

# Influencia del protocolo de control de transmisión (TCP) en el comportamiento autosimilar del tráfico en redes convergentes

“Tesis presentada para obtener el grado de Magister en Redes de Datos”



Autor: Emanuel Adrián Arias

Director: Luis Armando Marrone

Facultad de Informática – Universidad Nacional de La Plata

Mes y año: Mayo 2020

# RESUMEN

---

El trabajo aborda la problemática que presentan los modelos estocásticos clásicos debido al comportamiento autosimilar del tráfico existente en las redes convergentes. Las actividades realizadas buscan determinar la influencia de TCP en mencionado comportamiento.

El desarrollo presenta un análisis teórico y práctico detallado de la conducta estocástica del tráfico frente a escenarios de red estresada y relajada.

En particular, se propone tomar muestras de tráfico TCP y observar si existe relación entre el mecanismo de control de congestión de dicho protocolo con el grado de autosimilitud presente en el tráfico de una red convergente.

En ambos casos, red sin congestión y estresada, se iniciará el estudio a través de una simulación en un ambiente controlado y se contrastarán los resultados con muestras obtenidas de una red de datos operativa, concluyendo con una respuesta al objetivo planteado.

Esta tesis tiene como fin poder contribuir al proceso de investigación aportando conocimiento y evidencia que contribuyan a la creación de modelos más acertados, entendiendo que los avances tecnológicos implican cambios y dinamismo en los modelos científicos que representan un fenómeno determinado.

# ABSTRACT

---

This work deals with the problems presented by classic stochastic models due to the self-similar behavior of existing traffic in convergent networks. The activities carried out seek to determine the influence of TCP on said behavior.

The development presents a detailed theoretical and practical analysis of stochastic traffic behavior in stressed and relaxed network scenarios.

In particular, it is proposed to take samples of TCP traffic and observe whether there is a relationship between the congestion control mechanism of said protocol with the degree of self-similarity present in the traffic of a convergent network.

In both cases, a stressed network and a relaxed network, the study will start through a simulation in a controlled environment and the results will be compared with samples obtained from an operational data network, concluding with a response to the proposed objective.

The purpose of this thesis is to be able to contribute to the research process by providing knowledge and evidence that leads to the creation of more precise models, understanding that technological advances imply changes and dynamism in the scientific models that represent a given phenomenon.

1.	<u>INTRODUCCIÓN</u>	4
1.1	<u>ESTADO DEL ARTE</u>	6
1.2	<u>OBJETIVOS PLANTEADOS</u>	11
1.3	<u>METODOLOGÍA</u>	11
2.	<u>MARCO TEÓRICO</u>	12
2.1	<u>PROCESOS ESTOCÁSTICOS</u>	12
2.1.1	<u>ESTACIONARIEDAD</u>	12
2.1.2	<u>ERGODICIDAD</u>	13
2.1.3	<u>AUTOCORRELACIÓN</u>	14
2.1.4	<u>AUTOSIMILARIDAD</u>	15
2.1.5	<u>MODELOS CLASICOS</u>	16
2.1.6	<u>NUEVOS MODELOS</u>	16
2.2	<u>TRAFICO EN INTERNET</u>	16
2.2.1	<u>CONGESTIÓN EN REDES TCP/IP</u>	17
2.2.2	<u>PROTOCOLO DE CONTROL DE TRANSMISIÓN (TCP)</u>	18
2.2.2.1	<u>Descripción de TCP</u>	18
2.2.2.2	<u>Características de TCP</u>	18
2.2.2.3	<u>Control de flujo</u>	18
2.2.2.4	<u>Control de congestión</u>	19
2.2.2.6	<u>Equidad en TCP</u>	19
2.2.2.7	<u>Versiones de TCP</u>	19
2.2.3	<u>MECANISMOS DE CONTROL DE CONGESTIÓN DE TCP/IP</u>	20
2.2.3.1	<u>Algoritmo de arranque lento</u>	20
2.2.3.2	<u>Algoritmo de prevención de congestión</u>	20
2.2.3.3	<u>Algoritmo de retransmisión rápida y recuperación rápida</u>	21
2.2.3.4	<u>Detección temprana aleatoria</u>	21
2.2.3.5	<u>Notificación explícita de congestión</u>	21
2.3	<u>REDES CONVERGENTES</u>	22
2.3.1	<u>CONVERGENCIA DE LA RED EXPERIMENTAL</u>	23
3.	<u>TRABAJO EXPERIMENTAL</u>	24
3.1	<u>AMBIENTE CONTROLADO</u>	24

<a href="#">3.1.1</a>	<a href="#">CONDICION DE RED RELAJADA: ESCENARIO-1</a>	24
<a href="#">3.1.1.1</a>	<a href="#">Ámbito de realización</a>	24
<a href="#">3.1.1.2</a>	<a href="#">Calidad de servicio</a>	25
<a href="#">3.1.1.3</a>	<a href="#">Análisis de tráfico</a>	26
<a href="#">3.1.1.4</a>	<a href="#">Comportamiento de TCP</a>	28
<a href="#">3.1.1.5</a>	<a href="#">Recolección de las muestras</a>	30
<a href="#">3.1.1.6</a>	<a href="#">Autocorrelación</a>	31
<a href="#">3.1.1.7</a>	<a href="#">Autosimilaridad</a>	32
<a href="#">3.1.1.8</a>	<a href="#">Caracterización de las muestras</a>	35
<a href="#">3.1.1.9</a>	<a href="#">Interpretación de los resultados</a>	35
<a href="#">3.1.2</a>	<a href="#">CONDICION DE RED ESTRESADA: ESCENARIO-2</a>	36
<a href="#">3.1.2.1</a>	<a href="#">Ámbito de realización</a>	36
<a href="#">3.1.2.2</a>	<a href="#">Calidad de servicio</a>	36
<a href="#">3.1.2.3</a>	<a href="#">Análisis de tráfico</a>	36
<a href="#">3.1.2.4</a>	<a href="#">Comportamiento de TCP</a>	39
<a href="#">3.1.2.5</a>	<a href="#">Recolección de las muestras</a>	40
<a href="#">3.1.2.6</a>	<a href="#">Autocorrelación</a>	40
<a href="#">3.1.2.7</a>	<a href="#">Autosimilaridad</a>	42
<a href="#">3.1.2.8</a>	<a href="#">Caracterización de las muestras</a>	44
<a href="#">3.1.2.9</a>	<a href="#">Interpretación de los resultados</a>	52
<a href="#">3.2</a>	<a href="#">RED OPERATIVA</a>	53
<a href="#">3.2.1</a>	<a href="#">CONDICION DE RED RELAJADA: ESCENARIO-3</a>	54
<a href="#">3.2.1.1</a>	<a href="#">Ámbito de realización</a>	54
<a href="#">3.2.1.2</a>	<a href="#">Calidad de servicio</a>	54
<a href="#">3.2.1.3</a>	<a href="#">Análisis de tráfico</a>	57
<a href="#">3.2.1.4</a>	<a href="#">Comportamiento de TCP</a>	58
<a href="#">3.2.1.5</a>	<a href="#">Recolección de las muestras</a>	60
<a href="#">3.2.1.6</a>	<a href="#">Autocorrelación</a>	60
<a href="#">3.2.1.7</a>	<a href="#">Autosimilaridad</a>	61
<a href="#">3.2.1.8</a>	<a href="#">Caracterización de las muestras</a>	63
<a href="#">3.2.1.9</a>	<a href="#">Interpretación de los resultados</a>	70

<a href="#">3.2.2</a>	<a href="#">CONDICION DE RED ESTRESADA: ESCENARIO-4</a>	70
<a href="#">3.2.2.1</a>	<a href="#">Ámbito de realización</a>	70
<a href="#">3.2.2.2</a>	<a href="#">Calidad de servicio</a>	71
<a href="#">3.2.2.3</a>	<a href="#">Análisis de tráfico</a>	73
<a href="#">3.2.2.4</a>	<a href="#">Comportamiento de TCP</a>	73
<a href="#">3.2.2.5</a>	<a href="#">Recolección de las muestras</a>	75
<a href="#">3.2.2.6</a>	<a href="#">Autocorrelación</a>	75
<a href="#">3.2.2.7</a>	<a href="#">Autosimilaridad</a>	76
<a href="#">3.2.2.8</a>	<a href="#">Caracterización de las muestras</a>	78
<a href="#">3.2.2.9</a>	<a href="#">Interpretación de los resultados</a>	86
<a href="#">4.</a>	<a href="#">CONCLUSIÓN</a>	87
<a href="#">5.</a>	<a href="#">GLOSARIO</a>	89
<a href="#">6.</a>	<a href="#">BIBLIOGRAFÍA</a>	92

# 1. INTRODUCCIÓN

---

Las redes de telecomunicaciones tuvieron su inicio con Samuel F. B. Morse (Massachusetts, Estados Unidos 1791 - Nueva York, Estados Unidos 1872) a través de la creación del primer telégrafo en 1837 [1].

En 1854 Antonio S. G. Meucci (Florencia, Italia 1808 - Nueva York, Estados Unidos 1889) creó el primer teletrófono, posteriormente bautizado como teléfono [2].

En 1876 Alexander Graham Bell (Escocia, Reino Unido 1847 - Isla del Cabo Bretón, Canadá 1922) obtuvo la patente del teléfono en Estados Unidos, a pesar de que el aparato ya había sido desarrollado anteriormente por el italiano Antonio Meucci. La compañía que creó Bell para explotar la patente, Bell Telephone Company, fue la protagonista de los primeros pasos de la inserción del teléfono como medio de comunicación a escala internacional [3].

En 1877 Western Union, la mayor compañía de telégrafos de Estados Unidos, decidió comenzar a instalar teléfonos utilizando su ya desplegada red de telégrafos.

En enero de 1878 se pone en funcionamiento la primer Central Telefónica con 21 abonados, en New Haven, Connecticut, Estados Unidos [4]. En febrero del mismo año Western Union pone en funcionamiento otra Central Telefónica en San Francisco, Estados Unidos.

Los sucesos históricos citados, junto a otra serie de hitos, provocaron un crecimiento vertiginoso de las redes telefónicas y en consecuencia comenzaron a surgir los primeros problemas de desempeño en el funcionamiento de las mismas. Se origino la necesidad de disponer de modelos matemáticos para planificar las capacidades de los elementos del sistema, así como también predecir su rendimiento.

En 1909 Agner K. Erlang (Lønborg, Dinamarca 1878 - Copenhague, Dinamarca 1929) publico uno de sus trabajos más importantes denominado “La teoría de las probabilidades y las conversaciones telefónicas” [5], dando origen a un nuevo campo de estudio conocido como ingeniería del tráfico.

La ingeniería de tráfico se convirtió en una disciplina cuya finalidad es optimizar el desempeño de las redes, incrementando la eficiencia del uso de los recursos y minimizando la congestión.

Erlang demostró, a través de su trabajo, que la Distribución de Poisson es adecuada para modelar el tráfico telefónico aleatorio. Erlang consideró que los tiempos entre llegadas de las demandas de los usuarios eran independientes entre sí, así como también la cantidad misma de la demanda.

El tráfico se modela mediante procesos estocásticos que representan la demanda de los usuarios sobre los recursos de una red de telecomunicaciones. Por consiguiente, dichos modelos de tráfico permiten diseñar mecanismos para asegurar una determinada calidad de servicio a los usuarios.

En consecuencia, decimos que un modelo de tráfico solo puede considerarse correcto si las técnicas de inferencia estadísticas utilizadas sobre las trazas de tráfico real permiten concluir que dichas muestras de tráfico son consistentes con el modelo.

El éxito de la implementación del modelo de Poisson en redes telefónicas motivo su aplicación al tráfico de datos durante 1970, a pesar de que las predicciones del modelo no resultaban tan exactas. Hasta principios de 1980, modelar el tráfico de datos mediante procesos de Poisson era ampliamente aceptado.

Sin embargo, con la digitalización de las redes y el aumento de la capacidad de transmisión de los enlaces, surgieron nuevos tipos de servicios tales como el envío de imágenes y videos que se integraron a las redes de datos y voz. A causa de esto se vio la necesidad de incluir los efectos de la correlación existente entre las muestras, para lo cual se desarrollaron modelos más elaborados en los que la correlación decae exponencialmente con el tiempo (procesos con dependencia a corto plazo o por sus siglas en inglés short range dependent).

Desde entonces, se ha hecho uso de dichas premisas para gestionar el tráfico en los sistemas de telecomunicaciones.

A la familia de modelos cuya principal característica es la independencia o dependencia a corto plazo entre las muestras, se la conoce como Modelos Clásicos.

A inicios de 1990, con la integración del modelo OSI a la arquitectura de internet y la aparición de la Web, se inicio el auge que actualmente le conocemos a internet.

El concepto tradicional de disponer de redes separadas para soportar los distintos servicios de voz, datos y video quedo obsoleto. Los avances tecnológicos han permitido consolidar estas redes diferentes en una única plataforma denominada redes convergentes o redes multiservicio.

Muchos resultados predichos en las redes multiservicio, que consideran que la naturaleza del tráfico de datos responde a los modelos clásicos, difieren significativamente del rendimiento observado en la realidad. Es por esto después de 1990 se comenzaron a publicar documentos referidos a un nuevo paradigma, donde la correlación entre las muestras no decae rápidamente sino que lo hace de manera hiperbólica con el tiempo (procesos con dependencia a largo plazo o por sus siglas en ingles long range dependent).

Este último fenómeno, que afecta significativamente el desempeño de las redes modernas de comunicaciones, se puede representar adecuadamente mediante modelos de tráfico autosimilar.

En relación a lo expuesto, se pretende contribuir con la investigación de modelado autosimilares aportando conocimiento y evidencia que contribuyan a la creación de modelos más acertados, entendiendo que los avances tecnológicos implican cambios y dinamismo en los paradigmas científicos que representan un fenómeno determinado.

## **1.1 ESTADO DEL ARTE**

El presente trabajo relaciona dos temas de alto impacto en las redes convergentes, la arquitectura TCP/IP y el tráfico autosimilar.

El conjunto de protocolos TCP/IP cumple un rol determinante en las redes de telecomunicaciones, y si bien fue creado a principio de 1970, se consolido en 1990 a partir de su adopción como protocolo universal para conectar diferentes sistemas a través de internet.

El tráfico autosimilar aparece en las redes de telecomunicaciones a medida que se van unificando los servicios de datos, voz y video en una misma red. Es decir, el tráfico autosimilar se evidencia con el crecimiento de internet.

Al ser TCP/IP el conjunto de protocolos preestablecido para las comunicaciones a través de internet, gran parte de los trabajos realizados sobre autosimilaridad están basados en muestras provenientes de redes TCP/IP. En este punto se destacan aquellos trabajos que aportan a la causa de la presente tesis.

Se citan, de forma cronológica, algunas investigaciones que contribuyen al avance científico de la materia, resumiendo el aporte de cada uno de los trabajos:

[6]. (1994). *On the Self-Similar Nature of Ethernet Traffic*. Bellcore, Morristown, NJ, USA.

Es el primer artículo que reporto tráfico autosimilar en trazas reales sobre redes Ethernet. Refuto la idea de que las redes modernas podían ser analizadas mediante un simple análisis de tráfico basado en un modelo de Poisson.

[7]. (1998). *Introducción a los modelos de tráfico para las redes de banda ancha*. España.

El artículo realiza una revisión de los trabajos realizados hasta la fecha y constata que hay aspectos todavía abiertos, aparecen lagunas y surgen muchas preguntas:

- ¿Es realmente autosimilar el tráfico de las redes de banda ancha? Parece una buena aproximación, pero podría ser que el tráfico fuera simplemente no estacionario, y no necesariamente autosimilar (que es uno de los casos particulares de no estacionariedad).
- ¿La autosimilaridad es debida a las fuentes o a su agregación? Depende del servicio. Por ejemplo, el tráfico LAN es autosimilar sólo cuando hay agregación, mientras que en el caso del vídeo VBR el efecto aparece en el nivel de fuentes individuales.
- ¿Hasta qué punto es la agregación y la propia dinámica de la red la que provoca la autosimilaridad? Todavía está por estudiar a fondo la influencia de los protocolos en el comportamiento del tráfico.
- ¿Qué papel juega la presencia residual de dependencia a corto plazo o en inglés short range dependent (SRD) en las fuentes de vídeo VBR? ¿Cómo incluir sus efectos? ¿Es suficiente con los modelos F-ARIMA, o es necesaria una nueva familia de procesos que permitan una estimación de parámetros más fácil?
- Si el tráfico es realmente autosimilar, ¿son los modelos propuestos una buena solución para modelarlo? Ciertamente, son mejores que los modelos poissonianos, y proporcionan una explicación física del fenómeno (al menos en el caso del tráfico Ethernet, a través de la

agregación de terminales que se comportan según modelos on-off con distribuciones de Pareto). Pero incluso los modelos autosimilares no son capaces de capturar al 100% el comportamiento de las trazas reales, por la presencia de SRD y periodicidades (como el caso de las fuentes MPEG).

- ¿Son realmente importantes los efectos autosimilares en la práctica? Es un tema abierto, y se ofrecen conclusiones contradictorias. No está claro que a las escalas temporales en que trabajan los buffers aparezca LRD, aunque se habla de importantes errores de subestimación. Stallings [Sta98] ha propuesto el estudio a dos niveles: aplicación (auto similitud inherente a la fuente) y red (influencia de la dinámica de los protocolos como el TCP). A nivel de aplicación, la autosimilaridad tendría influencia en el control de acceso y la asignación de recursos de transmisión, mientras que a nivel de red, influiría en la congestión, el dimensionado de los buffers, etc. Tampoco queda claro qué influencia tiene cada nivel.

Las preguntas citadas muestran que, pese a que parece bastante claro que los modelos autosimilares son prometedores, hace falta mucho más trabajo en este campo para llegar a conclusiones sólidas y útiles.

[8]. (2001). *Introducción al tráfico autosimilar en redes de comunicaciones*. Colombia.

Concluyó que los modelos tradicionales de tráfico son insuficientes para explicar muchas de la característica observadas en el tráfico de las redes modernas de comunicaciones, especialmente en cuanto a las estructuras complejas de correlación que se extienden a muchas escalas. Los modelos de tráfico autosimilar capturan los efectos de estas estructuras de correlación. Finalmente mencionó muy brevemente cómo la transformada wavelet se convierte en una herramienta muy útil no sólo para detección, estimación y síntesis sino para el modelamiento mismo de tráfico.

[9] (2010). *Análisis del comportamiento autosimilar del tráfico ethernet en las redes de datos*. Mendoza, Argentina.

El trabajo verifica, a través de un estudio experimental y usando la prueba de Kolmogorov-Smirnov, que el tráfico de datos ethernet responde efectivamente a la Distribución de Pareto.

[10] (2010). *Estudio del tráfico autosimilar orientado a la simulación mediante la utilización de Wavelets u Onditas*. La Plata, Argentina.

La tesis desarrollada brinda un marco o estructura didáctica que permite el desarrollo del tema como una unidad de estudio y aprendizaje. También aporta una colección de scripts o aplicaciones, para abordar el estudio de distintas características y/o propiedades en el contexto de la autosimilaridad, y de esta manera, permite el análisis, detección de comportamientos, estimación de parámetros, y la producción de tráfico sintético, pudiendo utilizarse en el ámbito de la simulación.

[11] (2010). *Desarrollo de un modelo generador de tráfico autosimilar para la evaluación de los algoritmos de asignación de ancho de banda en WIMAX. La Plata, Argentina.*

A partir de la utilización de Simulink de Matlab, demuestra que los tráficos de la mayoría de las estaciones, consideradas individualmente, sean enviados o recibido, presentan un comportamiento no autosimilar. También corroboró que la autosimilaridad en el tráfico de red es un fenómeno robusto, natural y persistente. La superposición del tráfico de pocas estaciones, cuyos tráficos no son autosimilares, construye uno que si lo es.

[12] (2011). *Aplicaciones del modelo On/Off al tráfico agregado en las redes de comunicaciones. Colombia.*

Este artículo encuentra que el modelo de fuentes On/Off puede ser un modelo estacionario de rango corto o largo, esto con base en la implementación de la fuente, bien sea de una forma sencilla o multiplexada. Explica que el modelo On/Off es una herramienta útil al momento de estimar y predecir aproximadamente la aleatoriedad del comportamiento de diferentes tipos de tráfico que presenten características autosimilares de muestras independientes e idénticamente distribuidas (IID) y una distribución de cola pesada.

[13]. (2013). *Evaluación de desempeño de un modelo autosimilar para tráfico en redes 802.11. Colombia.*

Los autores del artículo alegan que la autosimilaridad para este tipo de tráfico fue cogiendo fuerzas a medida que el proceso estocástico se hacía extremadamente parecido a distintas escalas, pero tal y como se presenta en este artículo, esta técnica fue perdiendo popularidad debido al alto nivel matemático y probabilístico que existe detrás de la misma, además de que se hace un poco engorroso el proceso de determinar el grado de autosimilaridad obtenido a partir del parámetro de Hurst debido a

que hay que separar cada una de las tramas agregadas y hallar el parámetro de Hurst por separado para después poder unificarlo y obtener el grado de autosimilaridad total del tráfico agregado.

[14] (2012). *Estudio sobre la distribución de tráfico autosimilar en redes ethernet. Argentina.*

El trabajo relaciona los temas de autosimilaridad, con el tráfico ethernet, la distribución de Pareto y la prueba de Kolmogorov-Smirnov. Verifica a través de un estudio experimental y usando la prueba de Kolmogorov-Smirnov, que el tráfico de datos ethernet responde a una distribución de Pareto, estableciéndose los valores de los parámetros de dicha distribución. Demuestra que dos muestras ethernet, tomadas en distintos momentos, tienen distribución Pareto con parámetros prácticamente similares.

[15] (2017). *Impacto de la Dependencia de Rango Largo en el QoS del Tráfico de Video. Colombia.*

El trabajo analiza diferentes tráficos generados a partir de trazas de video. En todos los casos el parámetro de Hurst dio resultados mayores a 0.5, por lo que el autor infiere que el tráfico de video presenta un comportamiento autosimilar de dependencia de rango largo. La autosimilaridad la detecta a diferentes escalas, ya sea milisegundos a horas, por lo tanto la cantidad de valores que toma no afecta las diferentes simulaciones y cálculos. Afirma que los modelos correlacionados con LRD son más adecuados ya que presentan correlación en los modelos de tráfico, debido a que la llegada de paquetes se encuentra fuertemente correlacionada.

[16] (2018). *Predicting characteristics of self-similar traffic. International Conference on Information and Telecommunication Technologies and Radio Electronics. Ucrania.*

El trabajo aborda la problemática de predecir el tráfico autosimilar. Resulta que mediante el uso de métodos de extrapolación y aproximación de spline, se obtuvieron los resultados favorables. Esto permitirá proporcionar el volumen requerido de dispositivos de almacenamiento intermedio y, por lo tanto, evitar sobrecargas en la red.

Los trabajos citados demuestran el esfuerzo que han realizado diferentes investigadores sobre el tema de interés, y aportan evidencia suficiente para dar por sentado de que el tráfico presente en las redes multiservicio tiene una fuerte componente autosimilar que no puede ser obviada.

## 1.2 OBJETIVOS PLANTEADOS

El objetivo del presente trabajo de tesis es estudiar el comportamiento del protocolo de control de transmisión (TCP) con la intención de buscar patrones que influyan en la componente autosimilar de las redes multiservicio.

En particular, se tomarán muestras de tráfico TCP y se observa si existe relación entre el mecanismo de control de congestión de dicho protocolo con el grado de autosimilaridad presente en el tráfico de una red convergente.

Para analizar el escenario se tomarán muestras representativas de una red sin congestión, donde solo actuará el control de flujo, y muestras representativas de una red estresada con la finalidad de que se dispare el mecanismo de control de congestión de TCP. Obtenidas y caracterizadas las muestras, se analizará el comportamiento autosimilar de las mismas.

En ambos casos, red sin congestión y estresada, se iniciará el estudio en un ambiente controlado y se contrastarán los resultados con muestras obtenidas de una red de datos operativa, concluyendo con una respuesta al objetivo planteado.

## 1.3 METODOLOGÍA

La metodología a utilizar para desarrollar la tesis es la siguiente:

- Definir y acondicionar la red para cada escenario.
- Evaluar la fidelidad de las muestras.
- Recolectar y caracterizar diez (10) muestras por cada escenario.
- Analizar gráficamente la autocorrelación de las muestras con R.
- Analizar la autosimilaridad de las muestras calculando el coeficiente de Hurst con R.
- Analizar el comportamiento de las retransmisiones de cada muestra.
- Análisis de resultados y conclusión.

El trabajo se limita a un análisis práctico, focalizando en interpretaciones graficas y cálculos estadísticas a través de software.

## 2. MARCO TEÓRICO

---

### 2.1 PROCESOS ESTOCÁSTICOS

Las señales son función del tiempo y pueden ser clasificadas en dos grandes categorías: determinísticas y aleatorias.

Las señales determinísticas atraviesan una trayectoria predeterminada en el tiempo y en el espacio, y puede ser completamente descripta mediante una función del tiempo.

Las señales aleatorias no pueden ser caracterizadas mediante expresiones matemáticas; entonces se trata de establecer propiedades que satisfagan un conjunto de señales en lugar de analizar las señales individualmente. Se hace uso de herramientas de probabilidad y estadísticas para analizar su comportamiento.

Un proceso aleatorio es una familia o ensamble de señales que corresponden a cada posible resultado de la medición de cierta señal. Cada señal de este ensamble es denominado realización o función muestra del proceso.

Cuando las funciones muestras dependen del tiempo al proceso se lo denomina **Proceso Estocástico**.

En relación a las muestras asociadas al problema en cuestión, donde tenemos cantidad de paquetes por segundos, decimos que estamos frente a un proceso aleatorio discreto ya que toma valores de estado discretos observados a través de periodos continuos de tiempo.

#### 2.1.1 ESTACIONARIEDAD

Si bien las funciones miembros del proceso aleatorio pueden, individualmente, fluctuar en función del tiempo, ciertas propiedades de conjunto pueden permanecer constantes en el tiempo. En términos generales, se puede decir que un proceso aleatorio es estacionario si sus funciones de distribución o ciertos valores esperados permanecen invariantes respecto a una traslación del eje del tiempo.

Se dice que un proceso aleatorio es estacionario en el tiempo o estacionario en sentido estricto “SSS” (Strict Sense Stationarity) si todas las funciones de distribución que definen al proceso son invariantes ante una traslación en el tiempo [17].

Se dice que un proceso aleatorio es estacionario en sentido amplio o débilmente estacionario “WSS” (Wide Sense Stationarity) si su media es constante y su función de autocorrelación depende únicamente de la diferencia de tiempos [17].

## 2.1.2 ERGODICIDAD

Desde el punto de vista práctico, resulta muy atractiva la idea de poder hacer estimaciones a partir de datos registrados en una realización del proceso. En efecto; si se desea estimar la media  $\mu_x(t)$  del proceso aleatorio  $X(t)$ , el procedimiento usual es observar los valores de  $X(t)$  a través de cierto número de realizaciones y calcular la media de los mismos para este valor como estimación de la media del ensamble ya que, como sabemos, la media está definida como un promedio del ensamble.

