



DEPARTAMENTO
DE COMPUTACION

Facultad de Ciencias Exactas y Naturales - UBA

TP2: Rutas en Internet

Teoría de las Comunicaciones

Integrante	LU	Correo electrónico
Furman, Damián	936/11	damian.a.furman@gmail.com
Lambrisca, Santiago	274/10	santiagolambrisca@hotmail.com
Marottoli, Daniela	42/10	dani.marottoli@gmail.com
Vanecek, Juan	169-10	juann.vanecek@hotmail.com



Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires

Ciudad Universitaria - (Pabellón I/Planta Baja)

Intendente Güiraldes 2160 - C1428EGA

Ciudad Autónoma de Buenos Aires - Rep. Argentina

Tel/Fax: (54 11) 4576-3359

<http://www.fcen.uba.ar>

Contents

1	Introducción	3
2	Desarrollo	4
3	Resultados	5
3.1	Primera parte: selección de las universidades	5
3.2	Segunda Parte: Correr nuestro Traceroute	5
3.3	Segunda parte: búsqueda de enlaces transatlánticos	12
3.4	Mapas	14
4	Conclusiones	17

1 Introducción

Este trabajo práctico fue realizado con el objetivo de entender el protocolo ICMP y de implementar nuestra propia herramienta de traceroute utilizando este protocolo de control, para finalmente poder realizar un analisis sobre los resultados obtenidos que nos permita entender un poco mejor que es lo que realmente sucede.

En primer lugar se seleccionaron cuatro universidades situadas en diferentes continentes. Una de ellas es The University of British Columbia, ubicada en la ciudad de Vancouver en Canadá. La segunda es The University of Hong Kong, ubicada en Pokfulam, Hong Kong. Otra de las universidades seleccionadas es Lomonosov Moscow State University, ubicada en Moscú, Rusia. La ultima es The National University of Samoa.

La primera parte del trabajo consiste en implementar una herramienta que nos permita estimar el *RountTripTime* entre nuestro host y distintos hosts, calculando además el *valorstandard* de cada salto con respecto a la ruta global.

Luego, procedemos a gráficar y analizar los resultados obtenidos, intentando reconocer enlaces submarinos. Se tendra en cuenta para esto el *valorstandard* calculado por nuestra herramienta y se intentara encontrar un umbral que permita decidir de manera lo suficientemente acertada cuales son los enlaces submarinos.

2 Desarrollo

La implementación de nuestra herramienta consiste principalmente en una clase, la cual cuenta con los métodos necesarios para realizar el envío de paquetes *ICMP* a un host dado, recibir las respuestas, y calcular cuál es el tiempo entre el envío y la recepción de una respuesta, y estimar cuál fue la ruta de los paquetes.

Para poder estimar la ruta de los paquetes hasta el destino, la estrategia utilizada será incrementar el valor del campo *ttl* en 1, así empezando con un *ttl* = 1 iremos obteniendo una respuesta *ICMP* de tipo *timeexceeded* para cada salto. De esta manera, iremos guardando la información obtenida para cada valor del *ttl*, y rastreando la ruta de los paquetes.

Debido a que en cada medición los resultados obtenidos presentan variaciones, lo que hacemos es estimar los tiempos obtenidos realizando un promedio sobre cada tiempo. Por lo tanto, enviaremos más de un paquete por *ttl*, con el objetivo de evitar malas estimaciones causadas por circunstancias extraordinarias en un instante dado.

Por otro lado, utilizaremos herramientas de geolocalización, que nos permiten estimar la ubicación de los routers a través de los cuáles se forwardearon los paquetes y dibujar su recorrido en un mapa. Esto es de gran utilidad a la hora de corroborar qué saltos se corresponden con enlaces submarinos.

Nuestra herramienta nos permite correr el algoritmo para distintos destinos, cambiando la cantidad de paquetes por *ttl*, el umbral y también la herramienta de geolocalización utilizada, para así poder tener una mayor cantidad de posibilidades a la hora de experimentar. Además se implementó una herramienta de control de ejecución, la cual nos permite correr el algoritmo cada cierto intervalo de tiempo, y tener resultados para distintos momentos del día. Nosotros utilizaremos intervalos de 1 hora.

Los resultados son almacenados en archivos de texto, para poder contar con la información a la hora de graficar y analizar los resultados.

3 Resultados

A continuación presentaremos los distintos gráficos y análisis realizados, en relación a las 4 universidades elegidas. Merece nombrarse, que se había elegido una universidad más, llamada University of Pretoria.

En cuanto a esta última universidad nombrada, nos encontramos con que no pudimos rastrear su ruta. Lo que obtuvimos en nuestros intentos, fue siempre el mismo resultado. Incrementamos el *tll* hasta obtener un *ubyte* overflow. El valor de nuestro *tll* se fue del rango $0 < tll < 255$. Por este motivo no utilizamos este caso de estudio en los gráficos y análisis.

3.1 Selección de las universidades

Las 4 universidades de distintas partes del mundo seleccionadas para llevar a cabo el análisis requerido son las siguientes:

- **The University British of Columbia**
Distancia: 11302.14 km
IP: 137.82.130.49 (www.ubc.ca)
- **Lomonosov Moscow State University**
Distancia: 13481.01 km
IP: 93.180.0.18 (www.msu.ru)
- **The Chinese University of Hong Kong**
Distancia: 18511.04 km
IP: 137.189.11.73 (www.cuhk.edu.hk)
- **National University Of Samoa**
Distancia: 10771.06 km
IP: 23.229.137.67 (www.nus.edu.ws/)

Las distancias descriptas corresponden a la distancia lineal que hay entre la universidad y un punto en común en Buenos Aires.

3.2 Primera Parte: Correr nuestro Traceroute

Se ejecutó el traceroute implementado por nosotros utilizando Scapy a lo largo de un día, con intervalos de una hora. El traceroute envía 20 paquetes a cada hop, para lograr un promedio más ajustado del RTT.

Evaluamos distintas API's para calcular las coordenadas y/o la localización de las IP's con el objetivo de que nuestros datos sean lo más precisos posibles.

