ampliar el alcance de la red, o para conectar dos redes remotas. Se suelen emplear antenas directivas de gran ganancia (dependiendo de la distancia que separe los nodos).
- **Hot Spot** : Lugar donde existe un punto de acceso en una WLAN que ofrece servicio de banda ancha a usuarios móviles.
- **IEEE 802.x** : Conjunto de especificaciones de las redes LAN dictadas por el IEEE (the Institute of Electrical and Electronic Engineers). Existe un comité 802.11 trabajando en una normativa para redes inalámbricas de 1 y 2 Mbps. La norma tendrá una única capa MAC para las siguientes tecnologías: Frequency Hopping Spread Spectrum (FHSS), Direct Sequence Spread Spectrum (DSSS) e infrarrojos.
- **Infraestructura de red** : Red inalámbrica centrada en un punto de acceso. En este entorno los puntos de acceso no solo proporcionan comunicación con la red cableada sino que también median el tráfico de red en la vecindad inmediata.
- **PAN (Red de Área Personal)** : Sistema de red conectado directamente a la piel. La transmisión de datos se realiza por contacto físico. También se le llama así a la red de área personal, un conjunto de dispositivos que normalmente son de uso personal, ej: micrófono, teclado, impresora conectados con un computador.
- **Punto de acceso** : Dispositivo que permite comunicar a varios clientes wireless entre ellos e incluso con otras redes inalámbricas o de cable. Los puntos de acceso hacen las funciones de Bridges y algunos incluso de routers (encaminadores).
- **Router (Encaminador)** : Dispositivo hardware (o software) para redes informáticas dotado de capacidad para conmutación y con la principal finalidad de proporcionar un encaminamiento de paquetes IP.
- **SSID** : El primer paso para poder autenticar un cliente en una red wireless es el conocimiento del SSID (Service Set Identifier). Para obtener acceso al sistema es necesario conocer el SSID.
- **VPN (Red Privada Virtual)** : Configuración lógica de una serie de componentes hardware, que permite la utilización de redes públicas para establecer canales de comunicaciones privados a los que sólo pueden acceder usuarios autorizados.
- **WPAN (Red Inalámbrica de Área Personal)** : Redes locales-personales que utilizan tecnología Bluetooth.
- **WECA (Alianza para la Compatibilidad de Ethernet Inalámbrica)** : Alianza de fabricantes formada para mantener la compatibilidad entre dispositivos wireless. La WECA creó el estándar de dispositivos inalámbricos Wi-Fi, que cumplen la norma IEEE 802.11b.
- **Wi-Fi (Wireless Fidelity)** : Sinónimo del estándar IEEE 802.11b, protocolo de transmisión inalámbrica que logra alcanzar desde 2 Mbps hasta un máximo teórico de 11 Mbps. Este estándar fue creado por un grupo de fabricantes de dispositivos inalámbricos para mantener la compatibilidad entre sus productos. Permite crear redes de ordenadores sin que exista un cable de por medio, usando para ello ondas de radio.
- **Wireless (Inalámbrico)** : Es un sistema de comunicación que utiliza ondas de radiofrecuencia, ultrasonido o rayos infrarrojos (IR) para intercambiar datos entre dispositivos. Cada vez se está popularizando más el uso de este sistema para transferencia de datos entre cámaras digitales, PDAs, calculadoras, etc. con el ordenador. En Internet, este término es utilizado para indicar que la transmisión de información se efectúa prescindiendo de cables. Es el caso de los celulares con sistema WAP, o las conexiones a Internet.
- **WLAN (Red de Área Local Inalámbrica)** : Una WLAN es un tipo de red de área local (LAN) que utiliza ondas de radio de alta frecuencia en lugar de cable para comunicar y transmitir datos entre los clientes de red y los dispositivos. Es un sistema de comunicación de datos flexible implementado como una extensión, o como una alternativa para una LAN conectada. Al igual que una LAN, la red permite que

los usuarios de esa ubicación comparten archivos, impresoras y otros servicios. La mayoría de las redes WLAN utilizan tecnología de espectro distribuido. Su ancho de banda es limitado (generalmente inferior a 11 Mbps) y los usuarios comparten el ancho de banda con otros dispositivos del espectro; no obstante, los usuarios pueden operar dispositivos de espectro distribuido sin autorización de la FCC (Comisión Federal de Comunicaciones).

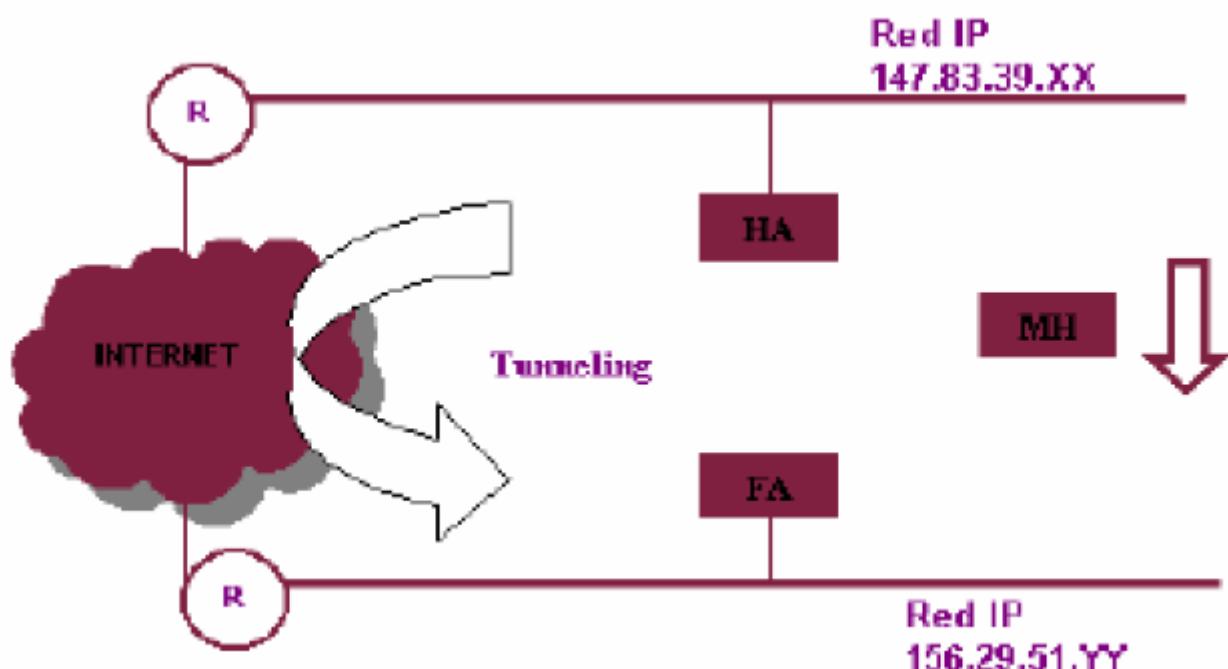
IP Móvil

Introducción

En los últimos años se han ido produciendo numerosos avances en el campo de las tecnologías de comunicación. Dos de los más relevantes son, sin duda, el rápido desarrollo de la informática portátil y la importante implantación de los sistemas de comunicación móviles. La conjunción de ambos factores permite a los usuarios acceder a una red en cualquier momento y en cualquier lugar, aún cuando se encuentren en movimiento.

Sin embargo, los actuales protocolos de internet working (TCP/IP, IPX o Apple Talk) presentan serias complicaciones a la hora de tratar con nodos que disponen de un cierto grado de movilidad entre redes. La mayoría de las versiones del protocolo IP (Internet Protocol) asumen de manera implícita que el punto al cual el nodo se conecta a la red es fijo. Por otra parte, la dirección IP del nodo identifica al mismo de manera única en la red a la que se encuentra conectado. Por consiguiente, cualquier paquete destinado a ese nodo es encaminado en función de la información contenida en la parte de su dirección IP que identifica la red en que está conectado. Esto implica que un nodo móvil que se desplaza de una red a otra y que mantiene su dirección IP no será localizable en su nueva situación ya que los paquetes dirigidos hacia este nodo serán encaminados a su antiguo punto de conexión a la red. El protocolo IP Móvil constituye una mejora del protocolo IP citado anteriormente. *Mobile IP* permite a un nodo circular libremente a través de Internet siendo éste siempre accesible mediante una única dirección IP.

Arquitectura IP Móvil



La *Internet Engineering Task Force* (IETF) propone una arquitectura *Mobile IP* que funciona, a grandes rasgos, bajo el siguiente concepto: un agente local, denominado

Home Agent (HA) y un agente externo, también denominado *Foreign Agent* (FA) colaboran para permitir que el nodo móvil o *Mobile Host* (MH) pueda moverse conservando su dirección IP inicial.

La Capa de Red Internet

En la capa de red, Internet puede verse como un conjunto de subredes, o sistemas autónomos (AS, Autonomous System) interconectados. No hay una estructura real, pero existen varios *backbone* principales. Estos se construyen a partir de líneas de alto ancho de banda y enrutadores rápidos. Conectadas a los *backbone* hay redes regionales (de nivel medio), y conectadas a estas redes regionales están las LAS de muchas universidades, compañías y proveedores de servicio Internet.

El pegamento que mantiene unida a Internet es el protocolo de capa de red, IP (*Internet Protocol* o *Protocolo de Internet*). A diferencia de la mayoría de los protocolos de capa de red anteriores, éste se diseñó desde el principio con la interconexión de redes en mente. Una buena manera de visualizar la capa de red es la siguiente. Su trabajo es proporcionar un medio de mejor esfuerzo para el transporte de datagramas del origen al destino, sin importar si estas máquinas están en la misma red, o si hay otras redes entre ellas.

La comunicación en Internet funciona como sigue. La capa de transporte toma corrientes de datos y las divide en datagramas. En teoría, los datagramas pueden ser de hasta 64 Kbytes cada uno, pero en la práctica por lo general son de unos 1500 bytes. Cada datagrama se transmite a través de Internet, posiblemente fragmentándose en unidades más pequeñas en el camino. Cuando todas las piezas llegan finalmente a la máquina de destino, son reensambladas por la capa de red, dejando el datagrama original. Este datagrama entonces es entregado a la capa de transporte, que lo introduce en la corriente de entrada del proceso receptor.

Objetivos de IP Móvil

IP Móvil (Mobile IP) es un concepto muy amplio, se puede aplicar a 3 formas distintas de movilidad:

- **Ordenadores portátiles** : que se transportan y conectan desde lugares remotos a Internet. Un buen ejemplo sería el ordenador de un hombre de negocios que a lo largo del día podría conectarse a Internet desde casa, la oficina, el centro de conferencias, Internet café, ... Se podría usar una configuración dinámica para conseguir direcciones temporales en una red remota, sin embargo, esto no es más que una solución parcial. Por poner un ejemplo, las conexiones TCP se identifican por un par de dirección IP y puerto, por lo que usando direcciones temporales implica que este tipo de conexiones no sobrevivirían.
- **Ordenadores móviles** : este tipo de ordenadores pueden mantener la conexión mientras se mueven de forma transparente al usuario análogamente a como lo hace un teléfono móvil. En cada celda hay una estación conectada a Internet que es capaz de intercambiar paquetes entre el canal de radio de la celda e Internet. Los ordenadores estarán escuchando constantemente las señales que proceden de las estaciones y escogerán aquella cuya señal sea más clara. El objetivo de la tecnología de IP móvil es permitir la transición de celda a celda (llamado 'roaming' o tránsito, vagabundeo...) mientras se mantiene a la unidad móvil conectada a Internet manteniendo la misma dirección IP, para que así las conexiones TCP puedan mantenerse. La tecnología utilizada a nivel físico va desde transmisión por infrarrojo hasta las ondas de radio de amplio espectro, cada una con sus ventajas e inconvenientes.
- **Redes móviles** : se trata de una red que en su totalidad es móvil. Por ejemplo, un portaaviones tiene su propia red interna y además necesita conectividad al exterior. Una de las primeras aplicaciones no militares de este tipo de redes ha sido en la fórmula uno con la conectividad de los coches (que tienen en la actualidad gran

cantidad de mini computadores) con los boxes. Se puede añadir una nueva dimensión en la movilidad de IP, añadiendo *hosts* móviles a estas redes móviles.

Cuando el grupo de trabajo de IP Móvil del IETF empezó a trabajar, una de sus primeras tareas fue delimitar los requerimientos del protocolo:

- Un nodo móvil debe ser capaz de comunicarse con otros nodos después de haber cambiado su punto de conexión a Internet, sin cambiar su dirección IP.
- Un nodo móvil debe ser capaz de comunicarse con otros nodos que no implementan las funciones de movilidad. De hecho el resto de *routers* y *hosts* que no implementan IP Móvil no tienen por qué mejorar su versión del protocolo para poder comunicarse con *hosts* móviles. Es impensable necesitar cambiar todos los *hosts* y *routers* de Internet para adaptarlos a IP Móvil.
- Todos los mensajes usados para actualizar la información de *host* móvil deben ser autenticados como medida de protección contra ataques remotos.

Estos son los objetivos fundamentales, hay algunos otros secundarios:

- Posibilidad *Multicast*. Presenta problemas nuevos a la ya mayor complejidad e *Multicast* (envío de un mensaje de un usuario a un grupo determinado de usuarios). El hecho de que los usuarios de un servicio *Multicast* puedan moverse hace que el árbol de expansión de la difusión *multicast* tenga que ser recalculado constantemente.
- Privacidad de información de localización. Se pretende esconder la localización de los *hosts* a otros *hosts*, para que nadie pueda 'rastrear' las celdas por las que un *host* ha estado moviéndose.
- Minimizar el número de mensajes administrativos que se envían por dos razones: en primer lugar las conexiones inalámbricas son de bajo ancho de banda y muy propensas a errores; en segundo lugar los nodos móviles están alimentados por baterías y minimizar el consumo de energía es importante.

