Ingeniería de Protocolos
Grados en Ingenierías Informáticas
E.T.S.I. Informática
Curso 2021/2022

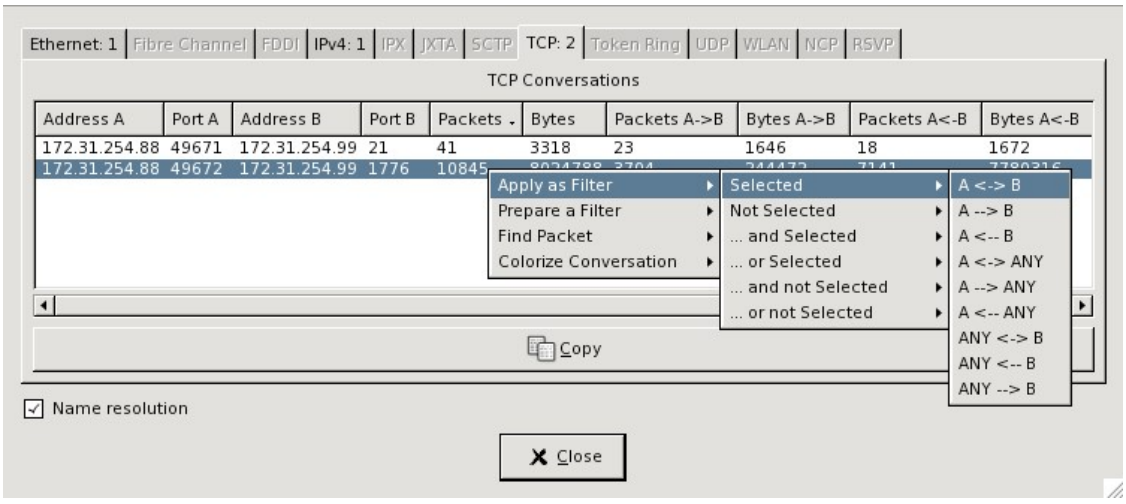
Práctica I. Análisis de rendimiento con Wireshark.

En esta sesión vamos a conocer algunas de las funcionalidades incorporadas en Wireshark para el análisis de rendimiento. En las sesiones anteriores nos habíamos centrado en los aspectos de Wireshark relacionados con el análisis estructural, y ahora nos centraremos en el otro tipo de funcionalidades.

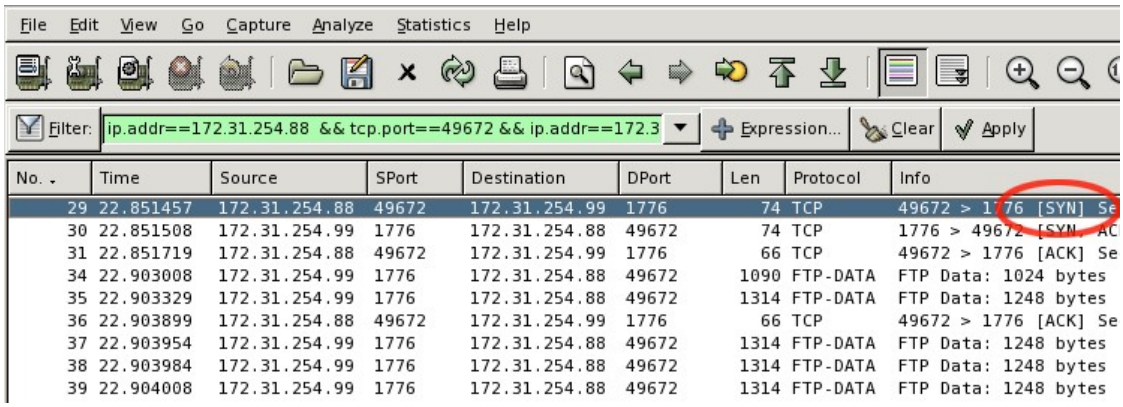
En Wireshark, aparte de las opciones de visualización de información de los paquetes y protocolos, que formarían parte del análisis estructural, también hay muchas funcionalidades estadísticas y de graficación, que en muchos casos son específicas para cada protocolo. Por ejemplo, en el caso de TCP, Wireshark ofrece gráficas para ver el rendimiento ("Throughput"), tiempo de retorno ("Round-Trip Time"), tamaño de ventana ("Window scaling"), etc.