Los resultados obtenidos se muestran en tres gráficos distintos por cada universidad, cada uno representando una corrida diferente con una API distinta. En cada gráfico hay una línea por cada hora en la que se ejecutó el traceroute marcando los distintos valores de los RTT's en función de los TTL's. A continuación se muestran los gráficos:

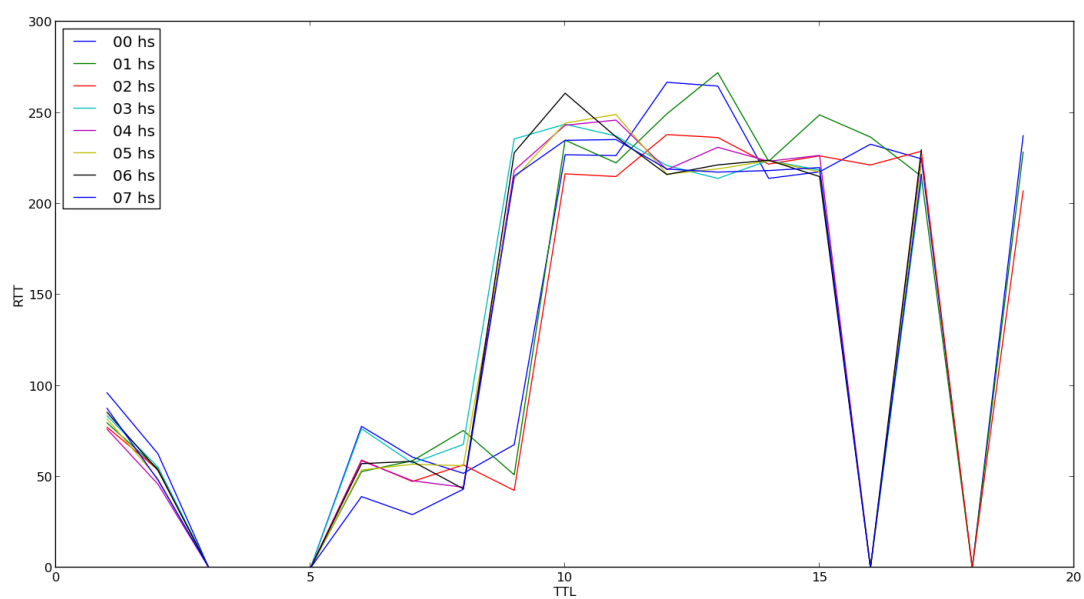


Figure 1: Canada Muestra 1

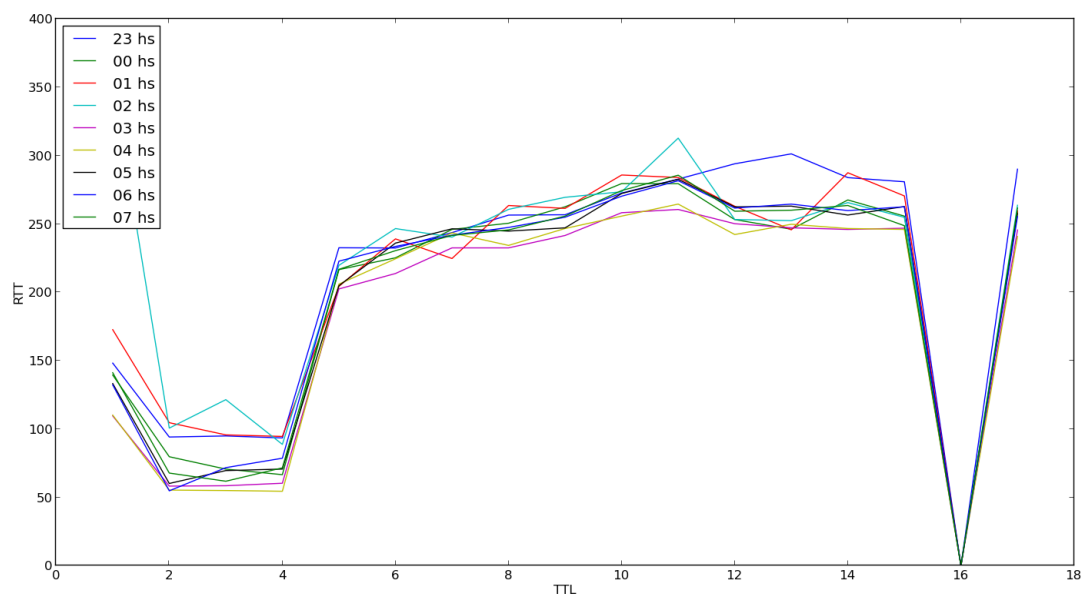


Figure 2: Canada Muestra 2

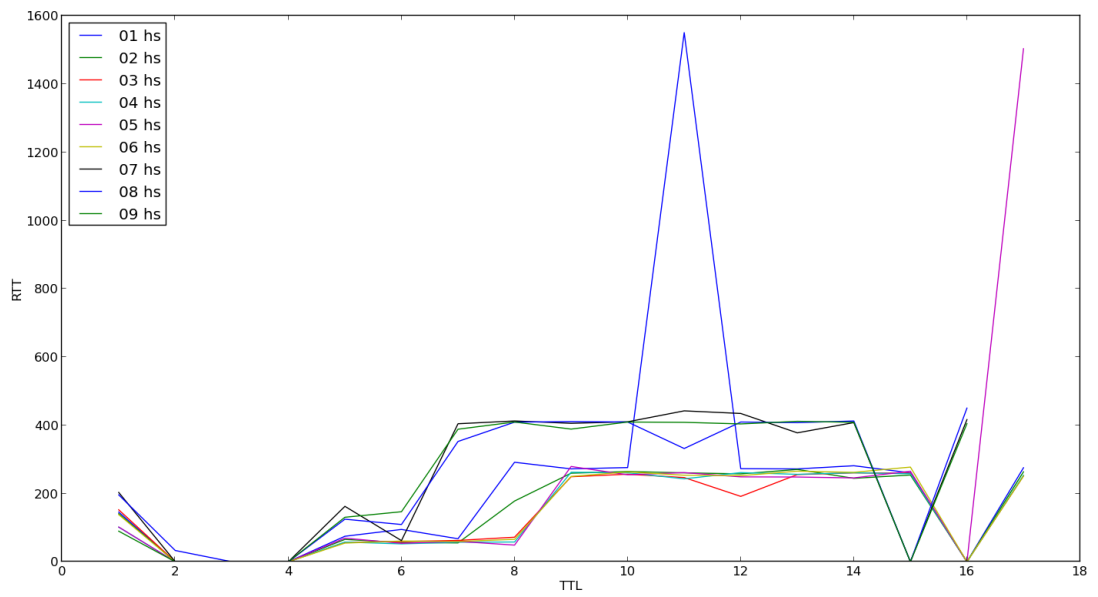


Figure 3: Canada Muestra 3

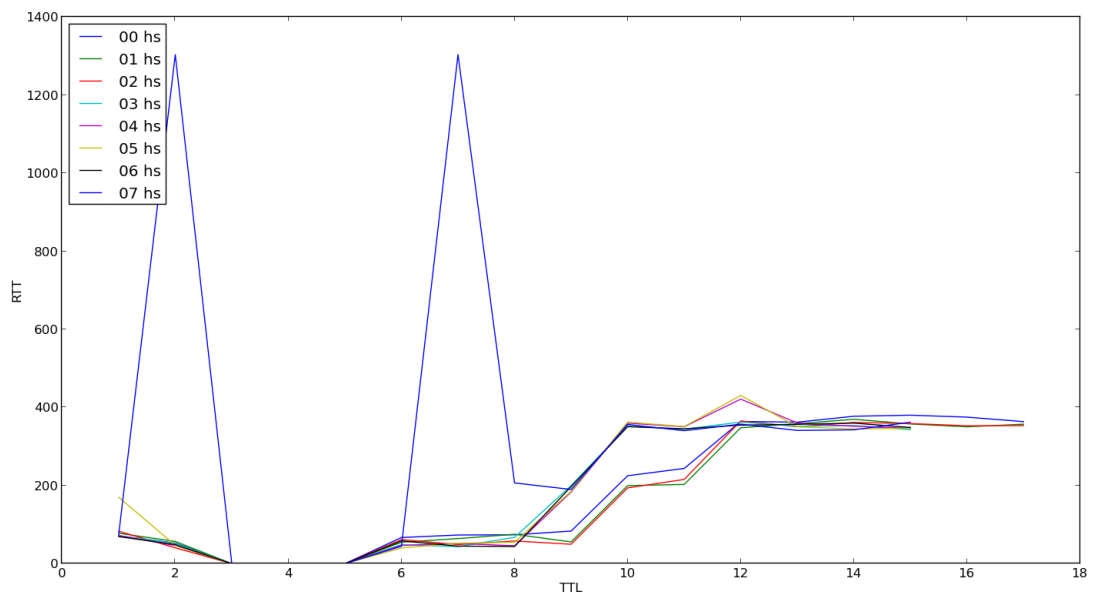


Figure 4: Canada Muestra 1

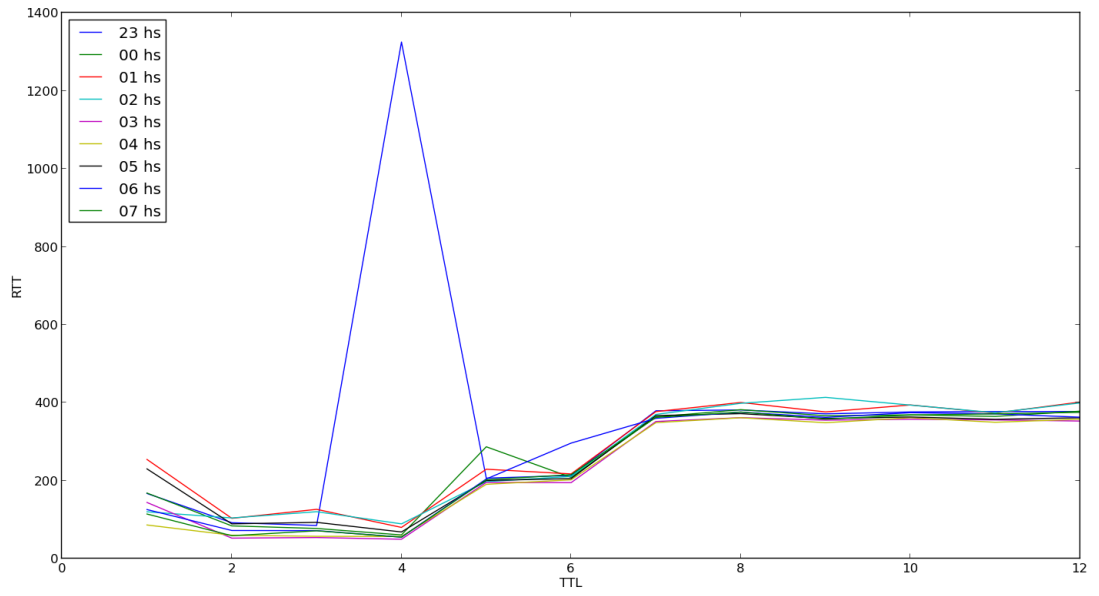


Figure 5: Canada Muestra 2

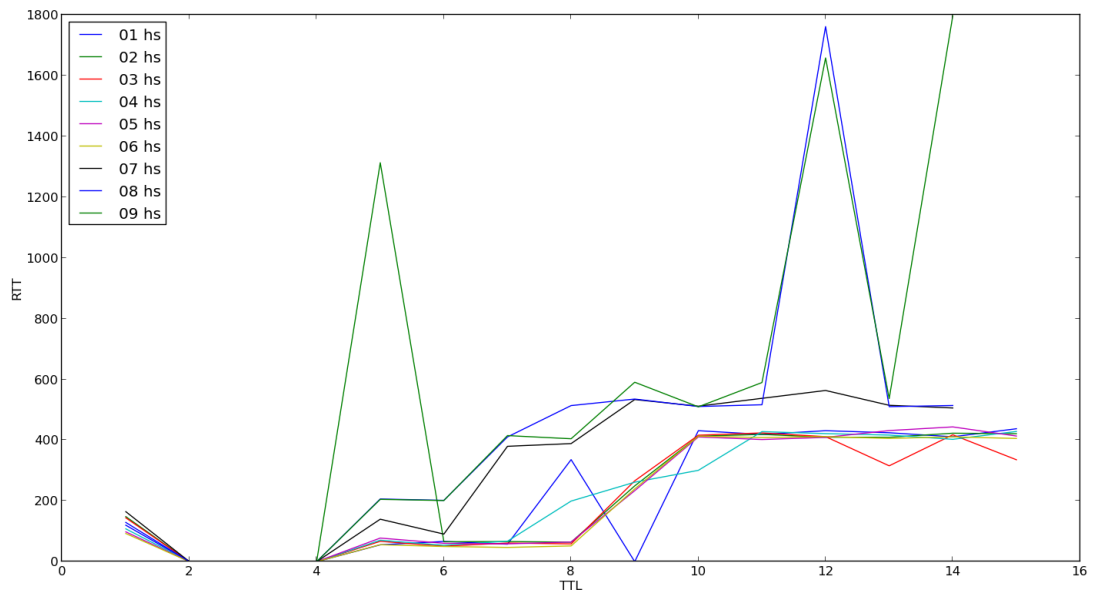


Figure 6: Canada Muestra 3

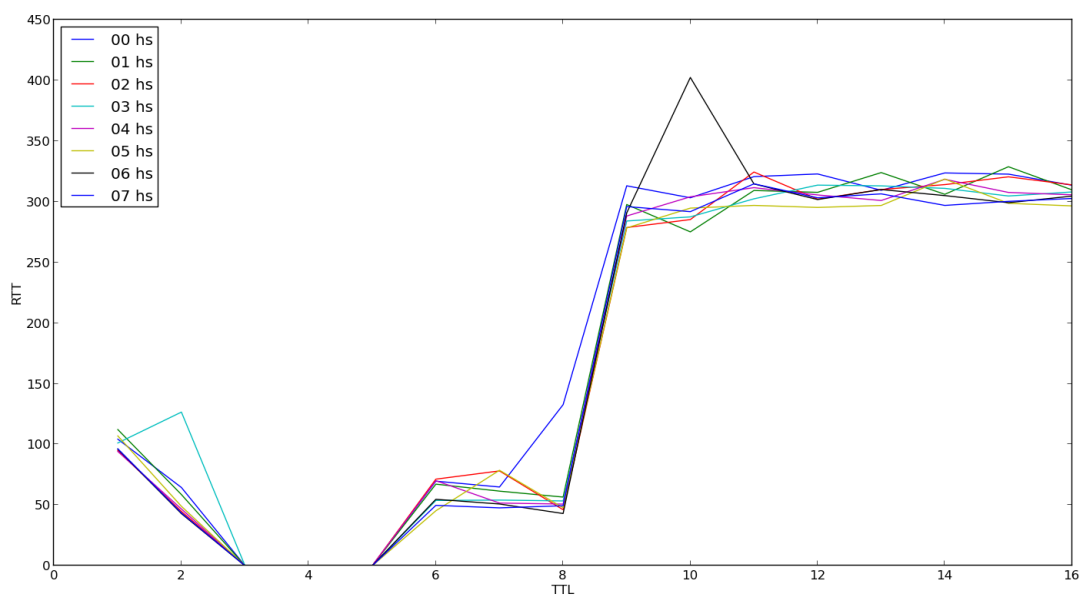


Figure 7: Canada Muestra 1

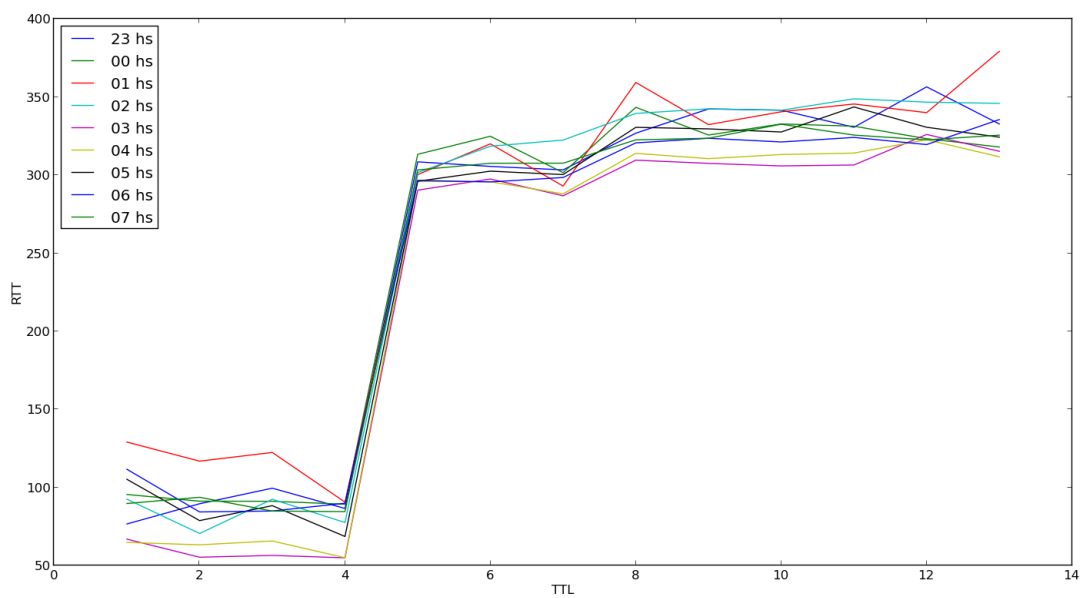


Figure 8: Canada Muestra 2

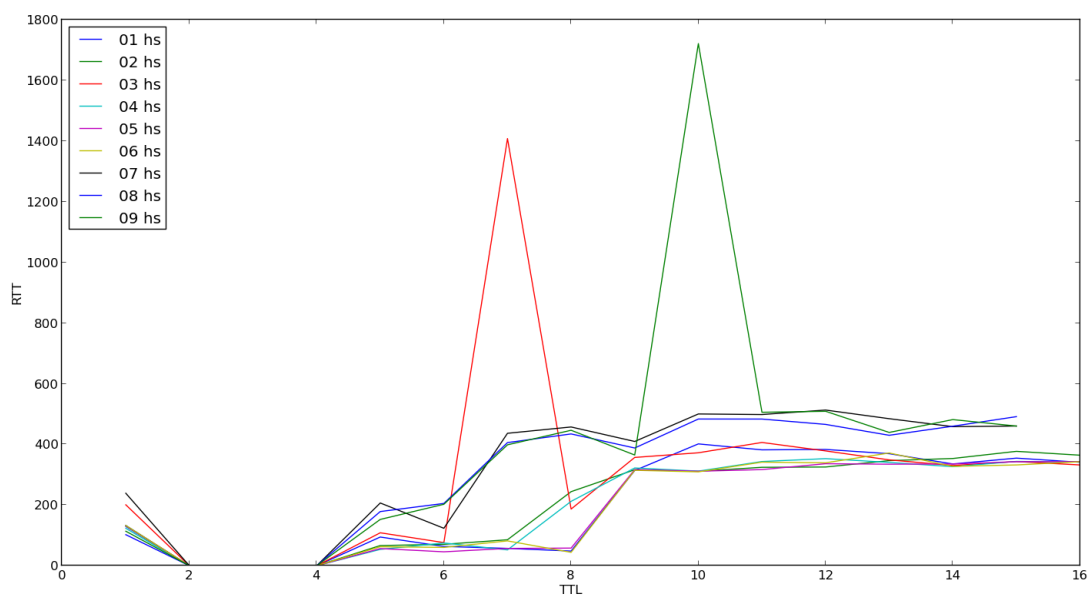


