2016

Práctica 3

Redes de Comunicaciones I

Memoria de la práctica 3 de REDES I sobre monitorización.

Oscar García de Lara Parreño

Santiago Gómez Aguirre

24/11/2016

CONTENIDO

INTRODUCCIÓN

En esta práctica hemos actuado como gestores de red haciendo una monitorización pasiva, ya que la hacemos con una traza que se nos proporciona de una red. Para ello hemos empleado una serie de scripts que hemos elaborado con el uso de:

- La herramienta **tshark** para la captura y análisis de tráfico de red. Esta herramienta posee unas prestaciones similares a las de **Wireshark**, con la diferencia de su manejo por la línea de comandos y la posibilidad de analizar archivos de mayor tamaño.
- Una serie de comandos generales de **shell scripting** para sistemas GNU/Linux en el ámbito del análisis de tráfico de red.
- El lenguaje de scripting **awk**, orientado al procesamiento por líneas en archivos de texto.
- Para poder analizar ciertos resultados, les aplicaremos una **ECDF** (Función de distribución acumulación empírica).
- La herramienta **Gnuplot** para la realización de las gráficas que empleamos para describir los resultados de los análisis del tráfico de red.

El informe está dividido en distintas etapas donde vamos analizando, de forma general a elementos más específicos, la traza de la red proporcionada.

PORCENTAJES DE PAQUETES

Lo que hemos hecho para poder obtener los porcentajes, ha sido usar tshark para que nos imprima en /dev/null todos los paquetes y a través de una tubería usar wc para contar las líneas e imprimir el valor en un fichero que sería el número de paquetes total, haciendo lo mismo, pero con un filtro de visualización para obtener solo los paquetes IP, y con una resta podemos sacar el número de paquetes IP.

Para los paquetes TCP, UDP hacemos algo parecido, filtramos por IP y por TCP o UDP según el caso y que no sean ICMP para evitar que nos lo cuente como TCO o UDP, después sumamos los TCP y UDP y los restamos al número de paquetes IP para que nos saque los demás tipos de paquetes que son la categoría.

PAQUETES IP-NOIP

Tipo	**Porcentaje**
IP	99.01 %
No-IP	0.99 %

Casi la totalidad de paquetes usan el protocolo IP ya que hoy en día es prácticamente el único que se usa a este nivel, por tanto, es lo esperado.

PAQUETES TCP-UDP-OTROS

Tipo	**Porcentaje**
TCP	89.59 %
UDP	9,73 %
Otros	1,68 %

La suma de porcentaje da 101% lo que no tiene sentido en porcentaje, esto se produce por el redondeo automático de los decimales. También se puede observar que la gran mayoría de paquetes IP usan TCP como protocolo de transporte, haciendo la dupla IP/TCP la más usada del mundo.

TOP 10 DE IP Y PUERTOS

Para obtener los distintos rankings hemos empleado el script2.sh. Empleamos tshark para obtener los campos de las direcciones IP, puertos y bytes mediante un filtrado en función de IP, TCP y UDP en cada caso. Los resultados de cada análisis los guardamos en distintos archivos de texto.

PAQUETES

DIRECCIONES IP

Posición	**Repeticiones**	**Dirección**
1	46449	32.22.91.4
2	19335	86.26.95.9
3	8662	14.7.119.221
4	5930	31.217.130.45
5	3889	77.74.5.48
6	3785	108.132.132.61
7	2929	120.24.138.137
8	2827	49.170.151.67
9	2795	17.238.193.226
10	2567	52.191.135.5

La dirección IP que está en primera posición se repite más del doble de veces que la segunda y la diferencia aumenta respecto al resto. Esto indica que la mayoría de paquetes de la traza establecen comunicación con esta dirección IP.

PUERTOS TCP

Posición	Repeticiones	Puerto
1	48982	80
2	6909	55934
3	5409	55860
4	3821	55865
5	2929	54615
6	2795	43585
7	2416	33896
8	2188	55173
9	1814	55848
10	1531	46371

El puerto más utilizado es el 80, lo cual ocurriría en cualquier otra traza seguramente, debido a que este puerto es el que por defecto se utiliza en la comunicación con protocolo HTTP usado en cada transacción de la web (www).

PUERTOS UDP

Posición	Repeticiones	Puerto
1	3785	48883
2	3785	14286
3	1183	53
4	190	5353
5	134	5355
6	124	547
7	124	546
8	54	1900
9	6	63423
10	6	58532

En el caso de los puertos UDP tenemos un empate respecto a los puertos que ocupan las dos primeras posiciones; en tercera posición tendríamos el puerto 53 que se corresponde con el de DNS, es lógico que este puerto ocupe una de las primeras posiciones de puertos

UDP ya que los mensajes DNS solo se envían mediante este protocolo. También podemos corroborar con estos datos que la mayoría de paquetes de la traza son TCP y no UDP debido a los valores de repetición tan discretos que obtienen este tipo de puertos.

