Reacción de TCP a errores ICMP

Fernando Gont
Grupo CEDI
Facultad Regional Haedo
Universidad Tecnológica Nacional

Primeras Jornadas de Divulgación Electrónica
Haedo, 26 de Octubre de 2006

Aislamiento de fallas y recobro de fallas en la Arquitectura de Internet

Dos funciones básicas en una red de computadoras son el **aislamiento de fallas** y el **recobro de fallas**.

- El aislamiento de fallas consiste en detectar condiciones de error en la red (sistemas inalcanzables, etc.)
- El recobro de fallas consiste en intentar sobrevivir esas condiciones de error.
- Así, una vez encontrado un problema de red (mediante la función de aislamiento de fallas), se realiza alguna operación (recobro de fallas) para sobrevivir tal error.

Internet Control Message Protocol

- La Arquitectura de Internet realiza gran parte del aislamiento de fallas mediante el protocolo ICMP.
- Conceptualmente, el protocolo ICMP forma parte de la capa de red, trabajando conjuntamente con el protocolo IP
- Se utiliza principalmente para la señalización de problemas de ruteo, aunque también puede ser utilizado para control de congestión y señalización de algunos errores en protocolos de transporte.
- También se lo utiliza para realizar algunas funciones de depuración de errores de red. Herramientas como ping y traceroute basan su funcionamiento en el protocolo ICMP.
- Originalmente, también se lo utilizaba para algunas funciones de configuración de sistemas (obtención de máscara de red, obtención de dirección de router, etc.)

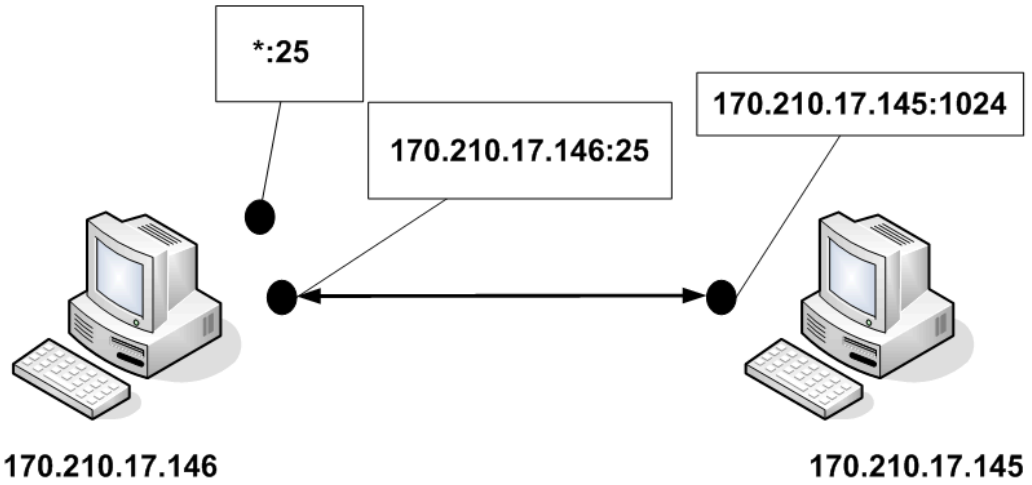
Mensajes ICMP

- ICMP define una serie de mensajes de error, para notificar diversas condiciones de error en la red.
- Algunos de estos mensajes pueden ser generados por sistemas intermediarios (routers), mientras que otros pueden ser generados por sistemas finales (hosts).