Figure 9: Canada Muestra 3

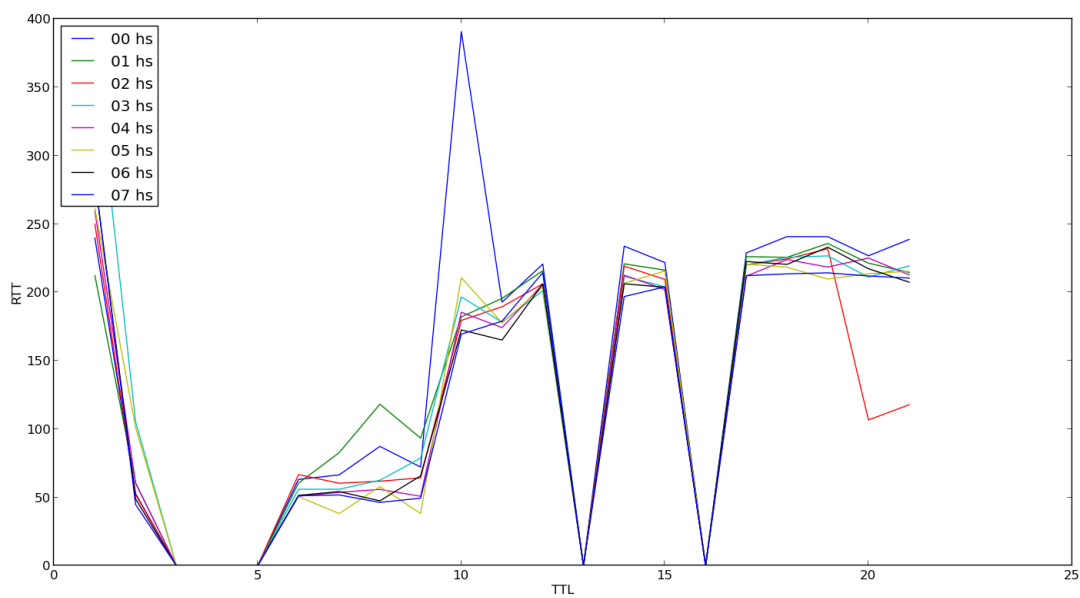


Figure 10: Canada Muestra 1

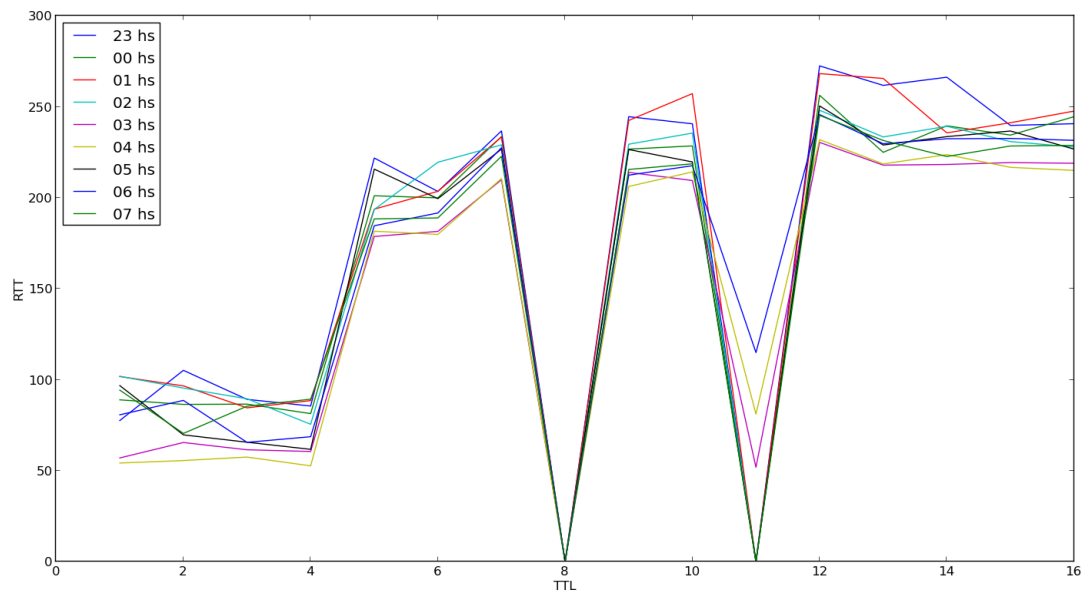


Figure 11: Canada Muestra 2

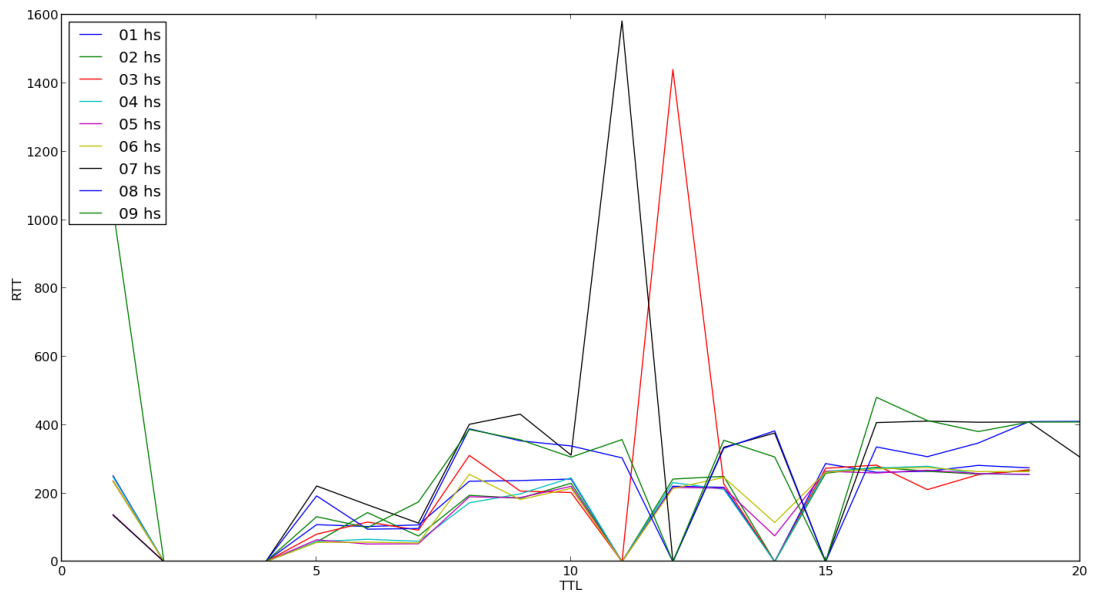


Figure 12: Canada Muestra 3

Para todas las API's se puede observar un aumento considerable del RTT entre dos TTL's específicas (distintas según cada gráfico) que representa una transmisión de gran escala.

Por ejemplo, en el gráfico 2, el RTT tiene un aumento vertiginoso entre el TTL 4 y el 5, que es cuando los paquetes llegan a Estados Unidos. Sucede algo similar en 1 sólo que se da entre el TTL 7 y el 8 en algunos casos y entre el 8 y el 9 en otros. También en ambos casos según cada corrida del traceroute coincide con el momento en que se envían paquetes a Estados Unidos. En 3 se da una situación similar entre los TTL's 6 y 7 y 8 y 9 también

coincidiendo con un salto de los paquetes entre Argentina y Estados Unidos