BYTES

Para hacer este apartado hemos considerado que el tamaño del paquete se contabilice tanto en el puerto/dirección de origen como el del destino.

DIRECCIONES IP

Posición	Dirección	Byte
1	32.22.91.4	51370740
2	86.26.95.9	23347683
3	31.217.130.45	6997269
4	14.7.119.221	4823709
5	77.74.5.48	4403688
6	17.238.193.226	3281612
7	49.170.151.67	3241463
8	52.191.135.5	3041083
9	120.24.138.137	2800279
10	40.222.156.38	2510891

Las primeras cinco direcciones son las mismas que en el caso de repeticiones, solo intercambiando la tercera con la cuarta. Es lógico que en las primeras posiciones casi no haya variaciones ya que la diferencia de repeticiones que había entre esas direcciones era muy grande por tanto los bytes transmitidos son mayores, aunque no tenga por qué, cómo se puede observar de la sexta a la décima dirección, que no se cumple esa máxima.

PUERTOS TCP

Posición	Puerto	Byte
1	80	53931024
2	55934	8324572
3	55860	6505361
4	55865	4849192
5	43585	3281612
6	54615	2800279
7	33896	2742973
8	55173	2594791
9	55848	2098831
10	46371	1778032

En este caso con los puertos pasa algo parecido, los primeros cuatro y los cuatro últimos puertos están en la misma posición, solo se intercambian el quinto con el sexto. También se puede observar que la cantidad de Byte que tiene el primero respecto con el segundo son 6.5 veces más.

PUERTOS UDP

Posición	Puerto	Byte
1	48883	1816645
2	14286	1816645
3	53	132111
4	5353	46634
5	1900	18462
6	547	18337
7	546	18337
8	5355	11460
9	63423	1080
10	58532	1080

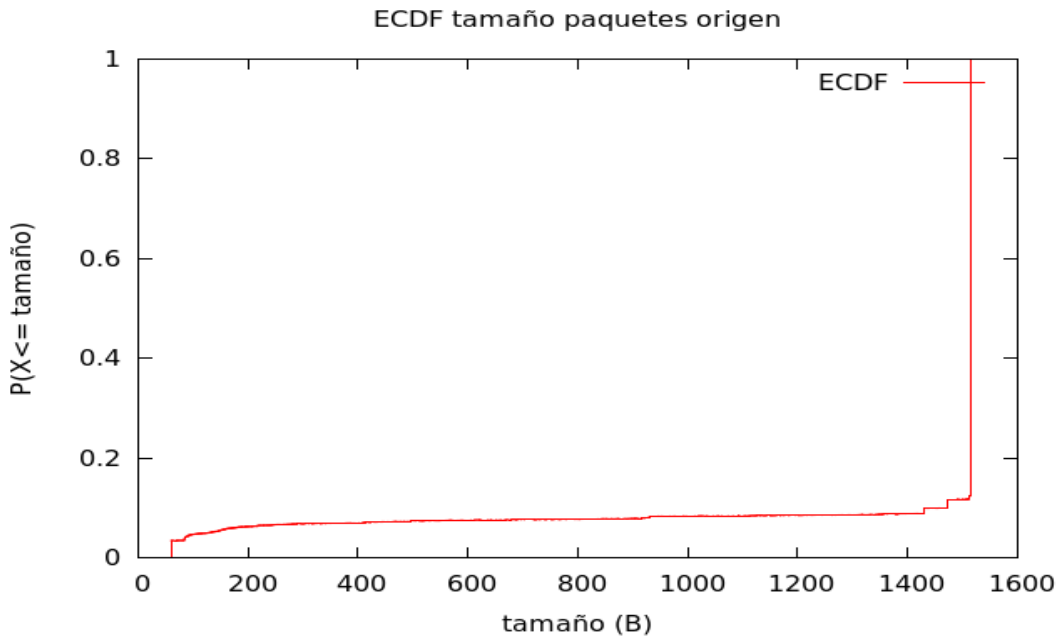
Esta tabla da un resultado muy curioso ya que hay repeticiones en el número de Byte entre varios puertos, esto se puede deberse a que las repeticiones de esos puertos son las mismas entre sí, y los paquetes deben tener los mismos tamaños. Podemos suponer que una misma aplicación, ya esté en distintos sistemas terminales, ha realizado la misma acción por puertos distintos. Sobre las diferencias entre las tablas son parecidas a los casos anteriores.