En cambio, si únicamente se tiene acceso a una única realización del proceso, digamos  $x(t)$ , se puede obtener un promedio temporal y es este el que sería deseable poder usar como estimación de la media del proceso.

Si el proceso subyacente es WSS (débilmente estacionario), la  $\mu_x(t)$  es constante (es decir, independiente de  $t$ ) y la bondad de la estimación temporal dependerá de la forma en que la esperanza de la media temporal tienda a  $\mu_x(t)$  y la varianza de la media temporal tienda a cero cuanto  $t \rightarrow \infty$ . Así, si:

$$\lim_{t \rightarrow \infty} E\{X(t)_T\} = \mu_x \quad \& \quad \lim_{t \rightarrow \infty} \text{Var}\{X(t)_T\} = 0$$

Podemos concluir que la media temporal converge a la media del ensamble. En general, las medias temporales y las medias del proceso aleatorio no son iguales, excepto en una clase especial de procesos, llamados ergódicos.

En resumen, se llama ergódico a un proceso aleatorio estacionario  $X(t)$  si sus propiedades de conjunto igualan a las homólogas temporales. Esto significa que cualquier media del proceso  $X(t)$  puede ser obtenida a partir de una función miembro cualquiera de  $X(t)$ .

Lo cierto es que las condiciones de ergodicidad son de uso limitado en aplicaciones prácticas dado que su verificación requiere el conocimiento previo de parámetros que, usualmente, no están disponibles.

En la práctica se está forzado a considerar el origen físico del proceso aleatorio para extraer conclusiones sobre ergodicidad.

En síntesis, una justificación intuitiva acerca de la ergodicidad de un proceso, reside en decidir si una función miembro luce como una “verdadera señal ergódica” cuyas variaciones en el tiempo puedan considerarse que representan una variante típica de las del ensamble.

### 2.1.3 AUTOCORRELACIÓN

Los procesos aleatorios estacionarios se dividen en dependientes a corto y largo plazo (short range dependent, long range dependent). Los primeros presentan correlación solo en intervalos cortos de tiempo. En cambio los modelos con dependencia a largo plazo, presentan correlación incluso en intervalos de tiempo grandes.

Se dice que un proceso que presenta dependencia a corto plazo la autocorrelación es sumable:

$$\sum_{k=-\infty}^{\infty} \rho_k < \infty$$

Equivalentemente, la varianza decae proporcionalmente a  $m^{-1}$  ( $m$  es el tamaño de la muestra), la densidad espectral de potencia tiene un valor finito en cero, y el proceso tiende a comportarse como un ruido blanco a medida que  $m \rightarrow \infty$ . Los procesos con autocorrelación que decaen exponencialmente se denominan dependientes a corto plazo.

Se dice que un proceso muestra dependencia a largo plazo si su autocorrelación no es sumable:

$$\sum_{k=-\infty}^{\infty} \rho_k \rightarrow \infty$$

Equivalentemente, la varianza decae a ritmo menor que  $m^{-1}$  ( $m$  es el tamaño de la muestra), la densidad espectral de potencia tiende a una singularidad en cero y el proceso promedio tiene la misma estadística que el proceso inicial. Los procesos con autocorrelación que decaen hiperbólicamente se denominan dependientes a largo plazo.

## 2.1.4 AUTOSIMILARIDAD

El tráfico se modela mediante un proceso estocástico que representa la demanda que los usuarios de una red de comunicaciones imponen sobre los recursos de la red.

Originalmente se consideró que los tiempos entre llegadas de las demandas de los usuarios eran independientes entre sí, así como la cantidad misma de la demanda.

Posteriormente se vio la necesidad de incluir los efectos de la correlación existente entre estas variables, para lo cual se desarrollaron modelos más elaborados en los que la correlación decaía exponencialmente con el tiempo (procesos short range dependent).

Sin embargo, recientemente se ha evidenciado que, en las redes modernas de comunicaciones, la correlación entre estas variables no decae tan rápidamente sino que lo hace de manera hiperbólica con el tiempo (procesos long range dependent).

Este último fenómeno, que afecta significativamente el desempeño de las redes de comunicaciones, se puede representar adecuadamente mediante modelos de tráfico fractal o autosimilar.

Un concepto que tiene vital importancia en el estudio de los procesos autosimilares es el coeficiente de Hurst. Dicho parámetro representa una medida del grado de autosimilaridad presente en un proceso estocástico.

En consecuencia, al calcular el coeficiente de Hurst se concluye lo siguiente [8]:

- H es un parámetro que se encuentra en el intervalo abierto (0.5; 1).
- Si  $H < 0.5$  el estudio no se relaciona con la autosimilaridad y la serie se considera impredecible.
- Si  $H = 0.5$  el estudio muestra que existe independencia entre las muestras.
- Si  $0.5 < H < 1$  el estudio muestra que la serie presenta autosimilaridad estadística.
- Si  $H = 1$  el estudio muestra que la serie presenta autosimilaridad determinística (la aleatoriedad desaparece).

Para el trabajo de tesis la definición presente se considera suficiente y permitirá sacar conclusiones sobre los procesos que modelan las muestras en estudio.

## 2.1.5 MODELOS CLASICOS

Se denominan como clásicos a la familia de modelos cuya principal característica es que no contemplan la posibilidad de que las trazas de tráfico presenten dependencia a largo plazo.

Tradicionalmente, se ha asumido que en las redes de conmutación de paquetes se daban las condiciones necesarias para suponer que el tráfico sigue un proceso de Poisson o de Bernoulli. Por esto, los modelos más utilizados son los procesos on-off markovianos con interrupciones de Poisson o Bernoulli.

## 2.1.6 NUEVOS MODELOS

El nuevo enfoque sobre los modelos contempla el hecho de que la correlación decae a un ritmo menor que en el caso de los modelos clásicos, presentando dependencia a largo plazo en las trazas de tráfico.

Investigadores han analizado el tráfico Ethernet de las redes modernas y constataron que en muchos casos el tráfico sigue un comportamiento autosimilar, que puede ser tratado como la agregación de muchas fuentes on-off con distribución de alta variabilidad (Pareto heavy tails).

## 2.2 TRAFICO EN INTERNET

La ingeniería de tráfico es una disciplina que tiene por finalidad optimizar el desempeño de las redes de datos, incrementando la eficiencia del uso de los recursos y minimizando la congestión.

Dado un conjunto de elementos de red, la intensidad instantánea de tráfico es el número de elementos ocupados en un instante dado.

El tráfico también puede ser descrito en términos de las características de los diferentes objetos que ocupan a los elementos de red, tales como paquetes, ráfagas, flujos, sesiones y conexiones.

También resulta útil distinguir entre tráfico elástico y no elástico:

Entendemos por tráfico elástico aquel que puede ajustarse a cambios en el retardo y rendimiento a través de un conjunto de redes, satisfaciendo las necesidades de sus aplicaciones. Este es el tráfico

tradicional admitido por las redes TCP/IP. Las aplicaciones que generan tráfico elástico son FTP, SMTP, TELNET, SNMP y HTTP.

Entendemos por tráfico no elástico a aquel que no acepta fácilmente los cambios en el retardo y el rendimiento a través de un conjunto de redes. Las aplicaciones que requieren de transmisión en tiempo real son el ejemplo principal de este tipo de tráfico.

### **2.2.1 CONGESTIÓN EN REDES TCP/IP**

La congestión es un fenómeno producido cuando a la red, o parte de ella, se le ofrece más tráfico del que puede cursar. Desbordamiento de buffer en los routers, pérdida de paquetes de datos y retransmisión son síntomas que pueden indicar congestión de la red.

Desde el punto de vista del tamaño de buffer, si evaluamos los extremos podemos decir que si el buffer es pequeño se consigue bajo retardo pero se pueden perder paquetes debido a ráfagas, mientras que si el buffer es grande reduce el número de paquetes perdidos pero aumenta el retardo y el jitter. Por otro lado, si se piensa en una cola de tamaño variable entonces se estará en presencia de la generación de jitter.

Por defecto los equipos de red utilizan colas simples para gestionar el tráfico, es decir, lo primero que entra es lo primero que sale o en inglés First In First Out (FIFO). Cuando la cola simple se llena los paquetes que llegan se descartan produciéndose la caída de la cola o en inglés tail-drop.

Si el esquema de gestión de colas que se utiliza es del tipo FIFO + TAIL DROP se presenta la desventaja de producirse sincronismo entre las sesiones TCP e inanición de TCP.

El sincronismo de TCP se produce por la pérdida de paquetes de muchas sesiones al mismo tiempo. Las sesiones TCP se reinician conjuntamente y quedan sincronizadas, cayendo en un comportamiento oscilatorio [18].

La inanición de TCP se produce cuando la cola se llena por exceso de tráfico y los paquetes se descartan sin aplicar una política de selección de flujos. Después de que comienzan las caídas de cola los flujos TCP se ralentizan simultáneamente, pero otros flujos no TCP como el protocolo de datagramas de usuario (UDP) y el tráfico no IP, no lo hacen. En consecuencia, el tráfico no TCP comienza a llenar las colas y deja poco o ningún espacio para los paquetes TCP [18].

## 2.2.2 PROTOCOLO DE CONTROL DE TRANSMISIÓN (TCP)

### 2.2.2.1 Descripción de TCP

TCP es un protocolo que reside en el nivel de transporte del modelo TCP/IP. Está situado entre IP y los niveles superiores y se ocupa de conservar la integridad de la transferencia de los datos entre los extremos [19].

El segmento TCP está compuesto por los datos enviados desde la capa de aplicación y la cabecera añadida por el protocolo de transporte. El segmento TCP después es encapsulado en un datagrama IP para ser enviado por la capa de red.

### 2.2.2.2 Características de TCP

Las características más relevantes del protocolo TCP son las siguientes [18]:

- Es un protocolo orientado a conexión.
- La conexión es full dúplex y extremo a extremo.
- Tiene control de secuencia de envío de segmentos.
- Los datos perdidos son reenviados.
- Controla el flujo y congestión adaptando el tamaño de la ventana de envío.
- Permite el monitoreo del flujo de los datos y así evitar la saturación de la red.

### 2.2.2.3 Control de flujo

TCP usa control de flujo para evitar que un emisor envíe datos de forma más rápida de la que el receptor puede recibirlos y procesarlos. El control de flujo es un mecanismo esencial en redes en las que se comunican computadoras con distintas velocidades de transferencia.

TCP usa una ventana deslizante, controlada dinámicamente por el receptor, para el control de flujo. En cada segmento TCP, el receptor especifica en el campo *receive window* la cantidad de bytes que puede almacenar en el buffer para esa conexión. El emisor puede enviar datos hasta esa cantidad. Para poder enviar más datos debe esperar que el receptor le envíe un ACK con un nuevo valor de ventana.

#### 2.2.2.4 Control de congestión

El protocolo TCP interpreta los paquetes perdidos como síntoma de congestión de la red y reacciona disminuyendo el tamaño de la ventana de transmisión. Mantiene el control de dos ventanas: ventana del receptor (control de flujo) y ventana de congestión (control de congestión). La cantidad de datos simultáneos que puede enviar sin reconocimiento está dado por el mínimo entre la ventana de recepción y ventana de congestión.

El protocolo TCP mantiene un temporizador de retransmisión, RTO (Retransmission Time Out), que al expirar produce el reenvío de los paquetes no confirmados. El temporizador debe ser suficientemente pequeño como para responder rápidamente a las pérdidas, pero no tanto como para forzar la retransmisión de datos que han sufrido un pico de retardo en la red sin haber llegado a perderse, como sería el caso de congestión. El tratamiento de la temporización y la retransmisión quedan sujetos a la utilización de algoritmos retransmisión adaptativo, algoritmo de Karn, y la utilización de marcas temporales.

#### 2.2.2.6 Equidad en TCP

La equidad en TCP es una particularidad del mecanismo de control de congestión que responde a lo siguiente: si disponemos de una capacidad de transmisión  $R$  y una cantidad de  $K$  conexiones, la velocidad media de cada conexión será  $R/K$ . Se dice que TCP es un protocolo justo. Sin embargo, nada previene que una misma aplicación pueda abrir conexiones en paralelos entre un par transmisor-receptor.

#### 2.2.2.7 Versiones de TCP

Existen varias implementaciones del protocolo TCP/IP, cada una con sus ventajas y desventajas. Las primeras implementaciones surgieron en los sistemas operativos BSD (Berkeley Software Distribution), que deriva de los sistemas UNIX. Luego fueron surgiendo nuevas versiones con la finalidad de conseguir mejoras en su rendimiento.

Entre las implementaciones de base, durante el inicio de TCP, tenemos: Tahoe, Reno, y Vegas. A partir de estas implementaciones de base, surgieron varias versiones mejoradas y cada una tuvo su

lugar en diferentes sistemas operativos, por lo que podemos afirmar que cada sistema operativo muestra un comportamiento distinto en cuanto a TCP.

No obstante, y más allá de las mejoras que cada versión pueda implementar, todas tienen en común una serie de componentes estándares que garantizan la compatibilidad y el correcto funcionamiento de internet. Las RFC-1122/1123/1127, entre otras, detallan los requisitos a cumplir y técnicas que deben incluir cada versión de TCP.

## **2.2.3 MECANISMOS DE CONTROL DE CONGESTIÓN DE TCP/IP**

### **2.2.3.1 Algoritmo de arranque lento**

El algoritmo de arranque lento (o en inglés Slow-start) actúa al comienzo de una sesión TCP y cuando se ha perdido algún paquete. Cuando por llenado del buffer de algún router empiezan a descartarse los paquetes, las sesiones TCP ejecutan slow-start reduciendo el ritmo de transmisión y se autocontrolan.

Si bien la ejecución de slow-start controla la congestión, tail-drop provoca un comportamiento oscilante bajando notablemente el rendimiento. Tail-drop debe ser evitado, y puede ser evitado si se previene la congestión.

### **2.2.3.2 Algoritmo de prevención de congestión**

Este algoritmo (en inglés Congestion-Avoidance) persigue adaptar la transferencia a la situación de la red en cada momento, reaccionando ante posibles estados de congestión.

Como se explico en el ítem 2.2.2.4 Control de Congestión, TCP interpreta una expiración del RTO como un síntoma claro de congestión. Cuando esto ocurre, se lleva a cabo un estrechamiento exponencial de la ventana que consiste en reducir su tamaño a la mitad.

Al reiniciarse la transferencia exponencial, cuando se llega al umbral de arranque lento, el algoritmo propone un ensanchamiento lineal de la ventana que permita adaptarse progresivamente a la nueva situación descongestionada aprovechando al máximo el ancho de banda disponible.

### 2.2.3.3 Algoritmo de retrasmisión rápida y recuperación rápida

Este algoritmo (en inglés Fast Retransmit / Fast Recovery) interpreta que al recibir 3 ACK duplicados relativos al mismo segmento, se dio una pérdida de paquete y no congestión de la red. Bajo este supuesto se produce la retrasmisión rápida del segmento sin esperar el vencimiento del RTO.

Además, el algoritmo pone el umbral de arranque lento a la mitad del valor actual de la ventana de congestión, elimina la fase de arranque lento, y sigue inmediatamente con Congestion-Avoidance. Este último conjunto de acciones se conoce como recuperación rápida.

### 2.2.3.4 Detección temprana aleatoria

Sabemos que la congestión puede prevenirse si las sesiones TCP (de hecho el 80% del tráfico de internet) pueden lentificarse. En consecuencia podemos pensar en lentificar las sesiones TCP si algunos paquetes son descartados ocasionalmente.

El descarte proactivo de paquetes antes de que se llene el buffer mejora notablemente el rendimiento global de los recursos. Esto es porque se obliga a realizar slow-start a algunas sesiones TCP pero no a todas a la vez, evitando el sincronismo.

El algoritmo que se basa en descartar paquetes al azar con probabilidad creciente a medida que la cola se llena se lo conoce como RED (por sus siglas en inglés Random Early Detect).

### 2.2.3.5 Notificación explícita de congestión

Permite la notificación de extremo a extremo de la congestión de la red, sin dejar caer los paquetes. Para su habilitación se requiere una infraestructura de red compatible con dicha funcionalidad.

La notificación explícita de congestión o en inglés Explicit Congestion Notification (ECN) funciona solo si es compatible con routers, remitente y destinatario. El algoritmo ECN queda definido mediante la RFC3168 (ampliada por 4301, 6040, 8311).

## 2.3 REDES CONVERGENTES

El concepto tradicional de disponer de redes separadas para soportar los distintos servicios de voz, datos y video quedo obsoleto. Los avances tecnológicos han permitido consolidar estas redes diferentes en una única plataforma denominada redes convergentes o redes multiservicio.

La convergencia en redes se está viendo propiciada en gran medida por la transición hacia las redes de la próxima generación (en inglés Next Generation Networking o NGN).

NGN, también conocida con el termino *all-IP*, encapsula los servicios multimediales en paquetes IP y permite la provisión de servicios integrados, explotando al máximo el ancho de banda del canal con el uso de Calidad del Servicio (en inglés Quality of Service o QoS) [20].

En la arquitectura NGN existe una separación bien definida entre la porción de red de transporte (conectividad) y los servicios que corren por encima de esa red. Esto quiere decir que siempre que un proveedor desee habilitar un nuevo servicio, puede hacerlo fácilmente definiéndolo desde la capa de servicio directamente sin tener en cuenta la capa de transporte. Los servicios son independientes de la infraestructura de red. La tendencia es que los servicios de voz, video y datos, sean independientes de red y que residan en los dispositivos de usuario.

La convergencia se manifiesta en múltiples dimensiones dentro del campo de la tecnología de la información y las comunicaciones, donde se destacan las siguientes:

Convergencia de redes: una misma infraestructura soportando múltiples servicios y contenidos.

Convergencia de terminales: un mismo terminal para acceder a múltiples servicios.

Convergencia de servicios: un mismo servicio adaptándose a distintas prestaciones y contenidos.

Convergencia de aplicaciones: una misma red y una misma aplicación en múltiples terminales, para múltiples contenidos y servicios.

Convergencia hacia la nube: una tendencia masiva de consumir la tecnología de forma ubicua, utilizando computación en la nube.

Convergencia de mercados: empaquetamiento de servicios con tarifas planas bajo el concepto de “todo en uno”.

La manifestación de la convergencia en las diferentes dimensiones del campo TIC, conjuntamente con la transformación digital de hogares, empresas, y organizaciones, tienen un fuerte impacto en telecomunicaciones, y conllevan a constantes cambios en el tratamiento del tráfico con el objeto de garantizar las comunicaciones.

Los acontecimientos históricos que dieron origen a la existencia de tráfico autosimilar en las redes de datos, tales como el nacimiento de internet y las innovaciones tecnológicas que han permitido unificar diferentes redes, son precedentes suficientes para pensar que la autosimilaridad llegó para quedarse, y con esto necesidad de desarrollar conocimientos y crear modelos de tráfico más acertados que contemplen el fenómeno de autosimilaridad.

### **2.3.1 CONVERGENCIA DE LA RED EXPERIMENTAL**

El trabajo experimental desarrollado en la presente tesis, y que se detalla en la sección siguiente, tiene el potencial de haberse realizado en gran parte sobre una red multiservicio por la que transitan de datos de diferente naturaleza.

La red operativa, sobre la cual se efectuaron capturas del orden de 1 millón de paquetes para su análisis, se corresponde con una red totalmente IP y con más de 500 usuarios accediendo en simultáneo a una gran variedad de servicios.

Dentro del conjunto de servicios disponibles en la red y de uso habitual se destacan: telefonía IP, salas de videoconferencias, WIFI corporativo, correo electrónico, transferencia de archivos, navegación WEB, seguridad basada en la nube, actualización automática de equipos, capacitación en línea, servicios para proveedores publicados hacia el exterior, entre otros.

Si nos remitimos a los diferentes tipos de convergencias en el campo TIC, podemos garantizar que la red operativa estudiada es una auténtica red convergente.

## 3. TRABAJO EXPERIMENTAL

---

### 3.1 AMBIENTE CONTROLADO

En este contexto la generación de las muestras se lleva a cabo mediante la utilización de una red aislada que permite manipular con precisión las variaciones del tráfico.

El escenario de red sin congestión se consigue al disponer de un enlace con ancho de banda mayor a los recursos requerido por la comunicación emisor/receptor y sin pérdida de paquetes.

El escenario de red congestionada se consigue al disponer de un enlace saturado que hace fluctuar los recursos requeridos por la comunicación emisor/receptor, provocando la pérdida de paquetes y en consecuencia activando el mecanismo de control de congestión del protocolo TCP.

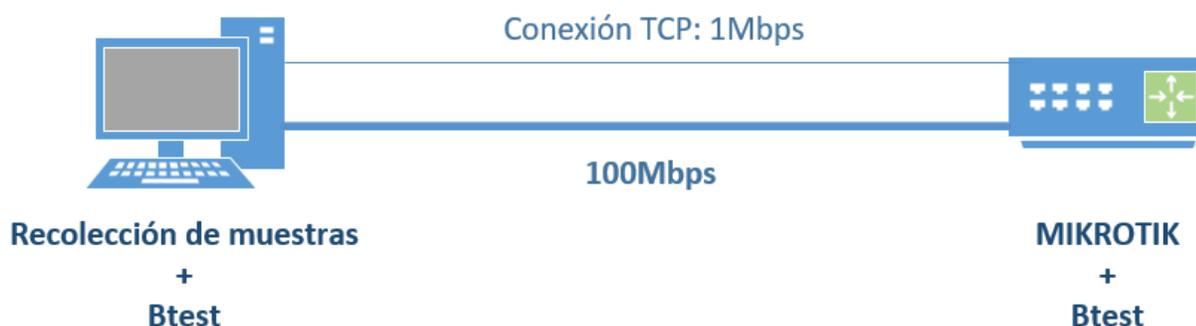
Durante el proceso de generación y recolección de las muestras, se monitorea el comportamiento del tráfico TCP con el afán de asegurar la conducta deseada y disponer de muestras fehacientes y representativas del escenario de interés.

#### 3.1.1 CONDICION DE RED RELAJADA: ESCENARIO-1

##### 3.1.1.1 *Ámbito de realización*

Las muestras se obtienen de un ambiente controlado, bajo una condición de red relajada y el trabajo experimental se efectúa sobre hardware dedicado.

Los equipos que intervienen son los siguientes:



El tráfico es generado por la herramienta *btest* de Mikrotik, con una ocupación de 1Mbps y un ancho de banda disponible de 100Mbps.

Basado en la propiedad de una muestra autosemejante, de que una pequeña sección suya puede ser vista como una réplica a menor escala de todo la muestra, se recolectarán paquetes durante un periodo de mil segundos y se dividirán en segmentos de cien segundos, obteniendo diez muestras. Luego, se analizarán las diez muestras por separados en busca de un denominador común que permitan concluir la influencia de TCP en el comportamiento autosimilar del tráfico.

Para capturar y analizar el comportamiento de los paquetes de datos se utiliza el software Wireshark. Una vez recolectados, y seguros de que es una muestra fidedigna del escenario de interés, son exportados y trabajados mediante un script con el objetivo de contabilizarlos y formatearlos para su procesamiento y análisis estadístico mediante el software R.

### 3.1.1.2 Calidad de servicio

Para la recolección de las muestras no se utiliza configuración de QoS ya que para el caso de estudio se utiliza una única conexión TCP de 1Mbps, y el vínculo disponible es de 100Mbps.

Se verifica la negociación a 100Mbps:

```
[admin@MikroTik] > interface ethernet monitor
numbers: 1
          name: ether2-master-local
          status: link-ok
    auto-negotiation: done
              rate: 100Mbps
          full-duplex: yes
    tx-flow-control: no
    rx-flow-control: no
          advertising: 10M-half,10M-full,100M-half,100M-full
link-partner-advertising: 10M-half,10M-full,100M-half,100M-full
```

Se verifica que no hay pérdidas de paquetes ni errores durante la transmisión:

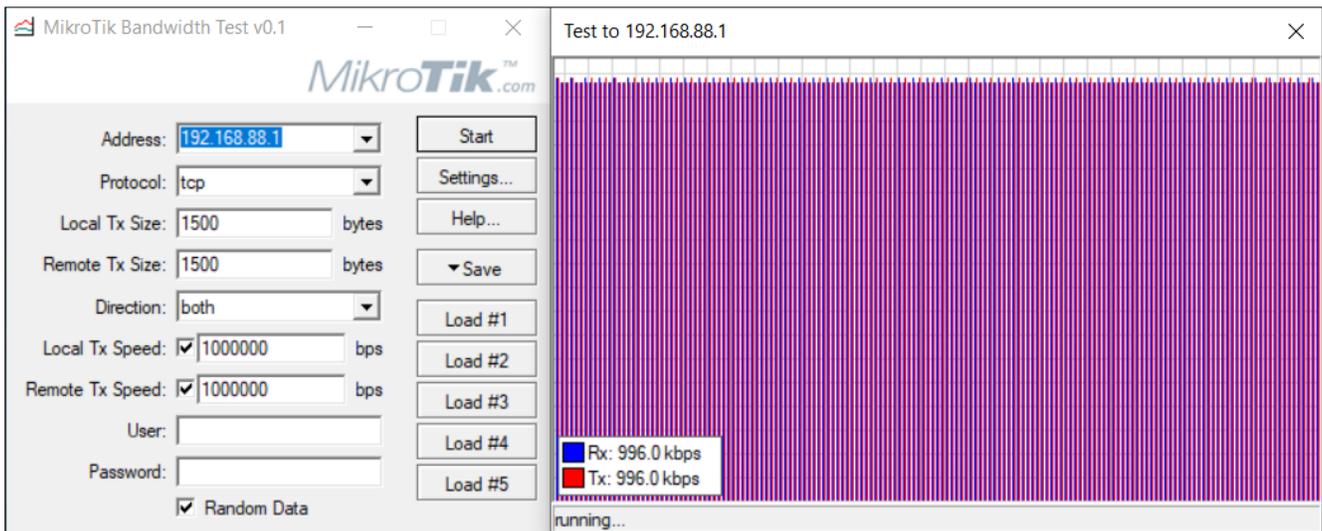
RX-BYTE	TX-BYTE	RX-PACKET	TX-PACKET	RX-DROP	TX-DROP	TX-QUEUE-DROP	RX-ERROR	TX-ERROR
0	0	0	0	0	0	0	0	0
380 515 080	410 014 518	718 582	661 349	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0

### 3.1.1.3 Análisis de tráfico

Se habilita la aplicación *btest* en la PC y en el router Mikrotik para que trafiquen en ambos sentidos, y se realiza la prueba de ancho de banda.

Las imágenes siguientes son representativas de proceso de generación del tráfico.