En esta guía nos vamos a centrar en analizar el rendimiento del protocolo TCP, ya que es un protocolo muy popular y es fácil conseguir capturas. Lo primero que haremos será realizar una captura de tráfico o cargar un fichero de captura. Lo más probable es que ahora tengamos tráfico de muchos tipos de protocolos, y diversas conexiones diferentes. Vamos a centrarnos en estudiar una sola conexión TCP. Usamos la herramienta "**Conversations**" (menú **"Statistics →→ Conversations"**) para ver una lista, organizada por protocolos, de todos los pares de origen/destino que hay en la captura, y seleccionamos la pestaña "**TCP**". Otra manera alternativa, es ir directamente a las conversaciones de este protocolo usando el menú "**Statistics →→ Conversation List →→ TCP (IPv4 & IPv 6)**".

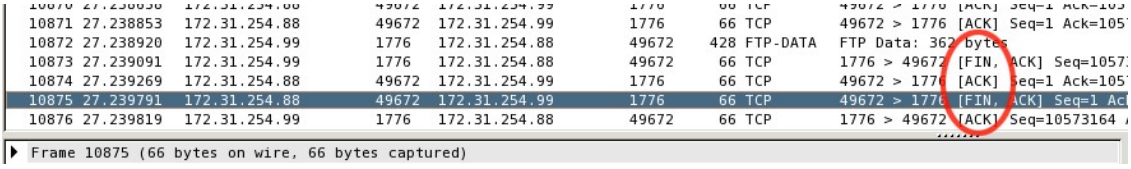
Ahora veremos todas las conversaciones de TCP, junto con las direcciones y puertos de origen y destino, el número de bytes y de paquetes trasferidos en ambas direcciones y en total. Podemos hacer click en cualquiera de las cabeceras de las columnas para ordenar la lista por ese campo en orden ascendiente o descendiente. Esto es muy útil para encontrar la conversación que estamos buscando (por ejemplo, la descarga de un fichero pesado). Una vez hemos identificado la conexión que queremos estudiar, le damos al botón derecho y seleccionamos **"Apply as Filter Selected →→ A<->B" y cerramos la ventana de las conversaciones.**



Ahora Wireshark deberá haber creado un filtro de visualización para la conversación seleccionada, y en el panel de la lista de paquetes solo aparecerá esta conversación. Si la captura que se ha hecho incluye toda la conexión TCP, entonces deberíampos poder ver el primer paquete de la conexión, el paquete "**SYN**", que inicia el "**three-way handshake**" en el protocolo TCP.

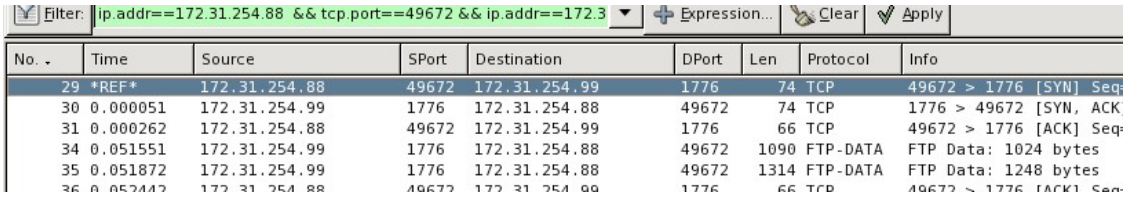


De la misma manera, al final de la conexión debe verse la secuencia típica para cerrar una conexión TCP ("**FIN; ACK; FIN; ACK**").