TAMAÑO A NIVEL 2

Para poder obtener las imágenes hemos usado el script3.sh. Generamos con tshark un fichero con el tamaño de los paquetes aplicando en cada llamada a tshark un filtro distinto para diferenciar el sentido del paquete. Después usamos el ejecutable de crearCDF que es el programa que nos proporcionaban completándolo para que después de ordenar nos cuente cuantas repeticiones tiene cada tamaño y así imprimir en un fichero el ECDF correspondiente. A continuación, usamos GNUplot para pintar las gráficas que mostramos a continuación.

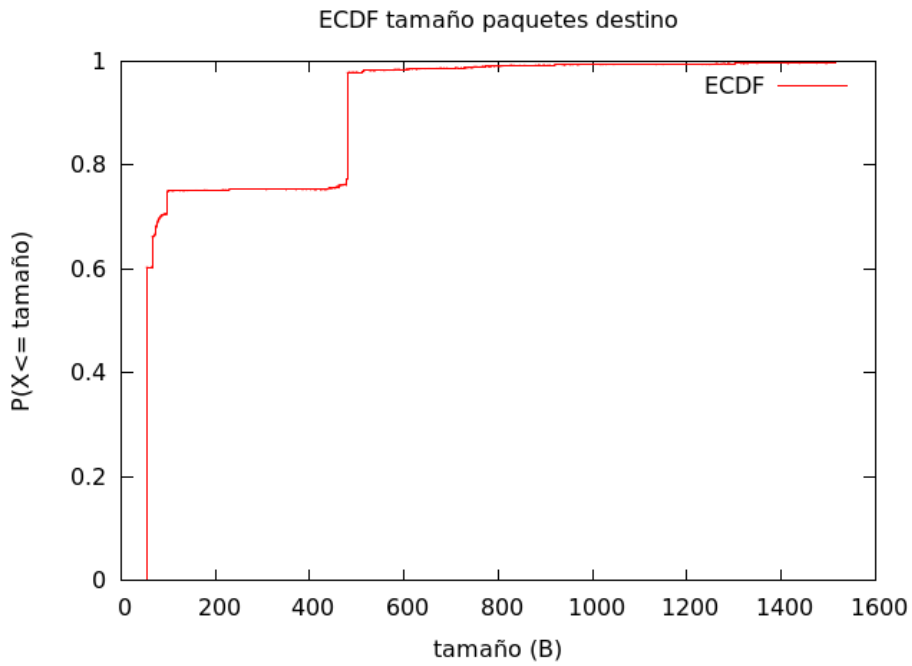
DE LOS PAQUETES DE LA TRAZA

ORIGEN



Los paquetes que salen tienen una probabilidad el 1% que el tamaño sea menor de 1430 B, con el crecimiento en vertical al 100% se confirma que la gran mayoría de paquetes tienen el tamaño máximo que soporta la línea que son 1514B

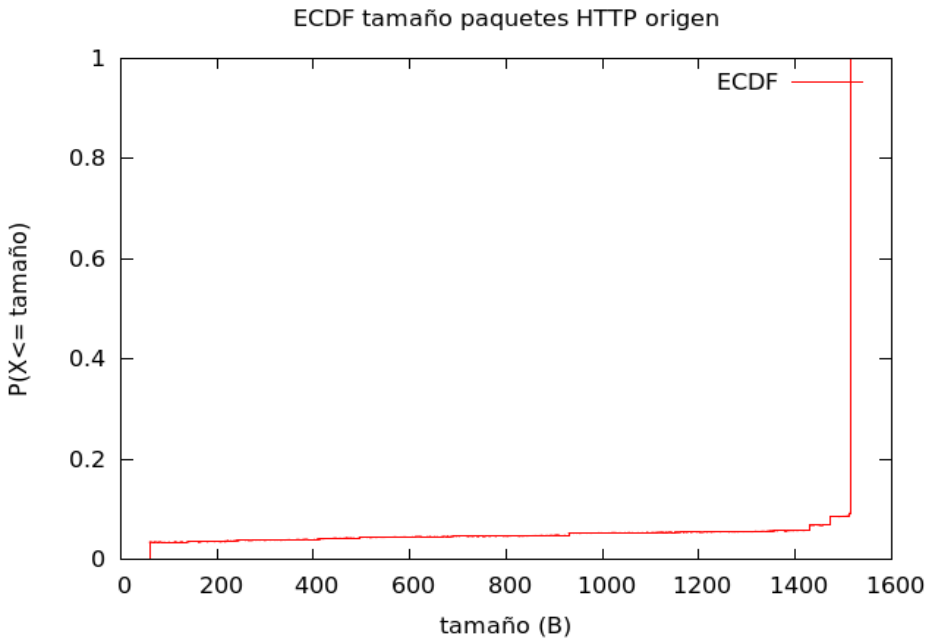
DESTINO



Los paquetes de entrada tienen más escalones el 75% es menor a unos 500B y unos 97% aproximadamente es menor de 600B, como en el caso anterior.