App *btest* corriendo en Windows-10:

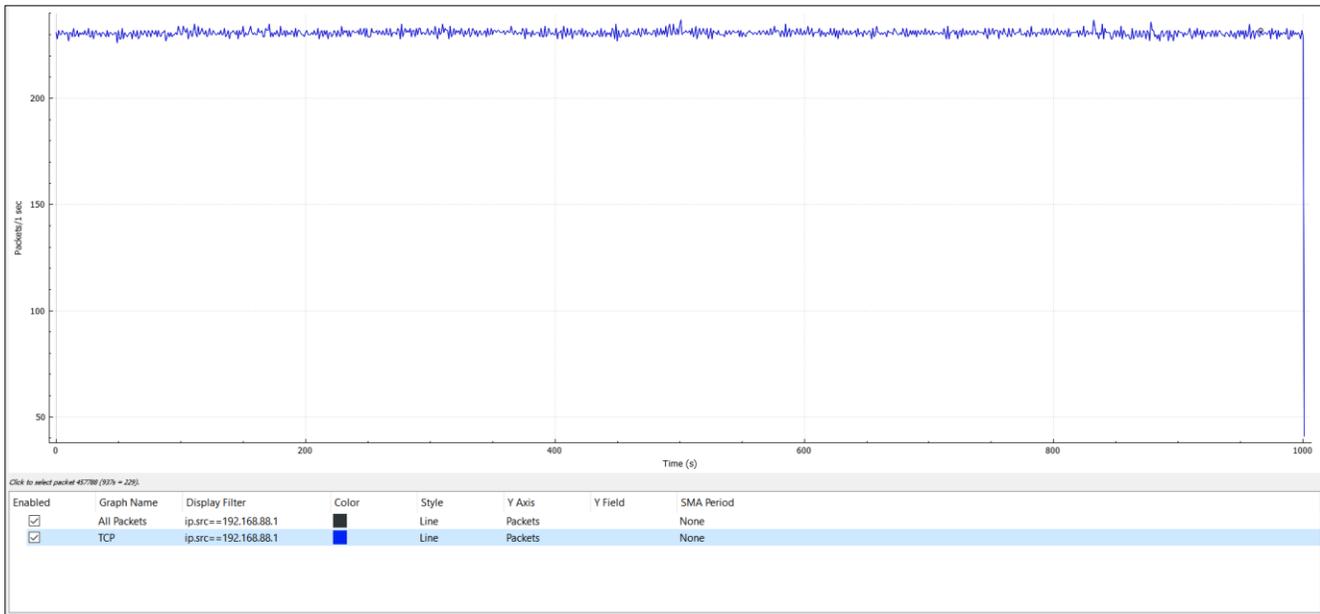


En la imagen de la aplicación *btest* se observa tráfico en ambos sentidos, en color azul el tráfico que recibe la PC y en color rojo el tráfico transmitido desde la PC hacia la Mikrotik, con una ocupación de 1Mbps, MTU de 1500 bytes y la opción de datos aleatorios activa para rellenar el cuerpo del segmento TCP.

Representación de datos aleatorios en el cuerpo del segmento TCP:

No. .	Time	Source	Destination	Protocol	Info
1	0.000000	192.168.88.10	192.168.88.1	TCP	59812 > cisco-sccp [ACK] Seq=1 Ack=1 win=8212 Len=1460
Frame 1 (1514 bytes on wire, 1514 bytes captured)					
Ethernet II, Src: d8:cb:8a:83:39:8f (d8:cb:8a:83:39:8f), Dst: d4:ca:6d:f1:ba:40 (d4:ca:6d:f1:ba:40)					
Internet Protocol, Src: 192.168.88.10 (192.168.88.10), Dst: 192.168.88.1 (192.168.88.1)					
Transmission Control Protocol, Src Port: 59812 (59812), Dst Port: cisco-sccp (2000), Seq: 1, Ack: 1, Len: 1460					
Data (1460 bytes)					
Data: 90A2CCCEF0225C5E70828C8E90C2DCDE10224C4F72A5E4F7... [Length: 1460]					
0000	d4 ca 6d f1 ba 40 d8 cb 8a 83 39 8f 08 00 45 00	..m..@...9...E.			
0010	05 dc 93 1a 40 00 80 06 00 00 c0 a8 58 0a c0 a8	...@...X...			
0020	58 01 e9 a4 07 d0 35 0d 86 68 c0 e6 d9 ca 50 10	X.....5..h...P.			
0030	20 14 37 2b 00 00 90 a2 cc ce f0 22 5c 5e 70 82	.7+... ..\p.			
0040	8c 8e 90 c2 dc de 10 22 4c 4f 72 a5 e4 f7 32 65	]......LOR...2e			
0050	6c 8f a3 d7 e7 1f 5b 77 b0 b1 d6 df 11 43 71 9f	].....[w....Cq.			
0060	da e5 fc 08 1c 28 4c 79 8e bb e5 03 42 75 a8 e4	.....(Ly....BU.			
0070	fc 11 32 4f 6d 93 be e1 09 2d 52 5b 7d 7f 8e 9d	...20m...-R[]...			
0080	a9 a9 d2 08 2e 4d 84 8c 98 b9 df e9 fc 13 13 2c	.....M.....			
0090	69 97 a6 b9 ed 1a 23 3d 50 67 77 80 9e ac eb 13	i.....#=Pgw....			
00a0	13 18 2e 61 90 b4 d1 e7 06 31 45 5a 74 a7 db df	...a....IEZt...			
00b0	10 1a 39 6d 9e b0 d7 0b 0b 34 72 b1 c1 ea f0 17	...9m....4r....			
00c0	47 80 9e bd dd f6 fc 23 44 4e 8d 91 92 c4 eb 0f	G.....#DN....			
00d0	40 6a 8a 9f d1 08 40 7d a0 a8 e9 03 37 48 82 c1	[j....@]...7H...			
00e0	e6 f3 fb 00 3d 48 84 a6 cd 09 33 3e 4f 71 9c cd	...-H...>0q...			
00f0	00 3b 5c 87 a7 ac b3 d7 09 14 15 2f 33 34 6f 6f	...v.../3400...			
0100	85 ac e1 f8 20 5e 91 ba d0 d8 fa 35 6a 81 8d b7	.....^...5]...			
0110	c2 e3 f6 03 15 45 5f 76 9f d6 f2 30 68 a2 dd fa	...E..v...0h...			
0120	01 26 51 85 9b d7 0d 29 47 5b 89 ac b1 bc bd f8	...&Q...G[.....			
0130	1d 20 21 54 79 94 b5 c0 c5 d8 f9 fc 01 2c 2d 48	...!Ty...-H...			
0140	6d 90 91 a4 c9 04 25 50 55 88 a9 cc d1 dc de 1a	m...%P U.....			
0150	40 48 5a 76 7c 95 98 b5 c9 d9 14 32 56 90 a8 b6	[GHZv]...2V...			
0160	bb d6 df 02 18 1c 56 66 7d b2 dd f2 fa 28 44 77	.....vT]....(Dw			
0170	a0 cf e1 15 27 5c 87 c5 c5 e3 f4 0b 35 41 5c 99	.....5A\...			
0180	bc f6 2a 49 6a 76 91 a2 c2 f4 01 30 52 6a 79 87	...*Ijv...0Rjy.			
0190	bb fa 1f 57 8a b3 ef 26 63 63 7e 7f b3 d2 d8 10	...w...&cc-....			
01a0	43 51 7a 89 c3 d0 d4 da e3 ff 16 34 54 8f b5 d6	CQZ...4T....			
01b0	ee 11 17 37 76 b0 c1 c1 f0 12 2b 3b 4a 84 95 a5	...7v...+;j...			
01c0	c4 c6 ff ff 1e 58 69 89 98 ba d3 03 32 6c 7d ad	...x1...2]...			

Representación de los paquetes por segundos recibidos por la interface de red:

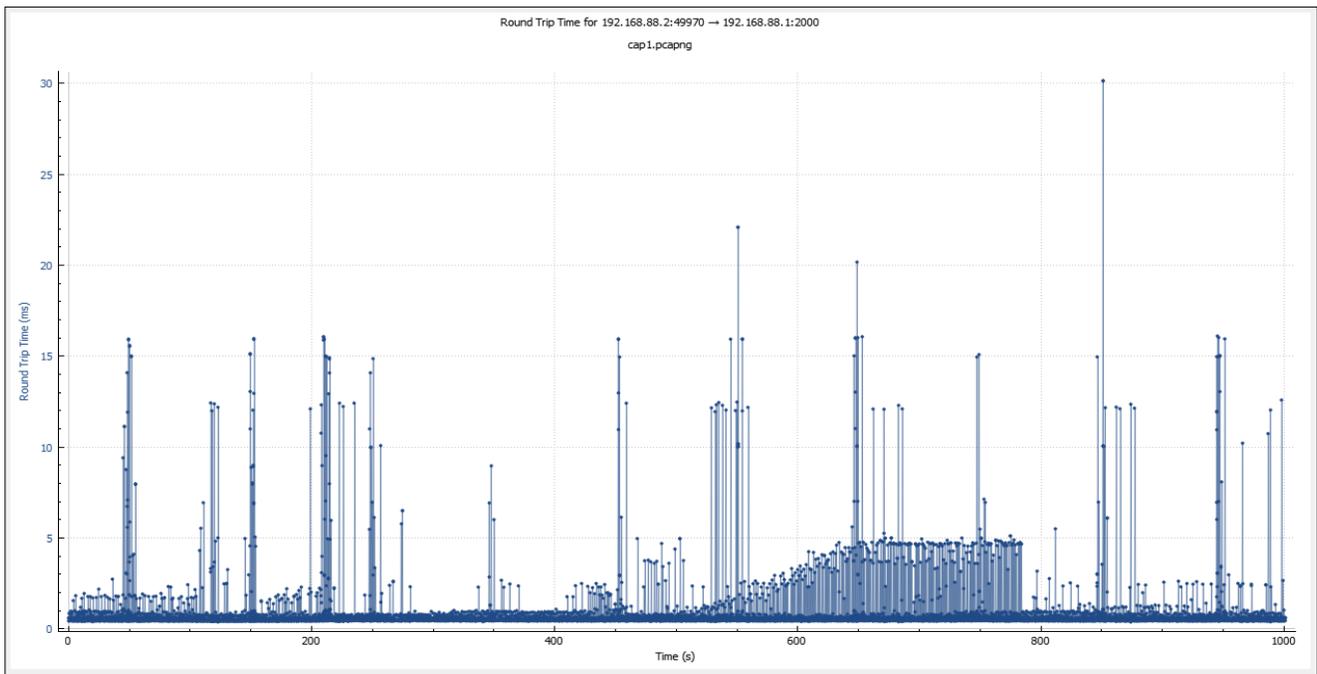


En las imágenes de cantidad de paquetes por segundos se presentan de color negro y todo el tráfico capturado, y de azul el tráfico TCP. Se verifica que la totalidad de datos transmitidos es TCP, dato que el tráfico TCP es equivalente a la totalidad del tráfico capturado.

Los datos son aleatorios, por lo que la longitud de paquetes también es variable, tal como se muestra en la siguiente imagen:

No.	Time	Source	Destination	Protocol	Length	Info
3	0.000438	192.168.88.1	192.168.88.2	TCP	60	2000 → 49970 [ACK] Seq=1 Ack=1501 Win=8010 Len=0
4	0.009361	192.168.88.1	192.168.88.2	TCP	1514	2000 → 49970 [ACK] Seq=1 Ack=1501 Win=8010 Len=1460
5	0.009361	192.168.88.1	192.168.88.2	TCP	94	2000 → 49970 [PSH, ACK] Seq=1461 Ack=1501 Win=8010 Len=40
9	0.015444	192.168.88.1	192.168.88.2	TCP	60	2000 → 49970 [ACK] Seq=1501 Ack=3001 Win=8010 Len=0
10	0.021373	192.168.88.1	192.168.88.2	TCP	1514	2000 → 49970 [ACK] Seq=1501 Ack=3001 Win=8010 Len=1460
11	0.021374	192.168.88.1	192.168.88.2	TCP	94	2000 → 49970 [PSH, ACK] Seq=2961 Ack=3001 Win=8010 Len=40
17	0.031566	192.168.88.1	192.168.88.2	TCP	60	2000 → 49970 [ACK] Seq=3001 Ack=6001 Win=8010 Len=0
18	0.033364	192.168.88.1	192.168.88.2	TCP	1514	2000 → 49970 [ACK] Seq=3001 Ack=6001 Win=8010 Len=1460
19	0.033365	192.168.88.1	192.168.88.2	TCP	94	2000 → 49970 [PSH, ACK] Seq=4461 Ack=6001 Win=8010 Len=40
21	0.045386	192.168.88.1	192.168.88.2	TCP	1514	2000 → 49970 [ACK] Seq=4501 Ack=6001 Win=8010 Len=1460
22	0.045388	192.168.88.1	192.168.88.2	TCP	94	2000 → 49970 [PSH, ACK] Seq=5961 Ack=6001 Win=8010 Len=40
26	0.047454	192.168.88.1	192.168.88.2	TCP	60	2000 → 49970 [ACK] Seq=6001 Ack=7501 Win=8010 Len=0
27	0.057365	192.168.88.1	192.168.88.2	TCP	1514	2000 → 49970 [ACK] Seq=6001 Ack=7501 Win=8010 Len=1460
28	0.057366	192.168.88.1	192.168.88.2	TCP	94	2000 → 49970 [PSH, ACK] Seq=7461 Ack=7501 Win=8010 Len=40
32	0.062448	192.168.88.1	192.168.88.2	TCP	60	2000 → 49970 [ACK] Seq=7501 Ack=9001 Win=8010 Len=0
33	0.069369	192.168.88.1	192.168.88.2	TCP	1514	2000 → 49970 [ACK] Seq=7501 Ack=9001 Win=8010 Len=1460
34	0.069371	192.168.88.1	192.168.88.2	TCP	94	2000 → 49970 [PSH, ACK] Seq=8961 Ack=9001 Win=8010 Len=40
38	0.078974	192.168.88.1	192.168.88.2	TCP	60	2000 → 49970 [ACK] Seq=9001 Ack=10501 Win=8010 Len=0

En lo que respecta al tiempo de ida y vuelta (o en inglés el RTT, round-trip time), tal como se muestra en la imagen, se tienen los siguientes tiempos:

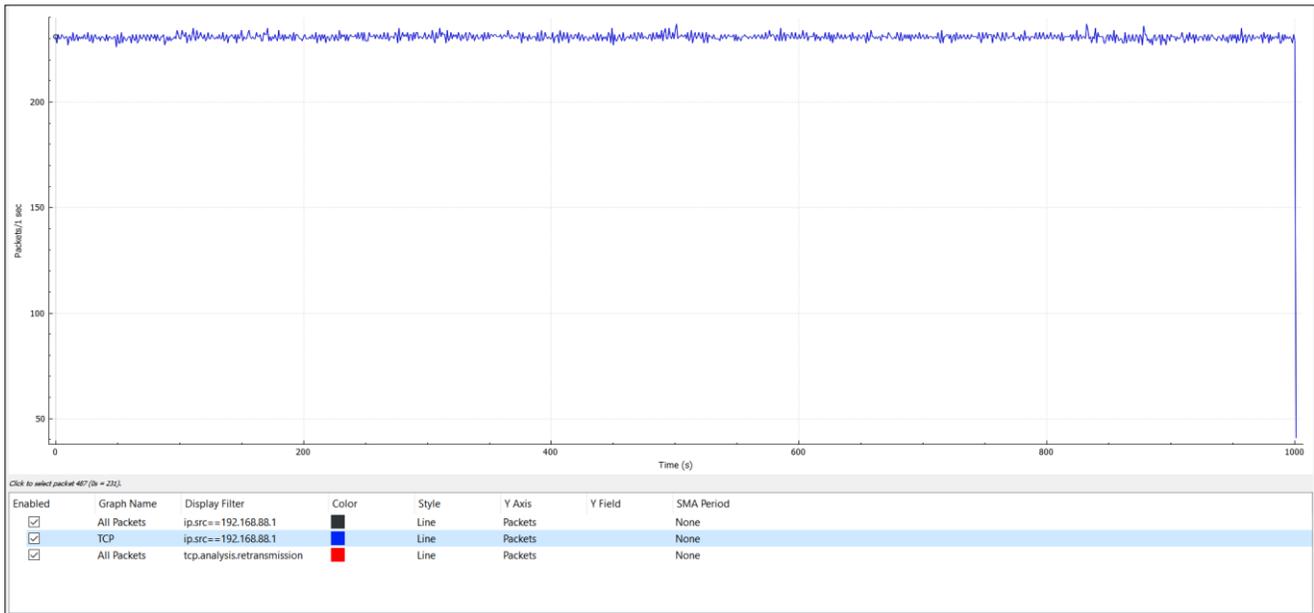


### 3.1.1.4 Comportamiento de TCP

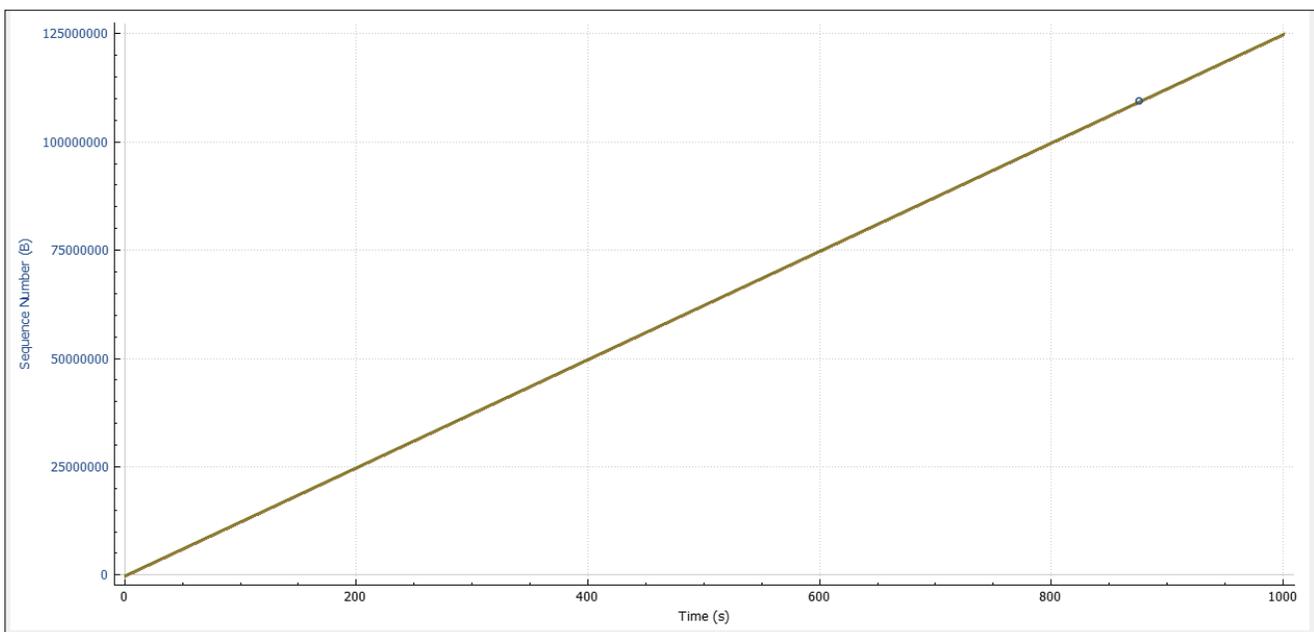
Tal como se indicó en el apartado de 2.2.2. Protocolo de Control de Transmisión (TCP), el algoritmo slow-start es ejecutado cuando vence el RTO porque se interpreta como congestión en la red.

La retransmisión también se da cuando se reciben 3 ACK repetidos, pero en este caso se asume pérdida de paquete y no congestión. En este caso hay una reducción de la ventana de transmisión causada por la ejecución del algoritmo de retransmisión y recuperación rápida.

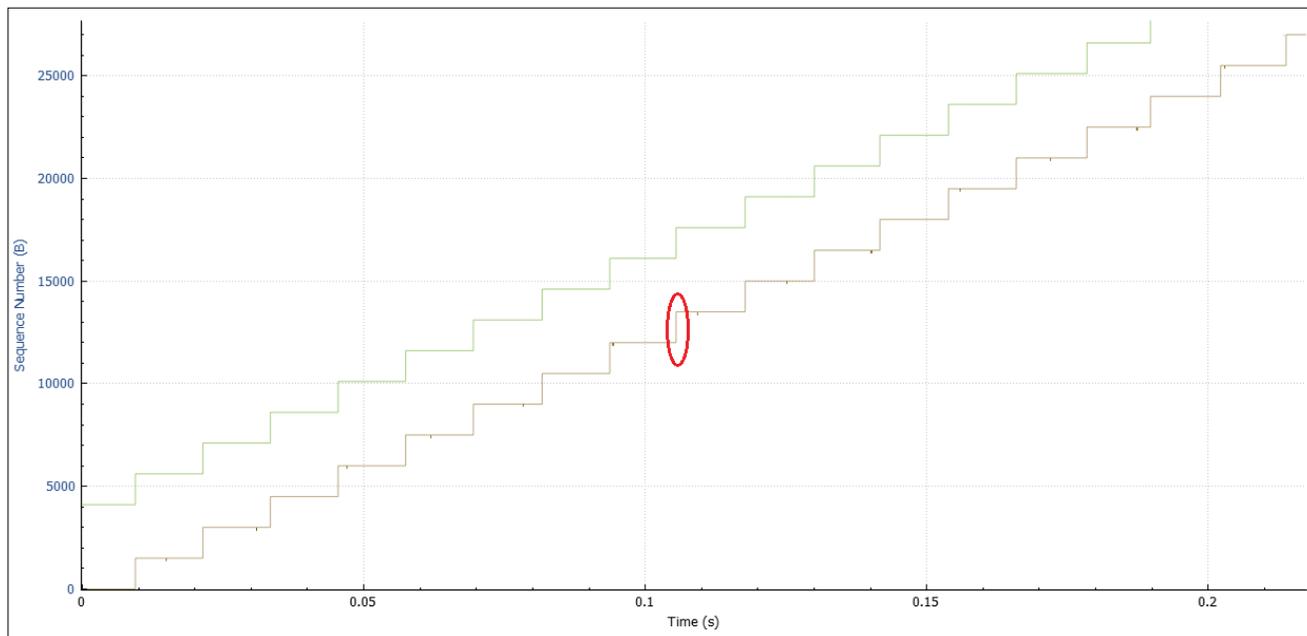
A Continuación se observa, través del filtro de Wireshark, que no hay retransmisiones de paquetes, lo que demuestra que la red está totalmente relajada.



Otro parámetro interesante de analizar es el número de secuencia contra el tiempo. Para este caso se utiliza la herramienta *tcptrace* de la opción estadísticas de Wireshark. Esta herramienta nos permite visualizar mediante un gráfico el crecimiento del número de secuencia a medida que pasa el tiempo, así como también el valor de ACK y la ventana de recepción anunciada por el otro extremo.



Al ampliar la imagen se puede observar con más detalle. Las líneas marrones verticales indican el rango de secuencias enviadas, las comas indican el ACK confirmado en el tiempo, y la gráfica verde superior el tamaño de ventana de recepción anunciado.



Tal como se observa en la imagen anterior, no hay presencia de ACK duplicados y no se produjeron retransmisiones. Las muestras se interpretan como una muestra fidedigna de un escenario sin congestión.

### 3.1.1.5 Recolección de las muestras

La recolección de muestras se efectúa capturando paquetes con Wireshark durante mil segundos, que luego se divide en segmentos de cien segundos, y se obtienen diez vectores de captura.

Las capturas muestran el instante de tiempo en que cada paquete fue recibido, por lo que es necesario procesarlo mediante un script para contabilizar los paquetes por segundos.

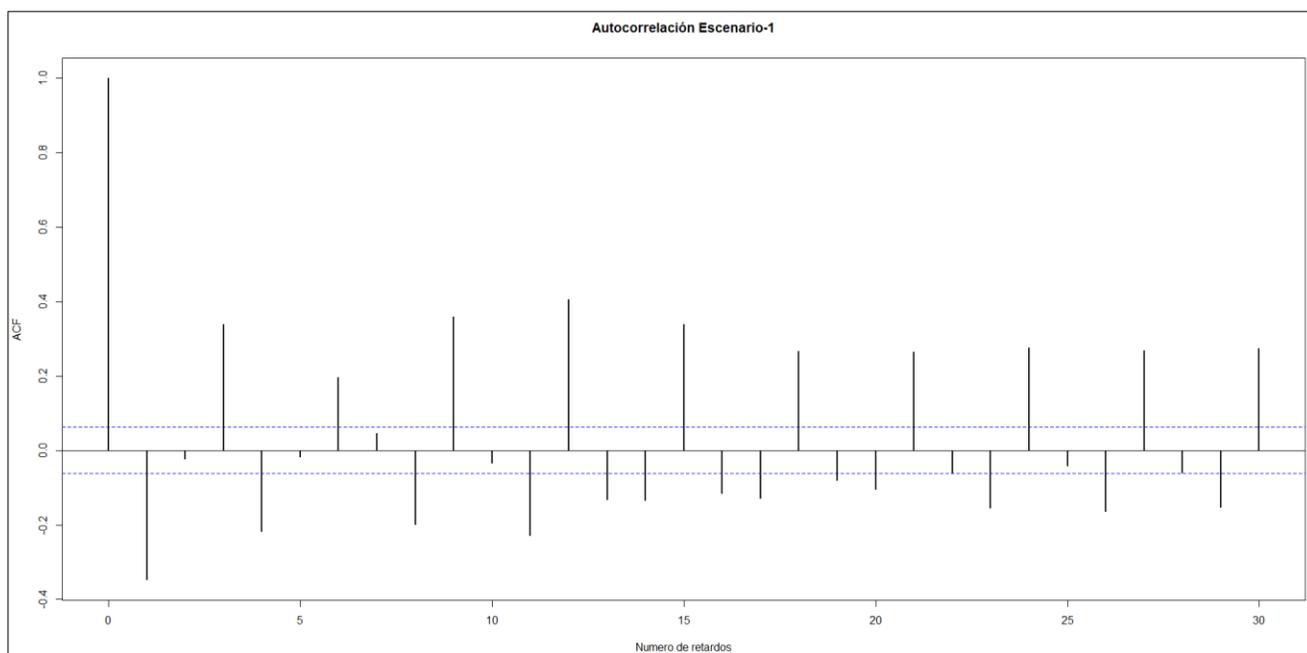
Una vez procesadas las capturas se obtiene el vector *muestra1* que almacena la contabilización de los paquetes capturados en cada segundo, durante los mil segundos. En los vectores *m11*, *m12*, *m13*, *m14*, *m15*, *m16*, *m17*, *m18*, *m19* y *m110* se almacenan segmentos de cien segundos consecutivos, permitiendo disponer de diez muestras.

En relación a lo dicho, se llama captura al vector que representa la recolección en crudo de los paquetes y muestra al vector que contiene la contabilización de los paquetes por segundos.

### 3.1.1.6 Autocorrelación

Analizar la autocorrelación nos permite observar la existencia de dependencia a corto o largo plazo. Tal como se expresa en el ítem 2.1.3 Autocorrelación, los procesos con dependencia a largo plazo pueden ser representados adecuadamente mediante modelos de tráfico autosimilar.

En consecuencia, evaluaremos la autocorrelación del proceso y los efectos respaldarán a los resultados que se obtengan luego en el análisis de autosimilaridad.