Terminología

A lo largo del documento usaremos frecuentemente estos términos:

- Nodo: un *host* o *router*.
- Nodo móvil: un *host* o *router* que cambia su 'punto de anclaje' de una red o subred a otra. Cuando cambia de lugar debe poder mantener la comunicación y su dirección IP.
- Agente doméstico (*home agent*): un *router* en la red doméstica del nodo móvil que envía datagramas al nodo móvil cuando está fuera de la red doméstica a través de un túnel. Además mantiene información sobre la localización del nodo móvil.
- Agente ajeno o exterior (*foreign agent*): un *router* en la red que está visitando el nodo móvil que da servicios de routing al nodo móvil mientras está registrado. Recibe la información del túnel que ha enviado el agente doméstico y la entrega al nodo móvil. Sirve también como *router* por defecto para los nodos móviles registrados en su red.
- Anuncio de Agente (*Agent advertisement*): un mensaje de anuncio construido añadiendo una extensión especial a un mensaje de anuncio.
- *Care-of-address* (COA): el nodo móvil recibe una dirección IP fija en su red local, esta es permanente. Cuando el nodo está fuera de su red local se le asigna una dirección de 'de reenvío' o care-of address (en inglés care of se usa cuando se envía un correo o fax a una dirección pero va dirigido a otra persona, se podría traducir como 'a la atención de ...'). Esta care-of address (COA) se asocia con el nodo móvil y representa su punto de anclaje a la Internet actual. El nodo móvil usa su dirección fija como dirección de origen en todos sus datagramas a excepción de algunos mensajes de administración de movilidad.

Fundamentos de IP Móvil

Hoy en día, millones de personas tienen ordenadores portátiles, y generalmente quieren leer su correo electrónico y acceder a sus sistemas de archivos normales desde cualquier lugar del mundo. Estos *hosts* móviles generan una nueva complicación: para enrutar un paquete a un *host* móvil, la red primero tiene que encontrarlo. El tema de la incorporación de *host* móviles en una red es muy nuevo, pero en esta sección plantearemos algunos de los problemas relacionados y sugeriremos una posible solución.

Se dice que son *estacionarios* los usuarios que nunca se mueven; se conectan a la red mediante hilos de cobre o fibra óptica. En contraste distinguimos otros dos tipos de usuarios. Los usuarios *migratorios* básicamente son los usuarios estacionarios que se mueven de un lugar fijo a otro de tiempo en tiempo, pero que usan la red sólo cuando están conectados físicamente a ella. Los usuarios *errantes* usan su ordenador en movimiento, y quieren mantener sus conexiones mientras se mueven. Usaremos el término *usuarios móviles* para indicar las dos últimas categorías.

Se supone que todos los usuarios tienen una localidad base que nunca cambia. Los usuarios también tienen una dirección de base permanente, que puede servir para determinar su localidad base, de manera análoga a como el número +034-93-3383694 indica España (código de país +034) y Barcelona (93). La meta de enrutamiento en los sistemas con usuario móviles es posibilitar el envío de paquetes a usuarios móviles usando su dirección base, y hacer que los paquetes lleguen eficientemente a ellos en cualquier lugar en el que puedan estar. Lo difícil, por supuesto, es encontrarlos.

El mundo se divide (geográficamente) en unidades pequeñas. Llamémoslas áreas, siendo un área típicamente una LAN o una célula inalámbrica. Cada área tiene uno o más agentes externos, que llevan el registro de todos los usuarios que visitan el área. Además, cada área tiene un agente de base, que lleva el registro de todos los usuarios móviles cuya base está en el área, pero que actualmente están visitando otra área. Al entrar un usuario nuevo en un área, ya sea al conectarse a ella (por ejemplo, conectándose a la LAN), o simplemente al entrar en la célula, su ordenador debe registrarse con el agente externo de ese lugar.

El procedimiento de registro, funciona de esta manera:

- Periódicamente, cada agente externo difunde un paquete que anuncia su existencia y dirección. Un *host* móvil recién llegado puede esperar uno de estos mensajes, pero si no llega ninguno, el *host* móvil puede difundir un paquete que diga: “¿hay agentes externos por ahí?”.
- El *host* móvil se registra con el agente externo, dando su dirección base, su dirección actual de capa de enlace de datos y cierta información de seguridad.
- El agente externo se pone en contacto con el agente de base del *host* móvil y le dice: “uno de tus *hosts* está por aquí”. El mensaje del agente externo al agente de base contiene la dirección de red del agente externo, así como la información de seguridad, para convencer al agente de base de que el *hosts* móvil en realidad está ahí.
- El agente de base examina la información de seguridad, que contiene una marca de tiempo, para comprobar que fue generada en los últimos segundos. Si está conforme, indica al agente externo que proceda.
- Cuando el agente recibe el reconocimiento del agente de base, hace una entrada en sus tablas e informa al *host* móvil que ahora está registrado.

Movilidad IP

Muchos usuarios de Internet tienen ordenadores portátiles y quieren mantenerse conectados a Internet en sus desplazamientos. Desafortunadamente, el sistema de

direcccionamiento IP hace que el trabajo lejos de casa sea más fácil de plantear que de hacer.

El verdadero problema es el esquema de direcccionamiento. Cada dirección IP contiene tres campos: la clase, el número de red y el número de host. Por ejemplo, considere la máquina con dirección IP 160.80.40.20. El 160.80 da la clase (B) y el número de red; el 40.20 es el número de host. Los enrutadores de todo el mundo tienen tablas de enrutamiento que indican la línea a usar para llegar a la red 160.80. Cuando llega un paquete con una dirección IP de destino de forma 160.80.xxx.yyy, sale por esa línea.

Si de pronto la máquina con esa dirección se lleva a algún lugar lejano, los paquetes para ella se seguirán enviando a su LAN (o enrutador) base. El dueño ya no podrá recibir correo electrónico, etc. Darle a la máquina una nueva dirección IP correspondiente a su nueva ubicación no es muy atractivo, pues habría que informar a una gran cantidad de gente, programas y BD sobre el cambio.

Otro enfoque es hacer que los enrutadores usen direcciones IP completas para el enrutamiento, en lugar de sólo la clase y la red. Sin embargo, esta estrategia requeriría que cada enrutador tuviera tablas de millones de entradas, con un coste astronómico para Internet.

Cuando la gente comenzó a exigir la posibilidad de tener hosts móviles, el IETF (Internet Engineering Task Force) estableció un grupo de trabajo para encontrar una solución. El grupo de trabajo pronto formuló varias metas deseables en cualquier solución. Las principales fueron:

- Todo host móvil debe ser capaz de usar su dirección IP base en cualquier lugar.
- No se permiten cambios al software de los hosts fijos.
- No se permiten cambios al software del enrutador ni a sus tablas.
- La mayoría de los paquetes para los hosts móviles no deben desviarse en el camino.
- No debe incurrirse en carga extra cuando un host móvil está en su base.

En síntesis, la solución escogida consiste en que cada instalación que quiera permitir la movilidad de sus usuarios debe crear un agente de base. Cada instalación que permita visitantes tienen que crear un agente externo. Al aparecer un host móvil en una instalación externa, se pone en contacto con el host externo y se registra. El host externo entonces se comunica con el agente de base del usuario y le da una dirección de encargado (care-of address), normalmente la misma dirección IP del agente externo.

Al llegar un paquete a la LAN base del usuario, llega por un enrutador conectado a la LAN. El enrutador entonces trata de localizar al host de la manera normal, difundiendo un paquete ARP que pregunta por ejemplo: “¿cuál es la dirección Ethernet de 160.80.40.20?”. El agente de base responde a esta solicitud dando su propia dirección Ethernet. En enrutador entonces envía los paquetes para 160.80.40.20 al agente de base. Éste a su vez, los envía a través de un túnel a la dirección de encargado, encapsulándolos en el campo de carga útil de un paquete IP dirigido al agente externo. El agente externo entonces los desencapsula y los entrega a la dirección de enlace de datos del host móvil. Además, el agente de base entrega la dirección de encargado al transmisor, para que los paquetes futuros puedan enviarse en túnel directamente al agente externo. Esta solución cumple con todos los requisitos indicados antes.

En el momento de moverse el host móvil, el enrutador probablemente tiene en caché sus direcciones Ethernet (que pronto dejarán de ser válidas). Para reemplazar esa dirección de Ethernet por la del agente de base, se usa un truco llamado ARP gratuito (gratuitous ARP). Éste es un mensaje especial al enrutador, no solicitado, que causa que reemplace una entrada específica de caché, en este caso la del host móvil a punto de desconectarse. Al regresar el host móvil, se usa el mismo mecanismo para actualizar la caché del enrutador.

Nada en el diseño impide que un host móvil sea su propio agente externo, pero este enfoque sólo funciona si el host móvil (en su capacidad como agente externo) está conectado lógicamente a Internet en su instalación actual. También, debe poder adquirir una dirección IP de encargado (temporal) para usarla. Esa dirección IP debe pertenecer a la LAN a la que está conectado actualmente.

La solución IETF (Internet Engineering Task Force) para hosts móviles resuelve otros problemas no mencionados hasta ahora. Por ejemplo, ¿cómo localizar a los agentes?. La solución es que cada agente difunda periódicamente sus direcciones y el tipo de servicios que está dispuesto a proporcionar (por ejemplo, base, externo o ambos). Al llegar un host móvil a algún lado, simplemente puede esperar estas difusiones, llamadas anuncios (advertisements). Como alternativa, puede difundir un paquete anunciando su llegada y esperar que el agente externo local responda.

Otro problema que tenía que resolverse es lo que había que hacer con los hosts móviles que salen sin decir adiós. La solución es hacer que el registro sea válido sólo durante un intervalo de tiempo fijo; si no se renueva periódicamente, termina su temporización, y así el agente externo puede limpiar sus tablas.

Un tema más es el de la seguridad. Cuando un agente de base recibe un mensaje solicitándole que redirija todos los paquetes de Pepe a alguna dirección IP, no debe acceder a menos que esté convencido de que Pepe es el origen de la solicitud, y no alguien que está tratando de hacerse pasar por él. Se usan protocolos específicos cifrados de verificación de autenticidad para este fin.

Alternativas a IP Móvil para Dotación de Movilidad a las Estaciones IP

Para dotar de movilidad a un nodo de la red, aparecen diferentes alternativas a IP Móvil que son estudiadas con más detalle para ver la viabilidad de su implementación en Internet. Así se concluirá que IP Móvil es la solución adecuada para proporcionar movilidad IP.

Algunas de las soluciones que apuntamos son las siguientes:

- Establecimiento de rutas específicas para terminales con movilidad.
- Cambio de la dirección IP de los terminales.
- Soluciones basadas en realizar cambios a nivel de la capa de enlace.

Rutas específicas para Nodos con Movilidad

La utilización de rutas específicas para los nodos a los que se les quiere dotar de movilidad implica la reconfiguración de las tablas de encaminamiento de los dispositivos de interconexión de red (routers) para permitir contactar con el nodo móvil en su nueva ubicación. Esta solución es extremadamente costosa debido al gran incremento de tráfico que se generaría en la red para soportar la movilidad de los terminales. Para ello sería necesario actualizar las tablas de encaminamiento de cómo mínimo todos los routers entre el enlace local y el nuevo punto de enlace.

Si se tiene en cuenta el número de nodos móviles en la red y la velocidad con que éstos cambian de ubicación, estas actualizaciones podrían llegar a colapsar la red. Por lo tanto es importante minimizar el número de routers a actualizar y esto a su vez limitará las posibilidades de encaminamientos alternativos propias del protocolo IP.

Cambio en la Dirección IP

Otra posible solución consiste en asignar al nodo móvil una nueva dirección IP acorde con su nuevo punto de conexión a la red. Esta solución no es en absoluto recomendable ya que

requiere que la entrada del nodo móvil cambia de dirección IP. Si esta operación no se realiza de forma instantánea, cualquier consulta de la dirección IP del nodo móvil puede ser errónea.

Por otra parte, y dada la velocidad a la cual el nodo móvil puede cambiar de dirección IP, se hace necesario un mecanismo para verificar la actualidad de la dirección IP devuelta por el servidor de nombres de dominio (DNS). El resultado es un gran número de consultas y actualizaciones que generan, al igual que en el caso anterior, un alto nivel de tráfico inyectado a la red.

Finalmente a nivel local un cambio de dirección IP provoca generalmente el cierre inmediato de todas las aplicaciones abiertas asociadas a la antigua dirección IP.

Soluciones a Nivel de la Capa de Enlace

Existen dos grandes soluciones a nivel de la capa de enlace que pretenden permitir la movilidad de los nodos. La primera de ellas se basa en el *Cellular Digital Packet Data* (células de paquetes de datos digitales CDPD), que se trata de un estándar diseñado para transmitir paquetes IP a través de canales de radio no utilizados por el servicio de voz en el sistema de telefonía celular norteamericano. El CDPD asigna a cada nodo móvil una dirección IP fija dentro de su área de cobertura. La segunda solución se basa en el estándar IEEE 802.11, realizado por el Institute of Electrical and Electronics Engineers (IEEE) para la comunicación de redes de área local inalámbricas.

Ambas soluciones presentan dos grandes inconvenientes. Por un lado, las soluciones a nivel de la capa de enlace proporcionan movilidad para un solo tipo de medio físico. Por lo tanto, para N tipos de medios físicos diferentes, se requieren N soluciones de movilidad diferentes. Por otro lado, estas soluciones proporcionan una movilidad más o menos limitada geográficamente, lo cual entra en directa contradicción con el afán expansionista de Internet. Como presentaremos en la siguiente sección, el protocolo IP Móvil es el único capaz de proporcionar movilidad en cualquier tipo de medio y en una extensa área geográfica.