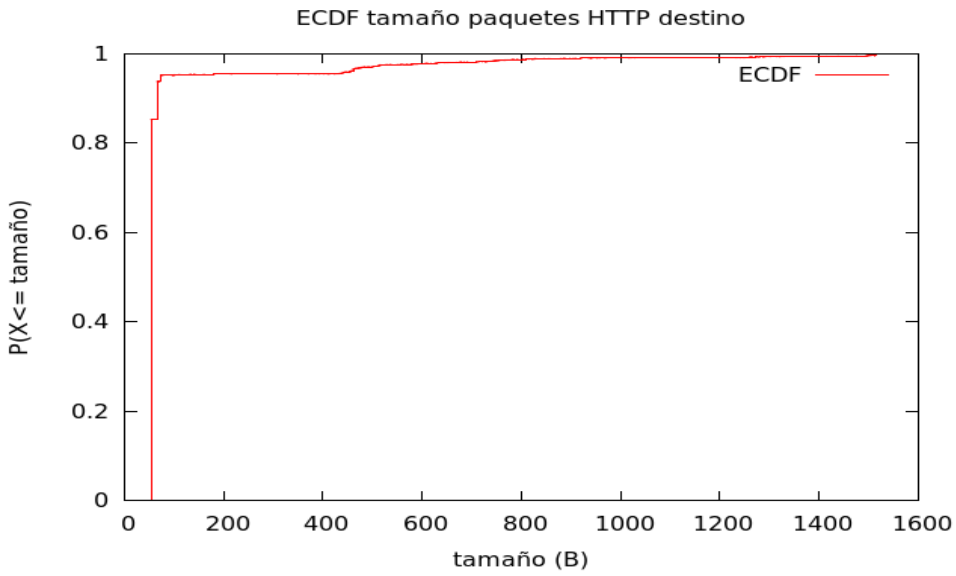
DE LOS PAQUETES HTTP

ORIGEN



Como se puede observar la gráfica es muy parecida a la anterior de origen, por lo que se puede deducir que una gran mayoría de paquetes son HTTP.

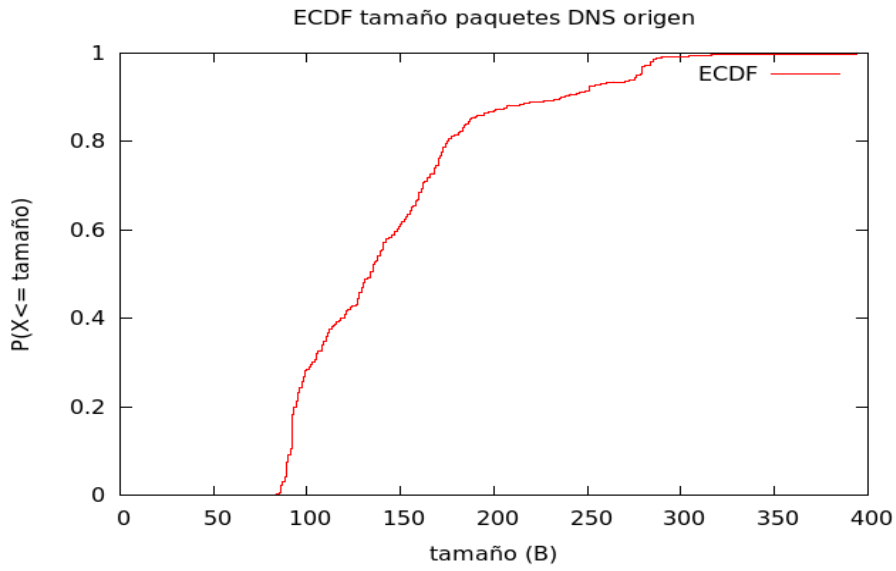
DESTINO



Esta no se parece a la gráfica destino anterior, es justo la inversa a su origen, un 95% es menor de 100B, después crece más rápidamente la probabilidad hasta los 800B que se vuelve otra vez plano hasta alcanzar el paquete máximo.

