Multihoming casero

Horacio J. Peña*

1. Resumen

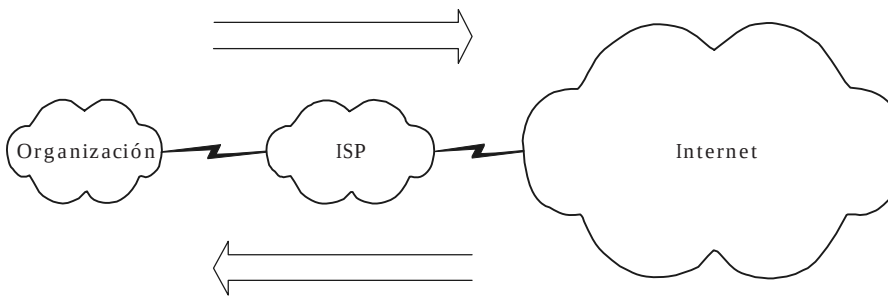
Las organizaciones actuales dependen cada vez más de su conectividad hacia Internet. Para disminuir el riesgo de la perdida de ésta se recurre a la paralelización de conexiones, llamada multihoming.

El método clásico de multihoming sirve únicamente para organizaciones de gran tamaño. Para las organizaciones pequeñas es necesario encontrar otro método que les permita tener la redundancia generada por la paralelización de manera económica.

En este trabajo se propone un posible método para permitir esto.

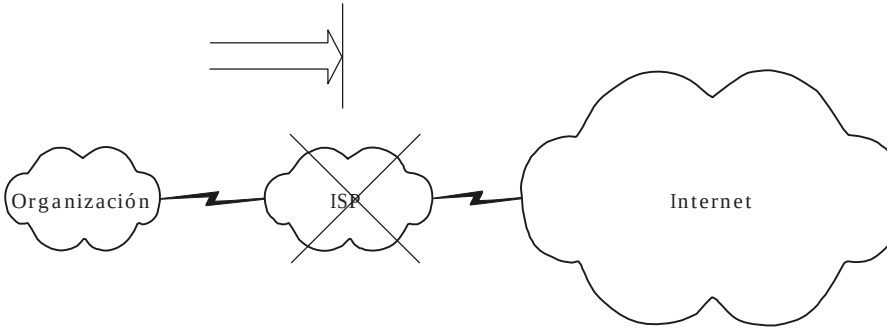
2. La necesidad del multihoming

La forma habitual de conexión a Internet es a través de un "proveedor de Internet'' (de aquí en más ISP) que está conectado a la gran red.

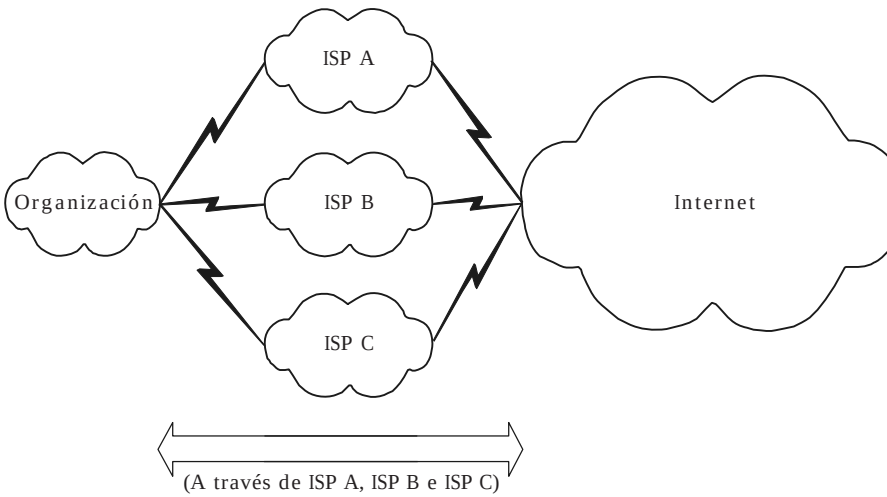


Cuando la conexión a este ISP o la conexión entre éste e Internet dejan de funcionar, la organización queda aislada de la red, con los costos que esto conlleva.