Funcionamiento del Protocolo IP Móvil

Igual que todo protocolo, éste también consiste en la consecución de una serie de operaciones:

- Los agentes local y externo anuncian su presencia mediante el nodo móvil mediante mensajes de anuncio, los cuales son generados periódicamente en la red. Opcionalmente el nodo móvil puede solicitar tales mensajes a un agente cercano a través de un *mensaje de solicitud de agente*.
- El nodo móvil recibe el mensaje de anuncio y determina si se encuentra en su red local o por el contrario al moverse ha ido a parar a una red externa.
- Si el nodo móvil detecta que se encuentra en su red local operará sin funciones de movilidad. Por otro lado, si regresa tras haber sido registrado en otra red procede a “des registrarse” a través de su agente local.
- Si el nodo móvil detecta que se encuentra en una red externa, obtiene su dirección de remite (*care-of-address*) en la nueva red. Esta dirección puede ser la del agente externo o una dirección de remite colocada (*colocated care-of-address*).
- Si el nodo móvil se encuentra fuera del alcance de ningún tipo de agente, el nodo móvil debe obtener su dirección de remite como una dirección IP local por algún método, como podría ser el DHCP (*Dynamic Host Configuration Protocol, configuración dinámica de host*). En este caso se trata de una dirección de remite “colocada”.
- El nodo móvil registra su dirección de remite con su agente local. Este proceso se realiza enviando una solicitud de registro al agente local y recibiendo de éste un mensaje de contestación.

- Todo paquete destinado al nodo móvil es interceptado por el agente local y enviado mediante *tunneling* por éste hacia la dirección de remite. Al otro lado del túnel el agente externo recibe el paquete y lo envía al nodo móvil. Si éste posee una dirección de remite colocada, el agente externo no interviene en la recepción del paquete.
- Por su parte, los paquetes originados por el nodo móvil pueden ser transportados a la dirección IP de destino sin pasar necesariamente por el agente local.

Fases

En los puntos que exponemos a continuación tratamos de forma más concreta los procedimientos que sigue el desarrollo del protocolo en cuestión, y que apuntábamos en el punto anterior.

Descubrimiento del Agente

Es un procedimiento por el que el nodo móvil determina si se encuentra en su red local o si por el contrario y debido a su movimiento se encuentra en una red externa. Asimismo se utiliza para obtener la dirección de remite necesaria para el nodo móvil.

Este procedimiento es sencillo y utiliza dos tipos de mensajes mencionados anteriormente: el anuncio de agente y de solicitud de agente. En primer lugar lo que vamos a necesitar es un anuncio por parte del agente local o bien por parte del agente externo de la disponibilidad para aceptar un nodo móvil en su red. El agente local deberá estar siempre dispuesto para servir a sus nodos móviles. Para evitar posibles saturaciones que le impidan cumplir con su compromiso se permite una configuración de un agente local a una determinada población de agentes móviles.

Puede ocurrir que un agente externo no pueda servir a un nodo móvil que no pertenece a su red. A pesar de ello el agente externo no puede parar de emitir los mensajes para que el nodo móvil identifique que se encuentra dentro de su red de cobertura.

Este mensaje de anuncio consiste en un mensaje ICMP (Internet Control Message Protocol) el cual ha sido extendido para permitir abarcar esta nueva funcionalidad.

Anuncio de Agente

La primera acción a realizar para permitir la movilidad de un nodo es la de anunciar, por parte del agente local o externo, la disponibilidad para aceptar al nodo móvil en su red. El nodo móvil utiliza mensajes de anuncio para determinar su punto de conexión actual a Internet. El agente local deberá estar siempre listo para servir a sus nodos móviles. Para evitar una posible saturación debida al exceso de móviles en una determinada red, es factible configurar múltiples agentes locales en una única red local, asignando a cada agente local una porción de la población de nodos móviles.

Por otro lado, es posible que un agente externo no tenga capacidad para servir a un nodo móvil no perteneciente a su red. Aún en ese caso, el agente externo debe continuar emitiendo mensajes de anuncio para que el nodo móvil sepa que se encuentra dentro de su área de cobertura o que no ha fallado. El mensaje de anuncio consiste en un mensaje ICMP de anuncio de router al cual se la ha añadido una extensión para permitir la gestión de los nodos móviles.

Los campos de la extensión de anuncio de agente son los siguientes:

- *Type* : 16
- *Length* : (6+4*N), donde N es el número de direcciones de cuidado anunciadas.
- *Sequence number* : número total de mensajes de anuncio enviados desde que el agente fue inicializado.

- *Registration lifetime* : tiempo de vida máximo (S) durante el cual este agente acepta una solicitud de registro.
- *R* : registro solicitado. Es conveniente registrar con un agente externo en vez de usar una dirección de remite colocada.
- *B* : el agente externo no puede aceptar nuevos registros, al estar ocupado (*Busy*).
- *H* : este agente ofrece servicios de agente local (*Home Agent*) en esta red.
- *F* : este agente ofrece servicios de agente externo (*Foreign Agent*) en esta red.
- *M* : el agente soporta encapsulado mínimo.
- *G* : el agente soporta encapsulado GRE.
- *V* : el agente soporta la compresión de cabecera Van Jacobson.
- *Reserved* : reservado.
- *Care-of addresses* : la dirección de remite anunciada por el agente externo.

Campos del mensaje de anuncio de agente

	0	1	2	3
IP header (RFC 791)	Ver=4	IHL	Type of Service	Total Length
			Identification	Flags Fragment offset
	Time to Live		Protocol = ICMP	Header Checksum
				Source Address = home and/or foreign agent address on this link
				Destination Address = 255.255.255.255 (broadcast) or 224.0.0.1 (multicast)
ICMP Router Advertisement (RFC 1256)	Type = 9		Code	Checksum
	Num Addrs		Addr. Entry Size	Lifetime
				Router address (1)
				Preference Level (1)
				Router address (2)
				Preference Level (2)
				...
Mobility Agent Advert. Ext. RFC 2002	Type = 16	Length	Sequence Number	
		(max.) Registration Lifetime	R B H F M G V RESERVED	
				Care-of address (1)
				Care-of address (2)
				...
Prefix-Length Ext. (option.)	Type = 19	Length	Prefix Length (1)	Prefix Length (2)
				...
				...

Para que un nodo móvil pueda averiguar si se encuentra en su red local o no, ha de verificar los bits F y H de alguno de los mensajes de anuncio que capture, y además sabrá si el agente actúa como agente local o externo. La obtención de su dirección de remite se obtendrá a partir del campo de datos *Care-of address* indicado anteriormente.

Solicitud de Agente

Estos son los mensajes que realiza el nodo móvil cuando no puede, o quiere, esperar la transmisión periódica de mensajes de anuncio de agente. Es decir, este mensaje busca forzar la transmisión de un mensaje de anuncio a cualquier agente ubicado en el mismo enlace. El formato de este tipo de mensajes es igual al que explicábamos al adjuntar la

figura en el apartado anterior, con la salvedad que los mensajes de solicitud de agente deben tener su campo de Tiempo de Vida a 1 (*Time To Live* -TTL).

Registro

Hay varias circunstancias bajo las cuales un nodo móvil debe registrarse. La primera de ellas es cuando detecta que su punto de conexión a Internet ha variado respecto al que tenía anteriormente. También deberá registrarse si su registro anterior está a punto de caducarse (aunque no haya cambiado el punto de unión). Y, por último, cuando el nodo móvil esté en una red externa y detecte que su nodo externo se ha reiniciado.

El procedimiento de registro sirve para pedir los servicios de un agente externo. A continuación el nodo móvil comunica a su nodo local su “dirección de remite” en la red. Por el contrario, si el nodo móvil detecta que ha regresado a su red local debe iniciar el proceso para desregistrarse con su nodo local, para así poder funcionar como cualquier otro nodo fijo.

Son tres los pasos que componen el procedimiento de registro:

- El nodo móvil envía un mensaje de petición de registro. Según el caso, este mensaje se puede enviar al agente local o al externo (previa aceptación del mismo).
- El agente recibe la petición de registro y envía al nodo móvil un mensaje de contestación de registro, para informar si su petición de registro ha sido aceptada o no.
- Si el nodo móvil no recibe la contestación de registro en un período razonable de tiempo procederá a retransmitir las peticiones de registro con intervalos cada vez más largos entre ellos, hasta que al fin reciba contestación.

Para poder llevar a cabo este procedimiento es necesaria la cooperación entre los agentes local y externo, intercambiando mensajes de petición de registro.

Petición de Registro

Es el mensaje que el nodo móvil envía a su agente local para poder registrarse, y así éste podrá crear o modificar la entrada del nodo móvil en su lista de nodos con movilidad. Su formato se presenta en la siguiente figura:

Campos del mensaje de petición de registro

	0	1	2	3
IP Header (RFC 791)	Ver=4	IHL	Type of Service	Total Length
			Identification	Flags
	Time to Live	Protocol = UDP		Fragment offset
			Source Address	
			Destination Address	
UDP header		Source port		Destination Port = 434
		Length		Checksum
Fixed Length Portion of Registration Request	Type = 1			Lifetime
			Mobile Node's Home Address	
			Home Agent Address	
			Care-of Address	
			Identification	
			Optional Extensions	
			...	
Mobility Home Auten.. RFC 2002	Type = 32	Length	Security Parameter...	
		Index (SPI)		
			Autenticator	
			Optional extensions	

Los diferentes campos que conforman el mensaje de petición de registro son los siguientes:

- *Type* : 1 (Petición de registro).
- *S* : El nodo móvil solicita que el agente local mantenga sus anteriores entradas de movilidad.
- *B* : El nodo móvil pide, solicita al agente local que mande hacia él los paquetes *broadcast* (mensaje uno a todos) que se reciban en la red local.
- *D* : El nodo móvil informa al agente local que desencapsulará los paquetes que le sean enviados a su “dirección de remite”. Esto implica que el nodo móvil está utilizando una “dirección de remite colocada”.
- *M* : el nodo móvil solicita que el agente local utilice encapsulado mínimo para los paquetes destinados a él.
- *G* : El nodo móvil pide al agente local que utilice encapsulado GRE para los paquetes destinados a él.
- *V* : El nodo móvil solicita que el agente local emplee la comprensión de cabeceras de Van Jacobson.
- *Reserved* : Reservado.
- *Lifetime* : Número de segundos restantes antes de la caducidad del registro actual.
- *Home Address* : Dirección IP del nodo móvil.
- *Home Agent* : Dirección IP del agente local del nodo móvil.
- *Care-of Address* : “dirección de remite” = dirección IP a la salida del túnel.

- *Identification* : Número de 64 bits creado por el nodo móvil para asociar peticiones de registro con respuestas de registro. También sirve para proteger contra respuestas de registro fraudulentas.
- *Extensions* : Extensiones.

Respuesta de Registro

Como ya hemos apuntado anteriormente, tras la recepción de una petición de registro el agente local devuelve al nodo móvil un mensaje de respuesta de registro. Si el nodo móvil solicita el servicio a través de un agente externo, será éste quién reciba la respuesta de registro y la envíe a continuación al nodo móvil. Por el contrario, si el nodo móvil está utilizando una “dirección de remite colocada” será él mismo quien reciba el mensaje de respuesta de registro.

Este mensaje informa al nodo móvil del resultado de su petición de registro y del tiempo de vida de tal registro, que puede ser igual o inferior al solicitado por el nodo móvil. El agente externo no puede modificar, en ningún caso, el tiempo de vida asignado por el agente local. En la siguiente figura se muestra el formato de este mensaje.

Formato del mensaje de respuesta de registro

	0	1	2	3
Fixed Length Portion of Reg. Reply (RFC 2002)	Type = 3	Code	Lifetime	
		Mobile Node's Home address		
		Home Agent Address		
		Identification		

Los campos del mensaje son los siguientes:

- *Type* : 3 (Contestación de registro).
- *Code*: Código indicador del resultado de la petición de registro.
- *Lifetime* : Tiempo de vida, en segundos, de la entrada del nodo móvil en la lista de movilidad del agente local.
- *Home Address* : Dirección IP del nodo móvil.
- *Home Agent* : Dirección IP del agente local del nodo móvil.,
- *Identification* : Número de 64 bits creado por el nodo móvil para asociar peticiones de registro con contestaciones de registro. También sirve para proteger contra contestaciones de registro fraudulentas.
- *Extensions* : Extensiones.

Posibilidades Opcionales del Procedimiento de Registro

Además de las acciones anteriormente descritas, el procedimiento de registro permite también llevar a cabo otras interesantes funciones que se enumeran a continuación:

- Descubrir la dirección de un agente local si el nodo móvil no ha sido configurado con esta información.
- Seleccionar diferentes tipos de encapsulado de los paquetes.
- Utilizar la compresión de encabezados de Van Jacobson.
- Mantener varios registros simultáneos para que cada dirección de remite activa reciba una copia de los paquetes destinados al nodo móvil.

- Des registrar ciertas direcciones de cuidado manteniendo otras activas.

Encaminamiento

En este apartado presentaremos los diferentes modos en que un paquete puede ser encaminado de su dirección IP de origen hasta la dirección IP de destino, distinguiendo entre las dos posibles opciones: que el nodo móvil esté conectado a su red local, o bien que se encuentre en una red externa. Si el nodo móvil se encuentra en su red local, actúa como si de cualquier otro nodo fijo se tratase. Por lo tanto, las reglas para el encaminamiento de paquetes en este caso son las mismas que para el encaminamiento de paquetes IP hacia cualquier nodo o *router* convencional.

En el caso que el nodo móvil se encuentre en una red externa, distinguiremos dos situaciones:

- Nodo móvil como destino.
- Nodo móvil como origen.

Nodo móvil como destino

El protocolo IP Móvil requiere que los paquetes enviados desde la red local hasta el nodo móvil sean encapsulados. Esto altera el encaminamiento habitual ya que los paquetes atraviesan un nodo intermedio antes de llegar a su destino. Este nodo realizará el desencapsulado y enviará el paquete original al destino.

Las operaciones que comprenden el envío de un paquete hacia un nodo móvil en una red externa son:

- Un *router* en la red local, normalmente el agente local, anuncia que existe conectividad hasta el prefijo de red equivalente al de la dirección local del nodo móvil. Es decir, todo paquete destinado al nodo móvil es encaminado hacia su red local y recibido por su agente local.
- El agente local intercepta el paquete y consulta su entrada en su lista de movilidad para conocer las “direcciones de remite” registradas.
- El agente local envía una copia del paquete, encapsulándolo, hacia cada “dirección de remite” a través de túneles (*tunneling*).