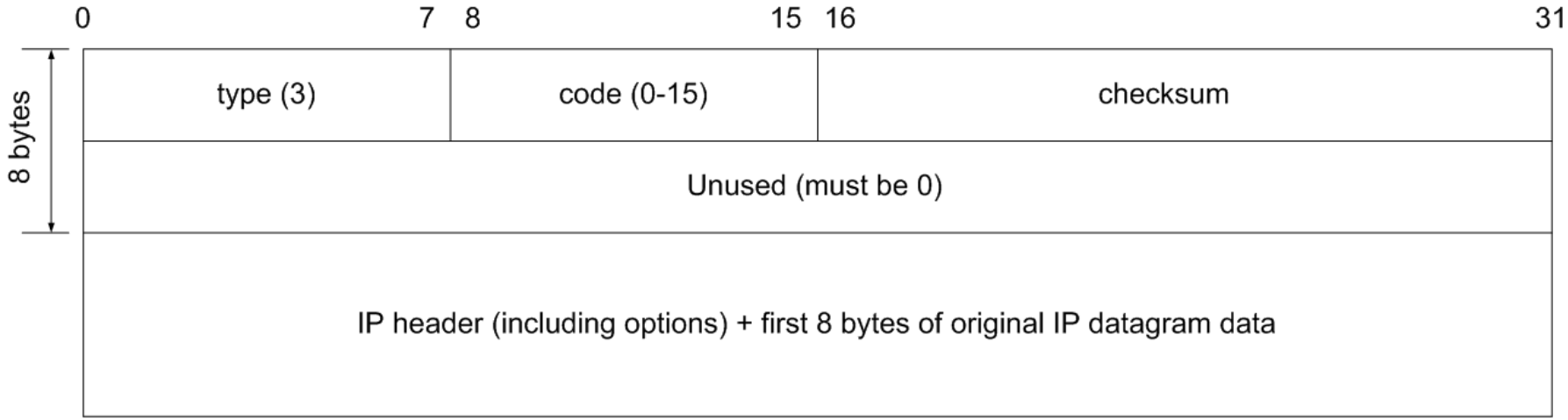
Type	Code	Descripción
3	0	Net unreachable
3	1	Host unreachable
3	2	Protocol Unreachable
3	3	Port unreachable
3	4	Frag. needed and DF set
3	5	Source route failed

Generación de mensajes de error ICMP

- Cuando un sistema detecta una condición de error al procesar un paquete IP, usualmente emitirá un mensaje de error ICMP, para notificaral sistema origen del paquete la condición de error detectada.
- Debido a que el sistema origen del paquete que causó la generación del mensaje de error ICMP podría tener varias instancias de comunicación activas (por ejemplo, varias conexiones TCP), es que se hace necesario poder demultiplexar el mensaje de error ICMP a la instancia de protoclo de transporte que lo generó.
- Con el fin de proporcionar dicha demultiplexación, se incluirá en el mensaje de error ICMP una porción del paquete que causó el error (el encabezado IP completo, y los primeros 64 bits del encabezado del protocolo de transporte), bajo la suposición que dicha porción del paquete original proveerá toda la información necesaria para realizar dicha tarea.