Con respecto a las pruebas realizadas para China, en el gráfico 5 se observan dos saltos claros entre los TLL's 4 y 5 y 6 y 7 que coinciden con dos saltos entre distintas regiones de Estados Unidos. Curiosamente, el salto entre Estados Unidos y Hong Kong no presenta un RTT mucho mayor que aquel que iba a dos regiones distintas de Estados Unidos. Se puede observar algo similar en el gráfico ??y 6 en el mismo salto (que tiene como destino una IP perteneciente a la ciudad de Miami) sólo que en estos casos es entre las TLL's 9 y 10.

La situación respecto a los gráficos de las pruebas realizadas para la universidad de Moscú presentan una anomalía similar a la de la universidad de Hong Kong. En todos se observa un salto más pronunciado que el resto (entre las TLL's 8 y 9 para 7, 9 y 10 para 9) y 4 y 5 para 8) que coincide con dos regiones de los Estados Unidos (las mismas para todas las API's). Interpretamos que esta situación puede ser causada porque si bien la IP pertenece a las direcciones IP's de los Estados Unidos el servidor puede encontrarse en verdad en una zona geográfica más lejana.

Finalmente, en Samoa se da una situación particular: la dirección IP final del dominio de la Universidad pertenece a los Estados Unidos. Producto de una investigación y del análisis del mapa, pudimos establecer una teoría acerca de esta anomalía. A saber, que las islas pertenecientes al archipiélago en el cuál se encuentra Samoa pertenecen a los Estados Unidos. Creemos que una posible explicación a este fenómeno es que Samoa no tiene ISP y utiliza los de las colonias vecinas.

Además, varios gráficos presentan picos de RTT injustificados para una hora determinada. Este puede ser el caso de los gráficos 9 y 3, y consideramos estos picos como producto de alguna anomalía, dado que sólo ocurrieron una única vez en todas las pruebas. Por otro lado, los paquetes que no obtuvieron respuesta se pueden apreciar como un pico decreciente que llega hasta el valor 0 (que es como se representa a los paquetes que nunca retornaron).

3.3 Segunda parte: búsqueda de enlaces transatlánticos

Es esperable que los enlaces transatlánticos tengan la mayor diferencia de rtt entre salto y salto, ya que son lo que probablemente tengan una mayor distancia entre ellos. Asíque en esta sección vamos a tratar de ubicar a estos hops distinguidos. El enunciado nos propone que utilicemos el z-score para tratar de encontrarlos.

El $ZRTT$ se define de la siguiente manera:

$$ZRTT_i = \frac{RTT_i - \overline{RTT}}{SRTT} \quad (1)$$

siendo \overline{RTT} y $SRTT$ el promedio y el desvío standard de los RTTs de la ruta, respectivamente, y RTT_i al RTT medido para el salto i .

Luego nos piden encontrar un umbral, que sea una cota inferior para los zrtt de los enlaces transatlánticos.