Se utiliza la función  $ACF$  de RGui para realizar el cálculo de autocorrelación. En el correlograma se observa que hay cambios pronunciados entre los desplazamientos, lo que indica que la *muestra 1* no presenta un decaimiento lento en el tiempo y, en consecuencia, no se corresponde con tráfico autosimilar. El valor negativo de la correlación, indica que una variable disminuye conforme la otra aumenta.

Cabe aclarar que, debido a la naturaleza del proceso y a las funciones utilizadas para la evaluación del presente ítems, se supone que estamos frente a un proceso al menos débilmente estacionario (WSS).

También, al disponer de una única realización del proceso para analizar, se considera que el proceso es Ergódico en la autocorrelación. Dicho esto, se observa la autocorrelación temporal suponiendo que iguala a la del proceso.

Los resultados obtenidos en este ítem serán utilizados para validar el cálculo de la autosimilaridad del presente escenario, y para contrastar la grafica con la autocorrelación del escenario-2.

### 3.1.1.7 Autosimilaridad

Haciendo uso de la función *hurstexp()* disponible en el paquete de funciones PRACMA (Practical Numerical Math Functions), del entorno de software RGui, obtenemos el coeficiente de Hurst que nos permitirá concluir si las muestras dadas contienen una componente autosimilar y en qué medida está presente.

Primero calculamos el coeficiente de Hurst para la muestra-1, de longitud mil (1000) segundos:

```
> H1<-hurstexp(muestral)
Simple R/S Hurst estimation:      0.6244604
Corrected R over S Hurst exponent: 0.7876259
Empirical Hurst exponent:        0.8806212
Corrected empirical Hurst exponent: 0.8121759
Theoretical Hurst exponent:      0.5388682
```

Coficiente de Hurst para la muestra en el intervalo 1-100 segundos:

```
> H11<-hurstexp(m11)
Simple R/S Hurst estimation:      0.4566289
Corrected R over S Hurst exponent: 0.5470967
Empirical Hurst exponent:        0.4212037
Corrected empirical Hurst exponent: 0.4713936
Theoretical Hurst exponent:      0.5229461
```

Coficiente de Hurst para la muestra en el intervalo 101-200 segundos:

```
> H12<-hurstexp(m12)
Simple R/S Hurst estimation:      0.4129353
Corrected R over S Hurst exponent: 0.4742544
Empirical Hurst exponent:        0.4676358
Corrected empirical Hurst exponent: 0.5101422
Theoretical Hurst exponent:      0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 201-300 segundos:

```
> H13<-hurstexp(ml3)
Simple R/S Hurst estimation: 0.48386
Corrected R over S Hurst exponent: 0.6445419
Empirical Hurst exponent: 0.4853886
Corrected empirical Hurst exponent: 0.525254
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 301-400 segundos:

```
> H14<-hurstexp(ml4)
Simple R/S Hurst estimation: 0.4672888
Corrected R over S Hurst exponent: 0.621268
Empirical Hurst exponent: 0.4640916
Corrected empirical Hurst exponent: 0.5071441
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 401-500 segundos:

```
> H15<-hurstexp(ml5)
Simple R/S Hurst estimation: 0.4971444
Corrected R over S Hurst exponent: 0.7073111
Empirical Hurst exponent: 0.4760553
Corrected empirical Hurst exponent: 0.5172896
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 501-600 segundos:

```
> H16<-hurstexp(ml6)
Simple R/S Hurst estimation: 0.534374
Corrected R over S Hurst exponent: 0.7134886
Empirical Hurst exponent: 0.493046
Corrected empirical Hurst exponent: 0.5318197
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 601-700 segundos:

```
> H17<-hurstexp(ml7)
Simple R/S Hurst estimation: 0.3971867
Corrected R over S Hurst exponent: 0.4639246
Empirical Hurst exponent: 0.3565412
Corrected empirical Hurst exponent: 0.4195222
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 701-800 segundos:

```
> H18<-hurstexp(ml8)
Simple R/S Hurst estimation: 0.3461506
Corrected R over S Hurst exponent: 0.300544
Empirical Hurst exponent: 0.3439975
Corrected empirical Hurst exponent: 0.4097707
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 801-900 segundos:

```
> H19<-hurstexp(ml9)
Simple R/S Hurst estimation:      0.548434
Corrected R over S Hurst exponent: 0.6668656
Empirical Hurst exponent:        0.4140716
Corrected empirical Hurst exponent: 0.4655476
Theoretical Hurst exponent:      0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 901-1000 segundos:

```
> H110<-hurstexp(ml10)
Simple R/S Hurst estimation:      0.3407147
Corrected R over S Hurst exponent: 0.3003457
Empirical Hurst exponent:        0.3706139
Corrected empirical Hurst exponent: 0.4305872
Theoretical Hurst exponent:      0.5229461
```

La función *hurstexp()* utiliza diversos enfoques de cálculo para obtener el coeficiente de Hurst de interés. Para el caso de estudio, con el objetivo de estandarizar el método de cálculo para todos los escenarios, se utilizarán el valor obtenido mediante el análisis R/S.

Si bien existen diversas formas para calcular el coeficiente de Hurst, más precisas que el análisis R/S, la finalidad es evaluar la existencia de autosimilaridad en las muestras y no así la exactitud del cálculo. Debido a esto, la utilización de la librería PRACMA del entorno de software RGui es suficiente.

Se observa que para el escenario-1 el coeficiente de Hurst muestra un valor que oscilan entre  $H=0.34$  y  $H=0.548$ , con un valor medio de  $H=0.448$ . También se calculo el valor del coeficiente para la muestra de 1000 segundos, la cual será comparada con los resultados que se obtengan de la red congestionada.

Los valores de  $H < 0.5$  no se relaciona con la autosimilaridad, en consecuencia desde el punto de vista del cálculo de  $H$  la serie se considera impredecible.

Los resultados arrojados por la autocorrelación muestran que un aumento en el pasado implica una tendencia a disminuir en el futuro, y viceversa.

Los valores de coeficientes de Hurst obtenidos, y en comparación con el escenario-2, representan evidencia necesaria para concluir si efectivamente existe influencia del comportamiento de TCP en la autosimilaridad de los datos.

### **3.1.1.8 Caracterización de las muestras**

La idea de caracterizar la muestra surge de la necesidad de disponer de datos que permitan ahondar en el análisis, con el afán de encontrar patrones que justifiquen los resultados anteriormente obtenidos.

Dicho esto, surge la libertad de elegir aspectos de la muestra que, a criterio de quien analiza, pueden tener relación con los resultados. En este contexto, y si bien lo habitual sería observar el tipo de datos que transporta la red (voz, video, texto, etc.), se ha optado por analizar las retransmisiones TCP.

Las retransmisiones TCP tienen un impacto determinante en el comportamiento del protocolo, tal como se detallo en el ítem 2.2.2 Mecanismos de control de congestión en redes TCP/IP, las retransmisiones tienen una relación directa con la ejecución de los algoritmos de control de congestión.

El trabajo experimental se sustenta de un conjunto de muestras que representan la cantidad de mensajes por segundos que llegan a la interface de un equipo de red. Si existe congestión, un subconjunto de dichas muestras corresponde a las retransmisiones.

Sabiendo que las retransmisiones ocasionadas por vencimiento del RTO producen la ejecución de Slow-start, y que las retransmisiones causadas por la recepción ACKs duplicados producen la ejecución de FR/FR, se analizarán estos aspectos de las muestras para caracterizarlas.

Particularmente, en el escenario-1 no hay presencia de retransmisiones. Este ítem no requiere de desarrollo práctico.

### **3.1.1.9 Interpretación de los resultados**

El trabajo experimental desarrollado en el escenario-1 muestra que no hay comportamiento autosimilar del tráfico bajo condición de red relajada.

En este escenario, el protocolo TCP transmitió a 1Mbps constante y de acuerdo a la ventana de control de flujo proporcionada por el receptor. En la transmisión no hubo presencia de pérdidas de paquetes, y por lo tanto nunca se ejecutaron los algoritmos Slow-start y Fast Retransmit / Fast Recovery, asociados al control de congestión.

En conclusión, el protocolo TCP no aporta autosimilaridad al tráfico si esta bajo condición de red relajada.

## 3.1.2 CONDICION DE RED ESTRESADA: ESCENARIO-2

### 3.1.2.1 Ámbito de realización

La recolección de la muestra en el escenario-2 se realiza utilizando la misma topología de red descripta para el escenario-1, con la salvedad de que en este caso la red se congestiona para conseguir el escenario de interés.

Se analizan los paquetes entrantes a la interface y se utilizan las herramientas *Bandwidth test* y *btest* como transmisor/receptor de tráfico.

Para este escenario se genera un flujo de tráfico TCP de 1Mbps y se inunda la red con tráfico UDP para provocar la pérdida de paquetes y fluctuaciones en el rendimiento del protocolo TCP, ocasionando la ejecución del mecanismo de control de congestión.

### 3.1.2.2 Calidad de servicio

Para la recolección de las muestras no se utiliza configuración de QoS ya que para el caso de estudio existe una única conexión TCP de 1Mbps, y el vínculo disponible es de 100Mbps.

La pérdida de paquetes se consigue saturando la red con paquetes UDP, y al no haber distintos tipos de tráfico en el vínculo TCP (solo se envían datos aleatorios), no es necesario de priorizar tráfico según su tipo.

### 3.1.2.3 Análisis de tráfico

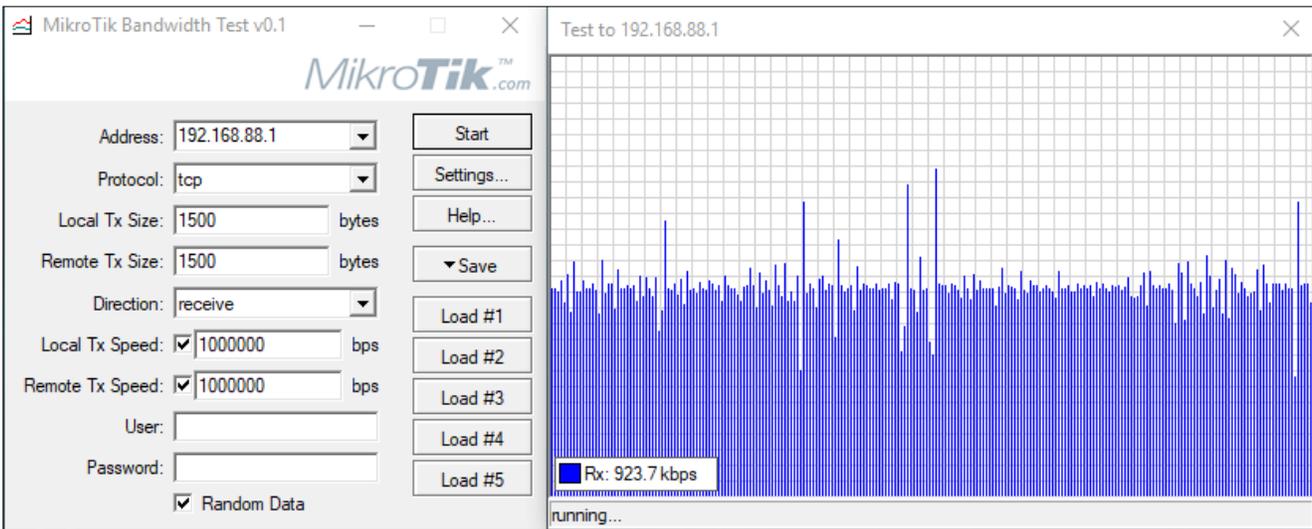
Se configura una transmisión TCP de 1Mbps en sentidos entrante a la PC, donde está instalado el software capturador de paquetes Wireshark, y al mismo tiempo se generan múltiples transmisiones UDP, en ambos sentidos y sin límite de ancho de banda, con el objetivo de hacer fluctuar la conexión TCP y provocar pérdidas de paquetes.

Se ejecuta la prueba durante un periodo de mil segundos, para luego dividir dicha captura en diez.

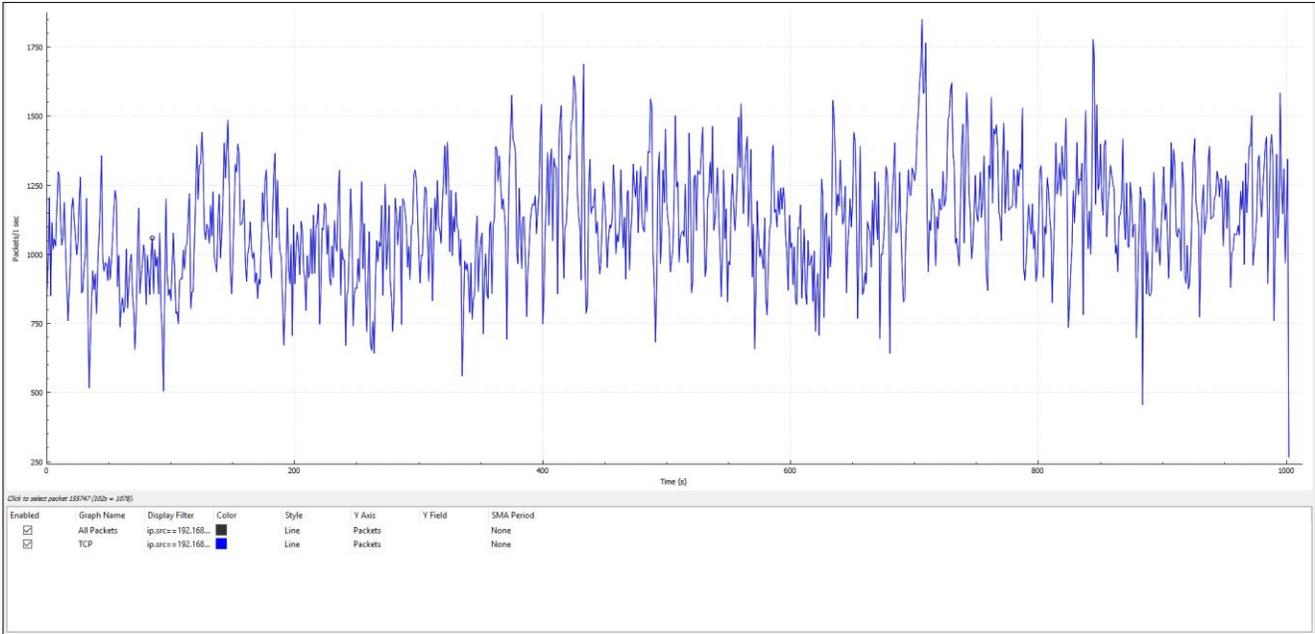
A continuación se muestran los múltiples tráficos UDP generados, para saturar el enlace:



Se puede observar, que dicha inundación de la red provoca la fluctuación de la conexión TCP:

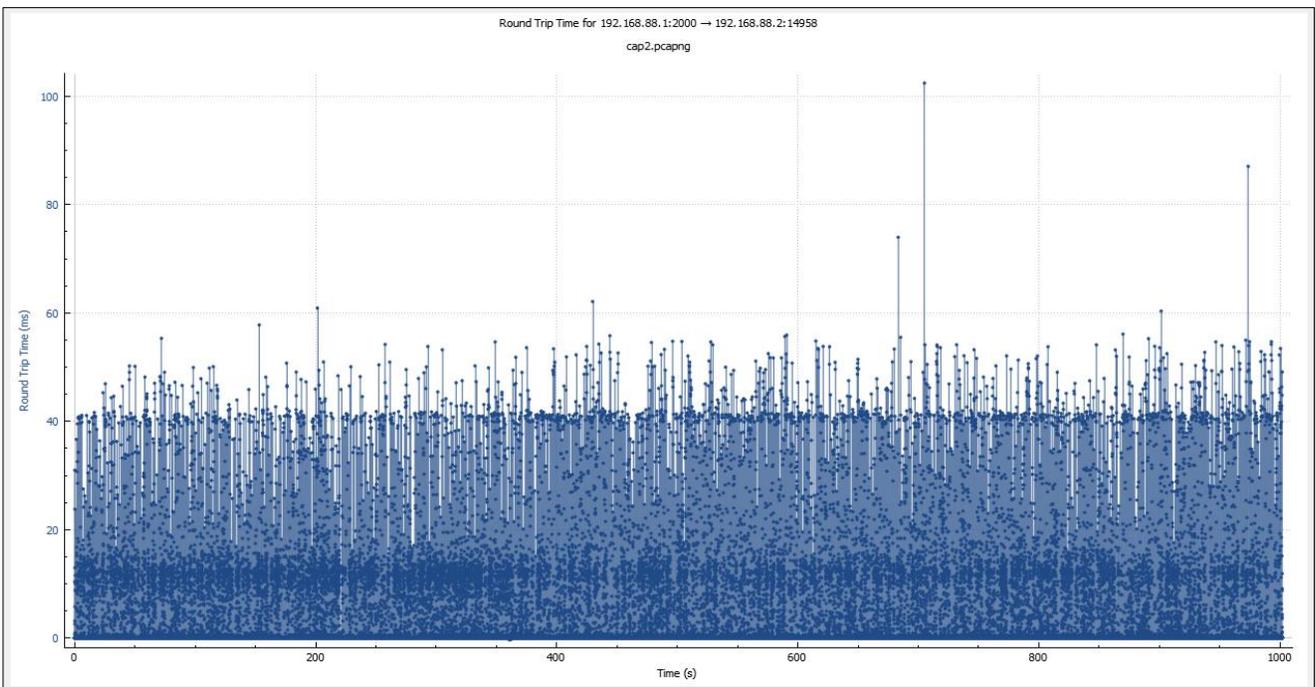


Representación del los paquetes por segundos recibidos por la interface de red durante los mil segundos:



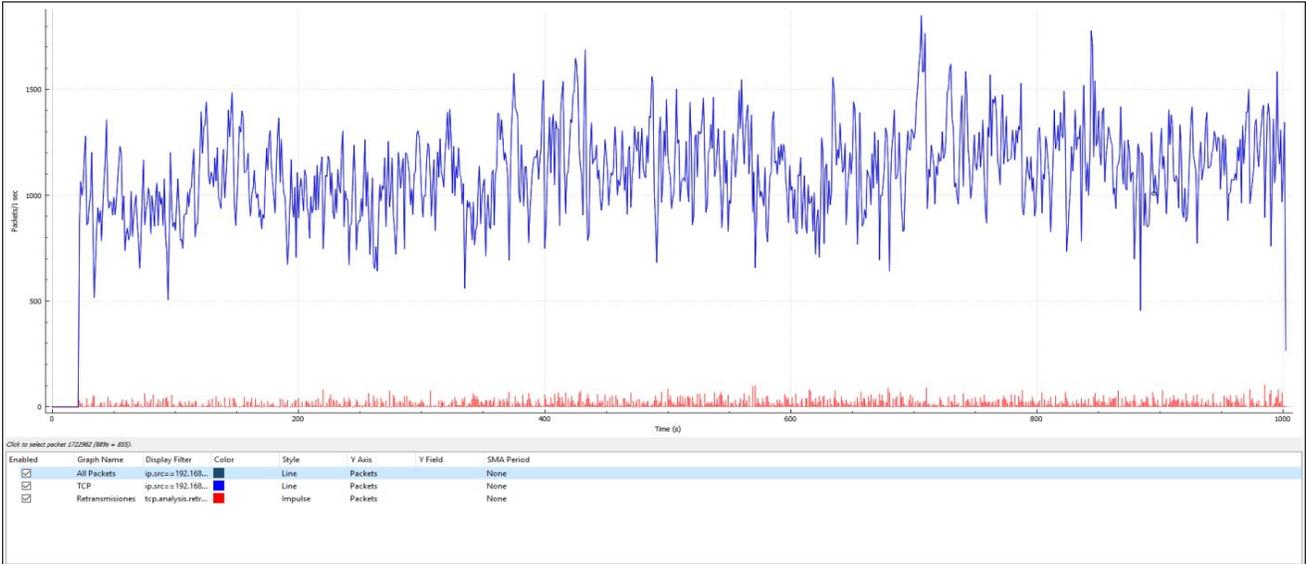
En las imágenes de cantidad de paquetes por segundos se presentan de color negro y todo el tráfico capturado, y de azul el tráfico TCP. Se verifica que la totalidad de datos transmitidos es TCP, dato que el tráfico TCP es equivalente a la totalidad del tráfico capturado.

En lo que respecta al tiempo de ida y vuelta (o en inglés el RTT, round-trip time), tal como se muestra en la imagen, se tienen los siguientes tiempos:



### 3.1.2.4 Comportamiento de TCP

A través del filtro de Wireshark, se observa en azul el tráfico TCP recibido y en rojo las retransmisiones de paquetes ocasionadas por vencimiento del RTO y por duplicidad de ACK. Esto es característico de una red esta estresada.



Para el escenario de estudio, los paquetes retransmitidos y duplicados representan el 1% de la totalidad de los paquetes entrantes a la interface.

ip.src==192.168.88.1 && tcp.analysis.retransmission

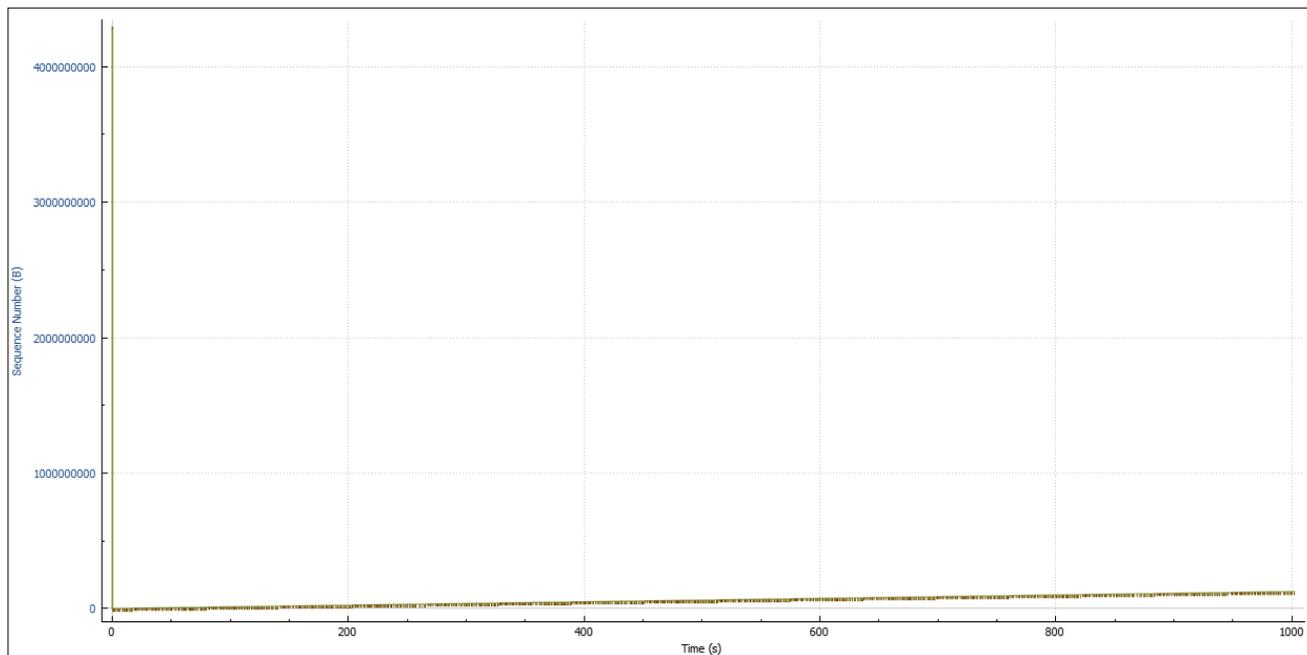
Time	Source	Destination	Protocol	Length	Info
33 0.014144	192.168.88.1	192.168.88.2	TCP	1514	[TCP Fast Retransmission] 2000 → 14947
71 0.050027	192.168.88.1	192.168.88.2	TCP	1514	[TCP Retransmission] 2000 → 14947
72 0.050027	192.168.88.1	192.168.88.2	TCP	254	[TCP Retransmission] 2000 → 14947
78 0.055848	192.168.88.1	192.168.88.2	TCP	1514	[TCP Fast Retransmission] 2000 → 14947
83 0.061417	192.168.88.1	192.168.88.2	TCP	1514	[TCP Retransmission] 2000 → 14947
85 0.061555	192.168.88.1	192.168.88.2	TCP	1514	[TCP Retransmission] 2000 → 14947
112 0.063904	192.168.88.1	192.168.88.2	TCP	1514	[TCP Retransmission] 2000 → 14947
114 0.064092	192.168.88.1	192.168.88.2	TCP	1514	[TCP Retransmission] 2000 → 14947
116 0.064344	192.168.88.1	192.168.88.2	TCP	1514	[TCP Retransmission] 2000 → 14947

Frame 33: 1514 bytes on wire (12112 bits), 1514 bytes captured (12112 bits) on interface \Device\NPF\_{9BD21AD...} Ethernet II, Src: Routerbo\_f1:ba:40 (d4:ca:6d:f1:ba:40), Dst: Micro-St\_83:39:8f (d8:cb:8a:83:39:8f)  
 Internet Protocol Version 4, Src: 192.168.88.1, Dst: 192.168.88.2  
 Transmission Control Protocol, Src Port: 2000, Dst Port: 14947, Seq: 4294954157, Ack: 1, Len: 1460  
 Data (1460 bytes)

58 02 07 d0 3a 63 80 59 b4 8c 10 63 99 27 50 10 X-...:c-Y ...c-'P-

Transmission Control Protocol (tcp), 20 byte(s) Packets: 1966227 · Displayed: 19114 (1.0%)

Graficando el número de secuencia contra el tiempo, a través de la herramienta *tcptrace* de la opción estadística de Wireshark, queda en evidencia que el tráfico está siendo limitado por venta de congestión, ya que es muy inferior a la ventana reportada por receptor:



Recordemos que la línea verde representa la ventana de recepción (control de flujo) y línea marrón representa la ventana de congestión (control de congestión).