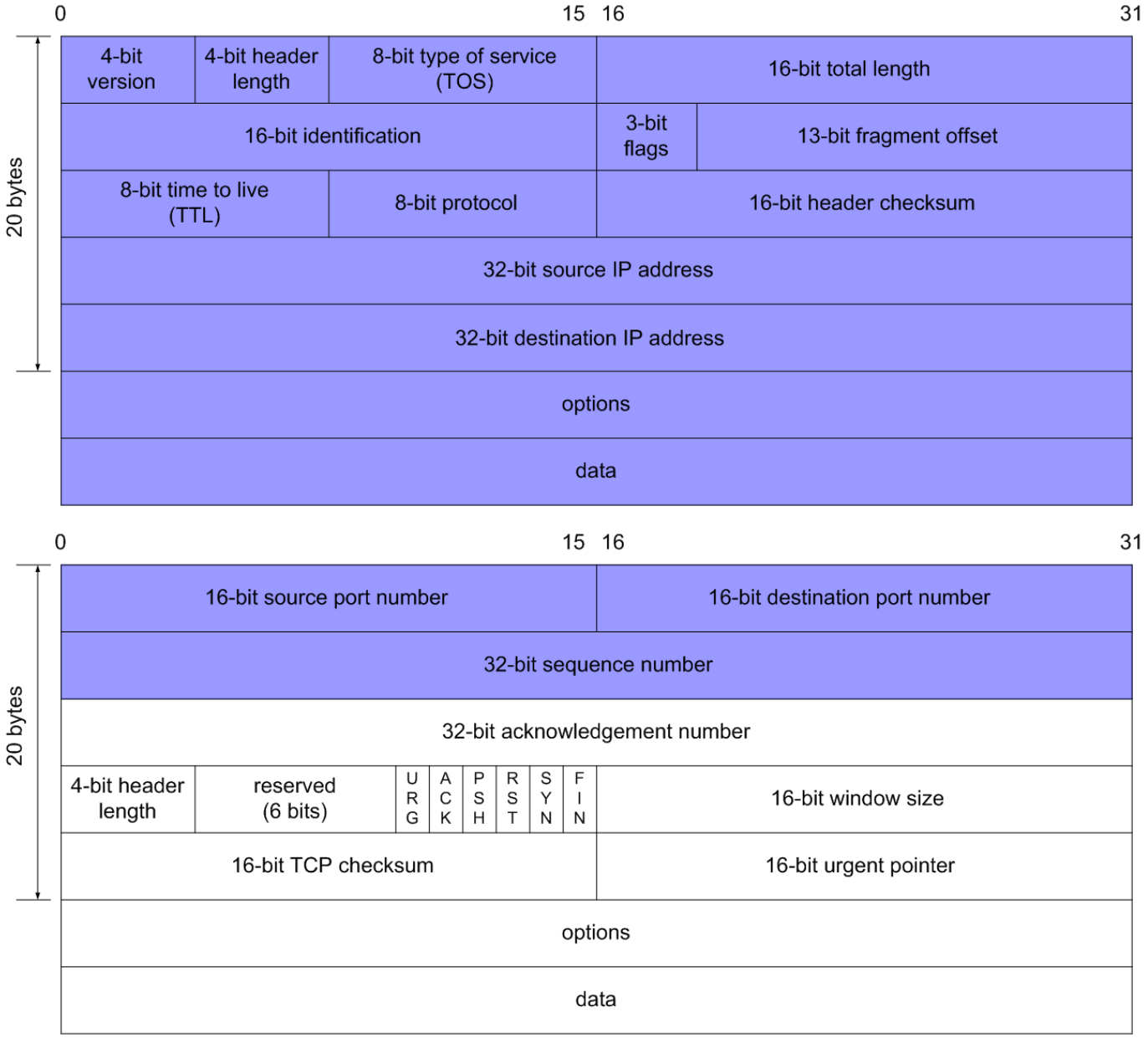


Formato de un mensaje de error ICMP “Destination Unreachable”