A continuación mostraremos los gráficos obtenidos para las universidades, a partir de una experimentación en la que calculamos el rtt y zrtt y los comparamos en un mismo gráfico:

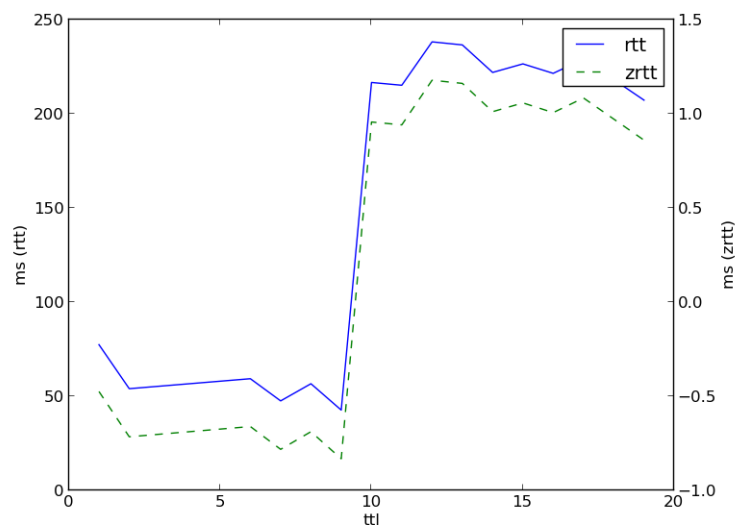


Figure 13: Canada

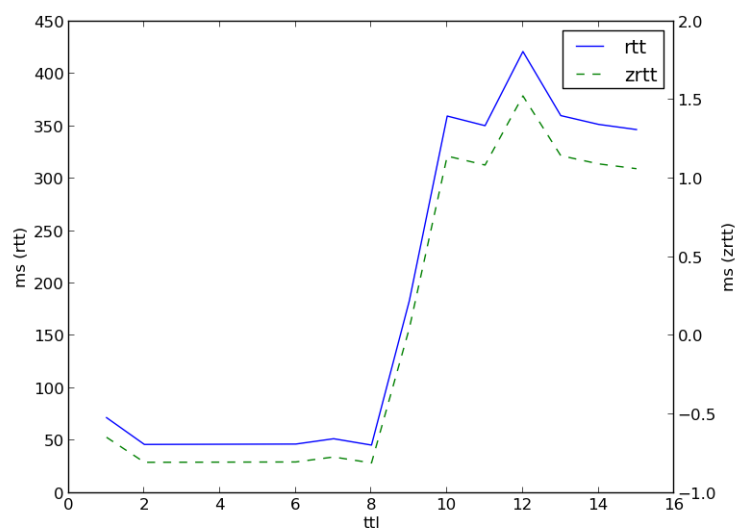


Figure 14: China

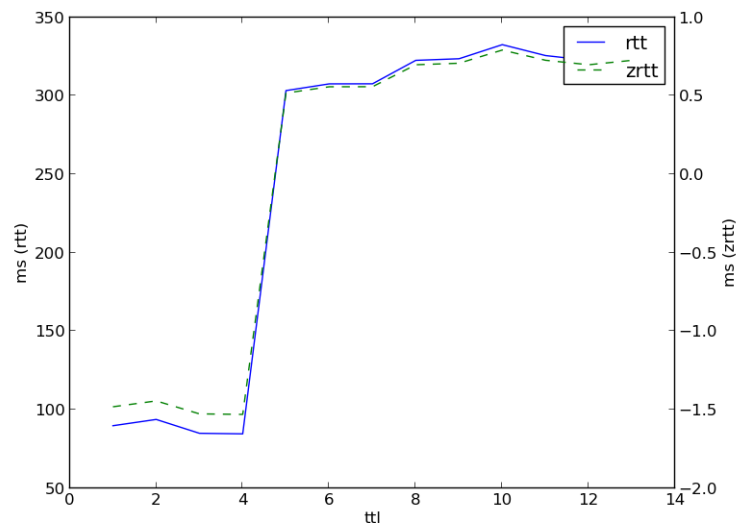


Figure 15: Rusia

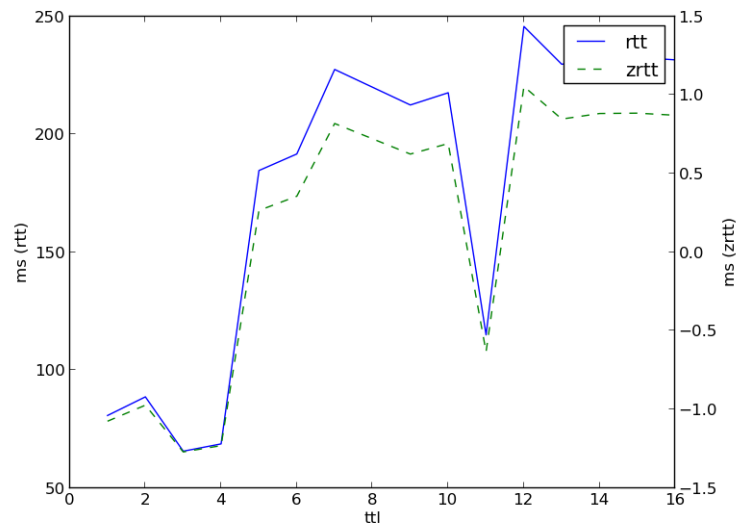


Figure 16: Samoa

De los mismos, podemos ver que el zrtt funciona como una normalización de los rtt tomados, y que en todos se cumple que los saltos transatlánticos tienen un z-score mayor a 0.5, siendo este el umbral que nos piden.