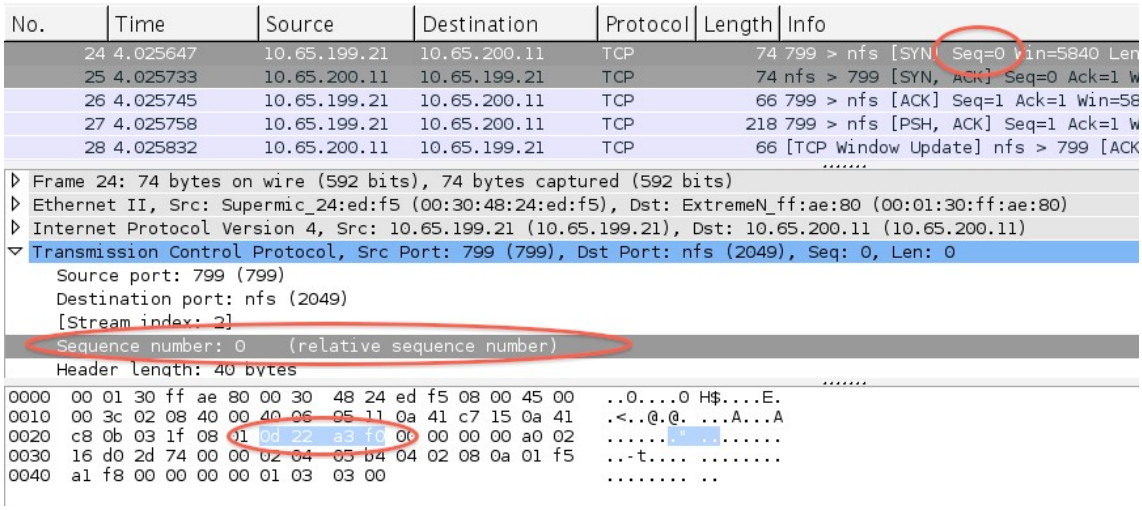


Ahora que tenemos aislada una sola conexión TCP, puede ser buena idea guardarla en un fichero de captura individual. Para ello hacemos uso del menú "**File→Export Specified Packets"** y activamos la opción "Displayed", para que guarde solo los paquetes que se están mostrando en ese momento (es decir, la conexión que hemos filtrado).

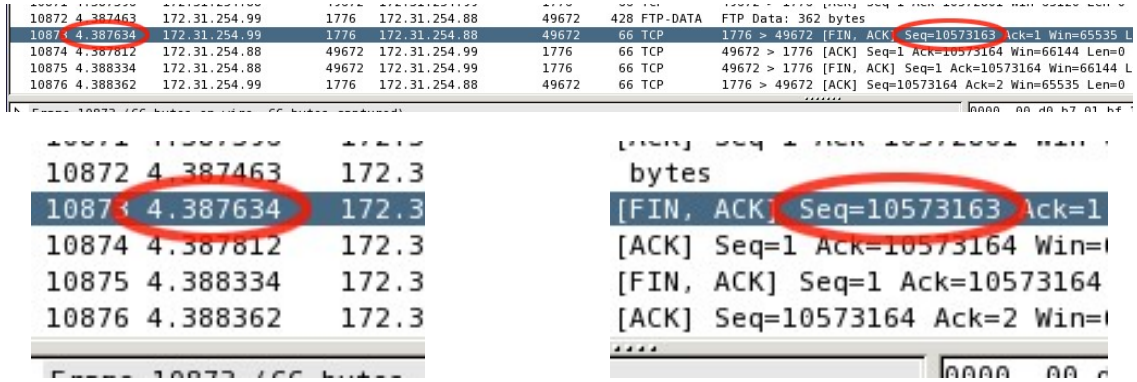
Queremos saber ahora cuál ha sido el rendimiento total de la conexión. Esto puede calcularse con precisión a partir de la información de los paquetes de la conexión. La gráfica de rendimiento de TCP que mencionamos al principio solo nos muestra el rendimiento instantáneo asociado a cada paquete, pero no el rendimiento global. Lo que queremos entonces es averiguar con exactitud cuántos datos se transmitieron y en cuánto tiempo, y dividir ambos valores. Lo primero que haremos será actualizar el tiempo de referencia al del primer paquete de la conexión (paquete "**SYN**"), que debería ser el primero de la lista. Para ello, lo seleccionamos, le damos al botón derecho sobre el paquete y elegimos la opción "**Set Time Reference**". Esto hará que el tiempo asociado a este paquete se cambie al valor "***REF***" y que el resto de tiempos sea calculado en función a este tiempo de referencia:



Por lo tanto, la marca de tiempo del paquete de final de la conexión (de tipo "**FIN**") va a coincidir con la duración total de la conexión. Ahora, para averiguar la cantidad de datos transferidos, analizaremos el número de secuencia de este paquete. Recordemos que los números de secuencia en TCP representan el número del primer byte de datos que hay en cada paquete. En teoría, el primer número de secuencia debe ser un valor aleatorio y el resto se asignan incrementalmente en referencia a este. Afortunadamente, Wireshark muestra todas las secuencias relativas a este valor, comenzando desde el 0. Por ejemplo, en la siguiente figura se puede ver que el número de secuencia 0 corresponde en realidad con el valor hexadecimal 0x0D22A3F0.