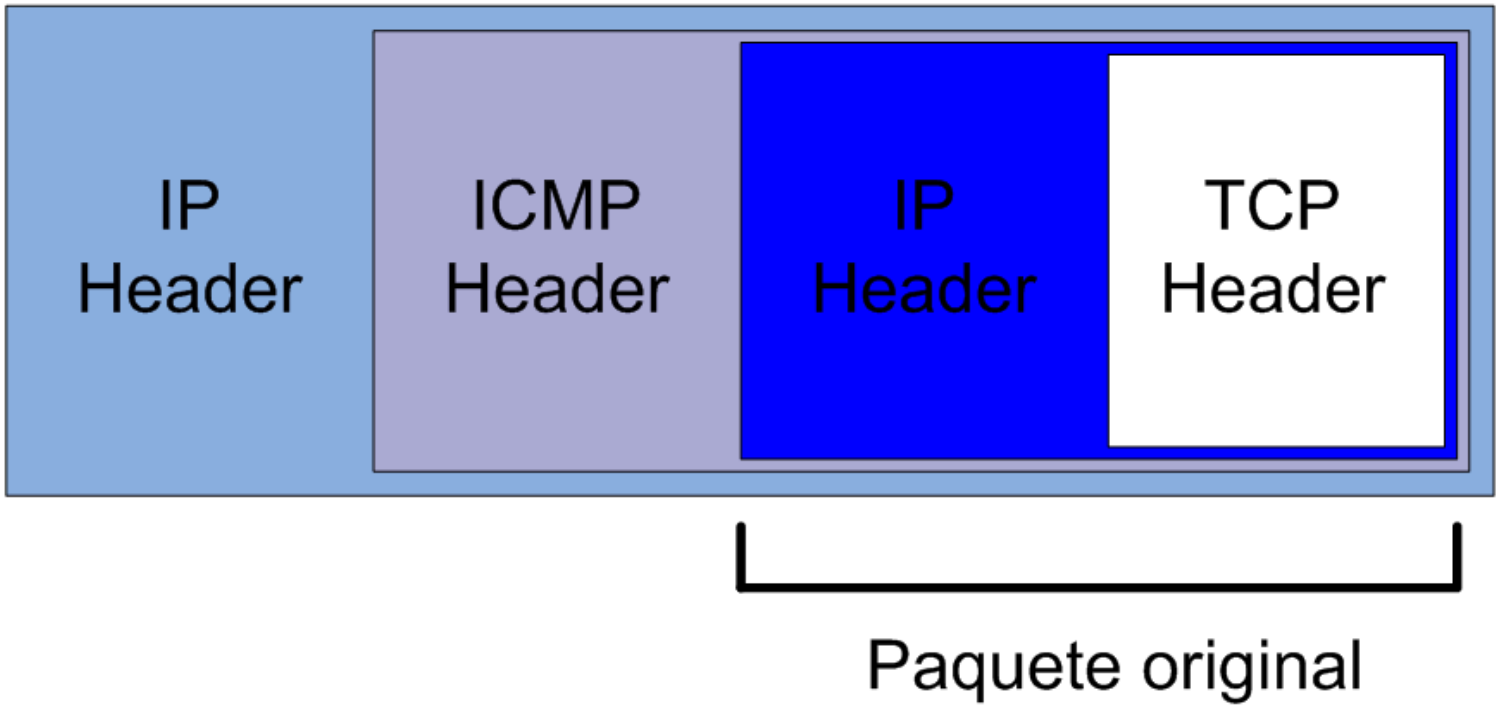


- Básicamente está compuesto por un checksum,”tipo” de mensaje,”código de mensaje”, e información para la demultiplexación del mensaje de error.

Información incluída en el mensaje ICMP



Estructura del paquete resultante



El paquete ICMP contiene en su “payload” parte del paquete original que causó el error. Dicho paquete ICMP se encapsula en un paquete IP, para ser enviado hacia el sistema que debe recibir el mensaje de error.

Clasificación de mensajes de error ICMP

Las especificaciones de la IETF hacen una clasificación grosera de los mensajes de error ICMP en aquellos que suponen indicar "**errores leves**" (soft errors), y aquellos que suponen indicar "**errores groseros**" (hard errors)

- Los errores leves son aquellos que se supone que se solucionarán en un corto plazo.
- Los errores groseros son aquellos que se supone que no desaparecerán en el corto plazo.

Type	Code	Descripción	Clasificación
3	0	Net unreachable	leve
3	1	Host unreachable	leve
3	2	Protocol Unreachable	grosero
3	3	Port unreachable	grosero
3	4	Frag. needed and DF set	grosero
3	5	Source route failed	leve

Reacción de TCP a los mensajes ICMP

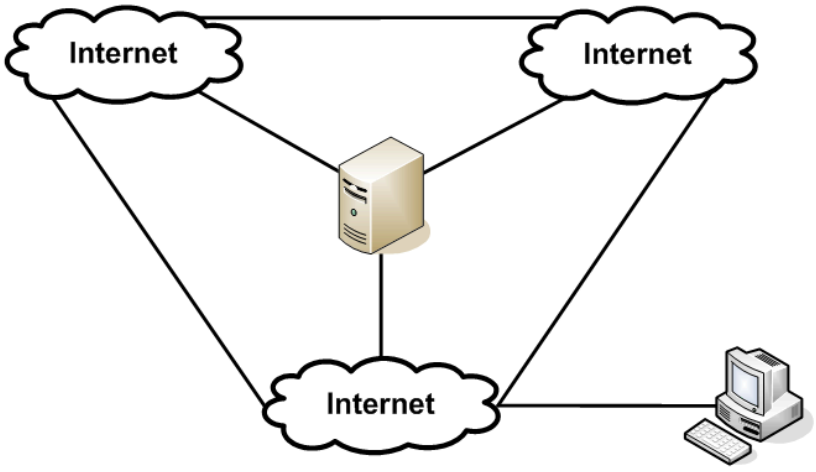
Cuando TCP es notificado de una condición de error, aplicará su política de recobro de fallos:

- Si el error notificado es un “error leve” (soft error), TCP recordará el error, y continuará retransmitiendo la información hasta recibir un acuse de recibo, o hasta que decida abortar la conexión (por haber reintentado muchas veces).
- Si el error notificado es un “error grave” (hard error), TCP abortará la correspondiente conexión inmediatamente.
- Esta política de reacción de TCP se corresponde con la semántica de los mensajes de error ICMP definida el estandar RFC 1122.

Ejemplos prácticos en los que la política de TCP no es adecuada

- Ejemplo I: Host accesible mediante varias direcciones IP.
- Ejemplo II: Host con pila doble (v4/v6) y v6 habilitado por defecto.

Ejemplo I: Sistema accesible a través de varias direcciones IP

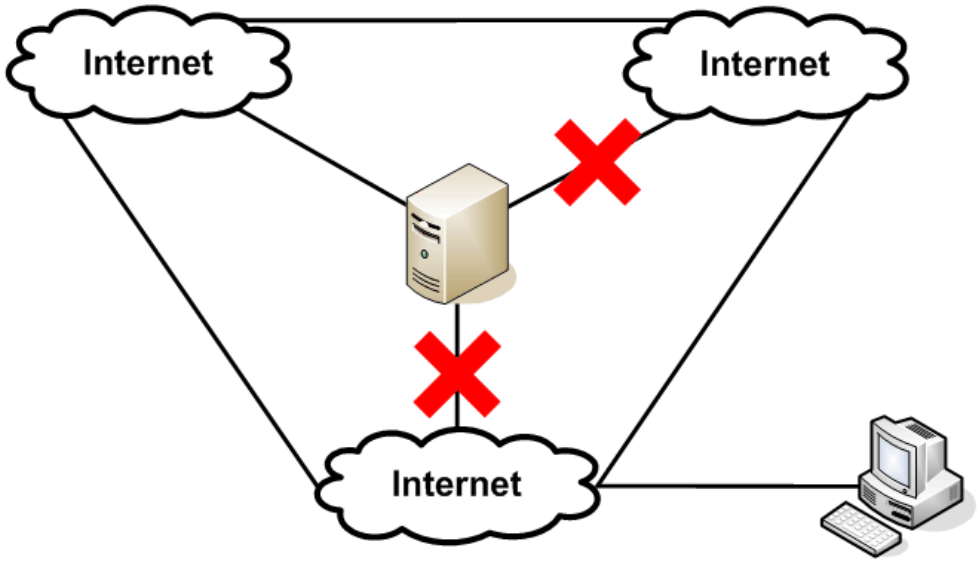


www.gont.com.ar

Record	Tipo
#1	A
#2	A
#3	A

- Con el fin de tener redundancia en su conexión a la red, un mismo sistema es accesible a través de varias direcciones IP.