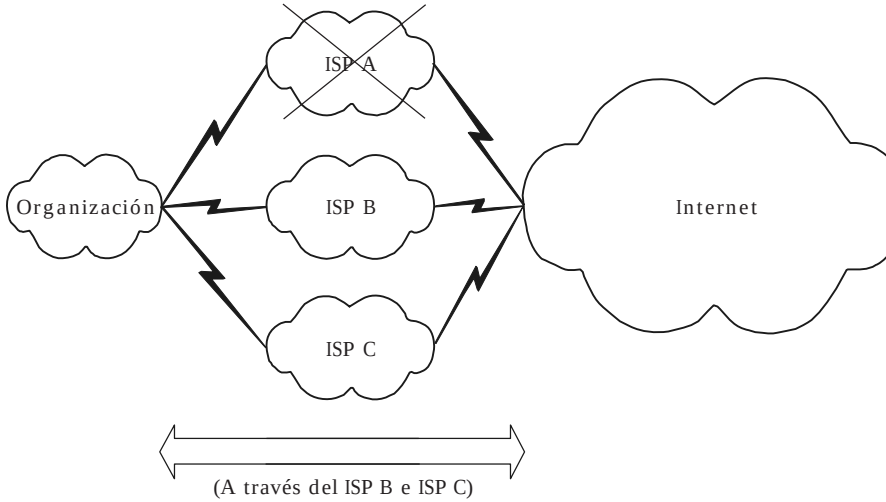
* Alumno de la Facultad de Ingeniería. Universidad de Palermo.



Una forma de evitar esto (o al menos de reducir la probabilidad de que se dé) es mediante la conexión a múltiples ISP (a esto se le llama “multihoming”)



De este modo si la conexión a través de uno de los ISP cae, la organización sigue estando conectada a la Red a través de los otros.



3. El método clásico de multihoming

Clásicamente para realizar multihoming se requiere lo siguiente:

- La organización debe tener asignado por los RIR un número de sistema autónomo.
- La organización debe tener asignado por los RIR un prefijo independiente de proveedor, o debe serle delegado un prefijo por alguno de los ISPs a los que se conecte.
- Debe utilizarse el protocolo de ruteo BGP con los ISPs y publicar su prefijo.

Esto conlleva una serie de costos:

- La cantidad de números de sistema autónomo que pueden ser asignados es limitada, además de que esta asignación debe pagarse.
- Los prefijos independientes de proveedor deben pagarse (son caros) y deben cumplirse una serie de requisitos acerca de la cantidad de direcciones utilizadas que hacen imposible a cualquier organización que no sea muy grande obtenerlos.
- Utilizar prefijos asignados por un ISP significa depender de ese ISP y no poder cambiarlo fácilmente.
- Utilizar BGP implica costos de configuración en los routers de los ISPs que estos normalmente no están dispuestos a afrontar sin cobrarlo a los clientes.
- La publicación del prefijo implica un costo no sólo para la organización y para sus proveedores, sino para toda la Internet, ya que cada router de "backbone"[1] debe agregar una entrada en sus tablas de rutas. Esto tiene grandes implicaciones en la capacidad de estos equipos y en el problema de computación distribuida que implica la convergencia de estas tablas enormes.

1. Técnicamente, todos los routers de la Default Free Zone.

Por todo esto, el sistema clásico de multihoming no es adecuado para ser usado por organizaciones pequeñas.

4. Multihoming para organizaciones pequeñas

Se necesita entonces encontrar un método para permitir a las pequeñas organizaciones acceder a la redundancia de ISPs que precisan para garantizar su permanente conectividad.

Consideramos como requisitos indispensables para cualquier método propuesto:

- No debe precisar asignación de número de sistema autónomo ni de prefijos independientes de proveedor.
- Debe evitar la dependencia de entidad externa alguna (en cualquier momento debe poderse cambiar cualquiera de los ISPs en uso por otro sin grandes problemas)
- Debe permitir que, en todo momento en que la conexión a través de cualquiera de los ISPs funcione, los servidores de la red de la organización sean accesibles desde Internet.
- Debe permitir que, en todo momento en que la conexión a través de cualquiera de los ISPs funcione, se puedan establecer conexiones desde la red de la organización hacia Internet.
- No debe precisar modificaciones en los sistemas operativos de los equipos de la red (esto limita el uso de técnicas basadas en selección de direcciones de origen en los hosts)
- No debe precisar cooperación por parte de los ISPs.
- Los costos generados por el uso de este modelo deben ser acotados (ie, no pueden generar costos a la red en su totalidad)
- Deben tener un bajo costo económico para la organización.

Como atributos deseables, pero no indispensables consideramos:

- Debería permitir que las conexiones establecidas sobrevivan a un cambio de estado en los enlaces a Internet a través de los ISPs.
- Debería evitar la complejidad.
- Debería poder implementarse en plataformas disímiles.
- Debería mantener el modelo de conexión entre extremos.

Aquí presentamos un método que cumple con todos los requisitos y con alguna de las características deseables.