Esto hace que calcular la cantidad de datos transferidos sea muy sencilla, ya que simplemente consiste en mirar cuál es el último número de secuencia. Hay que tener cuidado aquí y asegurarnos que este último número corresponde a un flujo en el sentido que estamos analizando (Recordemos que una conexión TCP permite transmitir datos en ambos sentidos). Lo habitual es que sea el emisor el que decida cerrar la conexión. En nuestro ejemplo, el último número de secuencia es 10573163, y el tiempo de conexión es 4.387634 segundos (aproximadamente 4.39 s.).



Por lo tanto, el rendimiento global de la conexión es:

(10573163 Bytes * 8 bits/Byte) / 4.39 s = 19267723 bits/s = **19.27 Mbps**

Antes de continuar, desharemos el tiempo de referencia, ya que esto puede interferir con la generación de gráficos más adelante. Simplemente seleccionamos el paquete de referencia, le damos al botón derecho sobre el paquete y elegimos la opción "**Set Time Reference**" otra vez para dejarlo todo como estaba.
Como mencionamos antes, Wireshark produce varios tipos de gráficos para analizar una conexión TCP:

- Gráfico del tiempo de retorno ("**Round-Trip Time**"): Aquí podemos ver el tiempo de retorno para cada paquete. Este tipo de gráfico suele tomar la forma de una nube de puntos.
- Gráfico de rendimiento ("**Throughput**"): Esta gráfica muestra el rendimiento medio (en bits/segundo) a lo largo del tiempo.
- Gráfico del flujo TCP: Aquí Wireshark nos ofrece dos tipos de gráficas ("**Time/Sequence (Stevens)**" y "**Time/Sequence (Tcptrace)**") en los que se representa la evolución del número de secuencia a lo largo del tiempo. El segundo tipo es el que ofrece más información.
- Gráfico del tamaño de ventana ("**Window Scaling**"): Muestra cómo varía el tamaño de la ventana a lo largo de la conexión.

Estos gráficos nos serán muy útiles para diagnosticar problemas en la conexión TCP. Para acceder a ellos seleccionamos el paquete de la conexión que queremos analizar y usamos el menú "**Statistics -> TCP Stream graph**" para elegir alguno de los tipos de gráficos.

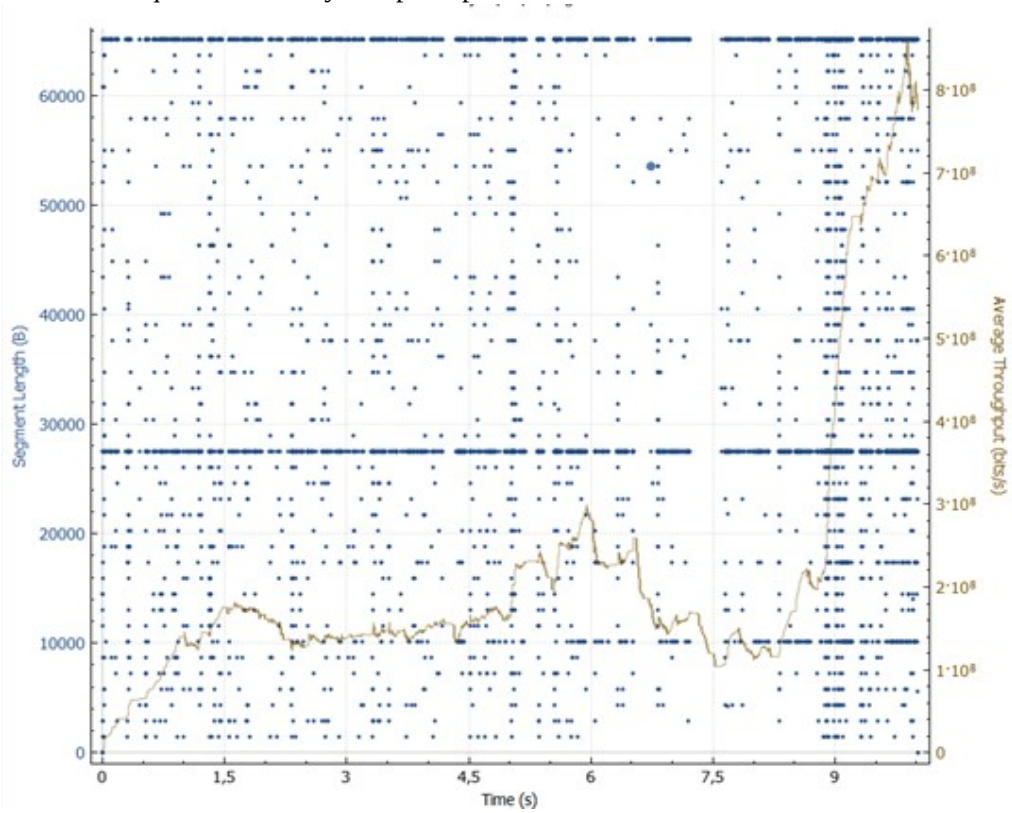
La ventana de control tiene 5 pestañas con diferentes opciones para controlar la visualización. Por ejemplo, en la pestaña "**Cross**", podemos activar que se muestre una cruz en el puntero del ratón para ayudarnos a medir sobre la gráfica (pulsar el botón secundario del ratón para visualizar todas las opciones). En la pestaña "**Graph type**" podemos ir cambiando el tipo de gráfico.