Ejemplo I: Escenario problemático (las primeras direcciones son inaccesibles)

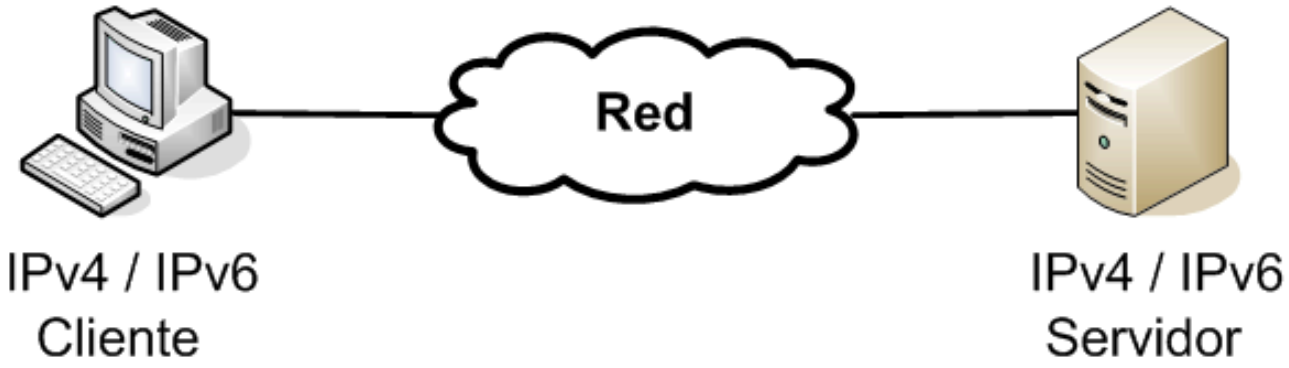


www.gont.com.ar

Record	Tipo
#1	A
#2	A
#3	A

- Si una dirección es inaccesible, TCP reintentará conectarse a cada dirección durante al menos 3 minutos.
- Si recién la tercera dirección de la lista estuviera accesible, estableceríamos nuestra conexión recién a los 6 minutos desde el primer intento.