En cada dirección de remite se extrae el paquete original y es entregado al nodo móvil. Si se trata de una dirección de remite de un agente externo, éste deshace el encapsulamiento del paquete. A continuación, consulta el campo de dirección IP de destino para comprobar si coincide con alguno de los nodos móviles a los que está prestando servicio, y si es así, el agente envía el paquete al nodo.

Si la dirección es *colocada*, el nodo móvil no recibe los servicios de ningún agente externo y, por lo tanto, efectúa el mismo las operaciones de desencapsulamiento.

Nodo móvil como origen

Si el nodo móvil depende de un agente externo, existen dos alternativas a la hora de determinar un router adecuado para dar salida a los paquetes:

- El propio agente externo, según especifica el campo *IP Source Address* del mensaje de anuncio de agente.
- Cualquier router cuya dirección IP aparezca en los campos Router Address del mensaje de anuncio de router. Siempre y cuando el nodo móvil sea capaz de determinar la dirección de la capa de enlace del router deseado, sin tener que enviar peticiones ARP (Address Resolution Protocol) que contengan su dirección local.

Si el nodo móvil posee una “dirección de remite colocada”, también tiene dos alternativas para elegir *router* :

- Escoger algún *router* que esté enviando mensajes de anuncio de *router* (no de agente) en la red en la que se encuentra.
- Mediante el mismo mecanismo por el que obtuvo su dirección de remite colocada puede obtener la dirección de un *router* adecuado. Por ejemplo, el protocolo DHCP ofrece todo tipo de información al nodo móvil, incluida la dirección de un *router*.

Contrariamente a los nodos móviles dependientes de un agente externo, los nodo móviles con una “dirección de remite colocada” pueden enviar peticiones ARP con su dirección local.

Resolución de Problemas: *Tunneling*

Los paquetes o datagramas son encapsulados (normalmente con datagramas IP) para alterar el normal enrutamiento (o encaminamiento), para que estos sean entregados a destinatarios intermedios que no constan en el campo de dirección de destino en la cabecera original IP. Una vez el datagrama encapsulado llega a este nodo intermedio se desencapsula, dejando el datagrama original IP, el cual es entonces entregado al destino, indicado en el original campo de dirección de destino. Este proceso de encapsular y desencapsular es frecuentemente llamado como “tunneling” del datagrama. Normalmente tenemos:

Fuente ----> encapsulador ----> desencapsulador ----> destino

El nodo encapsulador es generalmente considerado el punto de entrada al túnel, y el nodo desencapsulador es punto de salida. Pueden haber muchas parejas de fuente-destino usando el mismo túnel.

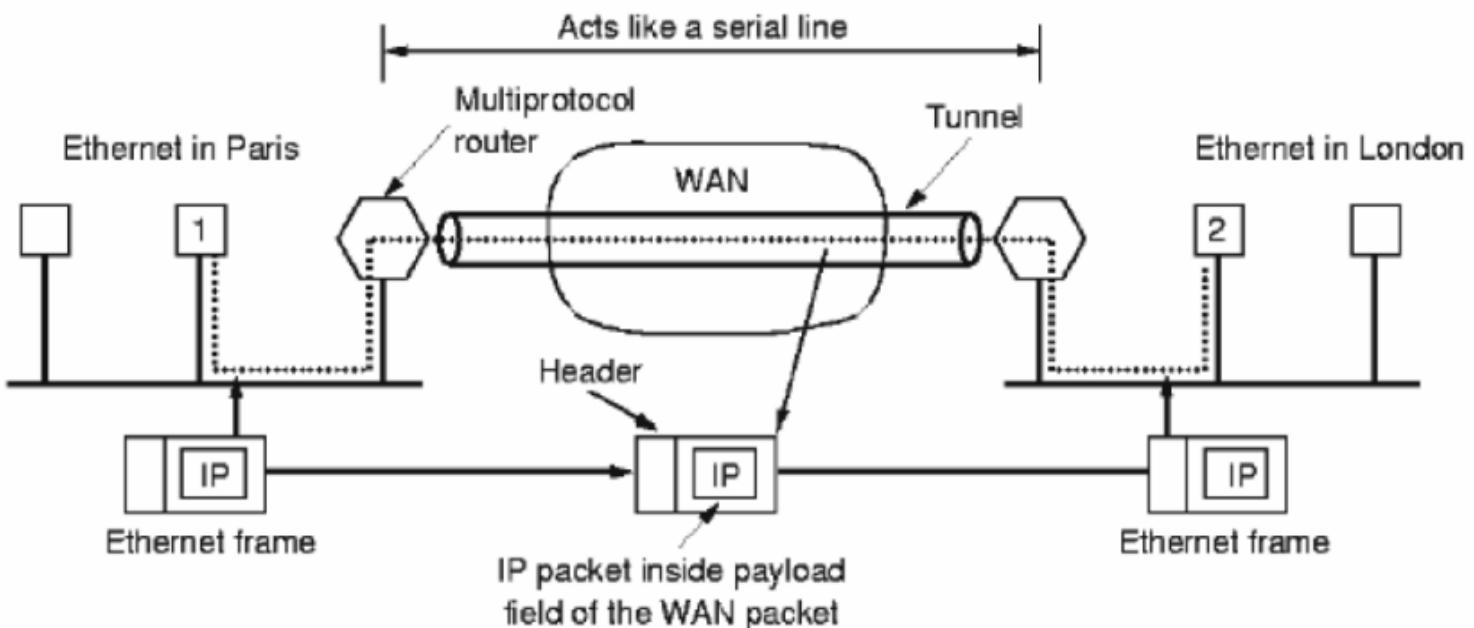
El protocolo Mobile IP ha especificado el uso del encapsulado como un modo de entregar datagramas desde un nodo móvil (“home network”) a un agente que puede entregar datagramas localmente cuando el nodo esté lejos de su red local. El uso del encapsulado también es conveniente cuando la fuente (o un router intermedio) de un datagrama IP quiere decidir la ruta por la cual será entregado a su destino. Por esto, las técnicas de encapsulado IP son especialmente útiles para realizar transmisiones multicast, e incluso llevar a cabo acciones de seguridad y privacidad en Internet.

El protocolo IP Móvil, requiere que los agentes locales, los agentes externos y los nodos móviles, que tengan una dirección de remite colocada, soporten el encapsulado IP-in-IP.

Ejemplo del Tunneling

El manejo del caso general de lograr la interacción de dos redes diferentes es extremadamente difícil. Sin embargo, hay un caso especial común que puede manejarse. Este caso se da cuando el host de origen y el de destino están en la misma clase de red, pero hay una red diferente en medio. Como ejemplo, piénsese en un banco internacional con una Ethernet basada en TCP/IP en París, una Ethernet basada en TCP/IP en Londres y una WAN PTT en medio, como se muestra en la siguiente figura:

Tunneling de un paquete de París a Londres

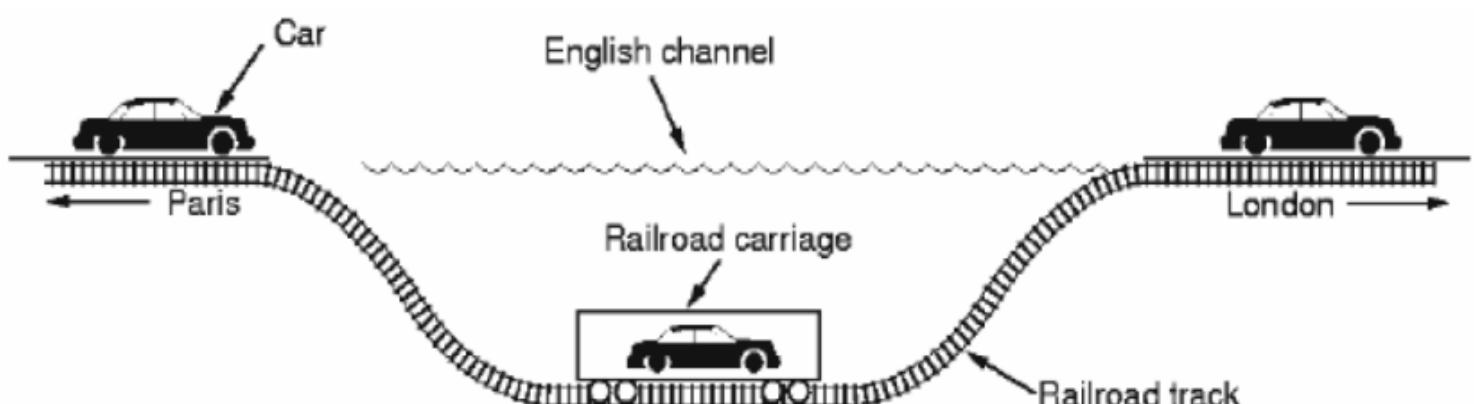


La solución a este problema es el proceso de túnel o tunneling. Para enviar un paquete IP al *host* 2, el *host* 1 construye el paquete que contiene la dirección IP del *host* 2, lo inserta en un marco Ethernet dirigido al enrutador multiprotocolo de París, y lo pone en el Ethernet. Cuando el enrutador multiprotocolo recibe el marco, retira el paquete IP, lo inserta en el campo de carga útil del paquete de capa de red de la WAN, y dirige este último a la dirección de la WAN del enrutador multiprotocolo de Londres. Al llegar ahí, el enrutador de Londres retira el paquete IP y lo envía al *host* 2 en un marco Ethernet.

La WAN puede visualizarse como un gran túnel que se extiende de un enrutador multiprotocolo al otro. El paquete IP simplemente viaja de un extremo del túnel al otro. No tiene que preocuparse por competir con la WAN. Tampoco tienen que hacerlo los *hosts* de cualquiera de los Ethernet. Sólo el enrutador multiprotocolo tiene que entender los paquetes IP y WAN. De hecho, la distancia completa entre la mitad de un enrutador multiprotocolo y la mitad del otro actúa como una línea serie.

Por ejemplo, considérese una persona que conduce su coche de París a Londres. En Francia, el coche se mueve por su propia energía, pero al llegar al Canal de la Mancha, se carga en un tren de alta velocidad y se transporta a Inglaterra a través del Channel, el túnel subterráneo que une a ambos países, ya que los coches no pueden conducirse a través del Channel. En efecto, al automóvil se transporta como una carga, como se muestra en la figura. En el otro extremo, se libera el coche en las carreteras inglesas y nuevamente continúa moviéndose con sus propios medios. El proceso de túnel de paquetes a través de una red externa funciona de la misma manera.

Paso de un coche a través del túnel Francia-Inglaterra



Las ventajas de la encapsulación queda claro que son muchas, aunque podemos observar los siguientes problemas:

- Los datagramas encapsulados son normalmente más largos que los datagramas originarios.
- La encapsulación, como ya hemos comentado antes, no se puede utilizar a menos que el nodo de la salida del túnel pueda desencapsular el datagrama.

Encapsulado IP-in-IP

El encapsulado IP-in-IP consiste en insertar una cabecera IP adicional (*outer IP header*) antes de la cabecera IP original (*inner IP header*) del paquete inicial. Si es necesario, también es posible insertar otras cabeceras entre las dos anteriores como por ejemplo la cabecera de autenticación IP. Hay que tener en cuenta que las opciones de seguridad de la cabecera original IP quizás afecten la elección de opciones de seguridad para la cabecera IP insertada en el encapsulamiento.

Los campos de la cabecera IP del encapsulado son puestos por el encapsulador:

- *Versión* : 4.
- *IHL (Internet Header Length)* : es la longitud de esta cabecera medida en palabras de 32 bits.
- *Total Length* : muestra la longitud total del IP datagrama encapsulado, incluyendo todas las cabeceras y la información o carga útil (payload).
- *Identification, Flags, Fragment Offset* : el bit “Don’t Fragment” indica si el paquete debe o no fragmentarse. Si este bit está a uno (no fragmentarse) en la cabecera IP original, éste se deberá poner a uno en la cabecera adicional; pero si el bit “Don’t Fragment” no está puesto a uno (fragmentarse) en la cabecera original, éste quizás se ponga a uno.
- *TTL (Time To Live)* : este campo de tiempo de vida se utiliza para que un datagrama no se quede erróneamente dando vueltas por Internet indefinidamente.
- *Protocol* : 4.
- *Header Checksum* (suma de comprobación de cabecera).
- *Source Address* : muestra la dirección del encapsulador, es decir, del punto de entrada al túnel.
- *Destination Address* : muestra la dirección del desencapsulador, es decir, del punto de salida del túnel.
- *Options* : cualquier opción presente en la cabecera original en general no es grabada en la cabecera insertada. Aunque, nuevas opciones específicas del camino de túnel sean añadidas.

La cabecera exterior contiene información sobre los extremos del túnel. La cabecera interior contiene información sobre los nodos origen y destino del paquete inicial y no puede ser modificada en ningún caso, salvo para decrementar el tiempo de vida (TTL) del paquete, aunque tan solo una vez dentro del túnel, a pesar de que pueda atravesar varios routers. Este campo como hemos visto tiene una copia en la cabecera exterior. Si el TTL vale 0, el datagrama será descartado, y un mensaje ICMP time exceeded message se enviará al nodo origen.

El encapsulador probablemente use cualquier mecanismo IP existente para la entrega de la información encapsulada al punto de salida del túnel. En particular, hace que las IP options sean permitidas, y hace que la fragmentación sea permitida a menos que el bit “Don’t Fragment” esté a uno en la cabecera IP original. Esta restricción es necesaria para que los nodos que utilizan el path MTU Discovery puedan obtener la información que buscan.

Después de que el datagrama encapsulado ha sido enviado, puede ser que el encapsulador reciba un mensaje ICMP desde cualquier router del túnel. La acción del encapsulador dependerá del tipo de mensaje ICMP recibido. Cuando este mensaje contiene suficiente información, el encapsulador deberá crear un mensaje similar ICMP, que será enviado de vuelta al que envió el mensaje original.