Estos gráficos de Wireshark son interactivos, y podemos hacer zoom sobre ellos y moverlos. Los controles varían dependiendo del sistema operativo usado y de la versión de Wireshark. En Windows son:

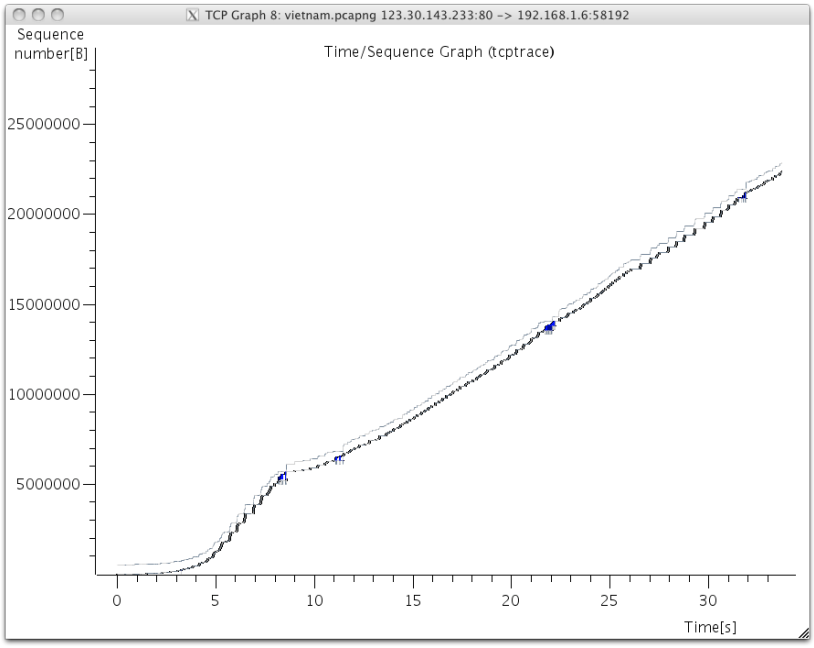
- Tecla "+": Hacer zoom
- Tecla "-" : Quitar zoom
- Espacio: Activar/desactivar cruz sobre el ratón
- Cursores: Desplazar el gráfico
- Tecla "1": Gráfico de RTT
- Tecla "2": Gráfico de rendimiento
- Tecla "3": Gráfico "Time/Sequence" (Stevens)
- Tecla "4": Gráfico "Time/Sequence" (tcptrace)
- Tecla "5": Gráfico de tamaño de ventana

También podemos hacer click sobre uno de los puntos de la gráfica, y se seleccionará automáticamente ese paquete en la lista de paquetes.