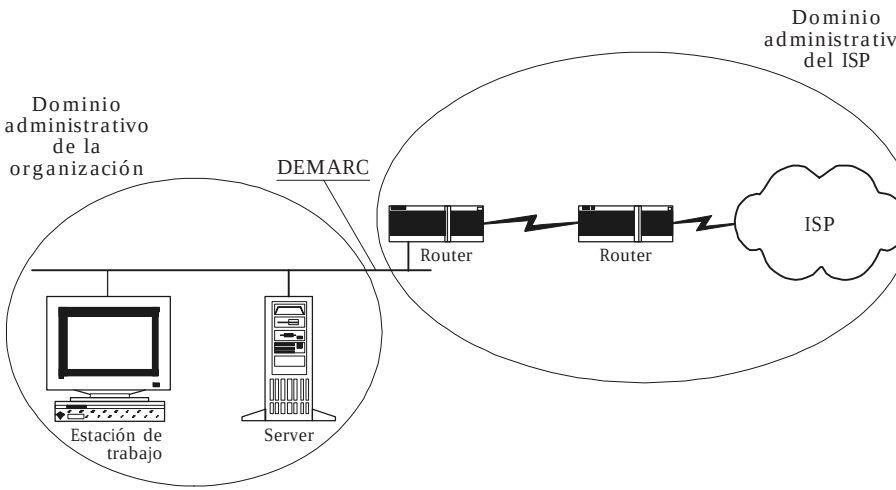
5. El método

Nota: los detalles de implementación de este sistema -basados en una plataforma Linux- están presentados en un anexo (que tiene alguna repetición con lo expuesto aquí dado que está pensado para poder ser leído independientemente de este artículo).

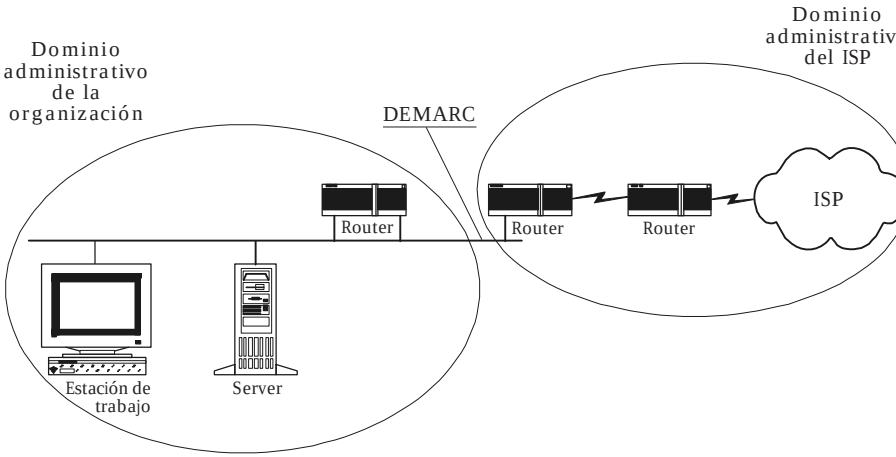
Para presentar el método utilizaremos la técnica de ir construyendo el sistema incrementalmente, generando sistemas intermedios, revisando qué falencias tienen y corrigiéndolas para generar así sistemas más depurados hasta llegar al sistema completo. Esto lo hacemos así para que sea comprensible la explicación, pero al momento de implementarlo puede ser más razonable hacerlo todo de una sola vez y no en etapas.

5.1 Conexión a un solo ISP

Normalmente los ISPs que nos dan conectividad instalan un enlace desde sus oficinas a las nuestras, donde lo hacen llegar a un router del que se reservan el control, haciendo así que el punto de demarcación de dominios administrativos (conocido como "demarc") sea la interfaz que une su router a nuestra red local. Consideramos esta política nociva ya que el demarc debería ser una interfaz WAN y el router que nos conecta nuestra red al enlace estar bajo control administrativo de la organización (esto es efectivamente así en otros países, donde incluso se llama al router CPE por Customer Provided Equipment) Pero, dado que éste es el modo de operación actual de los ISPs, nos acomodaremos a ello.

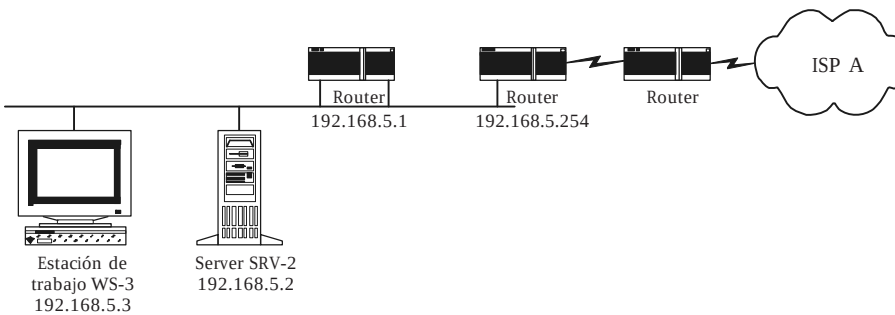


Como primera medida, necesitamos que entre nuestra red local y el demarc haya un router para poder definir nosotros la política de ruteos de la organización.