Ejemplo II: Sistema con dual-stack intenta acceder a servidor dual-stack

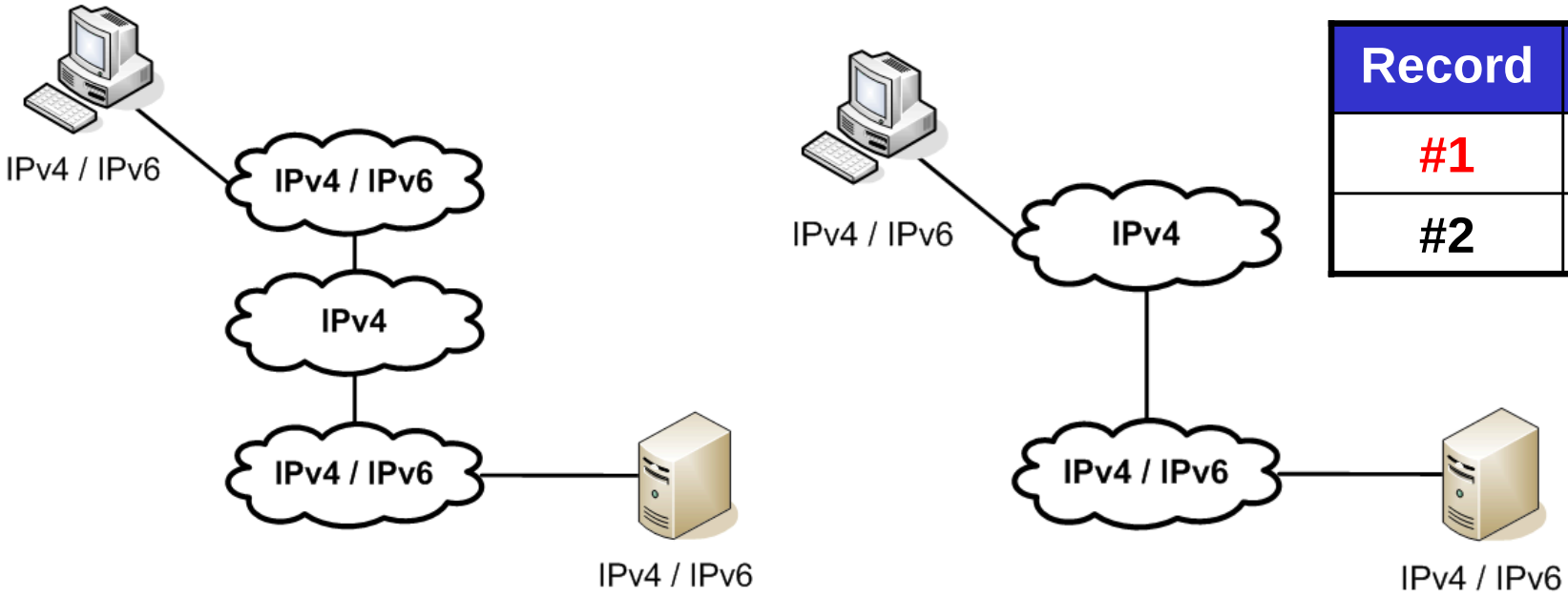


www.gont.com.ar

Record	Tipo
#1	AAAA
#2	A

- Un sistema dual-stack con v6 habilitado (por defecto), pero conexión v6 limitada, intenta acceder a un sistema con conectividad v4/v6

Ejemplo II: Escenario problemático (IPv6 no disponible como para alcanzar al servidor)

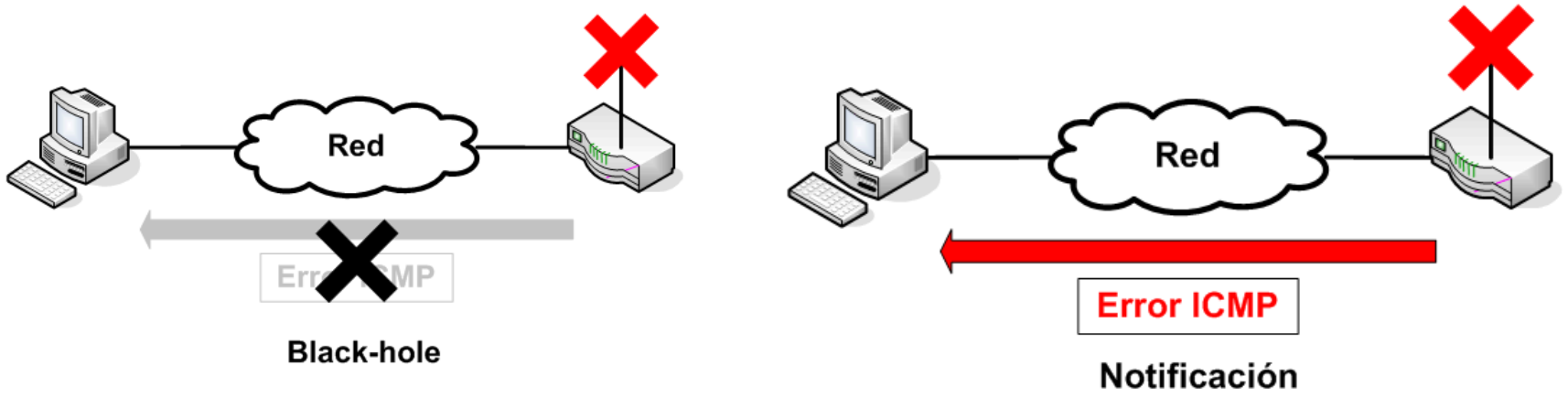


www.gont.com.ar

Record	Tipo
#1	AAAA
#2	A

- Un sistema tiene Dual-stack, con IPv6 habilitado por defecto. Sin embargo, no se dispone de conectividad como para alcanzar al servidor de la figura.
- Las reglas de DNS dictan que primero se debe intentar establecer la comunicación utilizando las direcciones IPv6.
- Esto generará una demora de 3 minutos por cada dirección IP que posea el servidor en cuestión.

Escenarios de error



- Si la condición de error no es señalizada (black-hole), no hay demasiadas alternativas de acción.
- Sin embargo, en una gran cantidad de casos los problemas de conectividad **sí** son señalizados, mediante un mensaje de error ICMP
- Sin embargo, debido a la política de respuesta a errores leves de TCP, se seguirá reintentando establecer la conexión con la misma dirección de destino.