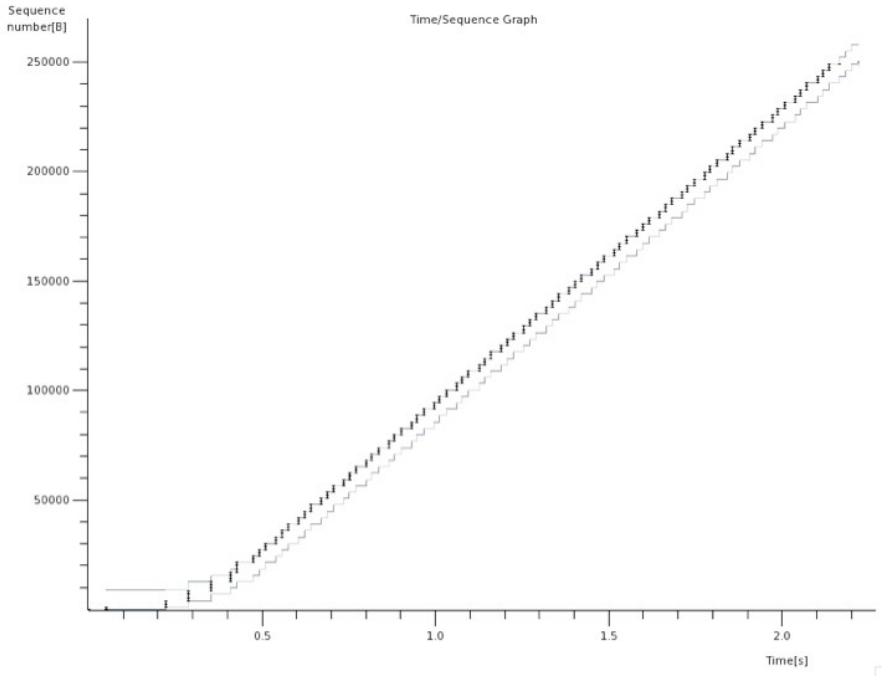
Seleccionemos por ejemplo el gráfico de rendimiento ("**Throughput**"). El gráfico de rendimiento muestra el rendimiento medio en diferentes instantes de tiempo. Por ejemplo, en el siguiente gráfico de ejemplo puede verse que hay una buena parte de la conexión oscila entre los 100 y los 300 Mbits/s, aunque también hay una parte por encima de 300Mbits/s.



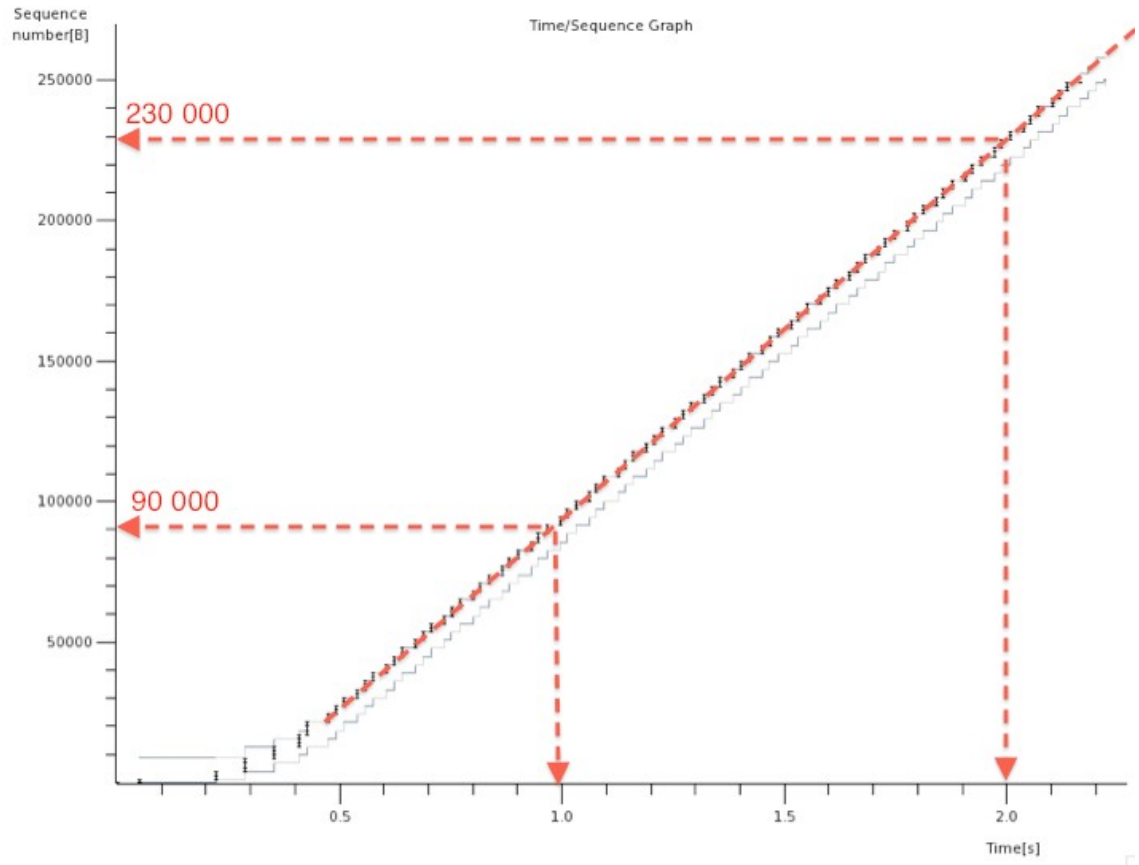
Ahora nos centraremos en las gráficos de tipo "**Time/Sequence (tcptrace)**". Estas gráficas son muy útiles ya que muestran mucha información sobre la conexión TCP al mismo tiempo.