### 3.1.2.5 Recolección de las muestras

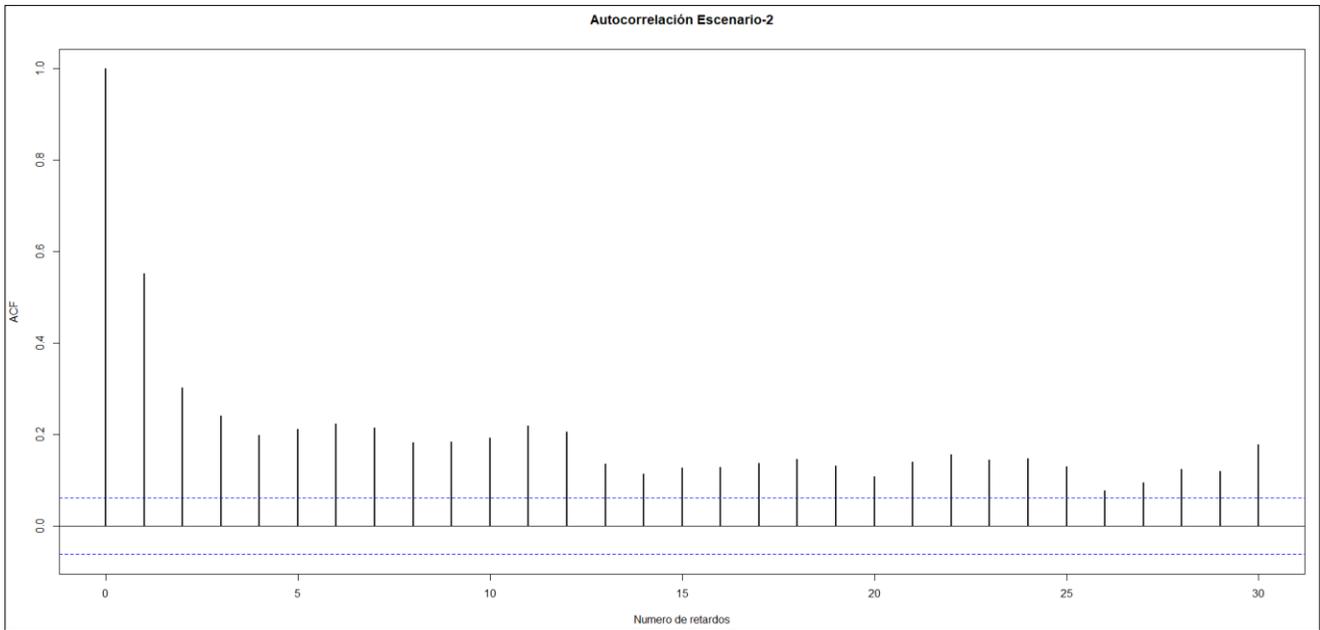
Equivalente al escenario-1, se capturan los paquetes y se le aplica un script para contabilizarlos.

El vector *muestra2* almacena la contabilización de los paquetes capturados en cada segundo, durante los mil segundos, y los vectores *m21*, *m22*, *m23*, *m24*, *m25*, *m26*, *m27*, *m28*, *m29* y *m210* almacenan segmentos de cien segundos consecutivos, permitiendo disponer de diez muestras.

### 3.1.2.6 Autocorrelación

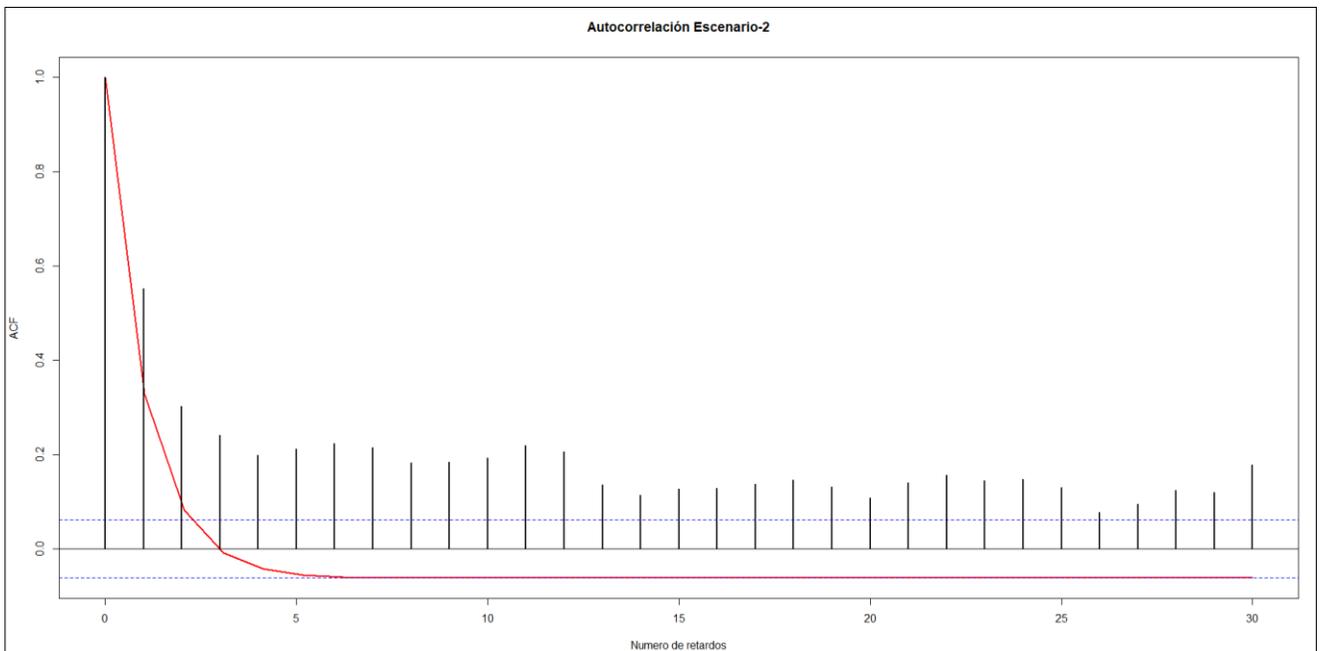
En el presente escenario también se supone que estamos frente a un proceso al menos débilmente estacionario (WSS) y Ergódico en la autocorrelación.

Se observa la autocorrelación temporal suponiendo que iguala a la del proceso:



A diferencia del escenario-1, donde el correlograma mostraba cambios pronunciados entre los desplazamientos, en el escenario-2 se observa que la autocorrelación decae lentamente entre los desplazamientos, lo que indica que la *muestra2* es más propensa a corresponderse con un comportamiento autosimilar.

Profundizando en la observación, la siguiente grafica compara la autocorrelación del escenario-2 con una curva representativa de un decaimiento exponencial:



Se evidencia, en la grafica anterior, que estamos frente a un proceso estocástico con dependencia a largo plazo. Al comparar la autocorrelación de la *muestra2* con un decaimiento exponencial (típico de procesos con dependencia a corto plazo), se observa una dependencia que persistente en el tiempo.

### 3.1.2.7 Autosimilaridad

Calculamos el coeficiente de Hurst que nos permitirá concluir si las muestras dadas contienen una componente autosimilar y en qué medida está presente.

Primero calculamos el coeficiente de Hurst para la muestra-2, de longitud mil (1000) segundos:

```
> hurstexp(muestra2)
Simple R/S Hurst estimation: 0.7429102
Corrected R over S Hurst exponent: 0.8538647
Empirical Hurst exponent: 0.8364729
Corrected empirical Hurst exponent: 0.8031184
Theoretical Hurst exponent: 0.5388682
```

Coficiente de Hurst para la muestra en el intervalo 1-100 segundos:

```
> hurstexp(m21)
Simple R/S Hurst estimation: 0.6328698
Corrected R over S Hurst exponent: 0.7250162
Empirical Hurst exponent: 0.5895436
Corrected empirical Hurst exponent: 0.6167474
Theoretical Hurst exponent: 0.5229461
```

Coficiente de Hurst para la muestra en el intervalo 101-200 segundos:

```
> hurstexp(m22)
Simple R/S Hurst estimation: 0.6969891
Corrected R over S Hurst exponent: 0.8406711
Empirical Hurst exponent: 0.6600676
Corrected empirical Hurst exponent: 0.6809724
Theoretical Hurst exponent: 0.5229461
```

Coficiente de Hurst para la muestra en el intervalo 201-300 segundos:

```
> hurstexp(m23)
Simple R/S Hurst estimation: 0.6276455
Corrected R over S Hurst exponent: 0.741815
Empirical Hurst exponent: 0.6022678
Corrected empirical Hurst exponent: 0.6282155
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 301-400 segundos:

```
> hurstexp(m24)
Simple R/S Hurst estimation: 0.6888198
Corrected R over S Hurst exponent: 0.8280453
Empirical Hurst exponent: 0.6631054
Corrected empirical Hurst exponent: 0.683773
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 401-500 segundos:

```
> hurstexp(m25)
Simple R/S Hurst estimation: 0.6168453
Corrected R over S Hurst exponent: 0.7159422
Empirical Hurst exponent: 0.6158315
Corrected empirical Hurst exponent: 0.6405009
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 501-600 segundos:

```
> hurstexp(m26)
Simple R/S Hurst estimation: 0.6137325
Corrected R over S Hurst exponent: 0.6904525
Empirical Hurst exponent: 0.6257969
Corrected empirical Hurst exponent: 0.6495655
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 601-700 segundos:

```
> hurstexp(m27)
Simple R/S Hurst estimation: 0.6394472
Corrected R over S Hurst exponent: 0.7797793
Empirical Hurst exponent: 0.6257597
Corrected empirical Hurst exponent: 0.6495316
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 701-800 segundos:

```
> hurstexp(m28)
Simple R/S Hurst estimation: 0.5945876
Corrected R over S Hurst exponent: 0.6219079
Empirical Hurst exponent: 0.6687871
Corrected empirical Hurst exponent: 0.6890179
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 801-900 segundos:

```
> hurstexp(m29)
Simple R/S Hurst estimation: 0.6592631
Corrected R over S Hurst exponent: 0.8617318
Empirical Hurst exponent: 0.583017
Corrected empirical Hurst exponent: 0.6108874
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 901-1000 segundos:

```
> hurstexp(m210)
Simple R/S Hurst estimation: 0.5670175
Corrected R over S Hurst exponent: 0.6254773
Empirical Hurst exponent: 0.570652
Corrected empirical Hurst exponent: 0.5998277
Theoretical Hurst exponent: 0.5229461
```

Para el caso de estudio, nuevamente el enfoque utilizado es el análisis R/S.

Se observa que para el escenario-2 el coeficiente de Hurst muestra un valor que oscilan entre  $H=0.567$  y  $H=0.697$ , con un valor medio de  $H=0.633$ . También se calculo el valor del coeficiente para la muestra de 1000 segundos, que dio como resultado  $H_2=0.743$  y que muestra un valor superior al coeficiente de Hurst de la *muestra1* que dio como resultado  $H_1=0.624$ .

Se observa que los resultados están en línea con los arrojados por la autocorrelación, y que los valores de coeficientes de Hurst obtenidos demuestran autosimilaridad de los datos.

### 3.1.2.8 Caracterización de las muestras

En el escenario-2, un subconjunto de las muestras corresponde a las retransmisiones ocasionadas por la congestión, y son precisamente estas retransmisiones en las que se profundizará el análisis.

Sabemos que las retransmisiones ocasionadas por vencimiento del RTO producen la ejecución de Slow-start, y que las retransmisiones causadas por la recepción ACKs duplicados producen la ejecución de FR/FR, por lo que se analizarán los dos tipos de retransmisiones.

En particular, el alcance del presente ítem se acota siguiente desarrollo:

- ✓ Tipos y cantidad de retransmisiones por cada muestra.
- ✓ Histograma de la dispersión de las retransmisiones.
- ✓ Efecto de las retransmisiones por cada muestra.

**Tipos y cantidad de retrasmisiones por cada muestra**

A continuación, la primera tabla representa una comparación entre la totalidad de los paquetes transmitidos contra las retrasmisiones presentes en cada muestra. La segunda representa las cantidades de retrasmisiones por tipo.

La finalidad de esta tabla de datos es visualizar y comparar las cantidades de retrasmisiones en las diferentes muestras con el objetivo de encontrar una relación con el coeficiente de Hurst calculado.

MUESTRAS	HURST	LARGO DE LA MUESTRA	RETRASMISIONES	RETRASMISIONES/MUESTRAS (%)
GENERAL	0.742	1110176	16773	1.51%
1	0.633	96731	1468	1.52%
2	0.697	106968	1077	1.01%
3	0.628	100411	1621	1.61%
4	0.689	109207	1632	1.49%
5	0.612	118073	1749	1.48%
6	0.613	114407	2112	1.85%
7	0.639	107674	2053	1.91%
8	0.594	124214	1449	1.17%
9	0.659	115573	1690	1.46%
10	0.567	116918	1921	1.64%

MUESTRAS	HURST	RETRASMISIONES	RTO (100s)	FRR (100s)	FRR / RTO (%)
GENERAL	0.742	19032	16773	2259	13.47%
1	0.633	1637	1468	169	11.51%
2	0.697	1228	1077	151	14.02%
3	0.628	1783	1621	162	9.99%
4	0.689	1847	1632	215	13.17%
5	0.612	2006	1749	257	14.69%
6	0.613	2374	2112	262	12.41%
7	0.639	2305	2053	252	12.27%
8	0.594	1709	1449	260	17.94%
9	0.659	1958	1690	268	15.86%
10	0.567	2184	1921	263	13.69%

En la primera tabla, si se toma las muestras 2 y 4 (que son las que presentan mayor autosimilaridad) y se las compara con las muestras 8 y 10 (que son las muestras con menor componente autosimilar) los datos no evidencian la causa de la autosimilaridad.

Si bien al comparar mas muestras 2 y 10 (que son las muestras de mayor y menor autosimilaridad) observamos una diferencia importante en las cantidades absolutas de retrasmisiones, y en consecuencia

podríamos sospechar que dicha diferencia puede estar asociada al comportamiento autosimilar, el supuesto pierde fuerza al comparar las muestras 4 y 8 (que son las segundas muestras con mayor y menor comportamiento autosimilar) ya que no se conserva dicha premisa en las cantidades absolutas de retransmisiones.

Respecto a las retransmisiones relativas al tamaño de la muestra, tampoco se observa un patrón que pueda inducir a pensar en una relación con la autosimilaridad de la muestra.

En la segunda tabla, si bien se observa que las dos primeras muestras con mayor coeficiente de Hurst coinciden con una menor cantidad de retransmisiones causadas por duplicidad de ACK, se considera la idea como inconsistente ya que al aplicar la lógica a la muestra 9, que tiene una componente autosimilar significativa en relación a las muestras 8 y 10, no cumple con dicho supuesto.

En conclusión, no se observa relación entre el grado de autosimilaridad con la cantidad de retransmisiones presentes en las muestras. Cabe aclarar, que todas las muestras se componen de aproximadamente un 1.5% de retransmisiones, por lo que no hay variaciones porcentuales relevantes para afirmar que dicho escenario se pueda traspolar a un ámbito de prueba con niveles de retransmisiones diferentes.

### **Histograma de la dispersión de las retransmisiones**

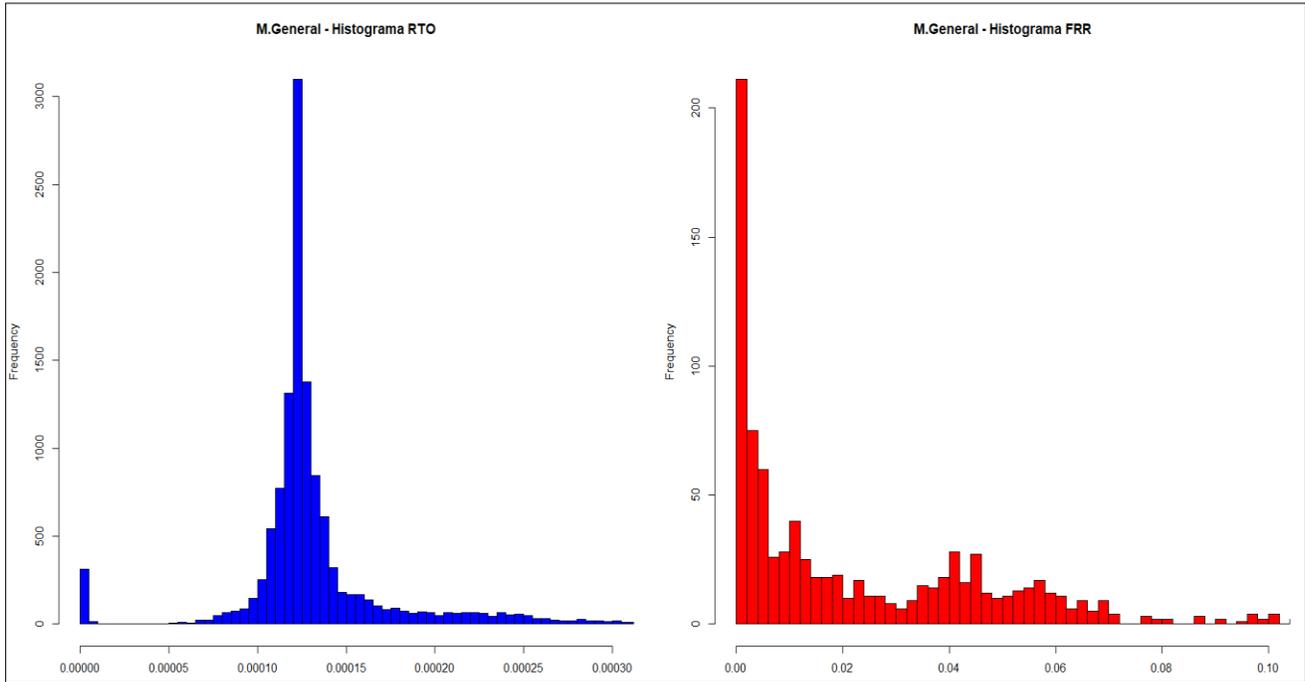
El histograma es una herramienta gráfica que nos presenta un panorama (primera vista) de la distribución de una muestra, y a través del cual podemos comprender fácilmente el comportamiento de datos estadísticos.

En particular, se calculan la dispersión de las retransmisiones. Se obtienen dos vectores, uno que contiene los intervalos de tiempos de las retransmisiones producidas por el vencimiento del RTO y otro que contiene los intervalos de tiempo de las retransmisiones producidas por duplicidad de ACKs.

Una vez calculado los intervalos de ocurrencias de las retransmisiones, a través del desarrollo de un script, se procede a graficar el histograma haciendo uso de RGui.

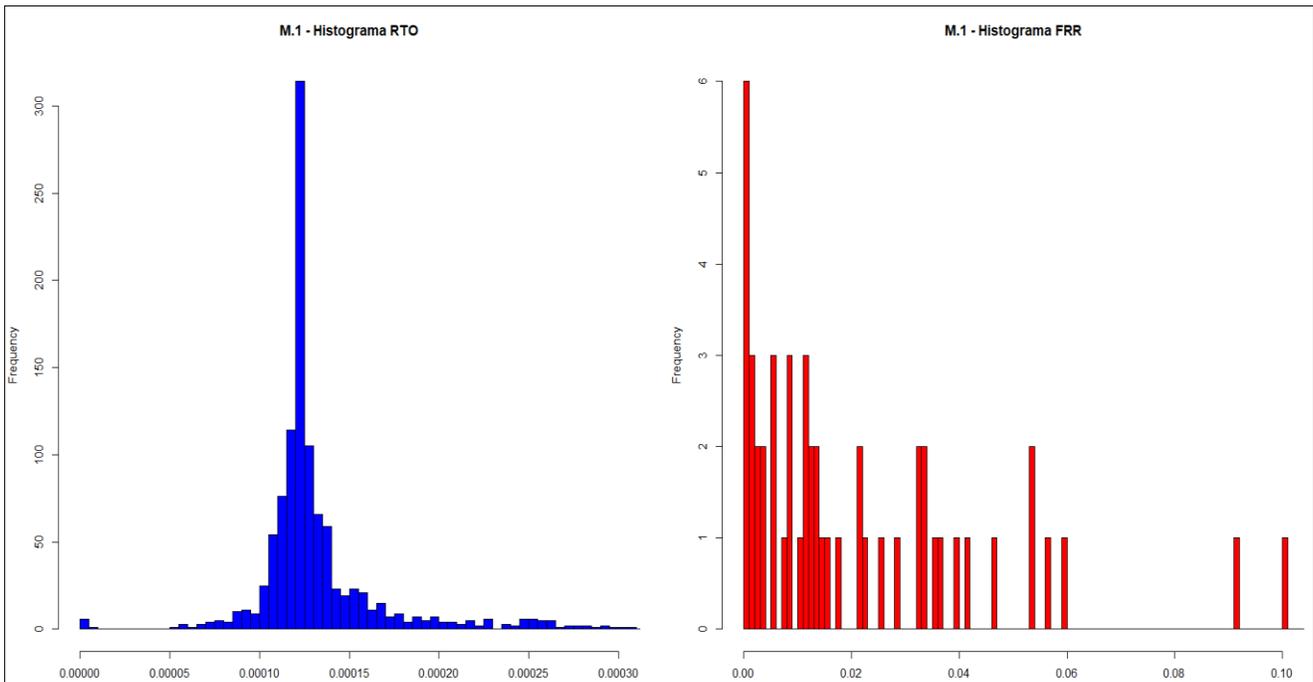
El histograma - Muestra general:

(Longitud 1000seg. - valor de Hurst=0.74)



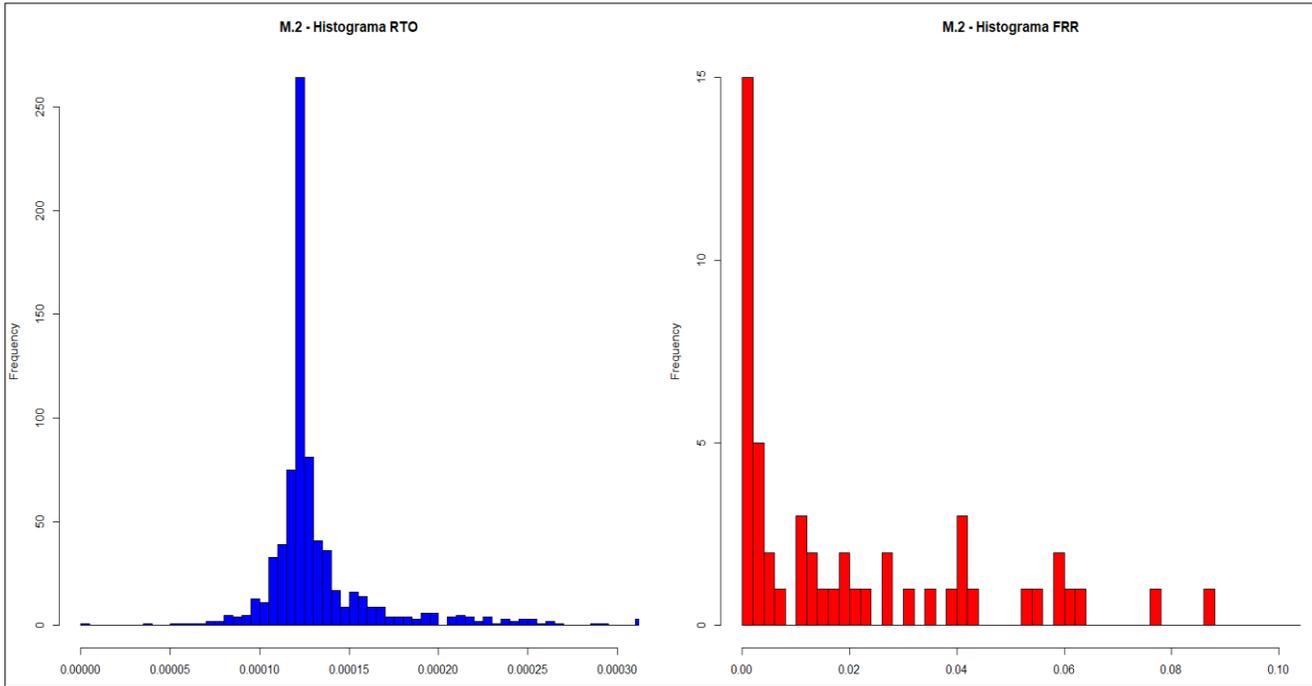
Histograma – Muestra-1:

(Longitud 100seg. - valor de Hurst=0.633)



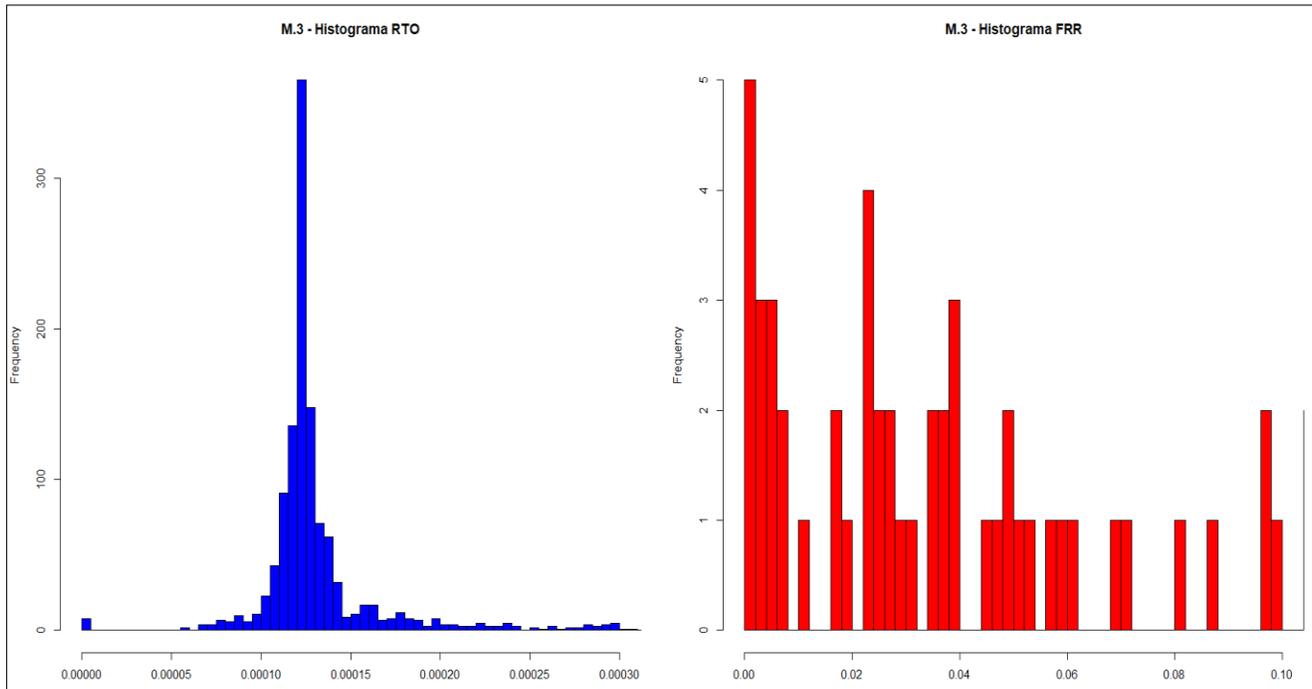
Histograma – Muestra-2:

(Longitud 100seg. - valor de Hurst=0.697)



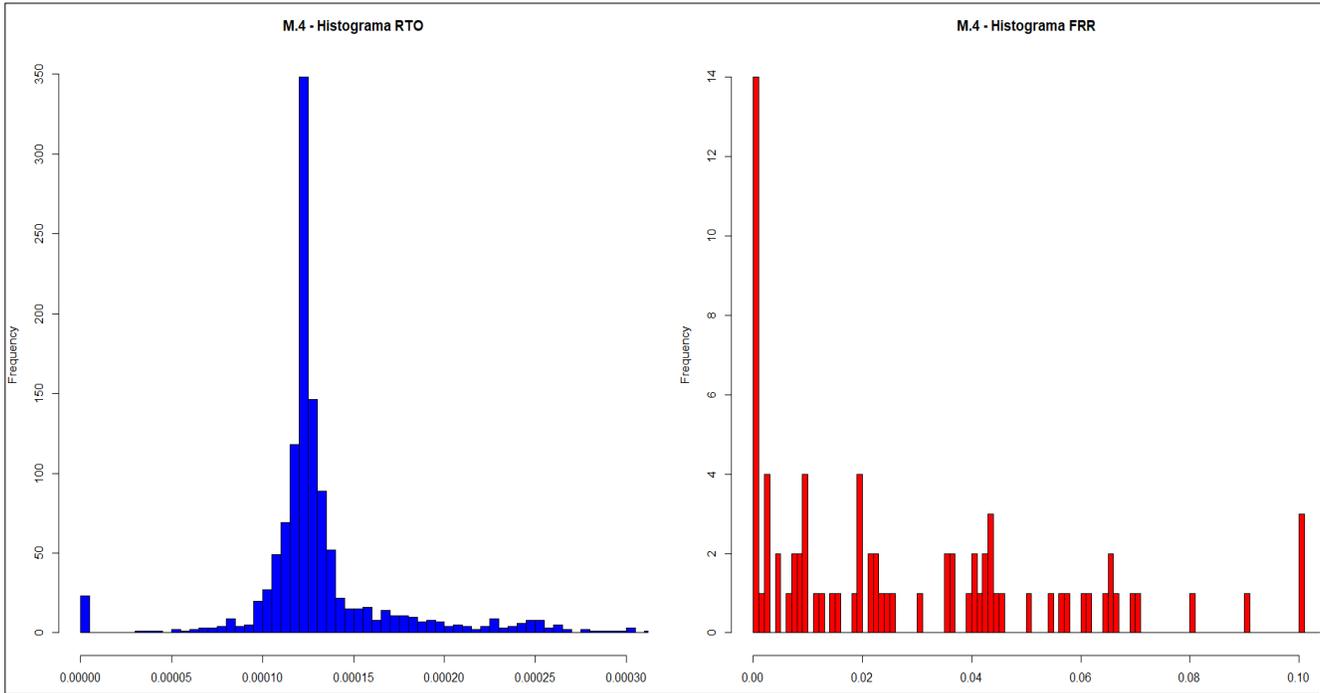
Histograma – Muestra-3:

(Longitud 100seg. - valor de Hurst=0.628)



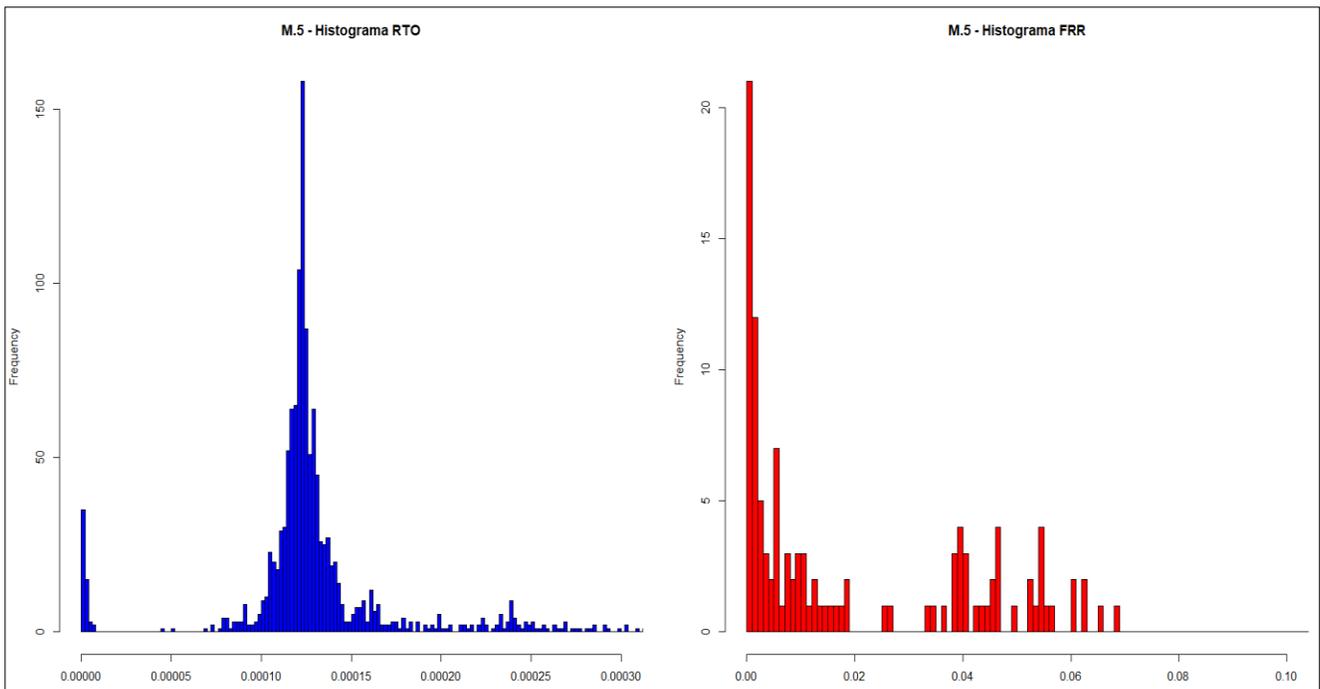
Histograma – Muestra-4:

(Longitud 100seg. - valor de Hurst=0.689)



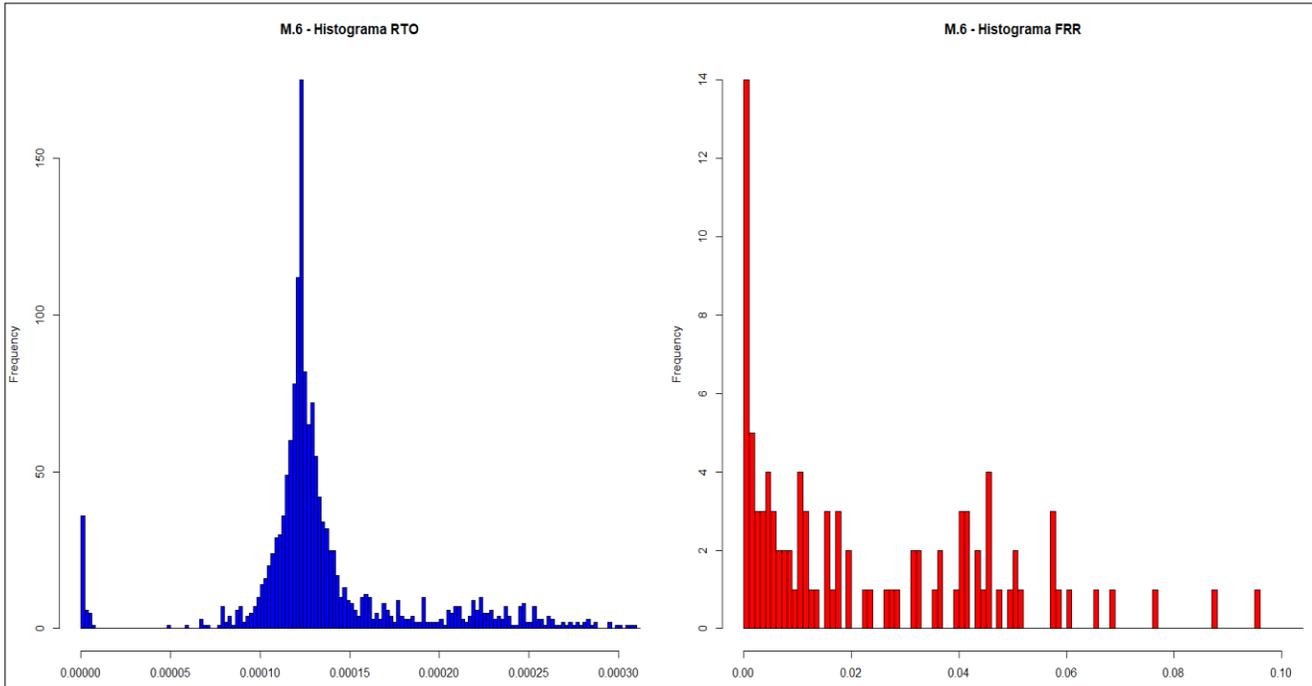
Histograma – Muestra-5:

(Longitud 100seg. - valor de Hurst=0.612)



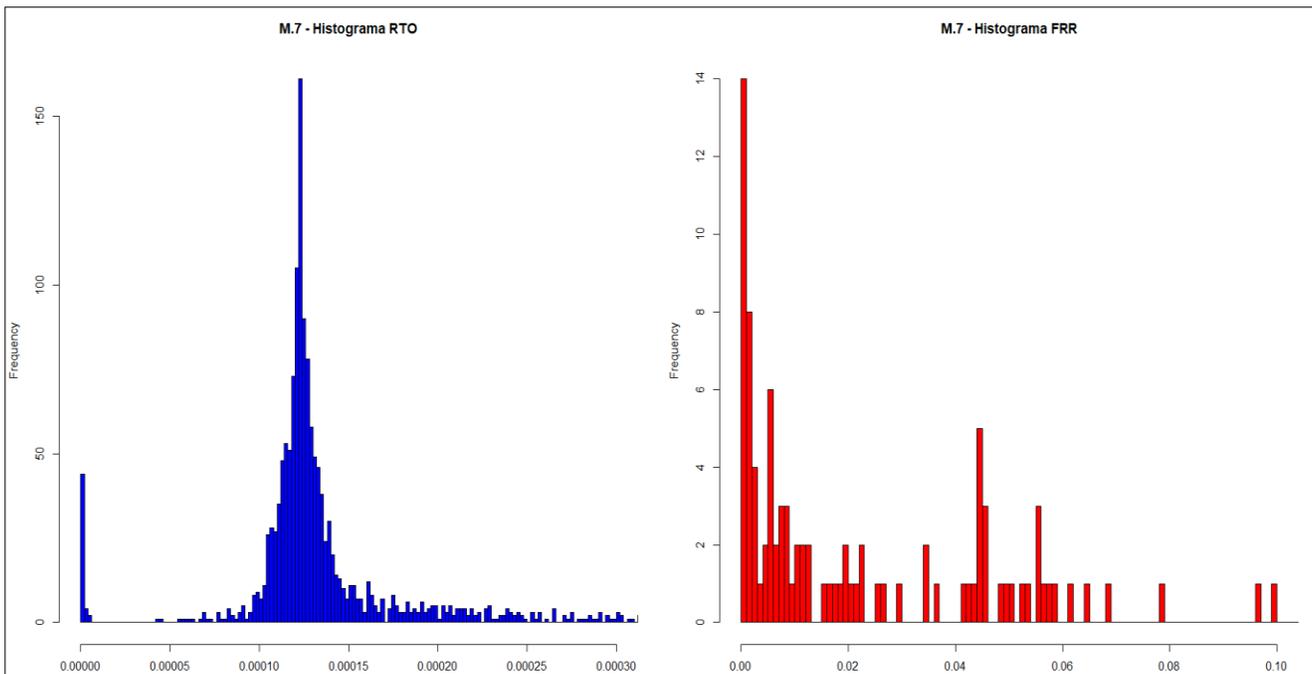
Histograma – Muestra-6:

(Longitud 100seg. - valor de Hurst=0.613)



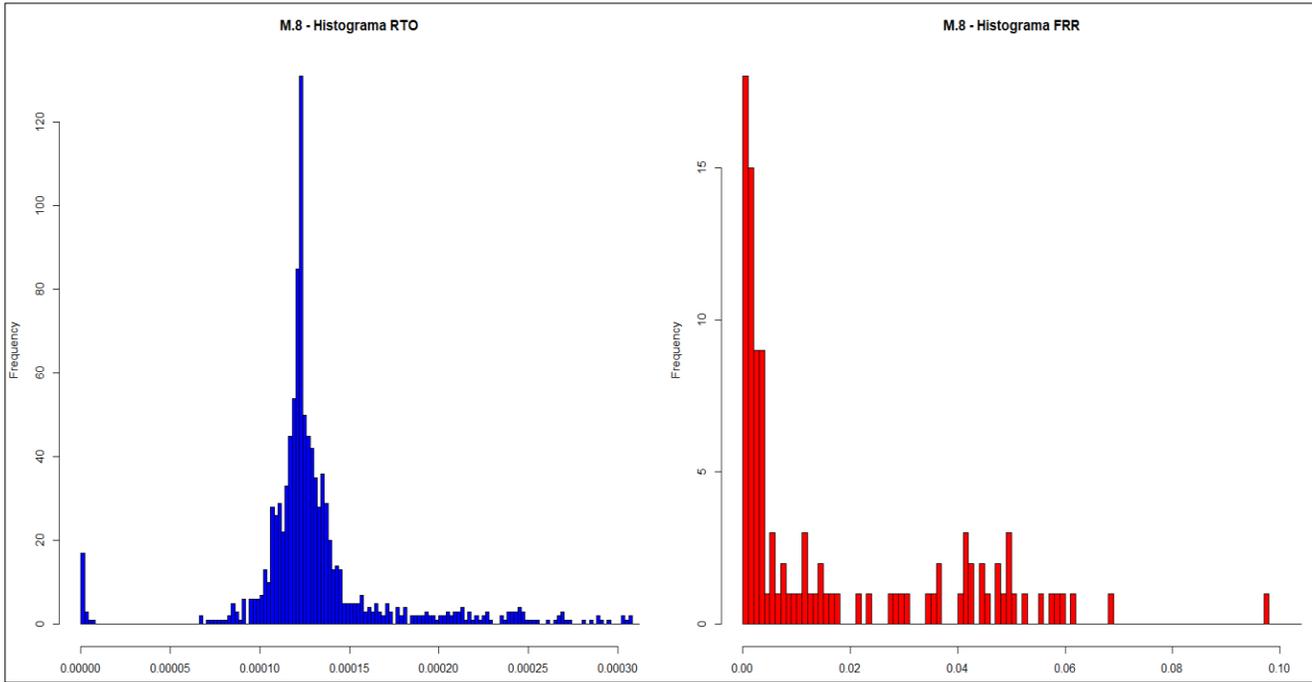
Histograma – Muestra-7:

(Longitud 100seg. - valor de Hurst=0.639)



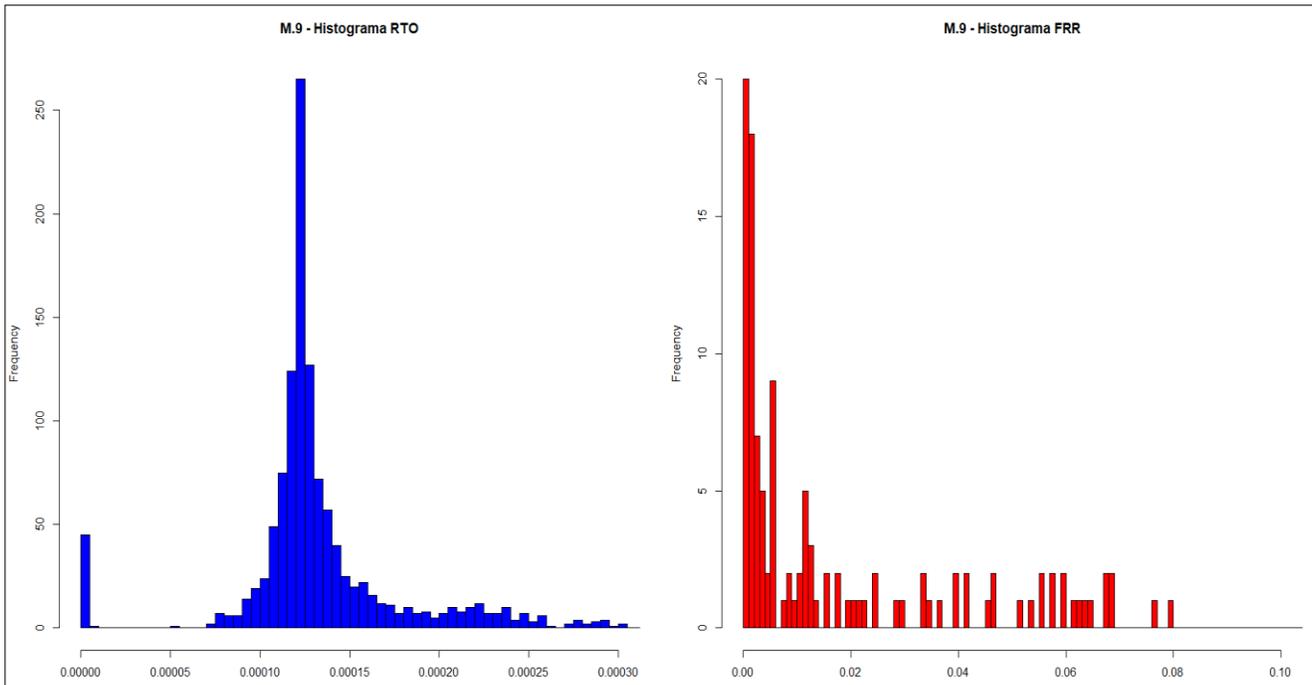
Histograma – Muestra-8:

(Longitud 100seg. - valor de Hurst=0.594)



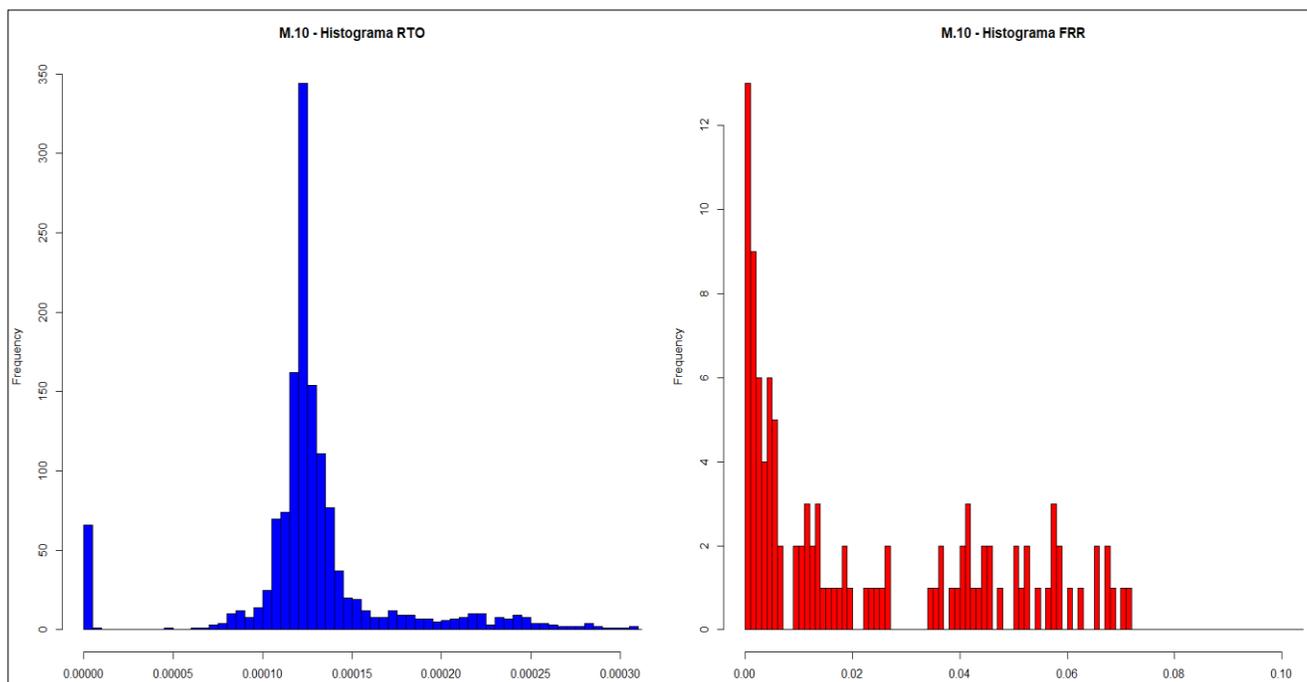
Histograma – Muestra-9:

(Longitud 100seg. - valor de Hurst=0.659)



### Histograma – Muestra-10:

(Longitud 100seg. - valor de Hurst=0.567)



Al analizar la distribución de los intervalos de ocurrencias, no se identifican patrones que puedan relacionarse con cambios en la componente autosimilar.

Si observamos las graficas, la distribución de la dispersión de las retrasmisiones causadas por vencimientos del RTO (que de hecho corresponden al 90% de las retrasmisiones) son muy parecidas, a tal punto que podría pensarse en que las retrasmisiones están sincronizadas aproximadamente cada 0.000123 segundos (valor obtenido a través del cálculo de la moda de las muestras).

Respecto a las retrasmisiones producto de ACK duplicados, las graficas tampoco evidencian relación con los cambios de la componente autosimilar.

#### 3.1.2.9 Interpretación de los resultados

El proceso de análisis llevado a cabo en el escenario-2 muestra cambios significativos en el comportamiento autosimilar del tráfico bajo condición de red estresada.

En este escenario, el protocolo TCP transmitió a 1Mbps con variaciones causadas por cambios en la ventana de control de congestión. En la transmisión se identificaron pérdidas de paquetes que son la causa de la ejecución de los algoritmos Slow-start y FR/FR, asociados al control de congestión.

El análisis realizado muestra que la congestión es al menos una condición necesaria para la existencia de una componente autosimilar en el tráfico de datos.

Me permití sospechar de que el grado de autosimilaridad podía estar vinculado con la cantidad de retransmisiones y su dispersión. Se estudiaron dichos supuestos en el escenario-2 pero no se encontró un patrón que permitiera confirmar dicha idea.

Es importante destacar que el escenario-2 presenta limitaciones prácticas, tales como poca variación en la cantidad de retransmisiones por muestra (entre el 1% y el 2%), poca variación del valor de RTT, envío de datos estandarizados (se utiliza una misma aplicación para generar de tráfico de todas las conexiones), una única conexiones TCP, entre otros aspectos.

El próximo paso es avanzar en el análisis de las cuestiones tratadas en los escenarios 1 y 2, pero utilizando muestras extraídas de una red operativa, intentando superar las limitaciones detalladas previamente y reconfirmando las afirmaciones conseguidas en esta primera etapa.

## **3.2 RED OPERATIVA**

En este contexto la recolección de las muestras se realiza desde un enlace de internet de 48Mbps disponible en la empresa Nucleoeléctrica Argentina S.A.

Durante el proceso de recolección de las muestras se monitorea el comportamiento del tráfico TCP con el afán de asegurar la conducta deseada y disponer de una muestra fehaciente y representativa del escenario de interés.

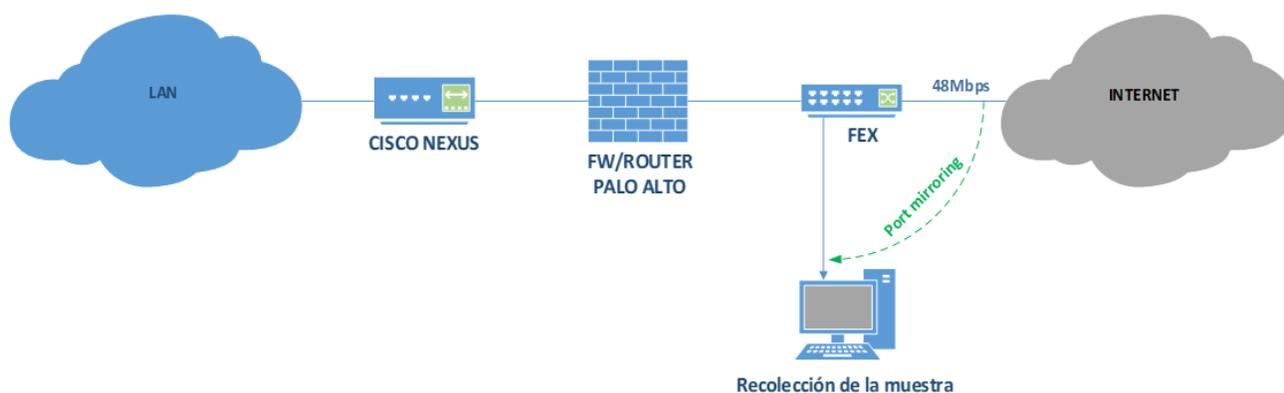
Al igual que en el ambiente controlado, nos basamos en la propiedad de una muestra autosemejante. Se recolectarán paquetes durante un periodo de mil segundos y se dividirán en segmentos de cien segundos, obteniendo diez muestras. Luego, se analizarán las diez muestras por separado en busca de un denominador común que permitan concluir la influencia de TCP en el comportamiento autosimilar del tráfico.

## 3.2.1 CONDICION DE RED RELAJADA: ESCENARIO-3

### 3.2.1.1 Ámbito de realización

Para recolectar la muestra se procede a configurar un puerto como espejo o en inglés port mirroring. En dicho puerto se espeja todo el tráfico que circula por el enlace de internet de 48Mbps.

Los equipos que intervienen son los siguientes:



El tráfico es generado por el consumo de servicios de internet que hacen los usuarios desde la red local. Se realiza la captura de la totalidad de los paquetes que atraviesan la interface, luego solo se filtran los paquetes que ingresan a la red local desde internet y en consecuencia se analizan solo los paquetes entrantes.

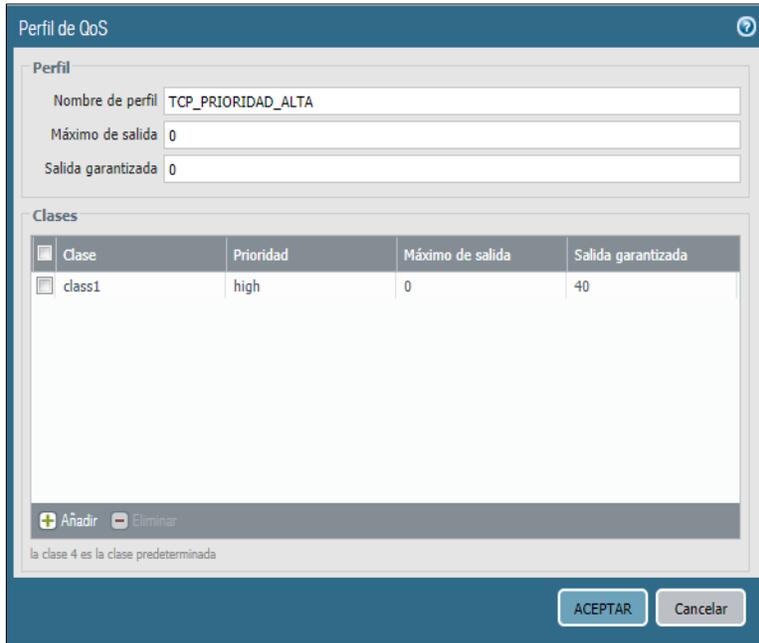
Se utiliza un perfil de calidad de servicio o en inglés quality of service (QoS) que prioriza las conexiones TCP y le garantiza más del 80% de ancho de banda disponible durante la recolección de las muestras, con la finalidad de evitar la ejecución del algoritmo slow-start.

Nuevamente, para capturar los paquetes de datos se utiliza el software Wireshark. Una vez recolectados son exportados en formato de texto plano y trabajados mediante un script con el objetivo de contabilizarlos y formatearlos para su procesamiento y análisis estadístico mediante el software R.

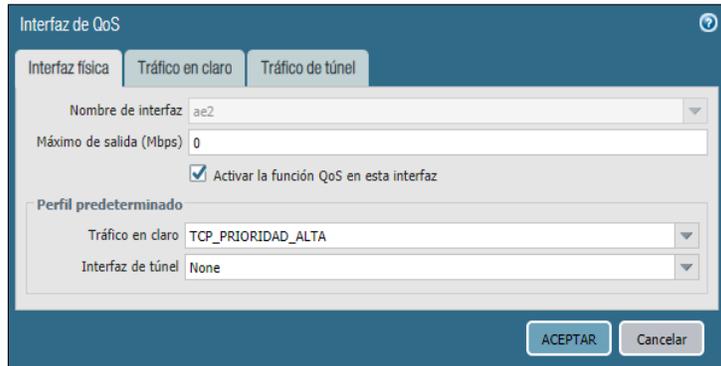
### 3.2.1.2 Calidad de servicio

La calidad de servicio es la capacidad que tiene la red para proveer diferentes niveles de servicios y asegurar distintos perfiles de tráfico.

En la recolección del escenario-3 se configura un perfil de QoS que prioriza el tráfico TCP, garantizando 40Mbps de ancho de banda.

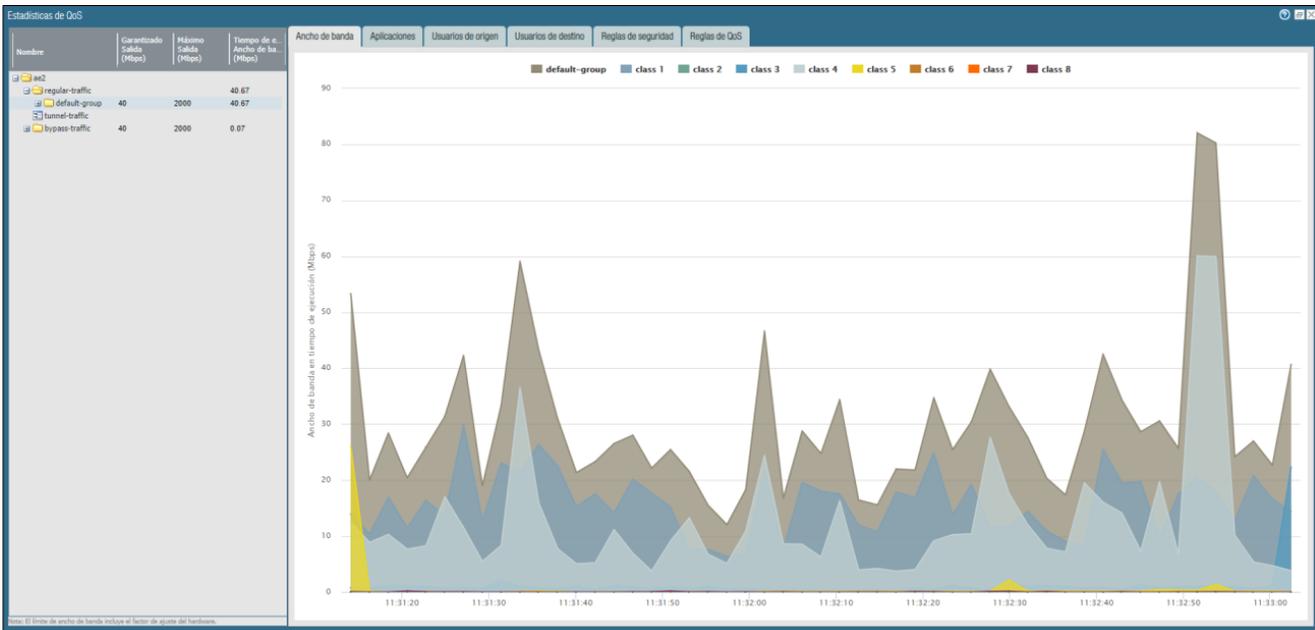
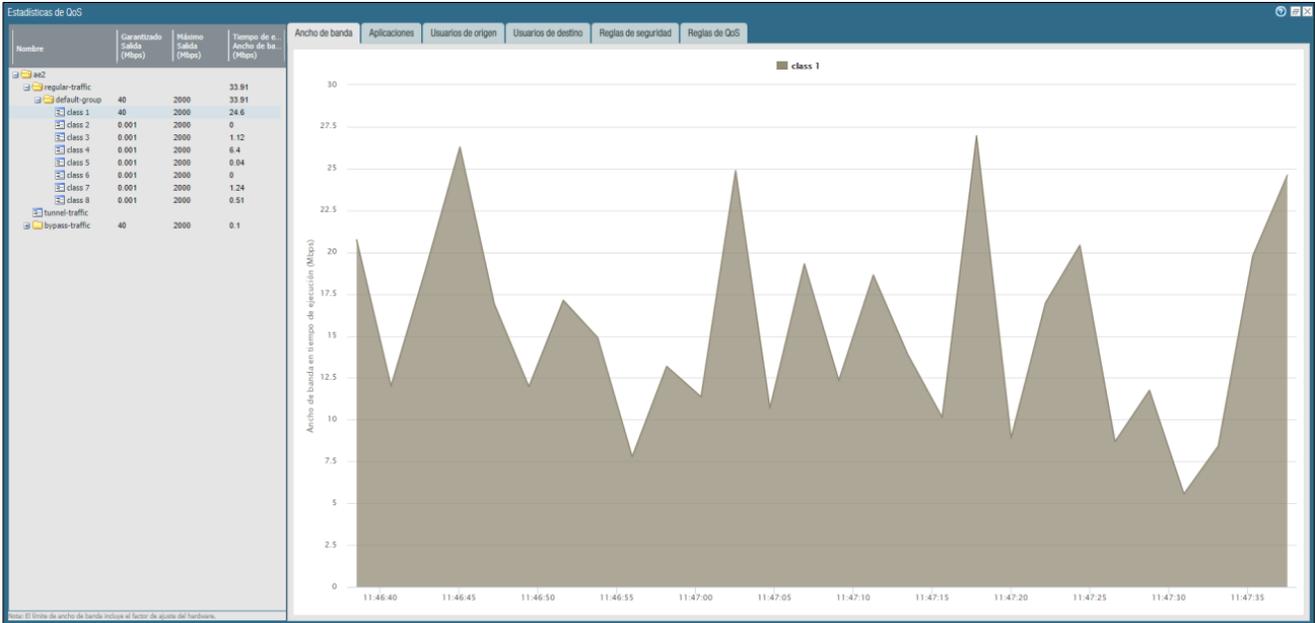


Dicho perfil es asignado a la interface de interés y se configura una política de QoS que hace efectivos los cambios.



Nombre	Etiquetas	IP Origen			IP Destino		Uso de reglas			Aplicación	Servicio	DSCP/ToS	Clase	Programación	
		Zona	Dirección	Usuario	Zona	Dirección	Recuento de aciertos	Último acierto	Primer acierto						
1	TCP	none	any	any	any	Internet	any	-	-	-	any	TCP	cualquiera	1	ninguno

La imagen siguiente muestra el tráfico TCP encolado con alta prioridad. Se observa que la red opera sin congestión, trabajando mayormente por debajo los 40Mbps garantizado y con picos que no superan los 48Mbps disponibles.



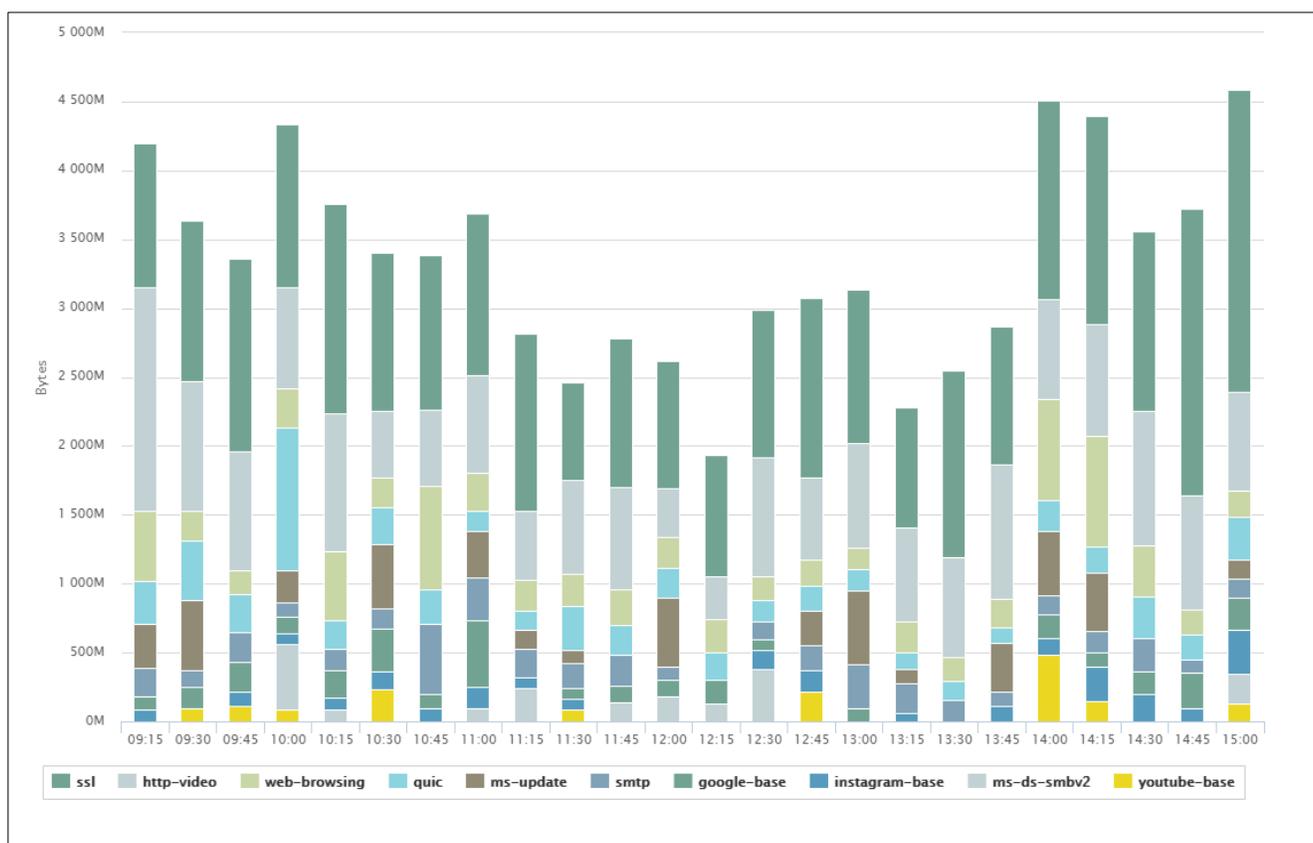
El monitoreo de la cola evidencia un escenario de red sin congestión. Esta interpretación es respaldada mediante los ítems 3.2.1.3 y 3.2.1.4 que se desarrollan a continuación.

### 3.2.1.3 Análisis de tráfico

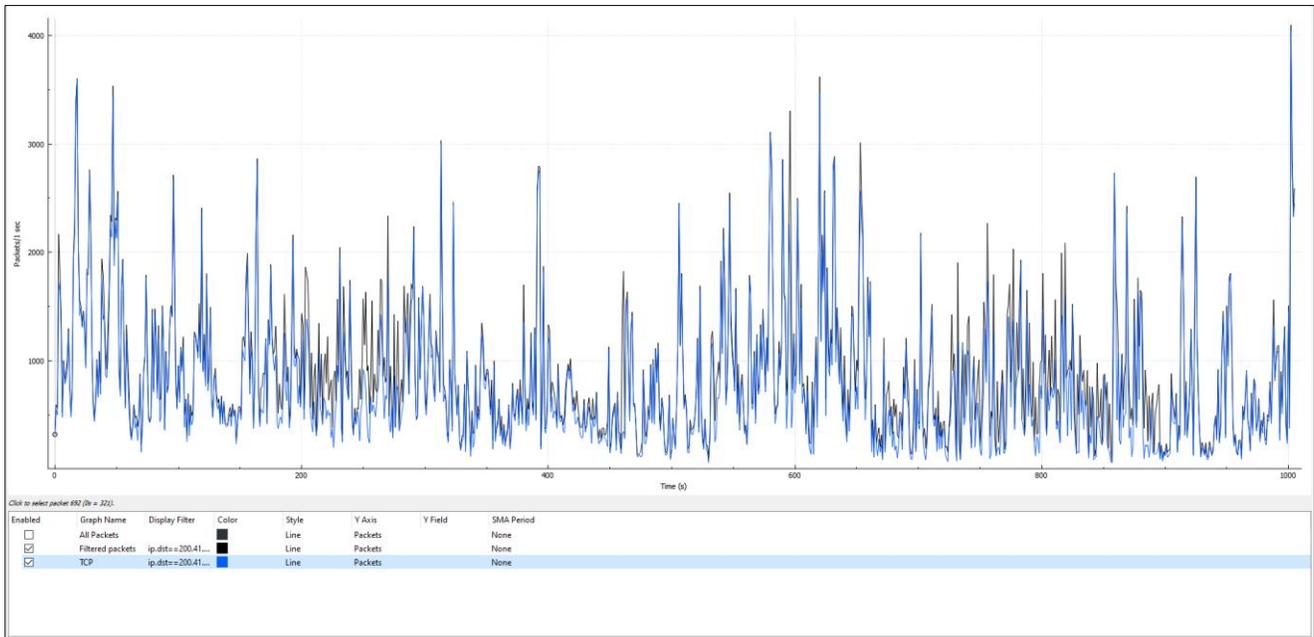
La supervisión de red muestra las TOP 10 aplicaciones de mayor consumo, donde se observa que el tráfico tiene una componente importante de TCP, siendo mayormente tráfico SSL, http-video y web-browsing los más utilizados.

El gráfico de supervisión de red es un estadístico tomado cada 15 minutos, por lo que es una representación histórica de los últimos consumos, y sirve para tener una idea general del tráfico que habitualmente circula por la red.

La supervisión de red detallada en este punto es válida para todas las muestras adquiridas de la red operativa, por lo que también aplica al escenario-4, que se analiza más adelante.



A continuación se observa el ancho de banda consumido para al momento de la recolección de las muestras.

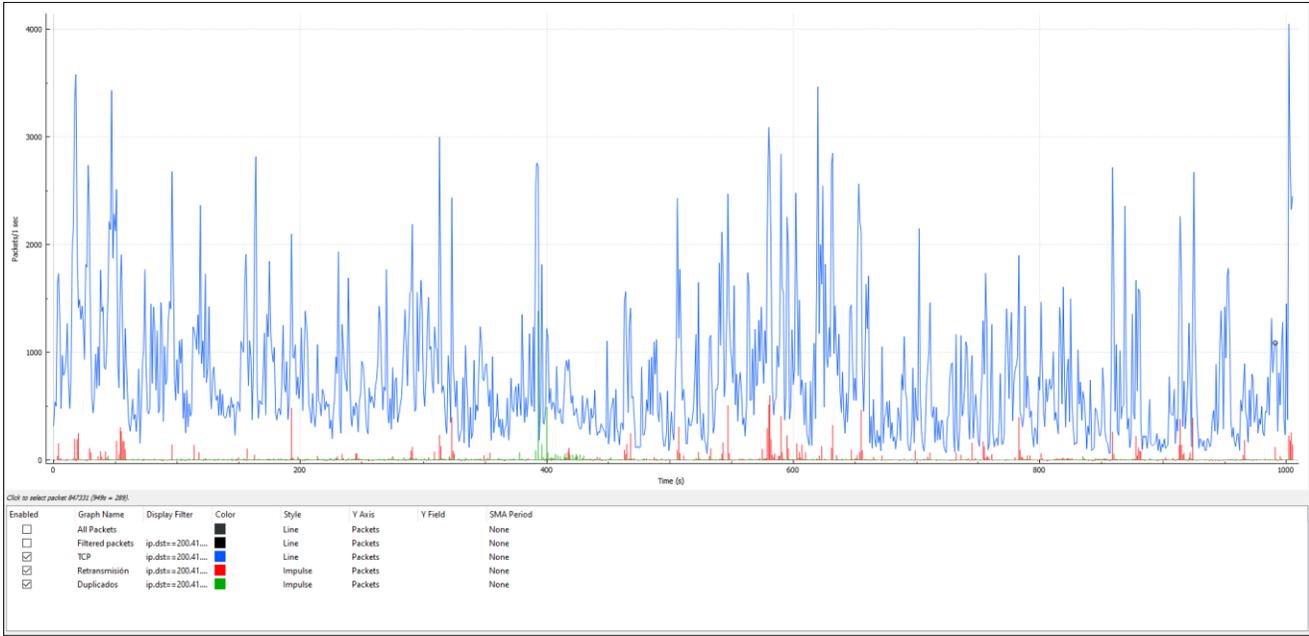


Se aprecia que, durante la recolección de las muestras, los 48Mbps disponibles en el enlace no se saturan y que el tráfico analizado es prácticamente TCP en su totalidad, garantizando desde el punto de vista del análisis de tráfico una colección de muestras representativa del escenario en cuestión.

### 3.2.1.4 Comportamiento de TCP

Si bien la ventana deslizante y su variación están disponibles en los extremos de la conexión, se puede tomar una referencia del comportamiento a través de los paquetes retransmitidos y duplicados, tal como se explicó en los escenarios previos.

En la gráfica siguiente se observan en azul todos los paquetes TCP que ingresan a la LAN, en rojo los paquetes retransmitidos y en verde los paquetes duplicados.



Los paquetes retransmitidos y duplicados representan el 2.4% de la totalidad de los paquetes entrantes a la LAN, para el escenario de estudio. Los paquetes duplicados solo se filtran para ver la ocurrencia de los mismos, pero no están necesariamente causados por retransmisiones, de hecho para que ocurra una retransmisión por duplicidad deben llegar 3 ACK consecutivos solicitando el mismo segmento. En este escenario no ha retransmisiones causadas por ACK duplicado, por lo que no hay eventos FR/FR.

Time	Source	Destination	Protocol	Length	Info
827	31.13.94.51	200.41.39.18	TCP	160	[TCP Retransmission] 443 → 443
1438	31.13.94.24	200.41.39.18	TCP	66	[TCP Retransmission] 443 → 269
1492	172.217.172.80	200.41.39.18	TCP	60	[TCP Retransmission] 443 → 259
1791	192.16.59.1	200.41.39.18	TCP	940	[TCP Retransmission] 443 → 264
2569	31.13.94.52	200.41.39.18	TCP	92	[TCP Retransmission] 443 → 167
2882	216.58.202.33	200.41.39.18	TCP	313	[TCP Retransmission] 443 → 113
3030	31.13.94.51	200.41.39.18	TCP	101	[TCP Retransmission] 443 → 246
3031	31.13.94.51	200.41.39.18	TCP	101	[TCP Retransmission] 443 → 246
3075	151.101.130.133	200.41.39.18	TCP	1494	[TCP Retransmission] 443 → 611

Frame 827: 160 bytes on wire (1280 bits), 160 bytes captured (1280 bits)  
 Ethernet II, Src: JuniperN\_e9:8f:2f (ec:3e:f7:e9:8f:2f), Dst: PaloAlto\_00:0a:01 (00:1b:17:00:0a:01)  
 Internet Protocol Version 4, Src: 31.13.94.51, Dst: 200.41.39.18  
 Transmission Control Protocol, Src Port: 443, Dst Port: 44368, Seq: 227, Ack: 1, Len: 94