Encapsulado IP mínimo

El encapsulado suele conllevar el duplicado innecesario de numerosos campos de la cabecera IP interna. El encapsulado mínimo intenta minimizar al máximo la información de encapsulado para disminuir el tamaño del paquete resultante. Esto se realiza añadiendo una cabecera IP más pequeña que en el encapsulado IP-in-IP, y la cabecera IP original es modificada como indicamos a continuación:

- El campo Protocolo es cambiado por el número 55 que indica que es encapsulado mínimo.
- El campo de dirección de destino es cambiado por la dirección IP del punto de salida del túnel.
- El *Total Length* (longitud total) es incrementado por el tamaño de la cabecera añadida al datagrama.
- El *Header Checksum* (suma de comprobación de cabecera) se actualiza por su nuevo valor correspondiente.

Al desencapsular un paquete con este encapsulado mínimo, se deberán restaurar los campos modificados en la cabecera original con los datos de la cabecera de encapsulado mínimo, actualizando los campos antes nombrados.

La cabecera añadida antes de la cabecera IP original consta de los siguientes campos:

- *Protocol* : campo copiado de la cabecera IP original.
- *Original Source Address Present* : (es un bit).
 - “0”: El campo dirección de la fuente original no está presente.
 - “1”: El campo dirección de la fuente original está presente.
- Bit reservado: puesto a cero. Ignorado en la recepción.
- *Header checksum* .
- Dirección del destino original.
- Dirección de la fuente original.

A pesar de todo, el encapsulado mínimo no está ampliamente difundido ya que presenta ciertas desventajas, concretamente, no funciona con paquetes ya fragmentados, ya que no hay sitio en la cabecera añadida para almacenar información de la fragmentación. Además, este encapsulado fuerza que el valor TTL sea decrementado en cada router dentro del túnel por lo que, puede suceder que los paquetes caduquen antes de llegar a su destino.

Encapsulado GRE

La encapsulación de enrutamiento genérico GRE (*Generic Routing Encapsulation*) es el más flexible de los tres presentados, ya que permite la encapsulación de cualquier tipo de paquete, incluidos los paquetes IP. El formato de un paquete GRE consta de una cabecera externa, seguida de la cabecera GRE, y de los datos IP.

Contrariamente a los encapsulados *IP-in-IP* y mínimo, este encapsulado ha sido diseñado para prevenir encapsulamientos recursivos (encapsular de nuevo lo encapsulado). Concretamente, el campo *recur* en la cabecera GE es un contador que informa del número de encapsulados adicionales que son permitidos. En el protocolo IP versión 6 se está estudiando implementar un mecanismo similar a éste.

Seguridad

Las redes y nodos móviles son particularmente propensos a recibir ataques que comprometan la seguridad. En la mayoría de los casos los nodos móviles se conectarán a Internet mediante conexiones inalámbricas. Este tipo de conexiones es especialmente vulnerable a escuchas silenciosas, ataques activos de respuesta y otro tipo de ataques activos.

Códigos de Autentificación de Mensajes

Los agentes domésticos y nodos móviles deben ser capaces de realizar autentificación. El algoritmo por defecto es el codificado MD5, con una clave de 128 bits. El modo operación por defecto es el uso ‘prefijo+sufijo’ en el que los datos son precedidos y seguidos por la clave de 128 bits. El agente externo también debe soportar la autentificación usando codificado MD5 y claves de 128 o más bits, con distribución explícita. También se permite el uso de otros algoritmos de autentificación y de distribución de claves.

El registro de un nodo móvil en el agente doméstico es un momento crítico en que se debe aplicar autentificación de mensajes, ya que si un usuario se registra en nombre de otro, le robará su tráfico, pues el agente doméstico se encargará de reenviarle todo el tráfico a la Care-of-Address que haya utilizado en el proceso de registro.

Privacidad

Los usuarios que tengan datos privados, que no desea que sean observados por nadie más, deben usar mecanismos de seguridad (encriptación) ajenos al protocolo de IP Móvil y que, por lo tanto, no están especificados.

Protección de Duplicado en las Peticiones de Registro

En la peticiones de registro hay un campo de identificación que permite al agente doméstico verificar que un mensaje de registro ha sido generado recientemente por el nodo móvil, y que no es una petición que haya sido escuchada con anterioridad por otro usuario que la está duplicando ahora. Existen dos algoritmos para este tipo de protección: marcas de tiempo (*timestamp*) y ‘*nonces*’.

Todos los agentes domésticos y nodos móviles deben implementar la protección contra duplicado basada en marcas de tiempo, la implementación del segundo algoritmo es opcional. El nodo móvil y el agente doméstico deben acordar qué método de protección van a usar. Independientemente del método usado, los 32 bits menos significativos de la identificación deben ser copiados sin cambiarse de la petición de registro a la contestación. El agente externo usa estos bits (y la dirección fija del nodo móvil) para emparejar las peticiones con las respuestas correspondientes. A su vez, el nodo móvil debe comprobar que los 32 bits menos significativos de la respuesta deben ser idénticos a los bits enviados en la petición de registro.

Protección de Duplicados mediante uso de Marcas de Tiempo

El principio en que se basa este método es que el nodo que genera el mensaje inserta la hora actual y el nodo receptor comprueba que la marca de tiempo es lo suficientemente cercana a su hora local como para ser cierta. Obviamente ambos nodos deben tener relojes correctamente sincronizados. Como cualquier otro mensaje los mensajes desincronización de tiempo deben estar protegidos contra manipulación por un mecanismo de autentificación determinado por ambos nodos.

Protección de duplicados usando ‘ nonces ’

Una posible traducción de ‘nonces’ es la de algo que sólo se produce una vez u ocasional, en este caso se le llama *nonce* a un número aleatorio.

El mecanismo es el siguiente: el nodo A incluye un número aleatorio o *nonce* nuevo en cada mensaje que envía al nodo B, y comprueba que el nodo B devuelve el mismo número en el siguiente mensaje al nodo A. Ambos mensajes usan un código de autentificación para protegerse de modificaciones producidas mediante un ataque. Al mismo tiempo B puede estar mandando su propios *nonces* en todos los mensajes que envía a A, de tal forma que los dos pueden comprobar que están recibiendo mensajes recientes.

El agente doméstico se supone que tendrá recursos para computar número pseudo-aleatorios que puedan ser usados como *nonces*. Así pues, inserta un nuevo número aleatorio en los 32 bits más significativos del campo identificación de cada respuesta de registro. El agente doméstico copia los 32 bits menos significativos de la identificación de la petición de registro en los 32 bits menos significativos de la contestación. Cuando el nodo móvil recibe una contestación de registro del agente doméstico, se guarda los 32 bits más significativos de la identificación para ser usados como los 32 bits más significativos de la siguiente petición de registro.

Normativa Reguladora

El crecimiento de las comunicaciones móviles y en especial el de los celulares que se está produciendo en los últimos años no tiene precedentes. Para el caso de España, la cuota de penetración ya supera el 50% y el número de líneas móviles ya supera al de las fijas. Por otra parte, la evolución de la redes celulares actuales de segunda generación (2G) hasta la tercera generación (3G) para por ofrecer velocidades más elevadas y acceso por conmutación de paquetes, como es el caso de GPRS (*General Packet Radio Service*).

Todo ello a fin de prepararse para un horizonte en que el tráfico de datos va a ser superior al de voz; apareciendo el acceso a redes IP en general y a Internet en concreto como principal artífice de esta situación. Si a todo ello añadimos el trabajo que se está realizando para soportar el transporte de voz sobre IP está bastante maduro, con soluciones comerciales disponibles, puede entenderse porque la opción de una red de acceso de comunicaciones celulares basada totalmente en IP va ganando peso.

El siguiente paso lógico sería que esta red de manera natural asumiera todas las funciones necesarias para el soporte de la movilidad. En esta dirección se están moviendo los dos grupos que desarrollan la 3G: 3GPP (*Third Generation Partnership Project*) y 3GPP2 (*Third Generation Partnership Project 2*). Cada uno siguiendo su camino, los grupos 3G estudian como convertir sus redes de acceso enredes totalmente IP basadas en *routers* y en como adoptar el trabajo que está realizando el IETF (*Internet Engineering Task Force*) para ofrecer movilidad mediante IP. El IETF ha estado trabajando en una solución universal para conseguir la movilidad en IP conocida como MobileIP.

Se trata pues, de una solución general no optimizada para ningún tipo de red de acceso y es aquí donde surgen los problemas. Uno de los requisitos clave que demandan las redes de 3G y de manera general las redes celulares es el soporte de la micro movilidad; entendiendo por micro movilidad, la posibilidad de cambiar de una manera frecuente y rápida de punto de acceso dentro de una red. Como se verá más adelante, MobileIP tiene serias limitaciones para cumplir con este requisito. Actualmente existe una sinergia importante entre los grupos de 3G y el IETF: por un lado se intenta ver cómo jemorar el protocolo MobileIP para cumplir con los requisitos de 3G sin perder de vista su carácter de solución universal y por otro se ha creado un nuevo grupo especializado en micro movilidad.

Dejando a un lado el IETF y los grupos de 3G también cabe mencionar al MWIF (*Mobile Wireless Internet Forum*). Creado en febrero del 2000, este foro está formado por empresas significativas en ámbitos como las comunicaciones móviles, las redes de datos, el software o la electrónica. Su propósito es el desarrollo de unas especificaciones clave que permitan la utilización de IP en cualquier tipo de redes sin hilos buscando aunar los esfuerzos de otros grupos como el IETF, 3GPP o 3GPP2.

Grupos de Trabajo Relacionados

Dejando a un lado las redes *ad-hoc* y la movilidad de usuarios entre ISPs (Internet Service Providers), temas que se tratan respectivamente en los grupos de trabajo MANET (*Mobile Ad-hoc Networks*) y ROAMOPS (*Roaming Operations*) el foco de trabajo en movilidad en IP ha sido y es el grupo MOBILEIP (*IP Routing for Wireless/Mobile Hosts*). El trabajo de MOBILEIP gira en torno al protocolo MobileIP (MIP).

Este protocolo ofrece el mantenimiento de la dirección IP independientemente de la localización de la máquina que la posea, con un encaminamiento transparente de los paquetes IP e intentando aumentar en lo mínimo los flujos de señalización. Todo esto manteniendo activas las conexiones TCP y las vinculaciones con los puertos UDP.

Existen dos versiones de este protocolo, una estandarizada para IPv4 y otra que todavía tiene el carácter de *draft* para IPv6 pero que se espera su propuesta como estándar a corto plazo. Dejando a un lado la estandarización de la versión para IPv6 de MobileIP y la revisión de la versión para IPv4, la actividad actual del grupo gira en torno a la solución de dos grandes problemas: la seguridad y la mejora de prestaciones para conseguir traspasos rápidos. En verano del 2000 se produjeron una serie de discusiones en torno a las diferentes, y muy numerosas, propuestas presentadas en forma de *draft* para conseguir un traspaso rápido.

Dentro de las propuestas se podían diferenciar dos grupos. El primero las dirigidas a solucionar el problema de la micro movilidad, con un grado de compatibilidad y cooperación respecto a MobileIP más o menos grande. En la mayoría de los casos la red de referencia era de tipo celular. Dentro de este grupo destacaron los protocolos HAWAII (*Handoff-Aware Wireless Access Internet Infraestructure*) y CellularIP, del que se hablará más adelante. En el segundo grupo figuraban las soluciones basadas completamente en MobileIP. Se trataba de soluciones que, respetando el carácter de solución universal no ligada a ninguna tecnología de Mobile IP, pretendían mejorar su rendimiento para conseguir traspasos más rápidos.

El resultado de estas discusiones fue una refundación del grupo de trabajo MOBILEIP que dejaba fuera las propuestas del primer grupo con objeto de mantener el carácter universal de MobileIP como solución para el soporte de la movilidad. Además se crearon dos equipos de trabajo para buscar una solución común para MIPv4 y otra para MIPv6. El trabajo de estos grupos ha cuajado de momento en sendos *drafts*. Todo esto no significa que MOBILEIP deje de lado toda la problemática de las redes celulares.

Existe una estrecha relación con los grupos de 3G y como ejemplo de ella puede verse el *draft* en el que se plantean las extensiones necesarias para que Mobile IP pueda administrar la movilidad en redes cdma2000. Otra consecuencia de la refundación fue la reciente aparición de un nuevo grupo, SEAMOBY (*Context and Micro-mobility routing*). Los objetivos de SEAMOBY son el desarrollo de un protocolo que soporte la micro movilidad, con traspasos rápidos en la red de acceso y *paging*, y la provisión de mecanismos que permitan el intercambio de información de estado, como pueden ser el nivel de calidad de servicio asociado al usuario o un contexto de seguridad.

Otros Grupos

Aparte de los grupos comentados en el apartado anterior, cuya dedicación es exclusiva de los temas de movilidad, existen toda una serie de grupos cuyo trabajo tienen una relación significativa con esta problemática. Destacaremos dos: ROHC (*Robust Header Compression*) y AAA (*Authentication, Authorization and Accounting*).

El objetivo de ROHC es conseguir un sistema de compresión que funcione correctamente sobre enlaces con tasas de error elevadas y retardos importantes. La motivación principal es el envío de información en tiempo real (voz o vídeo de baja calidad) sobre enlaces celulares. La combinación de protocolos IP/UDP/RTP/TCP utilizada para el transporte de tráfico *real-time* conlleva un alto *overhead*. Para trabajar eficientemente sobre enlaces de baja velocidad, como son los de las redes celulares, es necesario utilizar métodos de compresión.

Una posible solución pasaría por la utilización de los algoritmos tradicionales de compresión de cabeceras pero la elevada tasa de error así como los elevados retardos que se puedan dar en una red celular hacen que su comportamiento no sea el idóneo. De ahí la necesidad de un nuevo tipo de compresión.