Este tipo de gráfica nos va a permitir hacer una aproximación gráfica del rendimiento global de la conexión. Supongamos por ejemplo la siguiente conexión:



Para calcular rápidamente el rendimiento de una conexión podemos hacer las medidas que se muestran en la siguiente figura. El rendimiento de la conexión es la “velocidad” con la que aumenta el número de secuencia con respecto al tiempo. En este caso, como tenemos una línea prácticamente recta, podemos elegir dos puntos y calcular la pendiente entre ellos:



Los puntos que tenemos son:

Punto 1: X_1 = 1 segundo; Y_2 = 90.000 bytes
Punto 2: X_2 = 2 segundos; Y_2 = 230.000 bytes

Por lo tanto, el rendimiento de esta conexión es:

Rendimiento = Pendiente entre Punto 1 y Punto 2 = (Y_2 – Y_1)/(X_2 – X_1) =
= (230.000 – 90.000)/(2-1) = 140.000 bytes/segundo = 1.120.000 bits/segundo = **1,12 Mbps**

Ejercicios para clase

0. Usando el fichero “CapturaInicial”, seguir los pasos de la guía y documentarlo con capturas (los valores no tienen por qué coincidir con los indicados en esta guía).

1. Usando el fichero “Captura07”, analiza el rendimiento de la conexión que hay entre el nodo 128.3.164.249 (puerto 48805) y el nodo 128.3.38.201 (puerto SMTP). En concreto, se pide:

- Estimar el rendimiento “a ojo” con la gráfica “Thoughput”
- Estimar el rendimiento gráficamente usando la gráfica “Time/Sequence”
- Calcular el rendimiento de forma exacta

Dar todos los resultados en bits por segundo (bps, Kbps, Mbps, etc.), y no en bytes por segundo (Bps, KBps, MBps, etc.)

2. Selecciona un paquete de datos del flujo TCP y encuentra su correspondiente representación en la gráfica de tipo “**Time/Sequence (tcptrace)**” (tendrás que hacer mucho zoom). ¿Qué número de secuencia tiene este paquete? ¿Qué longitud? Confirma que esta información se corresponde con la gráfica.

3. La velocidad de conexión crecía muy rápido al principio, pero poco después del instante t=0.1 segundos, se mantiene constante. Analiza porqué sucede esto y explícalo.

4. Usaremos ahora el fichero de captura “tcp_stream_analysis.libpcap.pcap”. Analiza la conexión en la que se envían más datos, en el sentido principal. Podrás ver que en general es una conexión muy estable, pero que sucede algo en la mitad. Analiza qué sucede, y calcula el rendimiento global de la conexión y el rendimiento máximo que consigue de forma mantenida.

5. Ahora compararemos tres conexiones TCP. Hemos descargado el mismo fichero, pero de tres repositorios diferentes, uno en un servidor nacional de Rediris (“rediris.pcapng”), otro de un servidor de la Junta (“andalucia.pcapng”) y otro de Vietnam (“vietnam.pcapng”). Estudia las diferencias e intenta explicarlas. Para el caso de la conexión con Vietnam, vemos que se reciben paquetes en desorden en mitad de la conexión ¿Cómo se representa en la gráfica “**Time/Sequence (tcptrace)**” la llegada de paquetes en desorden?