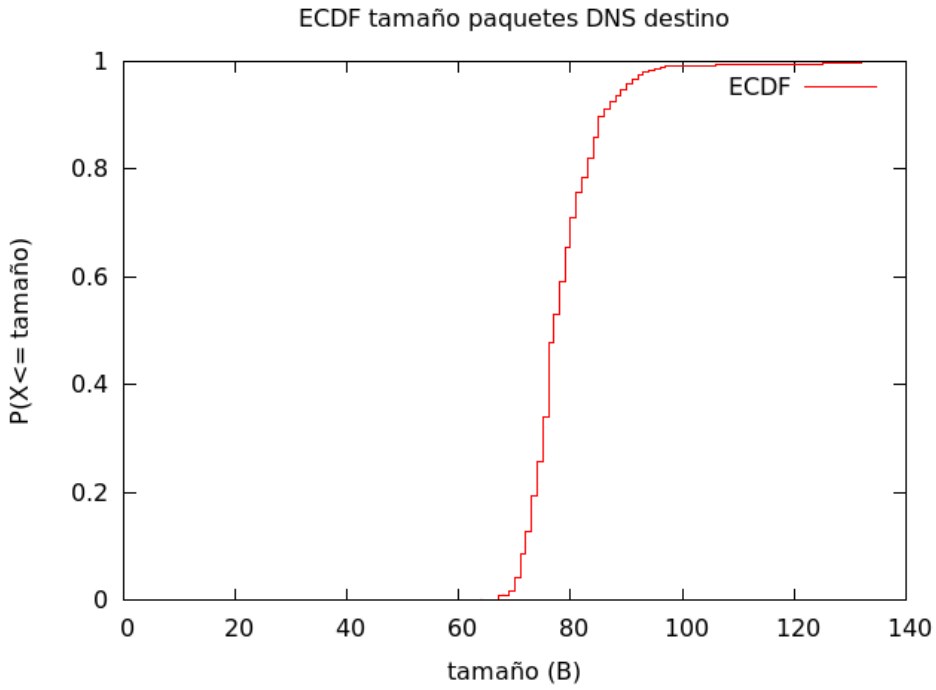
DE LOS PAQUETES DNS

ORIGEN



Es una gráfica que crece más lentamente, no tiene los grandes saltos verticales de la anterior, también hay que señalar que el 100% no supera el tamaño 400 y es un cambio de las tablas anteriores que llegaban al máximo de la red, esto se debe a que los paquetes DNS tienen como función obtener la dirección IP de un dominio.

DESTINO



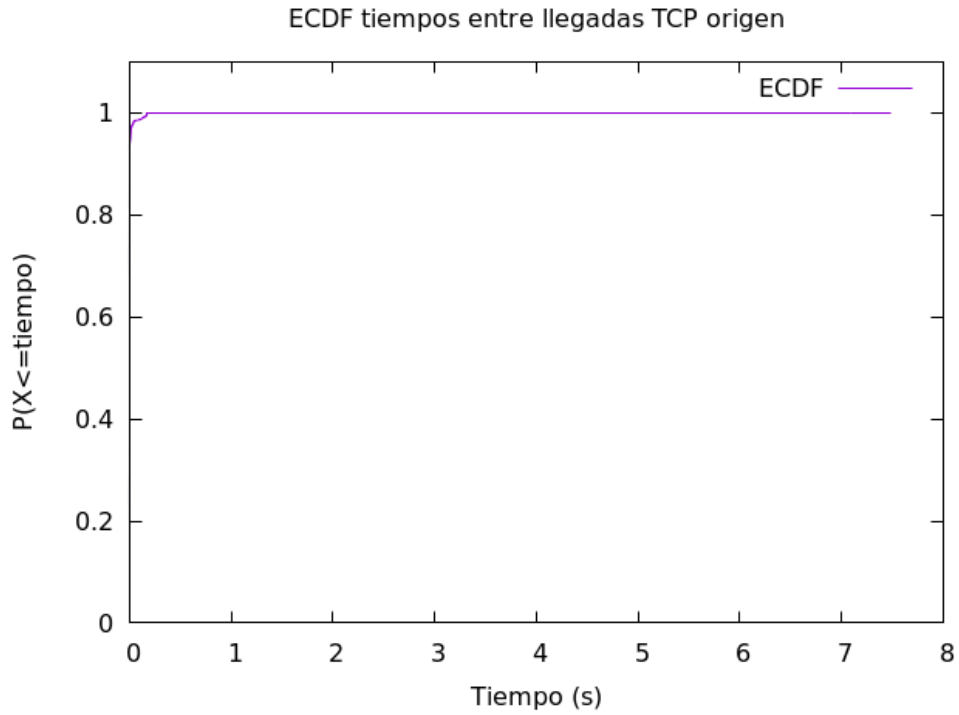
Una gráfica similar a su origen, pero con los escalones más marcados, también el tamaño de paquete es mucho más pequeño supera escasamente 130B el máximo, esto se deberá a que los paquetes de respuesta que traen la IP necesaria no necesitan meter mucha información.