En los *draft* se especifican los requerimientos que debería cumplir esta nueva codificación y se da una posible especificación de ella respectivamente. El AAA es el grupo encargado de desarrollar los requerimientos para la autenticación, autorización y contabilidad. Estas funciones son de vital importancia para control del acceso a cualquier sistema. El AAA trata el caso de un sistema con terminales como un caso particular con unas necesidades propias que requieren de unas extensiones determinadas.

Esto se ha traducido en un listado de requerimientos formulado por el grupo MOBILEIP y en *draft* del AAA sobre las extensiones a realizar para cumplir con estos requerimientos.

Integración de los Protocolos del IETF en 3G

Las aproximaciones realizadas por los dos grupos de 3G para integrar los protocolos desarrollados por el IETF están siendo diametralmente opuestas. Por un lado el 3GPP2 cuenta ya desde hace más de un año con un estándar de lo que ellos denominan *WirelessIP*. En este documento se describen los requerimientos para soportar redes e paquetes inalámbricas en las redes de 3G basadas en cdma2000; diferenciando dos alternativas: *SimpleIP*, basado en el protocolo PPP (*Point to Point Protocol*); y *MobileIP* basado en el protocolo del mismo nombre.

El documento también propone la utilización de servidores RADIUS (*R emote Authentication Dial In User Service*) para labores de AAA y la utilización de *Diffserv* para ofrecer calidad de servicio. Se trata pues de utilizar las soluciones ofrecidas por el IETF, aunque no estén optimizadas para sistemas celulares.

Paralelamente el 3GPP2 tiene abierto otro proyecto en fase de definición denominado *A//IP*, que consiste en el desarrollo de una red que se basa en IP como principal mecanismo para el transporte y la conmutación. El camino por el que ha optado el grupo 3GPP es mucho más ambicioso. En este *report* técnico el 3GPP propone una arquitectura basada totalmente en IP, *A//IP*, para el transporte de todos los datos de usuario y señalización. El documento tiene una doble vertiente: la identificación de los problemas clave a resolver y la proposición de un plan de trabajo para ofrecer una A//IP release 200 del estándar UMTS (*Universal Mobile Telecommunications System*).

Ventajas e Inconvenientes de Móvil IP

Ventajas

- No tiene limitaciones geográficas: un usuario puede usar una computadora *palmtop* o *laptop* en cualquier parte sin perder su conexión a su red.
- No requiere una conexión física: este nuevo protocolo encuentra los *routers* IP y se conecta automáticamente.
- No requiere modificación a otros *router* o clientes ya que deja otros protocolos intactos.
- Soporta autenticación, la cual se realiza para asegurarse que los derechos estén protegidos. El acceso de la red se asegura en todo momento desde todas las ubicaciones.

Inconvenientes

- Problemas de autenticación con el agente externo, cuando éste pertenece a otra organización.
- Falta de protocolos de manejo de claves que sean estándares Internet.
- El encaminamiento en triángulo es ineficiente.
- La creación de túneles es un coste añadido e incrementa el *overhead* por paquete.
- Cuello de botella en el Agente Local.
- Implicaciones de seguridad en cortafuegos.

PLC (Power Line Communications)

Introducción

Desde los años 40 se consideró la posibilidad de aprovechar la red eléctrica como red de comunicaciones. Finalmente, en 1997 se presentó un sistema que permitía el acceso a Internet desde la red eléctrica, que pasó a denominarse PLC (Power Line Communications). Los primeros problemas para la utilización de esta tecnología fueron debidos al ruido inherente a la red eléctrica de baja tensión, capaz de alterar la información transmitida, así como problemas legales de regulación del espectro de frecuencias en las que trabaja y de las emisiones electromagnéticas del sistema, dadas las potenciales interferencias sobre otros aparatos electrónicos.

En los últimos años la tecnología ha evolucionado muy rápidamente, permitiendo velocidades de acceso competitivas con tecnologías alternativas, y en la actualidad ya existen ofertas comerciales de servicios de telecomunicaciones basados en PLC en algunos países, así como multitud de experiencias piloto en otros muchos, entre ellos España. El sector eléctrico español se encuentra en un proceso de rápida evolución, tanto en las estructuras de capital de las empresas eléctricas como en el marco regulatorio, lo que ha introducido planteamientos totalmente innovadores en su funcionamiento, con el propósito de fomentar la competencia entre las empresas. Esta liberalización del sector eléctrico empuja a las empresas del sector a buscar nuevas oportunidades de negocio para compensar las pérdidas de cuotas del mercado.

El Sector de las Telecomunicaciones se encuentra también en un proceso de cambio acelerado, pasando en pocos años de una situación de monopolio a otra de amplia liberalización. La creciente orientación de las políticas económicas hacia la satisfacción de las demandas y necesidades de los usuarios, justifica la introducción de la competencia en un sector tan complejo y variado como el de las telecomunicaciones en constante y rápida

evolución tecnológica. No cabe duda de que la liberalización de las telecomunicaciones es ya uno de los motores del crecimiento económico y de la nueva economía de servicios basada en la sociedad de la información.

Tradicionalmente las Empresas Eléctricas han instalado, operado y mantenido redes privadas de telecomunicación fundamentalmente para el control u operación de la propia red eléctrica. El resto de los servicios demandados por las necesidades administrativas y societarias ha estado restringido por la legislación a ser prestado por el operador nacional que actuaba en condición de monopolio natural. Este hecho influyó en el desarrollo de la infraestructura y en el uso a veces de equipos especializados, previstos para estas condiciones de operación específicas. Los medios de transmisión empleados son muy diversos: equipos de ondas portadoras usando los propios cables de energía, cables pilotos, cables coaxiales, enlaces vía radio, fibras ópticas, satélites de comunicaciones, enlaces por infrarrojos, etc, y su utilización ha seguido los dictados de las propias necesidades de cada compañía y la oferta tecnológica existente en cada momento.

En cualquier caso, la ruptura de las barreras legales que supone la liberalización de las telecomunicaciones implica la posibilidad de poder usar la capacidad excedente de las redes privadas, proporcionada por la digitalización y los avances tecnológicos, para proporcionar servicios a terceros. Además la Disposición Adicional 14 a la Ley del Sector Eléctrico permite que la Empresa Eléctrica, que en principio tiene objeto social exclusivo, ponga en valor la infraestructura de que es titular con fines de telecomunicaciones. Y por supuesto nada impide que cedan el uso de dichas infraestructuras a un tercero para que las explote dado que además en la legislación de telecomunicaciones de España en la actualidad, existe la obligación de separación de cuentas por los operadores de telecomunicaciones que desarrollen actividades en otros sectores económicos.

Las Directivas armonizadoras del Consejo Europeo, han decidido promover el desarrollo de la Sociedad de la Información para todos, facilitando una mayor competencia en el segmento de acceso al hogar. El servicio Internet va a ser declarado universal y se deberán facilitar todas las infraestructuras que favorezcan su desarrollo e implantación. Esta consecuencia del rápido y universal desarrollo de la red Internet hacen que el acceso de banda ancha sea el negocio de más rápido crecimiento en las telecomunicaciones en los próximos años.

La tecnología Power Line Communications, "PLC", posibilita la transmisión de voz y datos a través de los cables eléctricos, convirtiendo cualquier enchufe de la casa en conexión potencial a todos los servicios de telecomunicaciones. El cliente sólo necesitará conectar un pequeño módem para acceder a Internet, telefonía y datos al mismo tiempo y a alta velocidad (banda ancha). La naturaleza y ubicuidad de la red de baja tensión permitirá también lograr una comunicación permanente y a bajo coste entre todos los aparatos electrónicos de la casa, dando lugar a nuevos y eficientes servicios de seguridad, control del consumo a distancia, domótica y teleasistencia, entre otros.

Actualmente muchas compañías desarrollan y comercializan sus propias soluciones para la creación de redes domésticas sin necesidad de tender cableados adicionales, y sin el todavía elevado gasto que supone las redes inalámbricas. Crear una red local en un domicilio es bastante sencillo.

Siguiendo el enfoque de estas soluciones, para conectar dos ordenadores sólo se necesitan dos enchufes. El siguiente paso es más difícil. Cientos de equipos de investigadores llevan varios años intentando estabilizar las tecnologías que permitan conectar estas pequeñas redes locales con la red de redes, o dicho de otra forma, se trata de conectarse a Internet a través del enchufe eléctrico de casa. Y también, de llamar por teléfono desde ese mismo enchufe. La solución, de tener éxito finalmente en las experiencias precomerciales que ya se están llevando a cabo, podría también penetrar -nada más fácil, en principio- los muros empresariales.

Características Técnicas

Existe interés generalizado en el mercado por los accesos a Internet de banda ancha, ya que este tipo de acceso es el que va a permitir que las diferentes compañías dejen de ser meros ISP para convertirse en auténticos proveedores de servicios multimedia.

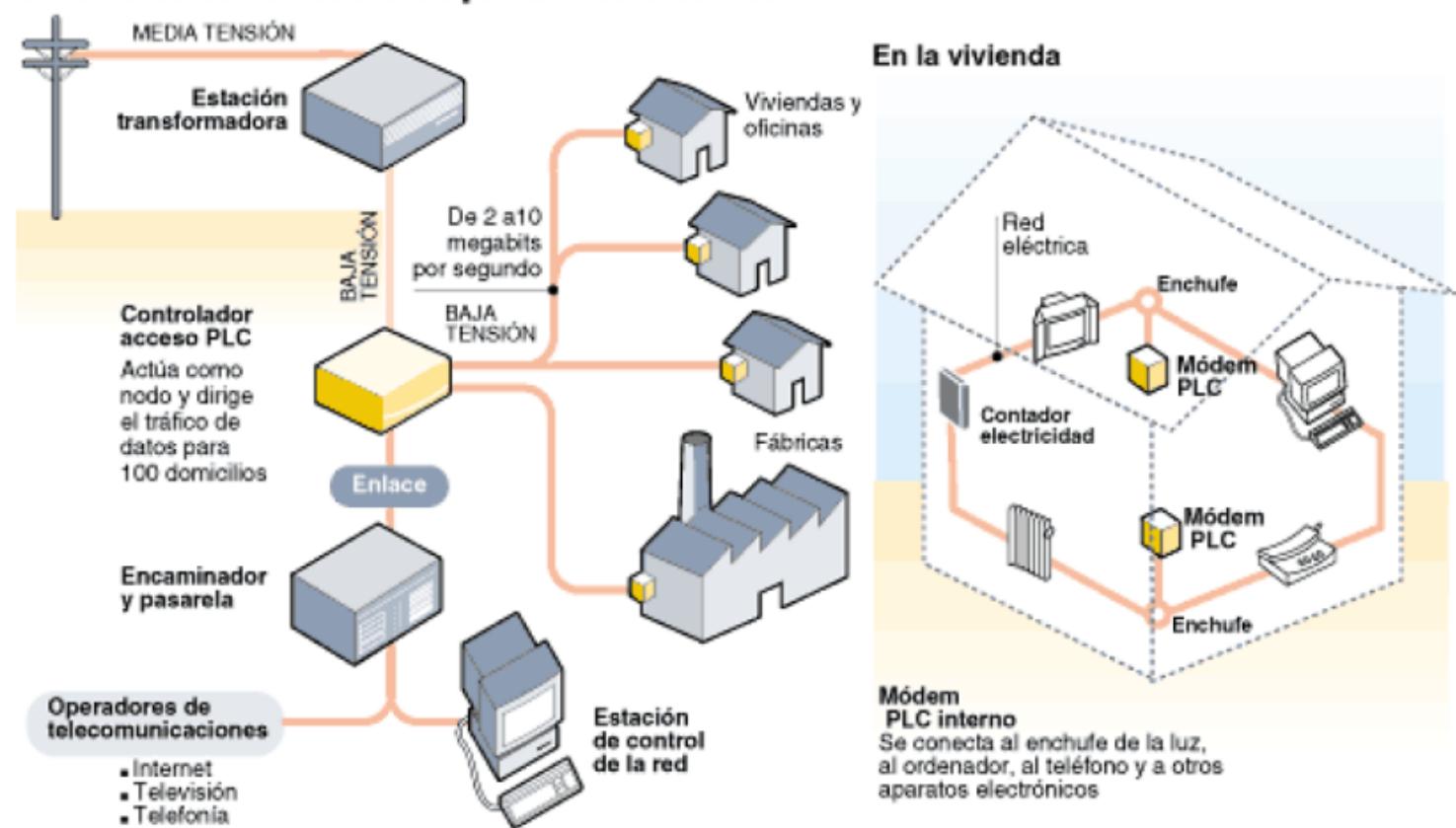
Power Line Digital (PLC) podrá alcanzar velocidades entre 1 y 1,5 megas de ancho de banda en la casa de cada usuario particular (en principio).

Esto hace posible que se ofrezcan servicios de Internet bajo un modelo de tarifa plana, así como otro tipo de transmisión de datos y hasta telefonía IP.

La técnica es bastante sencilla y tiene algunos puntos de similitud con los sistemas xDSL. Basta acondicionar parte de las actuales infraestructuras eléctricas para que puedan transmitir señales regulares de baja frecuencia y otras por encima de la banda de 1 MHz, sin que se vea afectado el rendimiento eléctrico.