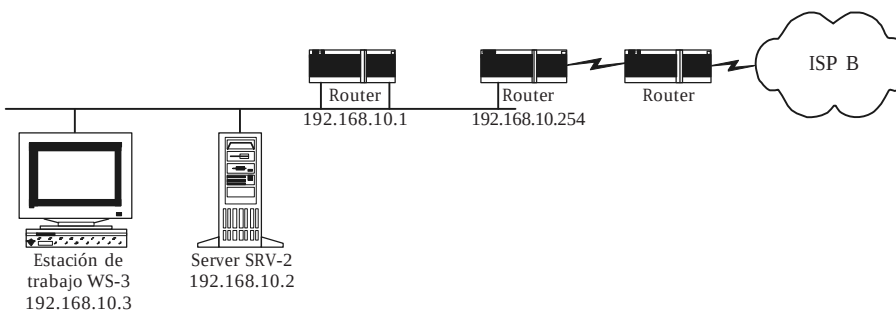


De aquí en más trabajaremos sobre el ejemplo de una organización conectada a dos ISPs. El ISP_A, que nos asignará el prefijo 192.168.5.0/24 y el ISP_B que nos asignará el prefijo 192.168.10.0/4.

Si estuvieramos conectados sólo al ISP_A tendríamos:

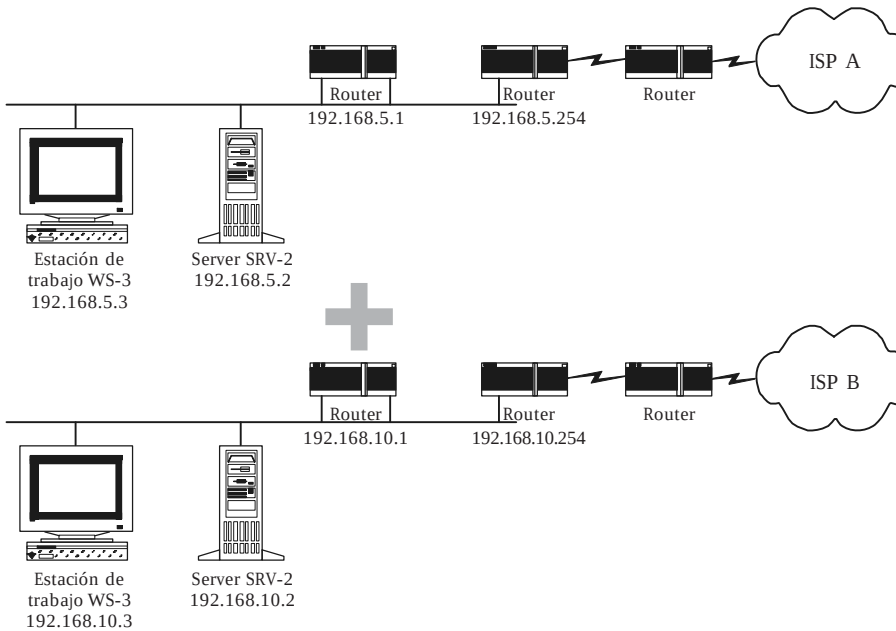


Si tuvieramos sólo a ISP_B sería en cambio:

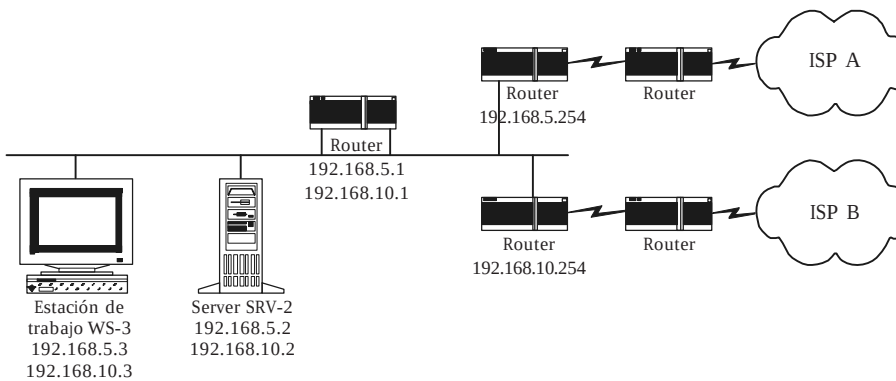


5.2 Conexión a múltiples ISPs

Fusionando los esquemas anteriores



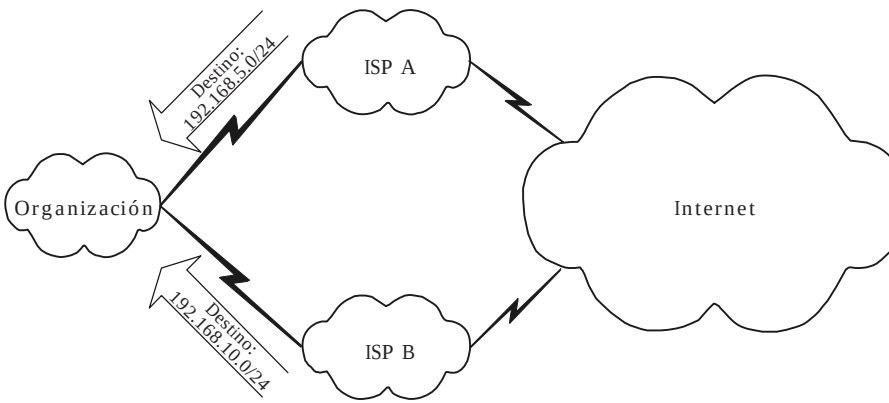
tenemos entonces una sola red física sobre la que existen dos redes lógicas:



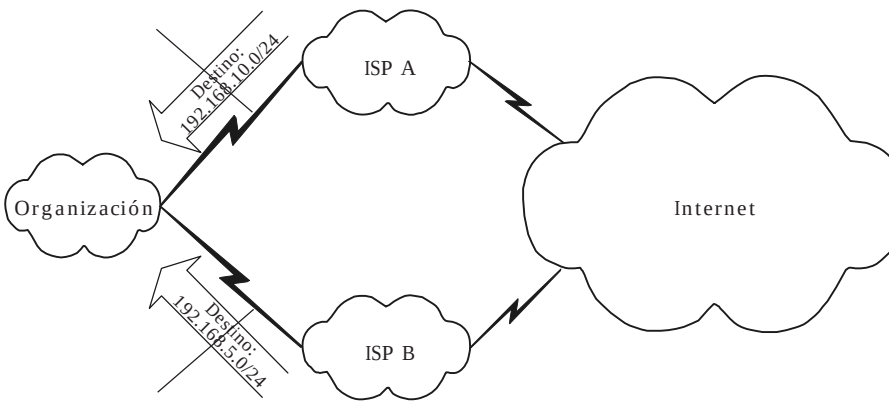
Este esquema de red permite que las conexiones iniciadas desde Internet hacia nuestra red puedan establecerse utilizando como destino las direcciones IP correspondientes al prefijo asignado por cualquiera de los ISPs indistintamente.

Con algo de inteligencia en el router, que le permita detectar cuando el enlace a través de un ISP hasta la internet deja de funcionar -luego veremos cómo hacerlo- esto permite que todo el tráfico generado en nuestra red sea enviado a Internet.

Nótese que el tráfico entrante hacia nuestra red vendrá siempre por el ISP al que corresponde la dirección utilizada. Por ejemplo: suponiendo dos conexiones iniciadas hacia el equipo SRV-2, una a cada una de sus dos direcciones, el tráfico correspondiente a la conexión que use la dirección 192.168.5.2 vendrá siempre por el ISP_A y el correspondiente a la que use la dirección 192.168.10.2 ingresará a través del ISP_B:

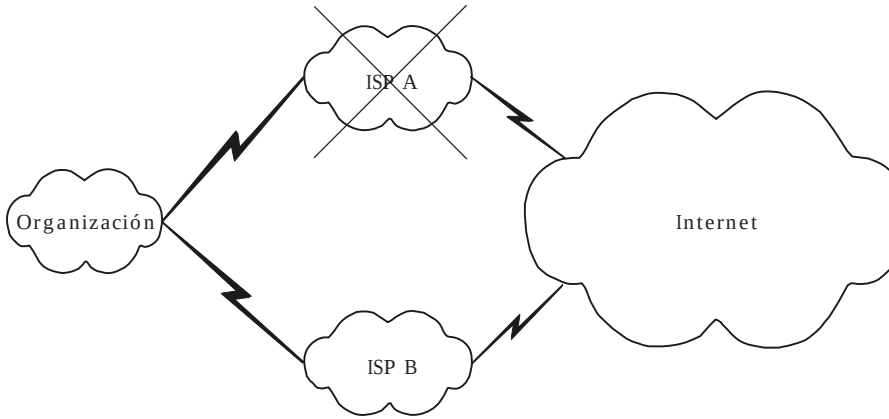


Nunca recibiremos por el ISP_A tráfico destinado a la red 192.168.10.0/24 ni por el ISP_B tráfico destinado a la red 192.168.5.0/24:

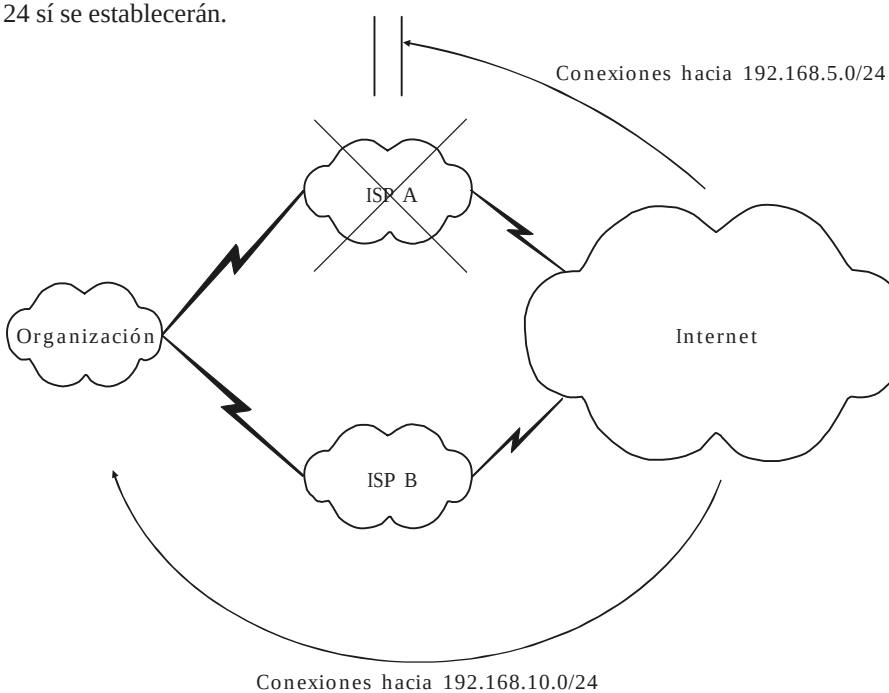


De esto surge un problema.

¿Qué sucede si un ISP cae? (Supongamos que ISP_A es el que cae)

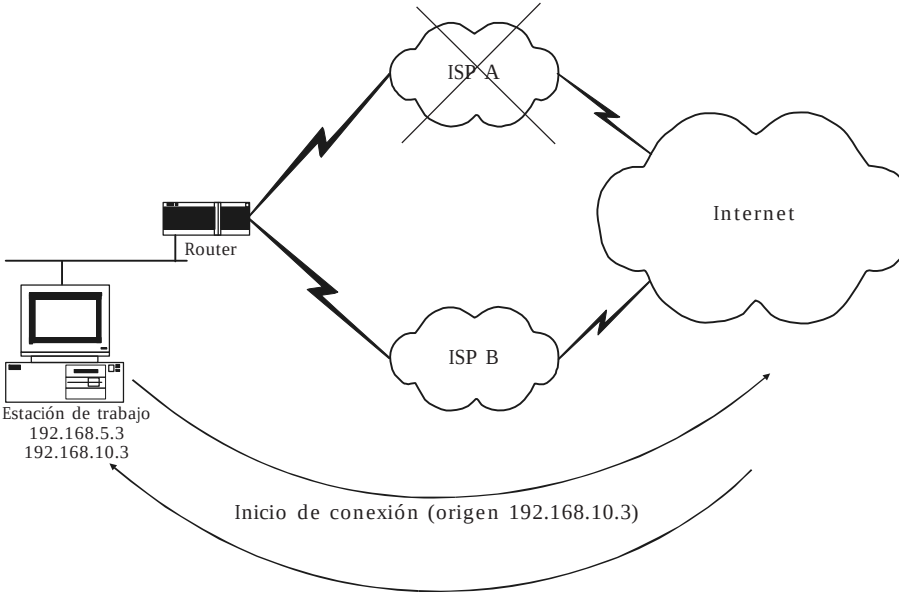


En este caso las conexiones iniciadas desde internet hacia equipos de nuestra red por medio de sus direcciones de la red 192.168.5.0/24 no podrán establecerse mientras que aquellas que hubieren sido iniciadas usando las direcciones del bloque 192.168.10.0/24 sí se establecerán.