TIEMPOS ENTRE LLEGADAS

Para la obtención de los tiempos entre llegadas del flujo UDP y TCP en ambos sentidos hemos empleado tshark para filtrar la traza principal en otras cuatro diferentes, dos por cada tipo (una para cada sentido). Después hemos obtenido el campo de los tiempos de cada traza y los hemos almacenado en un fichero para ejecutar el programa crearCDF y obtener las probabilidades que hemos representado en las siguientes gráficas mediante el uso de GNUplot.

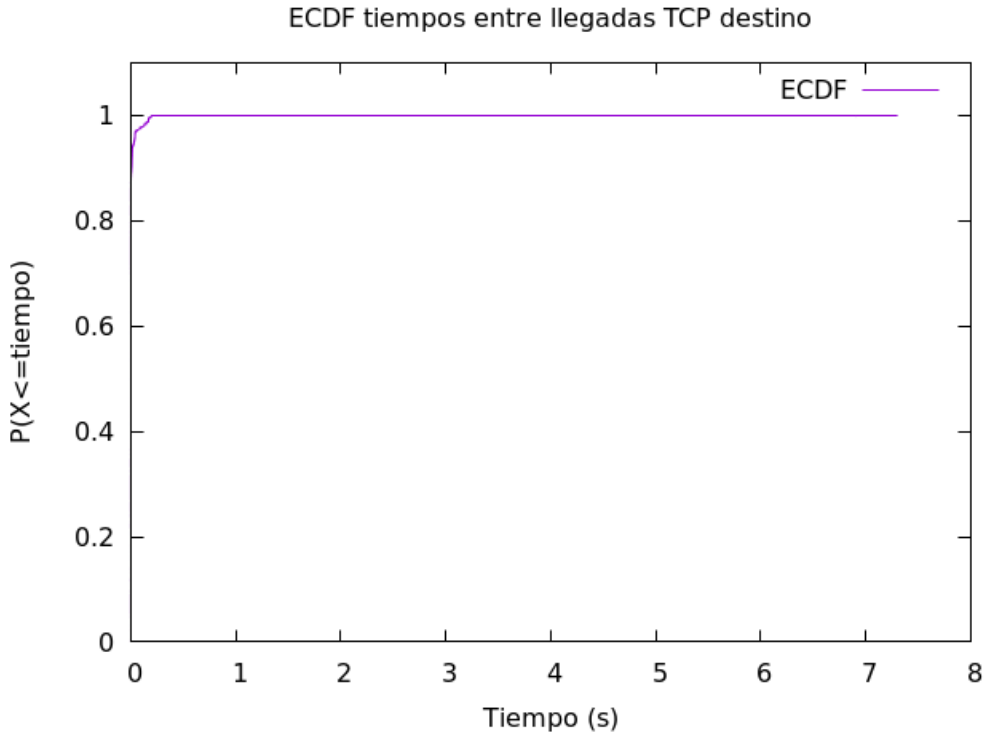
FLUJO TCP

ORIGEN



La mayoría de paquetes TCP tienen un tiempo entre llegada al origen de menos de 0.5 segundos, de tal forma que tenemos un 98% de probabilidad que sea de esta manera, y un 2% de que tarde más, pero podemos observar que no se llega nunca a más de los 7 segundos y medio aproximadamente en la traza suministrada.

DESTINO



En cuanto a los tiempos entre llegada al destino de paquetes TCP se puede observar que sigue el mismo comportamiento que hemos observado en la gráfica anterior.