Solucionando el inconveniente

- En una gran cantidad de casos la condición de error que no permite el estableciemiento de conexión **es** notificada al sistema emisor.
- Sin embargo, la politica de recobro de fallas de TCP determina que se siga reintentando el establecimiento de conexión co la misma dirección.
- Esto sugiere que hay que re-evaluar la semántica de los mensajes de error ICMP.

RFC 816 (Clark tenía razón)

- D. Clark señaló (RFC 816), en el año 1982, que establecer una política de reacción a mensajes de error no es algo simple.
- En principio, uno podría suponer que si se recibe un error durante la vida de una conexión, el mismo corresponde a un fallo temporal.
- Por el contrario, si se recibe un error como respuesta a un pedido de establecimiento de conexióno, probablemente sea que la conexión se intentó abrir de forma incorrecta.

Política de reacción a errores ICMP dependiente del estado de la conexión

- Si se recibe un error durante la vida de una conexión, el mismo no puede ser “grave” (de haberlo sido, nunca pudiéramos haber establecido la conexión en cuestión)
- Si se recibe un error como respuesta a una petición de establecimiento de conexión,no tiene demasiado sentido seguir intentando el mismo destino. Al fin y al cabo, no tenemos ningún conocimiento de que el destino en cuestión sea verdaderamente accesible.

En es decir….

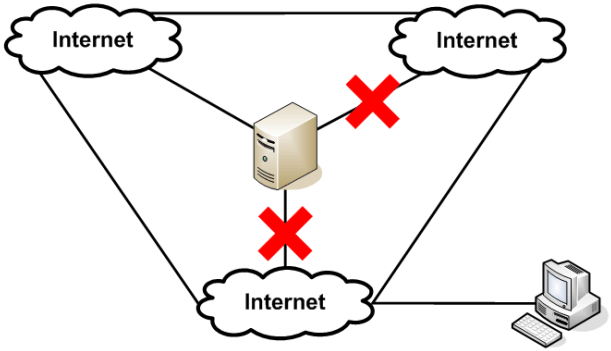
- Ningún mensaje indica “per se” la gravedad de un error.
- Mas bien,

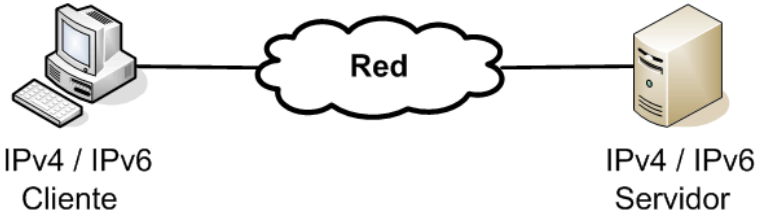
Los errores son “graves” o “leves” de acuerdo al estado de la conexión para la cual fueron recibidos

Modificación a la reacción de TCP a errores ICMP

- Si se recibe un mensaje de error ICMP destinado a una conexión que esté en cualquiera de los estados no-sincronizados, la conexión en cuestión debería ser abortada.
- Si se recibe un mensaje de error ICMP para una conexión en cualquiera de los estados sincronizados, entonces la conexión en cuestión **no** debe ser abortada.

Escenarios anteriores con la nueva política de reacción a errores ICMP





- Todas las direcciones “inalcanzables” fallan tan pronto como se recibe el mensaje ICMP correspondiente.
- Como consecuencia, virtualmente no existe demora en el establecimiento de conexión.

Documentos

- La reacción (nueva) de TCP a mensajes de error ICMP durante el establecimiento de conexión fue propuesta en “TCP’s reaction to soft errors” (draft-ietf-tcpm-tcp-soft-errors).
- La reacción de TCP a mensajes de error ICMP durante la vida de una conexión fue propuesta por “ICMP attacks against TCP” (draft-ietf-tcpm-icmp-attacks).

Ambos disponibles en

http://www.gont.com.ar/drafts

Preguntas y respuestas

Fernando Gont

fernando@gont.com.ar

http://www.gont.com.ar