Arquitectura de PLC

Sistema de comunicaciones por la línea eléctrica

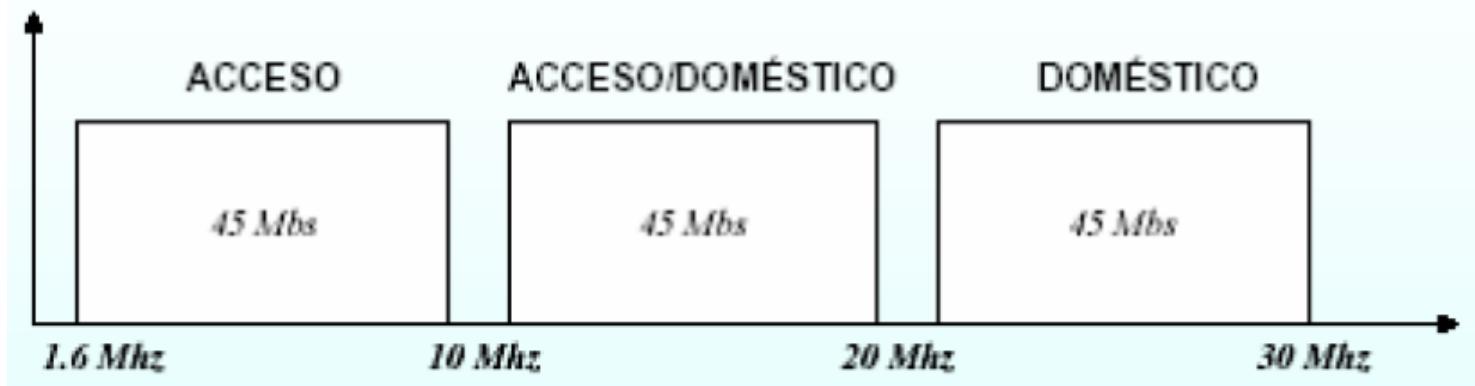


La red eléctrica transporta electricidad a una frecuencia de 50 Hz. En PLC se añaden frecuencias en la banda que va desde 1,6MHz hasta 30MHz para el transporte de los datos. Unos filtros instalados en el transformador de baja tensión separan las frecuencias altas de datos, de la frecuencia de 50Hz de la electricidad.

Por otro lado, en el enchufe del abonado, cuando se conecta un dispositivo de transmisión de datos (un PC, teléfono, etc) a la red, se hace a través de un módem adaptador.

Organización del espectro de PLC

Asignación de bandas del ETSI



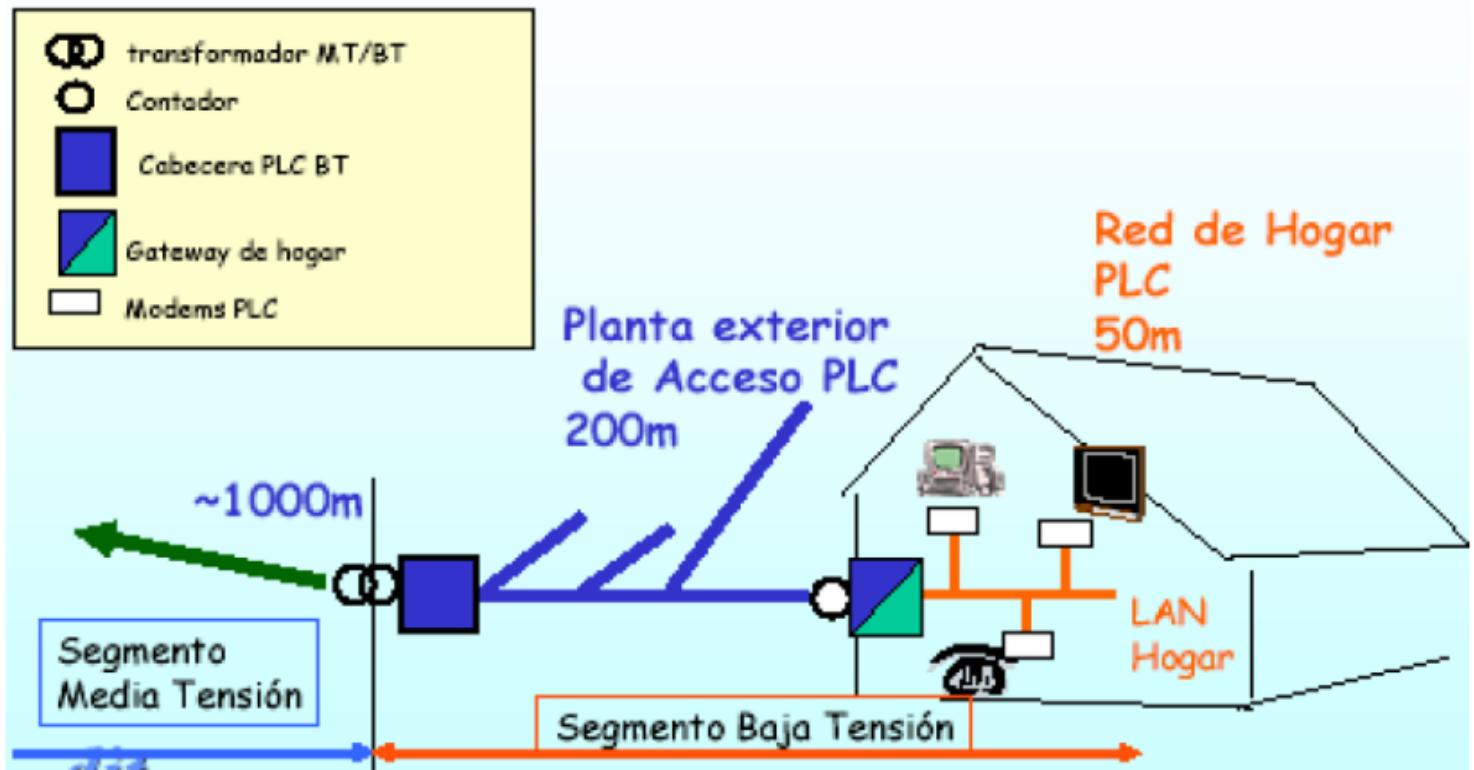
Power Line emplea una red conocida como High Frequency conditioned Power Network (HFCPN) para transmitir simultáneamente energía e información. Una serie de unidades acondicionadoras son las que se encargan del filtrado y separación de ambas señales.

Así pues estas unidades acondicionadoras separarían la electricidad, que alimenta a los electrodomésticos, de las señales de alta frecuencia, que van a un módulo o unidad de servicio, donde se reconvierten en canales de vídeo, datos, voz, etc. En las subestaciones eléctricas locales hay servidores de estación base que se conectan a Internet generalmente a través de fibra óptica. Esto quiere decir que nos utiliza toda la red eléctrica para la transmisión de datos.

La red eléctrica consta de tres partes bien diferenciadas: los tramos de baja tensión, los de media y los de alta tensión. Los de baja tensión – equivalentes a la “última milla” o bucle de abonado en las redes telefónicas- conecta los hogares con las subestaciones de distribución local. Es precisamente este tramo el único que se utiliza en PLC.

Tramos en la implantación de PLC

PLC características



Las estaciones base de PLC tienen una estructura típica de rack. Una localización puede llegar a contener unas doce unidades emisoras del tipo estación base, cada una capaz de comunicar un canal. Los datos llegan a estas estaciones que las incorporan a la señal eléctrica. Una estación estándar sirve a unos cincuenta usuarios, ofreciéndoles un espectro cercano a los 20 MHz en el caso de clientes próximos, o entre 6 y 10 MHz para clientes lejanos. El servidor opera con un sistema basado en IP para crear redes LAN en cada área de servicio.

Las unidades acondicionadoras situadas en los hogares de los abonados, que también pueden recibir el nombre de módem eléctricos, tienen en su interior dos filtros. El primero de ellos, el de baja banda, libera la corriente eléctrica de 50 Hz para su distribución a todos los enchufes de la casa. Este filtro además sirve para limpiar los ruidos generados en la red por los electrodomésticos conectados en casa del usuario. Si se dejaran pasar esos ruidos, al unirse a los procedentes de otros usuarios de la red, acabarían por introducir distorsiones muy significativas. En segundo lugar, el filtro de banda es el que libera los datos y facilita el tráfico bidireccional entre el cliente y la red.

En la actualidad no existen estándares tecnológicos para el PLC de acceso. Éste es uno de los principales problemas de esta tecnología, al no permitir la interoperabilidad entre los equipos suministrados por los distintos fabricantes.

Tampoco existe una regulación en cuanto a la utilización de frecuencias, aunque CENELEC y ETSI tienen previsto publicar este año una recomendación conjunta acerca del uso de frecuencias por los sistemas de última milla y los sistemas domésticos que al menos garantice la coexistencia de ambos tipos de sistema, dado que la red eléctrica es continua y las señales de ambos se mezclan y extienden por toda la red que depende de un mismo transformador.

La situación actual de la tecnología PLC queda reflejada en el siguiente cuadro:

	ASCOM	MAINNET	DS2
Posicionamiento	<ul style="list-style-type: none"> Solución en acceso (no contempla MT) Solución in-home 	<ul style="list-style-type: none"> Solución en acceso (contempla MT) Solución in-home 	<ul style="list-style-type: none"> Solución en acceso (contempla MT) Solución in-home
Diseñadores	<ul style="list-style-type: none"> Sí 	<ul style="list-style-type: none"> No 	<ul style="list-style-type: none"> Sí
Fabricantes	<ul style="list-style-type: none"> Sí 	<ul style="list-style-type: none"> Sí 	<ul style="list-style-type: none"> No
Características técnicas	<ul style="list-style-type: none"> AB máximo: 4,5 Mbps Modulación: GSMK 	<ul style="list-style-type: none"> AB máximo: 4,5 Mbps Modulación: DSSS 	<ul style="list-style-type: none"> AB máximo: 45 Mbps Modulación: OFDM
Baja tensión (BT)	<ul style="list-style-type: none"> Chipset PROPIO 	<ul style="list-style-type: none"> Chipset ITRAN 	<ul style="list-style-type: none"> Chipset PROPIO
Media tensión (MT)	<ul style="list-style-type: none"> No 	<ul style="list-style-type: none"> Disponible con chipset ITRAN 	<ul style="list-style-type: none"> Disponible con chipset PROPIO
VoIP	<ul style="list-style-type: none"> Disponible 	<ul style="list-style-type: none"> Disponible 	<ul style="list-style-type: none"> Disponible
Roadmap de producto	<ul style="list-style-type: none"> MT disponible con chipset DS2 en 1Q2003 BT disponible con chipset DS2 en 2004 	<ul style="list-style-type: none"> MT disponible con chipset DS2 en 2004 	<ul style="list-style-type: none"> Nuevo chipset 200 Mbps en 2Q2003 para MT y BT
Pruebas PLC en Europa con:	<ul style="list-style-type: none"> EDF EEF EnBW ENDESA ENEL TIWAG IBERDROLA 	<ul style="list-style-type: none"> EDF LINZ STROM MW NUON ENEL UNIÓN FENOSA VATTENFALL 	<ul style="list-style-type: none"> EDP ENDESA ENEL UNIÓN FENOSA IBERDROLA

En la práctica se están consiguiendo velocidades de hasta 45 Mbps compartidos en algunas de las pruebas realizadas, aunque las velocidades realmente disponibles a nivel comercial todavía sean muy inferiores. La mayor limitación actualmente para conseguir velocidades de transmisión mayores sigue siendo el ruido inherente a la red eléctrica de baja tensión. En el caso de la red de media tensión, se están obteniendo velocidades de hasta 135 Mbps para la conexión de centros transformadores.

Características del Chip DS2

- Flujo de datos de 45 Mbps; 27 Mbps en Bajada y 18 en Subida.
- Full dúplex, punto a multipunto, paquete orientado a enlace de comunicaciones.
- Cumple los estándares del ETSI y CENELEC para acceso y LAN.
- Eficiencia de la modulación hasta 7,25 bps/Hz.
- 1280 portadoras OFDM adaptativas para conseguir el máximo flujo sobre cualquier red.
- Monitorización de la SNR (relación señal ruido) del canal continuamente.
- Ratio adaptativo por portadora dependiendo de las condiciones SNR del canal.
- Empleo de control de errores mediante códigos bloque Reed Solomon.
- Opera por debajo de -1 dB de SNR.
- Ratios de error optimizados para TCP/IP, y programable para cualquier aplicación.
- Mensajes Broadcast disponibles.
- Soporte de encriptación y SNMP.
- Control de QoS.

- Hasta 254 usuario, (512 en versiones posteriores).

Comparativa con otras Tecnologías

La comparativa da una idea de las posibles diferencias entre las distintas tecnologías. Se espera que el precio de la tecnología PLC sea bastante inferior al de los actuales ADSL y Cable en el mismo rango de velocidades, que junto con su mayor velocidad de conexión la convierten en una llamativa tecnología. En el cuadro se reflejan las principales características:

		TIPOS DE CONEXION				
		MODEM	RDSI	ADSL	CABLE	PLC
CARACTERISTICAS	Tipo de línea que la RTB	RTB	RTB	RTB	Línea propia (fibra óptica)	Línea eléctrica
	Velocidad de conexión	56 Kbps (bajada) 33.6 (subida)	128 Kbps	1,5 - 2 Mbps 16 - 640 Kbps	10 - 38 Mbps 128 Kbps - 10 Mbps	puede llegar hasta 20 Mbps
	Calidad	Media	Alta (digital)	Alta (digital)	Alta (óptico)	Alta
	Coste mensual orientativo	Acceso gratuito/Tarifa plana 18.03 euros/mes	30.05 euros/mes	45.07 euros/mes	30.05 euros/mes	39.00 euros/mes
	Distancia máxima a la central	No hay límite	5,8 Km Ampliable	5,5 Km	48,3 Km Ampliable	No hay límite
	Implantación de la tecnología	Completa	Completa	Completa con fallos	Completa con fallos	No instalada, en pruebas

Las categorías anteriormente mencionadas sólo son indicativas, ya que las fronteras entre las tecnologías están cambiando.

Es también significativo el hecho de que las tecnologías de banda ancha actualmente en el mercado pueden dar resultados comparables o mejores que la PLC, pero en general no tienen una infraestructura totalmente implantada con la que poder alcanzar un mercado de masas en un corto período de tiempo.

Modos de Operación y Servicios

Dentro de la tecnología PLC se distinguen la red externa de transmisión y la red interna de comunicación dentro del hogar o del negocio del usuario final. La red externa o tecnología de acceso a la “última milla” permite el transporte de señales hasta el usuario final vía el centro de transformación local y la red eléctrica. Los servicios típicos de telecomunicaciones que podrían proporcionarse son:

- Telefonía: Incluye la prestación de servicios de voz y fax sobre el protocolo de Internet IP.
- Acceso a Internet: Dependiendo de las diversas tecnologías empleadas es posible alcanzar velocidades aseguradas de unos 20 Mbps.
- Vídeo bajo demanda: Aunque es posible, esta opción por las necesidades tan elevadas de ancho de banda que requiere, parece algo más lejana de implantarse.