FLUJO UDP

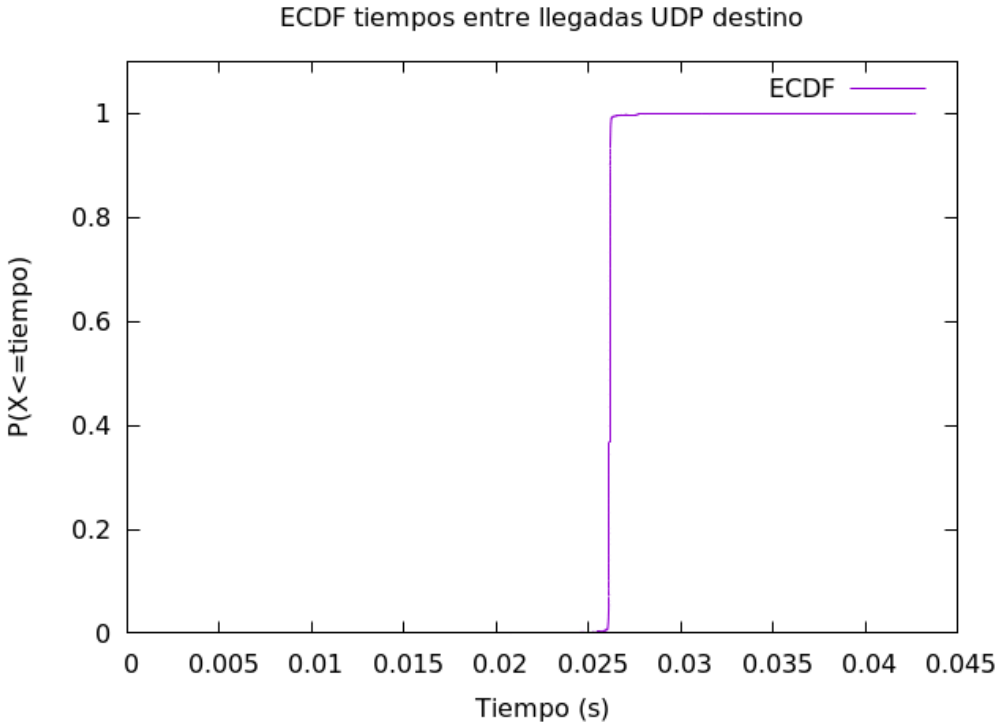
ORIGEN

En esta traza hemos comprobado que no hay paquetes UDP que se envíen desde el origen por lo que no tenemos tiempo entre llegadas del flujo UDP del origen.

```
santinix@santinix-X556UJ ~/Documentos/UAM/REDES I/practica 3 $ tshark -r traza.pcap -Y 'udp && !icmp && udp.srcport eq 14286'
santinix@santinix-X556UJ ~/Documentos/UAM/REDES I/practica 3 $
```

Mediante este comando hemos podido realizar la comprobación de que no hay ningún paquete que cumpla ese filtro, ya que al ejecutarlo como se puede observar no obtenemos ninguna salida.

DESTINO

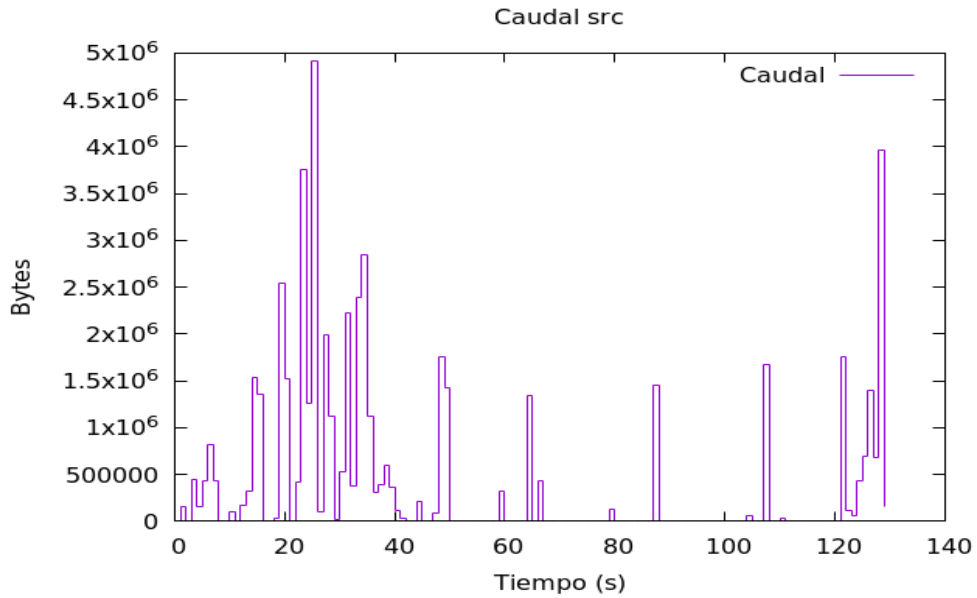


Por otro lado, sí que tenemos tiempo entre llegadas del flujo UDP destino, teniendo un 99% de probabilidad de que tarde 0.026 segundos. Se puede observar que los tiempos entre llegadas del flujo UDP son menores que los del flujo TCP, entendemos que esto puede ser a que TCP es un protocolo orientado a la conexión ya que comprueba la integridad de los datos.