3.4 Mapas

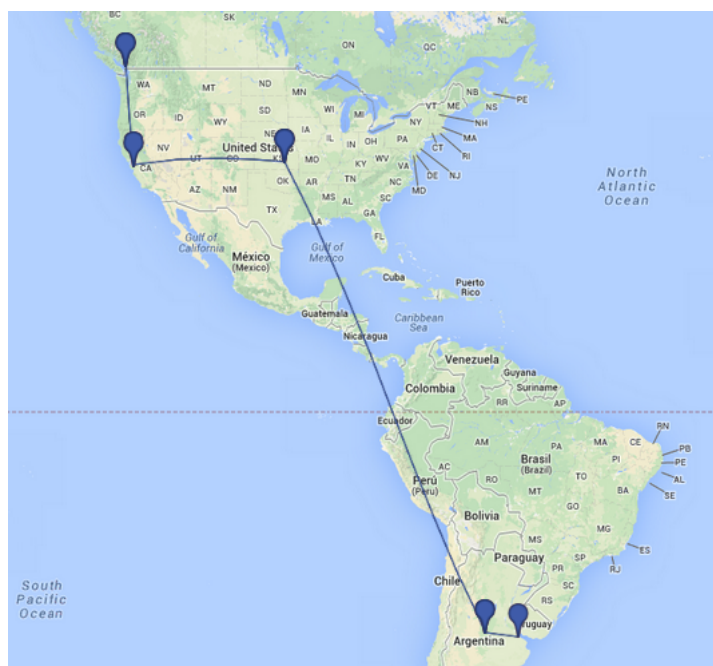


Figure 17: Canada

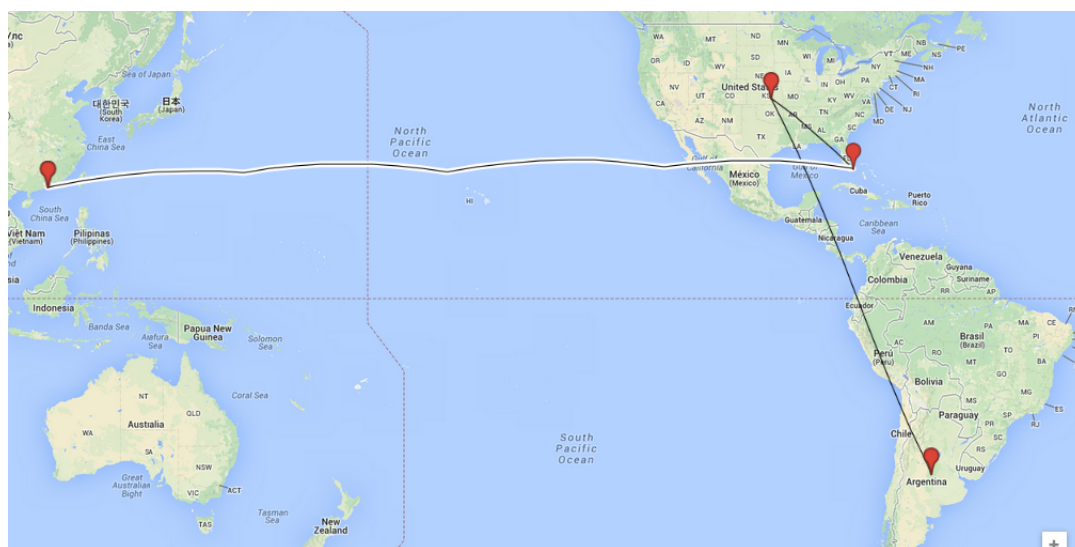


Figure 18: China



Figure 19: Rusia

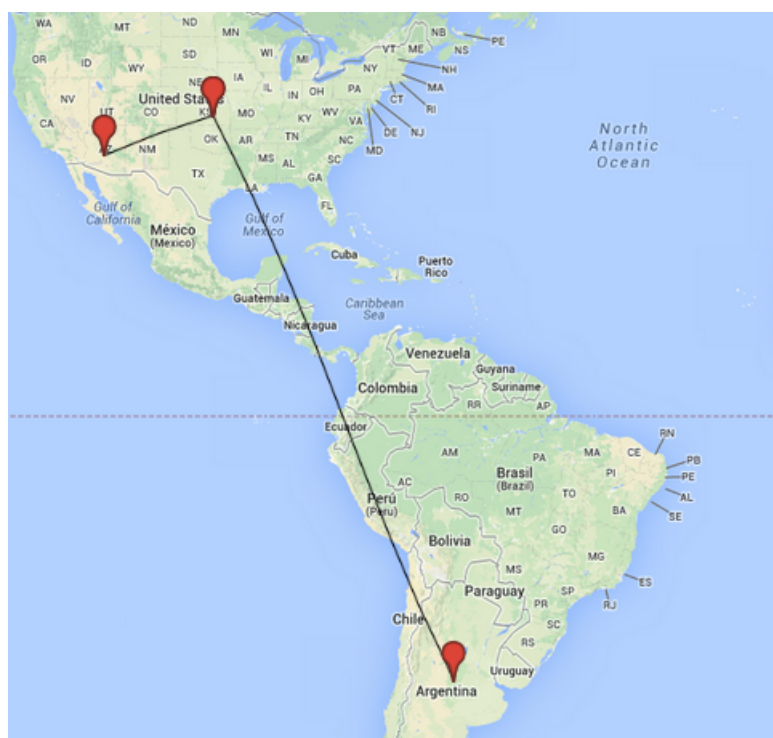


Figure 20: Samoa

4 Conclusiones

Luego de realizadas las observaciones sobre una gran cantidad de resultados, podemos corroborar, como es de esperarse, que en general los mayores tiempos se producen en los enlaces submarinos.

Pero decir esto, no es realmente pensar en lo que pudimos observar. Si bien nos aseguramos en líneas generales que eso ocurre, nos encontramos con algunos resultados que no estaban dentro de lo esperado. Sin dudas, esto se debe a la naturaleza de Internet, por ser una red tan grande y diversa, podemos encontrar grandes delays provocados por congestión, podemos encontrarnos también con nodos que no van a darnos respuestas, ya sea por ser obsoletos o por su configuración. También, vimos como las rutas de los paquetes cambian, no son algo estático.

Sin embargo, pudimos notar, que aún realizando los experimentos desde distintos lugares, con distintos proveedores de Internet, y en distintos horarios, las rutas seguidas por los paquetes, son geográficamente similares. Esto se debe, en caso de estar enviando paquetes hacia otro continente, a que no existen muchos enlaces submarinos, sino que para poder llegar al otro lado del océano, nuestros paquetes atravesaron seguramente un mismo enlace, o conjunto de enlaces.