La red interna de comunicaciones o tecnología de uso doméstico integra la conexión y el control de dispositivos mediante un único interfaz dentro del edificio. Esta red interna es utilizada para la transmisión de la señal a alta velocidad proveyendo soluciones de comunicación interna.

Como ejemplos de los posibles servicios que se pueden obtener están:

- La implantación de una red e área local de ámbito doméstico.
- Control de seguridad remoto a través de dispositivos dotados de cierta inteligencia.
- Gestión y control remoto de electrodomésticos.

Seguridad

Cualquier línea conductora es, por definición, una antena lo que nos lleva a pensar en seguridad.

La seguridad es uno de los aspectos menos investigados de PLC. Los problemas técnicos se traducen en dinero. Para filtrar y limpiar las líneas hacen falta equipos costosos, y aún así siempre hay un equilibrio entre la velocidad y el aislamiento: cuanto más se filtre la línea, más difícil es transmitir a altas velocidades. Las soluciones a estos problemas de confidencialidad, ya que todavía no hay estándares al respecto, pasan por soluciones propietarias de cifrado implantadas por las empresas que proporcionan los servicios, como el cifrado por hardware propuesto por la empresa DS2 en su chip.

Por último, decir que esta tecnología es totalmente compatible con las tecnologías de cifrado IPSEC y también se ha tratado el problema de compatibilidad con VLANs basadas en el protocolo 802.1q. Este protocolo consiste en añadir un encabezado a la trama Ethernet para identificar la VLAN que le corresponde. Por supuesto, dicho protocolo es soportado por los switches actuales, siendo dicha información adicional manejada e intercambiada entre ellos exclusivamente y no por el usuario final, por lo que no se necesitan tarjetas Ethernet especiales.

Normativa Reguladora

Disponer de un marco regulatorio estable, es esencial para el desarrollo y la aplicación práctica de la tecnología PLC. Además, como toda nueva tecnología necesita estándares (bandas de frecuencia, potencia, límites EMC, etc), para poder desarrollarse comercialmente de una forma competitiva al permitir la interoperabilidad entre distintos fabricantes.

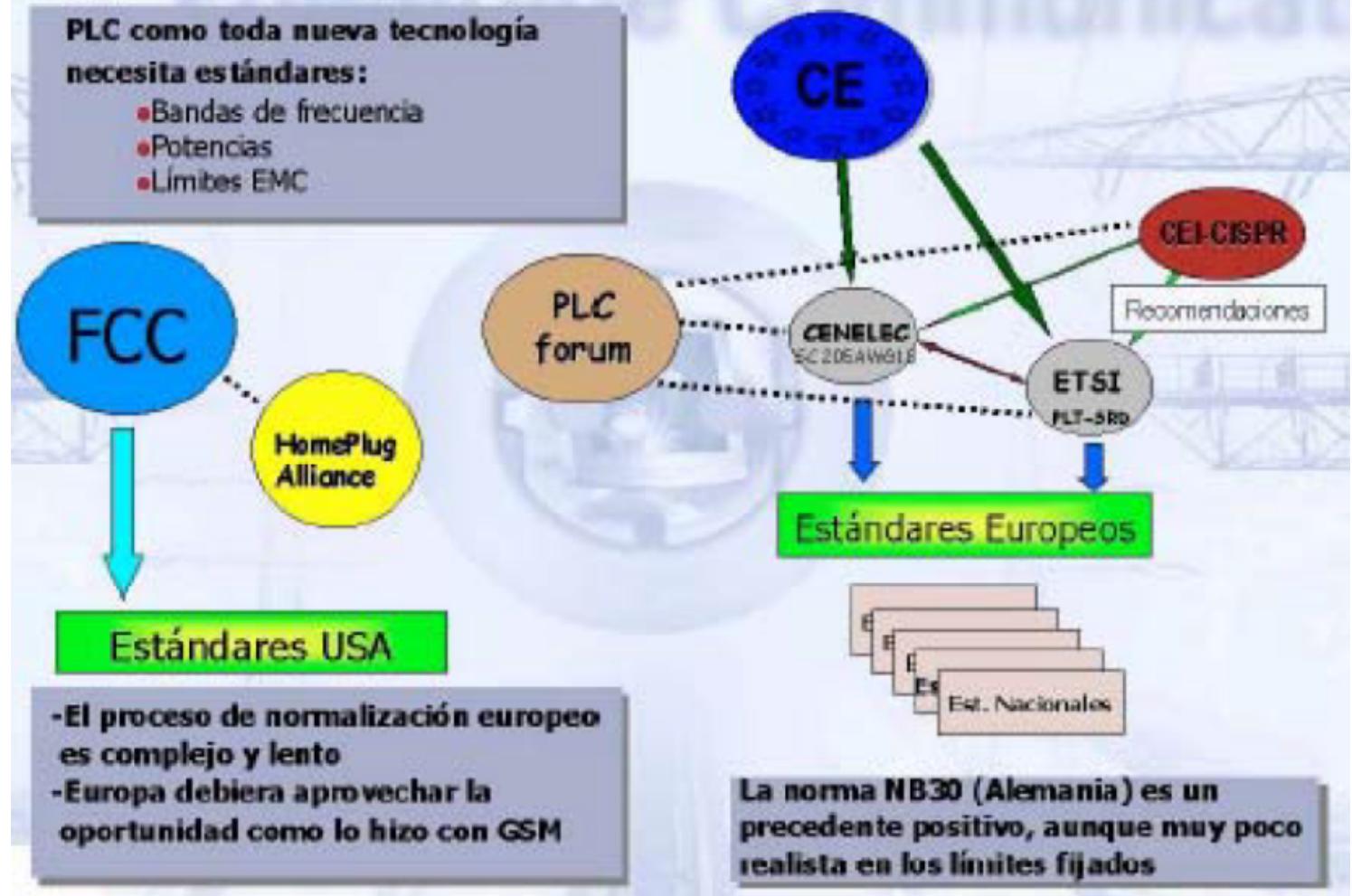
En Europa no existe una regulación unificada al no haberse aprobado ninguna Norma sobre el PLC de banda ancha. El principal problema está surgiendo con la regulación del espectro que evite problemas de interferencias. Se trata de fijar los límites a imponer, sobre la posible acumulación de las emisiones EMC (sobre todo en condiciones atmosféricas adversas) según PLC vaya implantándose.

Organismos encargados de la estandarización de PLC

Panorama estandarizador del PLC

PLC como toda nueva tecnología necesita estándares:

- Bandas de frecuencia
 - Potencias
 - Límites EMC



En los aspectos regulatorios/normativos se pueden anotar las siguientes consideraciones:

- El Parlamento Europeo ha aprobado el texto de compromiso propuesto por el Consejo de la UE sobre el nuevo marco regulador de las comunicaciones electrónicas (Directiva Marco, Directiva de Acceso a Interconexión, Directiva de Autorizaciones, Directiva de Servicio Universal y Decisión sobre el Espectro). Estos textos, una vez aprobados formalmente por el Consejo, deberán publicarse en el DOCE. El texto de compromiso sobre la Directiva relativa a un “marco regulador común de las redes y servicios de comunicaciones electrónicas”, incluía en su artículo 2 relativo a la definición de “red de comunicación electrónica”, la tecnología “Power Line communications” (PLC).
 - El proceso de normalización europeo es complejo y lento. Existe un borrador de mandato de la Comisión Europea a CEN (European Committee for Standardization), CENELEC (European Committee for Electrotechnical Standardization) y ETSI (European Telecommunication Standardization Institute) para la elaboración de normas armonizadas que cubran los aspectos de EMC de las redes de telecomunicaciones que usan cables coaxiales, pares de cobre, líneas eléctricas o cualquier otro tipo de medio físico.
 - La ETSI TS (101867 recomienda las condiciones de separación de las bandas de acceso y domésticas: como bandas a utilizar por el PLC de las aplicaciones de acceso (PLC Access) establece el espectro de frecuencias comprendidas entre 1,6 - 10 MHz, reservando al PLC de las aplicaciones de uso doméstico (PLC Inhouse) la banda de 10 - 30 MHz.
 - El SC205A de CENELEC está trabajando en la ES 59013, que define el espectro de frecuencias a utilizar por las aplicaciones de acceso y las domésticas, fijando la frecuencia de separación en 13,5 MHz en lugar de los 10 MHz de la ETSI. Esto ha

motivado la creación de un grupo conjunto de CENELEC/ETSI/CISPR para alcanzar un consenso final que resuelva los problemas de coexistencia entre los equipos PLC de acceso y domésticos.

Las particularidades del PLC como red de telecomunicaciones en cuanto aprovecha la infraestructura eléctrica existente, implica la aplicación de la doctrina relativa a la correlación entre derechos y obligaciones de Servicio Público. Además, por dicha particularidad, la CMT podrá imponer una cláusula específica en la licencia a otorgar.

Ventajas e Inconvenientes de PLC

Ventajas

Podemos destacar los siguientes puntos en relación a los beneficios que nos puede aportar esta tecnología:

- Como la PLC se ha posicionado como un servicio de tipo IP utilizará routers de paquetes en vez de los de conmutación de circuitos típicos, de los suministradores de telecomunicaciones tradicionales, manteniendo así los costes de los equipos de IT bajos.
- Como la electricidad se suministra a través de una conexión permanente, los servicios de transmisión de datos ofrecidos por la infraestructura eléctrica también están conectados permanentemente (no es necesario marcar el número de conexión) convirtiéndose en el ideal para el número creciente de servicios en línea. Las compañías eléctricas podrían pues comercializar un servicio básico de conexión a Internet con una suscripción mensual de tarifa plana, al igual que algunos operadores de cable. Pagar una tarifa estándar, sin tener en cuenta el nivel de utilización, será un gran atractivo para los clientes.
- Al dar a los clientes de las compañías eléctricas acceso a Internet mediante la red que ya les suministra la electricidad, esta tecnología se pone virtualmente al alcance de cualquiera, con un potencial mercado de masas sin necesidad de las inversiones necesarias para enterrar el cableado hasta los hogares.
- Ya existen varias tecnologías que transforman los cables eléctricos existentes en un cableado LAN (Local Area Network). Lo que hace diferente a la PLC es la alta velocidad de transmisión de datos que se puede conseguir y el hecho de que esté diseñada para trabajar en el exterior del hogar o del edificio. Por tanto, podrían instalarse sistemas sofisticados de automatización doméstica que permitiesen el acceso y el control remotos de aparatos electrodomésticos, alarmas antirrobo, etc.
- PLC podría también facilitar a las compañías eléctricas la oportunidad de ofrecer servicios de valor añadido orientados sectorialmente, tales como la gestión de la energía (enlazando contadores "inteligentes", controladores programables y dispositivos "inteligentes" de control de la demanda/suministro, de modo que la empresa eléctrica suministradora del servicio pudiera introducir tarifas innovadoras que premiasen el uso sensato de la energía, la información remota (la conexión permanente ofrecida por PLC se podría optimizar para proporcionar información en tiempo real o indicadores de estado en apoyo de algunas aplicaciones de seguridad para sistemas de alarma/vigilancia) y la automatización de la distribución (la lectura remota automática de los contadores mejoraría el control y ayudaría al proveedor en la gestión de los picos de demanda eléctrica). Algunas empresas eléctricas han empezado a usar recientemente estas técnicas.

Inconvenientes

La red eléctrica no ha sido diseñada para transmitir datos, sólo para transmitir energía, y esto hace que presente varias limitaciones y problemas de seguridad. Los obstáculos son fundamentalmente técnicos:

- En primer lugar, hay que elegir un tipo de modulación que sea el más adecuado para la red eléctrica. En PLC se emplea la modulación OFDM (Orthogonal Frequency Division Multiplexing). Otro de los problemas reside en el número máximo de hogares por transformador. Como las señales de datos de Power Line no pueden sobrevivir a su paso por un transformador, sólo se utilizan en la última milla. El modelo europeo de red eléctrica suele colocar un transformador cada 150 hogares aproximadamente.
- Si se juntan estos dos factores, se comprueba que es necesario que todos los transformadores vengan dotados de servidores de estación base Power Line. Y cuanto menor es el número de usuarios por cada transformador, más se elevan las inversiones necesarias.
- En tercer lugar, están las interferencias. Al poco tiempo de realizarse las primeras pruebas se comprobó que algunas de las frecuencias no se podían usar porque generaban interferencias en otros servicios preexistentes. Por ejemplo, el uso de determinadas frecuencias en las cercanías de un aeropuerto podía interferir, y de hecho interfería, con las frecuencias de la torre de control y las de los radares de aproximación. También se puede llegar a interferir con las transmisiones convencionales de radio en FM o incluso En DAB, o con las de los servicios de emergencia, como bomberos o policía.
- En la actualidad muchas compañías eléctricas están realizando intensivas pruebas de campo.
- La suciedad electromagnética de los cables es otro de los problemas significativos. Si piensa que el famoso espacio radioeléctrico está lleno de ondas de radio en constante peligro de interferencia, tendría que ver cómo está el tendido eléctrico. Es un problema de aislamiento. Compare un cable eléctrico con un cable de antena de televisión. El primero sólo está recubierto de plástico. El cable de antena tiene varias capas de plástico y una malla metálica intermedia que lo aísla de posibles interferencias.
- Cualquier línea conductora es, por definición, una antena. Eso quiere decir que la instalación eléctrica de una casa actúa como tal, y es muy sensible a las interferencias que se produzcan en las frecuencias de transmisión de datos, alrededor de los 30 MHz. La red eléctrica no está protegida contra las ondas de radio, pero tampoco contra el ruido electromagnético que puede introducir una afeitadora, la televisión o el propio PC. Todos estos aparatos se protegen a sí mismos de lo que pueda venir de la línea eléctrica (sobre una subida de tensión) con filtros y fusibles, pero nadie se preocupa de lo que vierten en ella.