CAUDAL A NIVEL 2

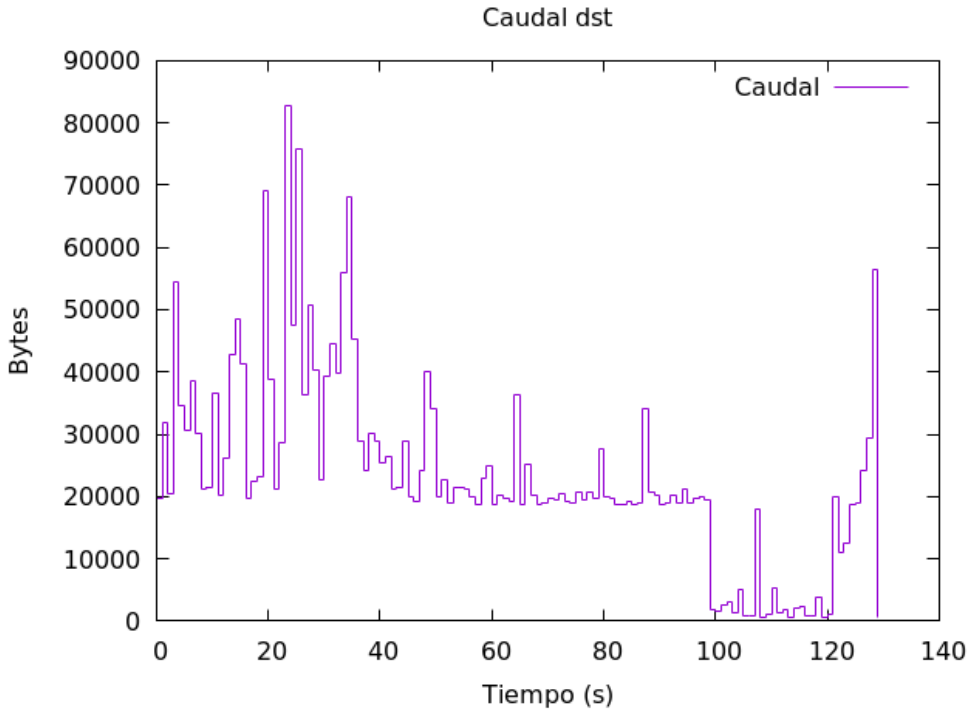
Para la obtención del caudal a nivel 2 hemos empleado tshark para filtrar la traza suministrada en otras dos, una para cada sentido a nivel ethernet. Después hemos analizado la conversación de estas trazas empleando la herramienta para conseguir la granulidad de 1 segundo. Los datos recogidos los hemos almacenado en dos ficheros que luego hemos utilizado para hacer las gráficas que representan el caudal con GNUplot.

ORIGEN



La gráfica representa como al inicio tenemos el mayor envío de bytes que en el resto de la conversación. Después de estos primeros 30 segundos se puede apreciar como hay picos periódicos cada 20 segundos en los que se envía medio millón de bytes aproximadamente. Al final de la conversación también se obtiene un pico de envió de bytes pero que no llega al máximo alcanzado al inicio.

DESTINO



En cuanto al caudal de bytes recibidos podemos observar que mantiene unos valores más constantes que el caudal de bytes enviados. Las dos principales diferencias entre ambos caudales es que se envían más bytes que los que se reciben y que el caudal de llegada de bytes que se reciben en ningún momento obtiene el valor 0 a diferencia del caso anterior. En esta gráfica también se puede apreciar que el mayor número de bytes recibidos se produce al inicio y que luego hay picos periódicos cada 20 segundos otra vez.

CONCLUSIÓN

Para concluir vamos a considerar el orden seguido y por qué creemos que se ha elegido este para realizar esta monitorización pasiva.

El motivo de empezar por saber la cantidad de paquetes que siguen distinto protocolo de red o de transporte influye en la forma de trazar las distintas pruebas, ya que si no se usara el protocolo IP tendría sentido despreciar los paquetes que lo usan, también se puede aplicar a los protocolos de la capa de transporte, tal como hemos despreciado en el análisis de los paquetes NoIP o IP/Otros.

Obtener un top de direcciones IP nos puede ser útil para saber qué servicios son los más utilizados no solo en cantidad de paquetes que circulan sino su tamaño ya que nos puede interesar ampliar el tamaño de paquete máximo que puede transportar la red para evitar una posible congestión en los router. Con los puertos TCP y UDP pasa algo parecido ya que nos permite saber que puertos están siendo más utilizados, ya que ciertos puertos están reservados para ciertas aplicaciones FTP, HTTP, DNS…; y la carga en bytes que manejan.

A continuación, vemos el tamaño de paquetes que recibimos entre dos MAC, lo que permite saber la cantidad de paquetes de un determinado tamaño que fluyen en los dos sentidos. Esto nos puede servir para decidir si la velocidad de subida o bajada es suficiente para el uso que damos a nuestra red.

Por último, observamos los tiempos entre llegadas de los flujos TCP y UDP, así como el caudal en bytes a nivel 2, para conocer la velocidad, frecuencia y el ancho de banda que tiene la comunicación en la red. De esta manera podemos detectar si es necesario aplicar alguna mejora para mejorar el tiempo de las comunicaciones y aliviar la carga en bytes que se soporta con algún otro sistema.