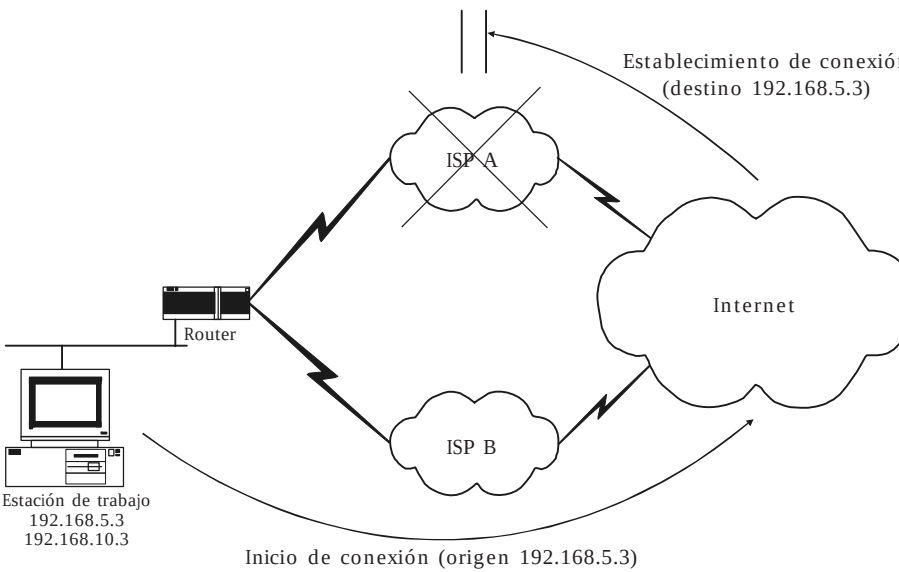


En cuanto a las conexiones iniciadas desde nuestra red la posibilidad de establecerse depende de la dirección utilizada como origen por nuestros equipos.

Si usan la dirección de algún ISP que esté activo, la conexión podrá establecerse.

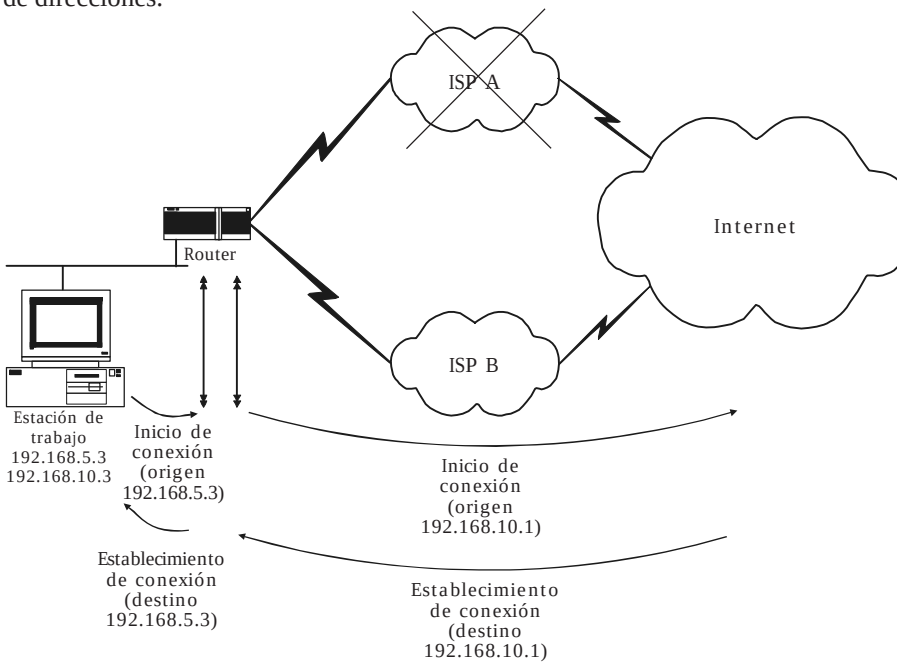


Si usan la dirección de un ISP que no funciona, la conexión no se establecerá.



5.3 Conexión a múltiples ISPs refinada

Dado que por los requerimientos debemos garantizar que estas conexiones puedan establecerse y que no podemos influir en cómo deciden los equipos qué dirección de origen utilizar debemos interceptar las conexiones y convertirlas de modo que usen direcciones con las que sea factible iniciarlas. Hacemos esto por medio de traducción de direcciones.



5.4 Publicación de información acerca de nuestros servidores

Dado todo esto, sería razonable pensar que no quedan ya problemas y que con sólo anunciar en el sistema de DNS todas las IPs correspondientes a cada uno de nuestros servidores estaría todo listo. Para ello bastaría por ejemplo agregar en nuestra zona de DNS los registros: (debería hacerse esto para todo servidor)

SRV-2 IN A 192.168.5.2
SRV-2 IN A 192.168.10.2

En un mundo ideal, donde todas las aplicaciones estuvieran bien programadas tendría razón quien así pensara. Las aplicaciones deberían pedir al DNS todas las direcciones de un servidor y probarlas todas hasta contactar alguna o dar error si a ninguna se puede conectar.

Lamentablemente no estamos en ese mundo y son muchas las aplicaciones que si no consiguen conectar a la primer dirección devuelta por el DNS dan error en vez de intentar con las otras, por ello es necesario un sistema que -utilizando la misma inteligencia mencionada antes para que el router reconozca cuando un ISP está funcionando o no- regenere las tablas de DNS para que sólo figuren registros correspondientes a direcciones a las que se pueda acceder desde Internet. Para ello usamos un sistema de templates que por cada línea de tipo

SRV-2 IN A $ISP.2

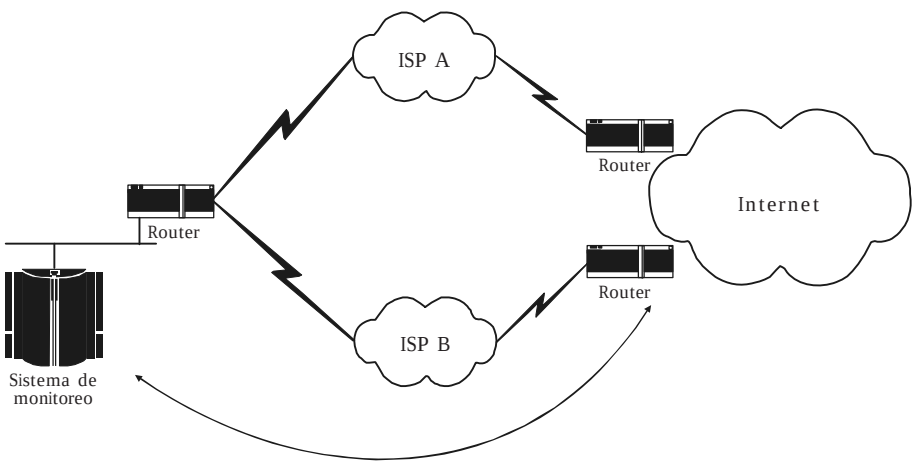
en el modelo genere tantas líneas como ISPs haya activos, reemplazando $ISP por el prefijo correspondiente a cada uno de los ISPs activos.

Para que esto sea eficaz debe mantenerse el tiempo de vida de los registros de DNS bastante bajos (cuán bajos depende de la confiabilidad de nuestros ISPs)

5.5 Detección del estado de los ISPs

Existen sistemas de monitoreo de red que permiten verificar constantemente mediante pruebas ICMP ECHO la conectividad a ciertos puntos de la red y permiten reaccionar a los cambios de estado de esa conectividad en forma programable. Algunos ejemplos de estos sistemas son Netsaint (para unix) y Whatsup (para windows)

Usando uno de estos sistemas mediremos la conectividad hacia los puntos dentro de la topología de nuestros ISPs que nos permitan detectar si su conexión hacia la red es adecuada (por ejemplo, estos puntos pueden ser los routers de proveedores del backbone).



Pruebas ICMP ECHO para
verificar conectividad de ISPs

El sistema de monitoreo al detectar modificaciones debe disparar procesos que alteren la tabla de rutas de nuestro router principal para redirigir el tráfico a los ISPs que siguen funcionando, y otros procesos que regeneren las zonas DNS según se mencionaba antes.

Debe notarse que en el router ha de fijarse que el camino a esos puntos que usamos para medir si existe conectividad debe ser siempre por el ISP que querramos medir, dado que si no podemos al redirigir el tráfico volver a tener conectividad a esos puntos a través de internet, creer erróneamente que el ISP cuya conectividad hemos medido erróneamente ha vuelto a funcionar, volver a mandar tráfico por ese ISP, volver a detectar el problema y entrar en un loop que genere un estado de inconsistencia que impida a la red funcionar adecuadamente.

6. Conclusiones

Hemos presentado aquí un método de obtener redundancia de proveedores a bajo costo económico, pero con una alta complejidad (y con algunas falencias)

Este método no es ideal, pero es eficiente y permite alcanzar esa redundancia hasta que sean desarrollados métodos más eficientes. El grupo de trabajo de la IETF multi6 está trabajando en desarrollar métodos de multihoming que no generen crecimiento en las tablas de rutas, estos métodos probablemente sean la base para que muchas pequeñas organizaciones puedan hacer multihoming de una manera más prolija y menos compleja.

Mientras tanto, este es nuestro pequeño aporte para permitir a las pequeñas organizaciones utilizar multihoming ahora sin tener que esperar a que estos métodos sean desarrollados.