```

000  00 1b 17 00 0a 01 ec 3e f7 e9 8f 2f 08 00 45 00  ...->.../..E-
010  00 92 d4 d5 40 00 58 06 e1 14 1f 0d 5e 33 c8 29  ...@.X. ....^3.)
020  27 12 01 bb ad 50 6d f0 29 3d bd cc 3d 14 80 18  '....Pm. )=-=...
030  01 da cd 0f 00 00 01 01 08 0a 18 1c 7c fc 01 26  ....|..&
    
```

cap3in.pcap | Packets: 894310 · Displayed: 21311 (2.4%)

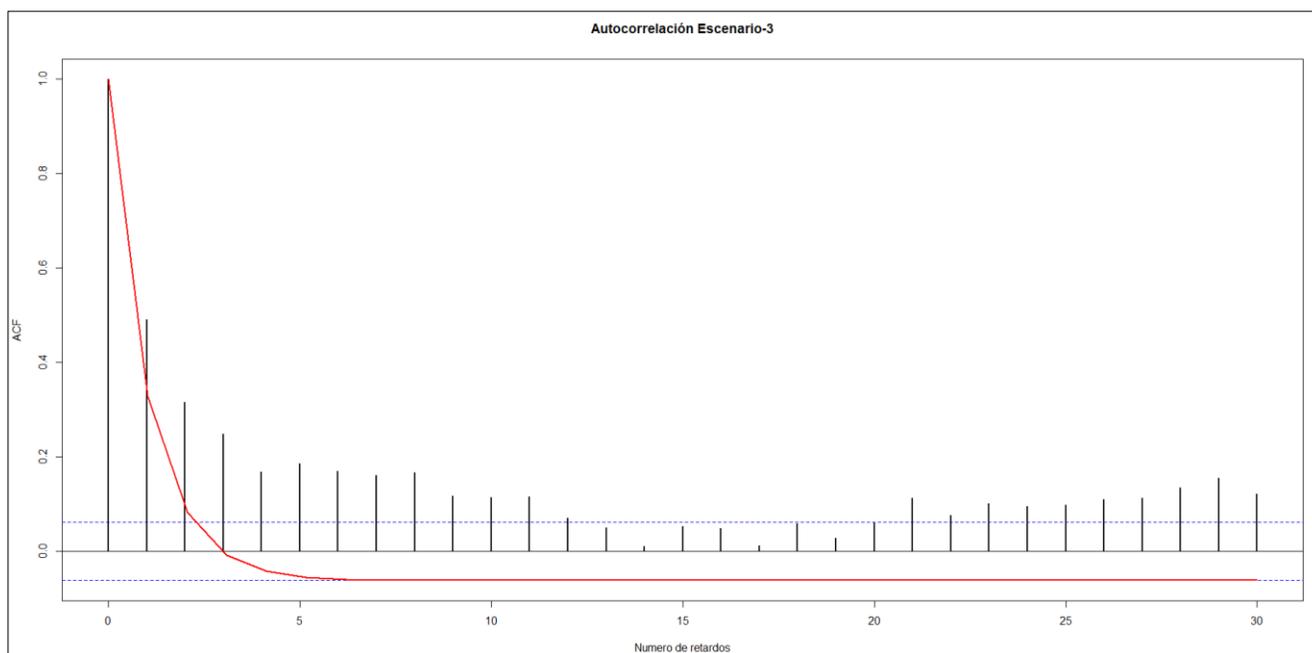
### 3.2.1.5 Recolección de las muestras

Se procesan los paquetes capturados, de la misma forma que en los escenarios previos, y luego se contabilizan los paquetes por segundos utilizando un script.

En el vector *muestra3* se almacena la contabilización de los paquetes capturados en cada segundo. Luego se secciona el vector *muestra3* en los vectores *m31*, *m32*, *m33*, *m34*, *m35*, *m36*, *m37*, *m38*, *m39* y *m310* y se obtienen las diez muestras a analizar.

### 3.2.1.6 Autocorrelación

Nuevamente se utiliza la función *ACF* de RGui, para realizar el cálculo de autocorrelación y se la compara con una curva representativa de un decaimiento exponencial:



Si bien estamos frente a una muestra obtenida de un escenario de red sin congestión, el escenario en cuestión se corresponde con una red operativa y es de esperar que existan retrasmisiones en el medio que provoquen ajustes en las ventanas deslizantes. Las retrasmisiones en este escenario deberían ser menores que en el escenario-4, correspondiente a una red operativa estresada.

El correlograma muestra que la autocorrelación decae más lento la función exponencial, por lo que es de esperar que el coeficiente de Hurst tenga valores mayores a 0.5.

### 3.2.1.7 Autosimilaridad

Calculamos el coeficiente de Hurst que nos permitirá concluir si las muestras dadas contienen una componente autosimilar y en qué medida está presente.

Primero calculamos el coeficiente de Hurst para la muestra-3, de longitud mil (1000) segundos:

```
> hurstexp(muestra3)
Simple R/S Hurst estimation: 0.6697019
Corrected R over S Hurst exponent: 0.7734603
Empirical Hurst exponent: 0.9015387
Corrected empirical Hurst exponent: 0.8652667
Theoretical Hurst exponent: 0.5388682
```

Coefficiente de Hurst para la muestra en el intervalo 1-100 segundos:

```
> hurstexp(m31)
Simple R/S Hurst estimation: 0.6995621
Corrected R over S Hurst exponent: 0.8123518
Empirical Hurst exponent: 0.6187229
Corrected empirical Hurst exponent: 0.6431276
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 101-200 segundos:

```
> hurstexp(m32)
Simple R/S Hurst estimation: 0.645133
Corrected R over S Hurst exponent: 0.7625578
Empirical Hurst exponent: 0.6364956
Corrected empirical Hurst exponent: 0.6593322
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 201-300 segundos:

```
> hurstexp(m33)
Simple R/S Hurst estimation: 0.6105514
Corrected R over S Hurst exponent: 0.6937115
Empirical Hurst exponent: 0.5532093
Corrected empirical Hurst exponent: 0.5843247
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 301-400 segundos:

```
> hurstexp(m34)
Simple R/S Hurst estimation: 0.6558835
Corrected R over S Hurst exponent: 0.8023654
Empirical Hurst exponent: 0.6312343
Corrected empirical Hurst exponent: 0.6545248
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 401-500 segundos:

```
> hurstexp(m35)
Simple R/S Hurst estimation: 0.5944035
Corrected R over S Hurst exponent: 0.6696473
Empirical Hurst exponent: 0.6685566
Corrected empirical Hurst exponent: 0.6888049
Theoretical Hurst exponent: 0.5229461
```

Coficiente de Hurst para la muestra en el intervalo 501-600 segundos:

```
> hurstexp(m36)
Simple R/S Hurst estimation: 0.6344032
Corrected R over S Hurst exponent: 0.7355596
Empirical Hurst exponent: 0.6353533
Corrected empirical Hurst exponent: 0.6582877
Theoretical Hurst exponent: 0.5229461
```

Coficiente de Hurst para la muestra en el intervalo 601-700 segundos:

```
> hurstexp(m37)
Simple R/S Hurst estimation: 0.7151837
Corrected R over S Hurst exponent: 0.9104837
Empirical Hurst exponent: 0.6815681
Corrected empirical Hurst exponent: 0.7008486
Theoretical Hurst exponent: 0.5229461
```

Coficiente de Hurst para la muestra en el intervalo 701-800 segundos:

```
> hurstexp(m38)
Simple R/S Hurst estimation: 0.5787682
Corrected R over S Hurst exponent: 0.6767111
Empirical Hurst exponent: 0.5706075
Corrected empirical Hurst exponent: 0.599788
Theoretical Hurst exponent: 0.5229461
```

Coficiente de Hurst para la muestra en el intervalo 801-900 segundos:

```
> hurstexp(m39)
Simple R/S Hurst estimation: 0.6060777
Corrected R over S Hurst exponent: 0.7363917
Empirical Hurst exponent: 0.6557309
Corrected empirical Hurst exponent: 0.676979
Theoretical Hurst exponent: 0.5229461
```

Coficiente de Hurst para la muestra en el intervalo 901-1000 segundos:

```
> hurstexp(m310)
Simple R/S Hurst estimation: 0.5674416
Corrected R over S Hurst exponent: 0.5831255
Empirical Hurst exponent: 0.6812594
Corrected empirical Hurst exponent: 0.7005623
Theoretical Hurst exponent: 0.5229461
```

Para el caso de estudio, nuevamente el enfoque utilizado es el análisis R/S.

Se observa que para el escenario-3 el coeficiente de Hurst muestra un valor que oscilan entre  $H=0.567$  y  $H=0.715$ , con un valor medio de  $H=0.629$ . También se calculo el valor del coeficiente para la muestra de 1000 segundos, que dio como resultado  $H_2=0.669$ .

### 3.2.1.8 Caracterización de las muestras

El alcance del presente ítem se acota siguiente desarrollo:

- ✓ Tipos y cantidad de retransmisiones por cada muestra.
- ✓ Histograma de la dispersión de las retransmisiones.
- ✓ Efecto de las retransmisiones por cada muestra.

#### Tipos y cantidad de retransmisiones por cada muestra

La tabla siguiente representa una comparación entre la totalidad de los paquetes transmitidos contra las retransmisiones presentes en cada muestra.

MUESTRAS	HURST	LARGO DE LA MUESTRA	RETRANSMISIONES	RETRANSMISIONES/MUESTRAS (%)
GENERAL	0.669	880435	20486	2.32%
1	0.699	124202	3069	2.47%
2	0.645	93319	1216	1.3%
3	0.61	101987	841	0.82%
4	0.655	79039	1467	1.85%
5	0.594	60956	1359	2.22%
6	0.634	100935	4731	4.68%
7	0.715	96670	2470	2.55%
8	0.578	77840	1802	2.31%
9	0.606	82374	1493	1.81%
10	0.567	63572	2039	3.2%

El escenario-3 todas las retransmisiones son causadas por vencimiento del RTO, es decir, las retransmisiones por duplicidad de ACK son cero para todas las muestras. Esto implica que en el control de congestión solo se ejecuto el algoritmo slow-start.

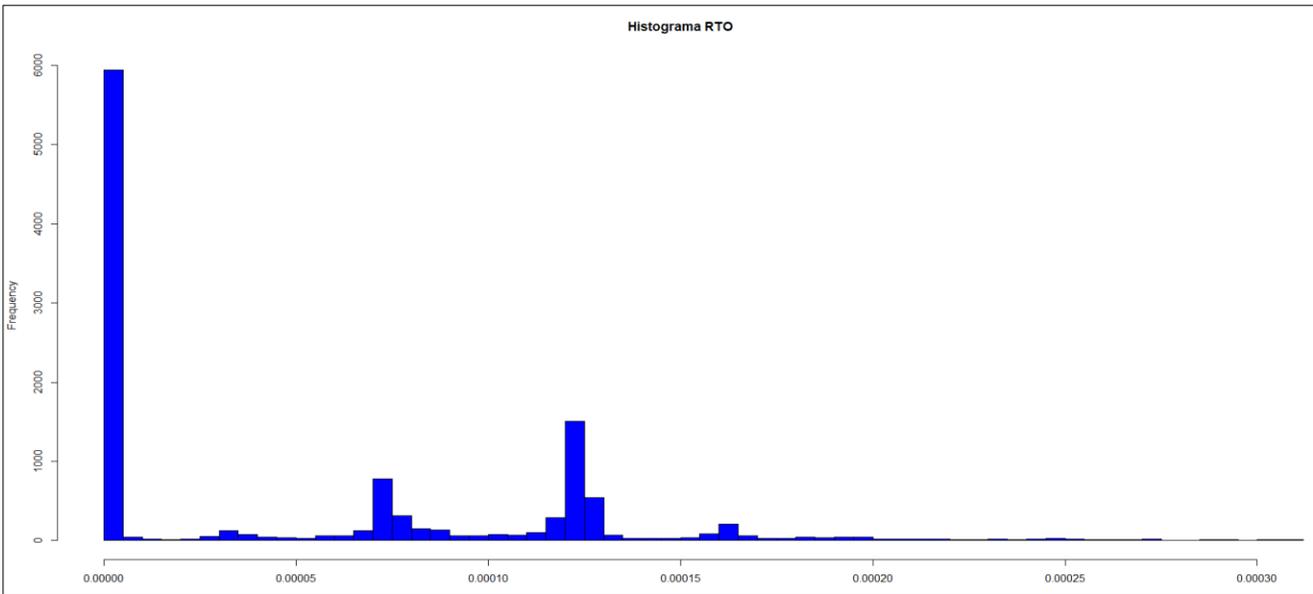
Al analizar la tabla anterior, se resalta en color verde las dos muestras con mayor autosimilaridad y en color rojo las dos muestras con menor autosimilaridad. No se observa relación entre el grado de autosimilaridad con la cantidad de retransmisiones presentes en las muestras.

### Histograma de la dispersión de las retransmisiones

En línea con el análisis realizado en el escenario-2 y con la intención de encontrar un patrón que podamos relacionar con el grado de autosimilaridad, se procede a graficar el histograma de la dispersión de las retransmisiones haciendo uso de RGui.

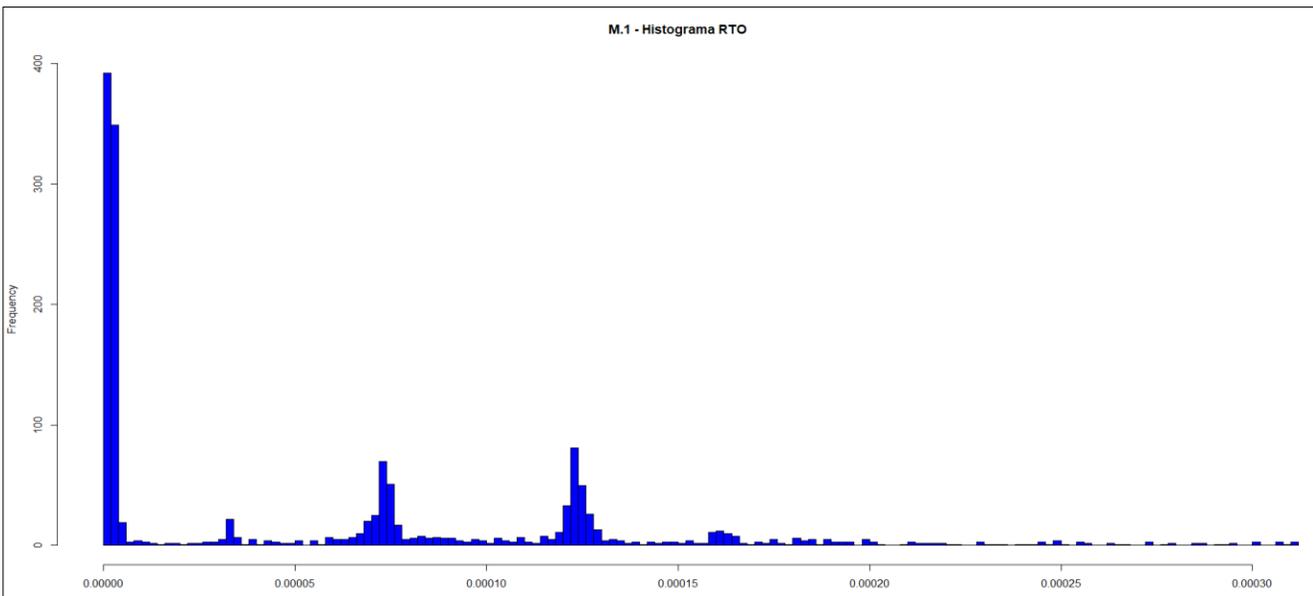
#### El histograma - Muestra general:

(Longitud 1000seg. - valor de Hurst=0.669)



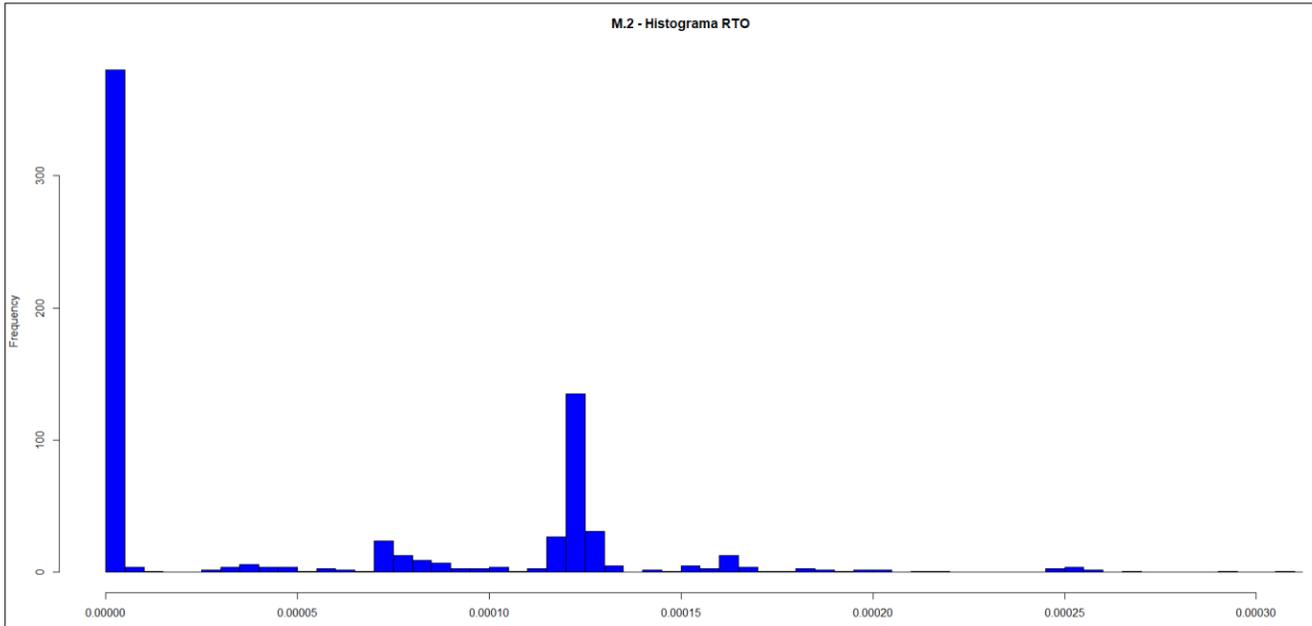
#### Histograma – Muestra-1:

(Longitud 100seg. - valor de Hurst=0.699)



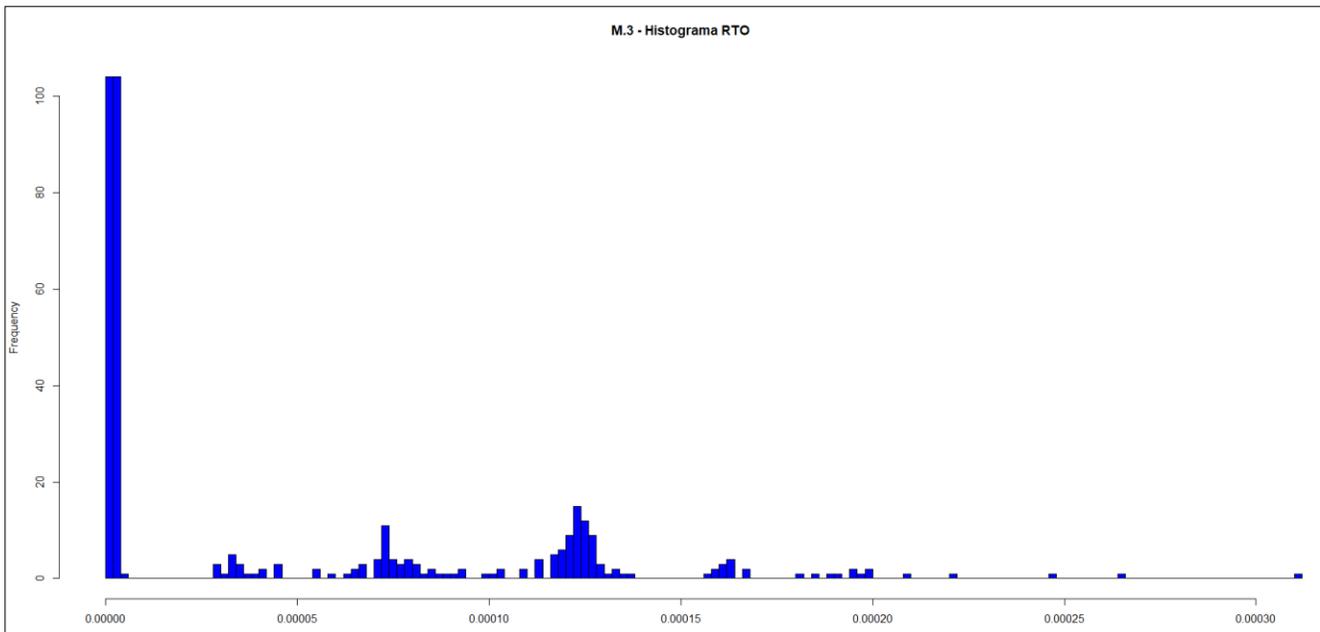
Histograma – Muestra-2:

(Longitud 100seg. - valor de Hurst=0.645)



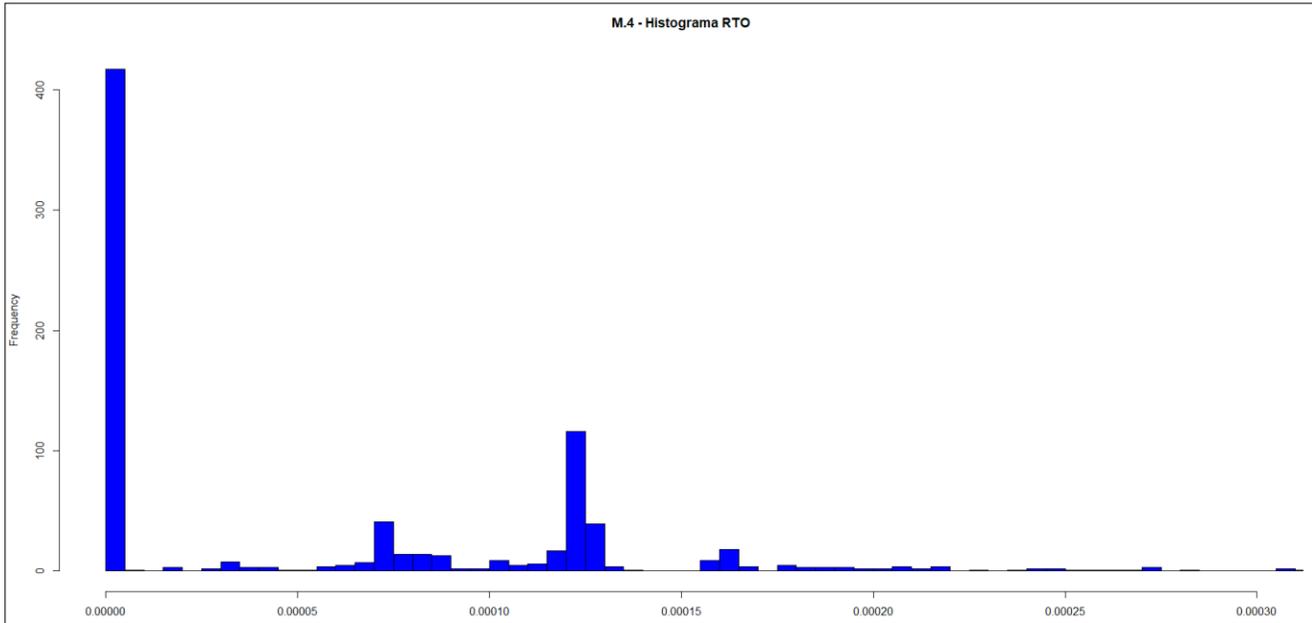
Histograma – Muestra-3:

(Longitud 100seg. - valor de Hurst=0.61)



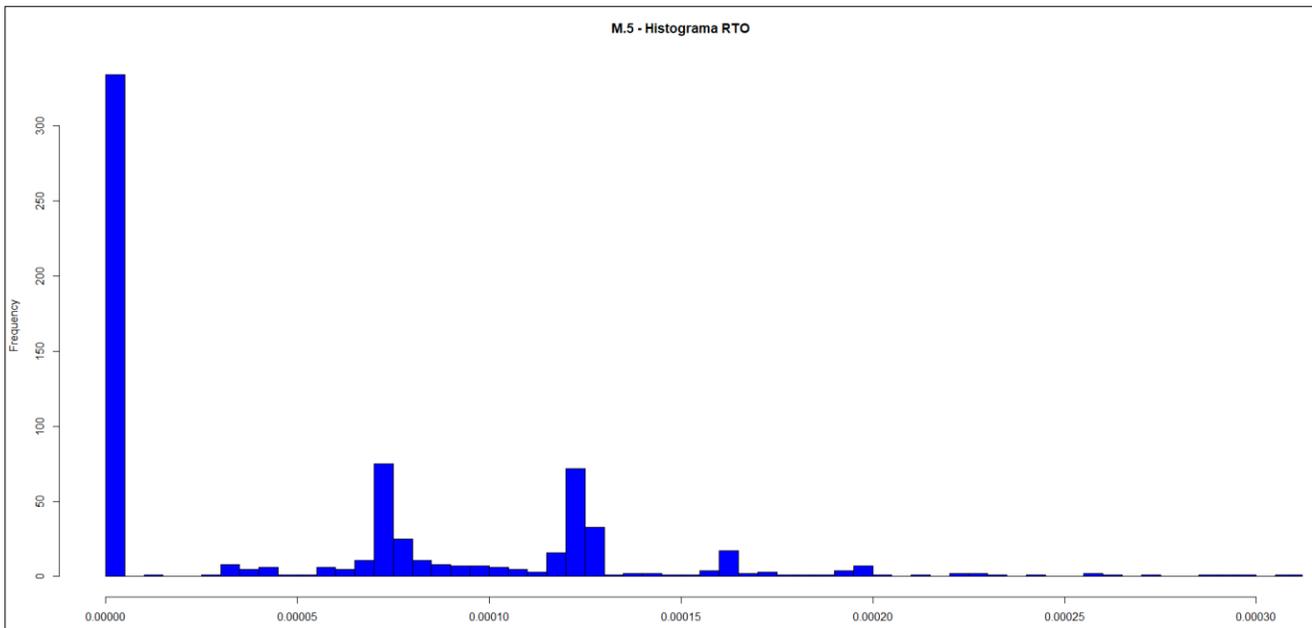
Histograma – Muestra-4:

(Longitud 100seg. - valor de Hurst=0.655)



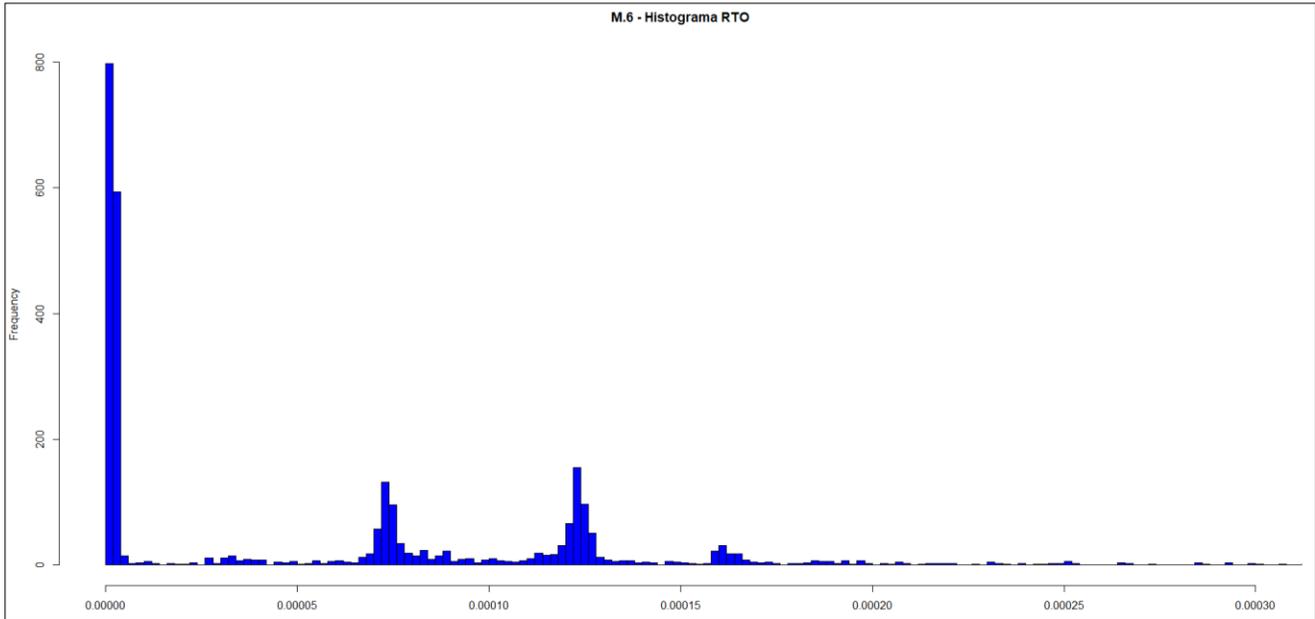
Histograma – Muestra-5:

(Longitud 100seg. - valor de Hurst=0.594)



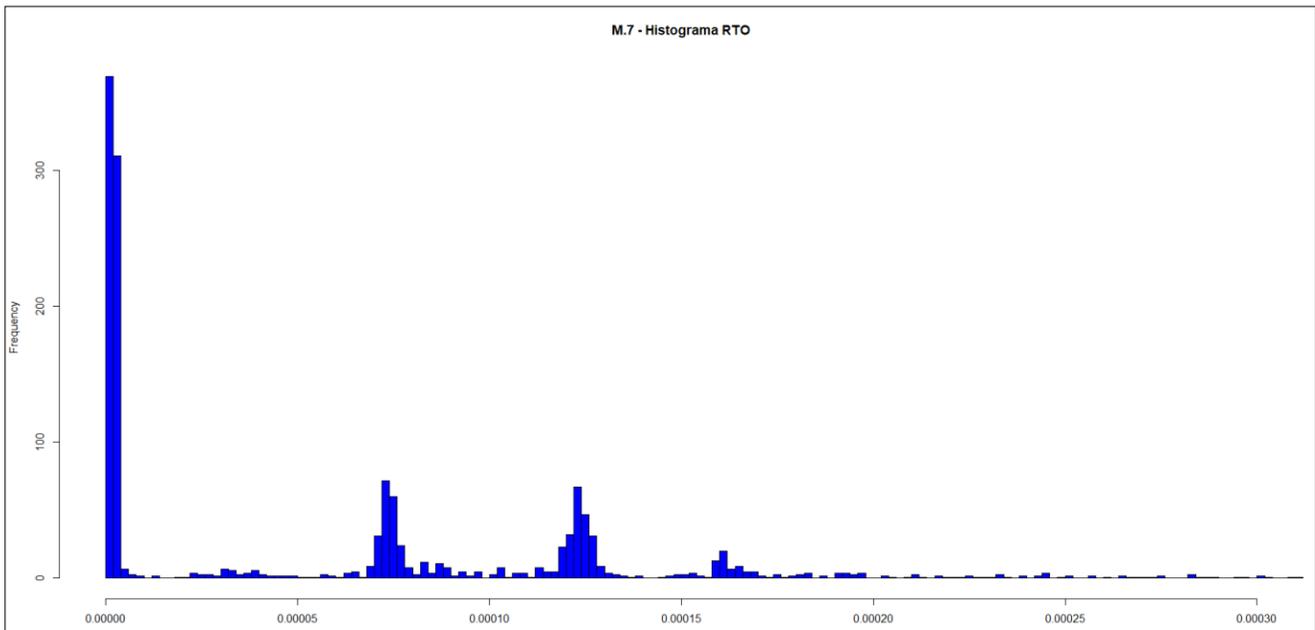
Histograma – Muestra-6:

(Longitud 100seg. - valor de Hurst=0.634)



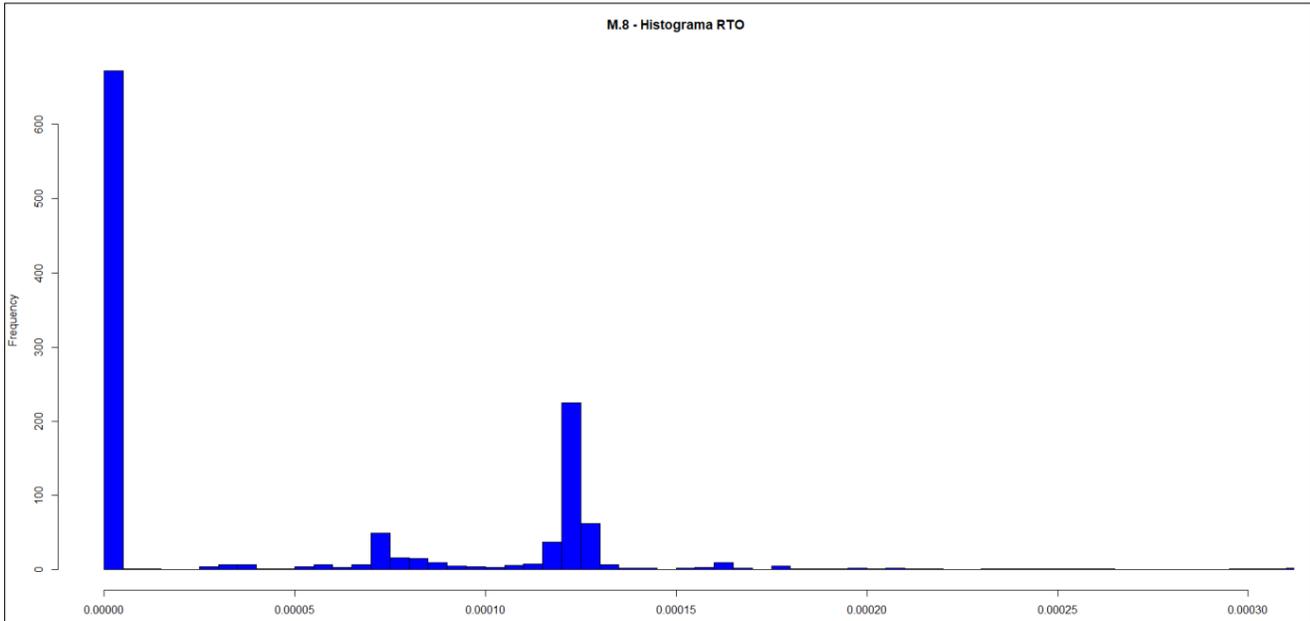
Histograma – Muestra-7:

(Longitud 100seg. - valor de Hurst=0.715)



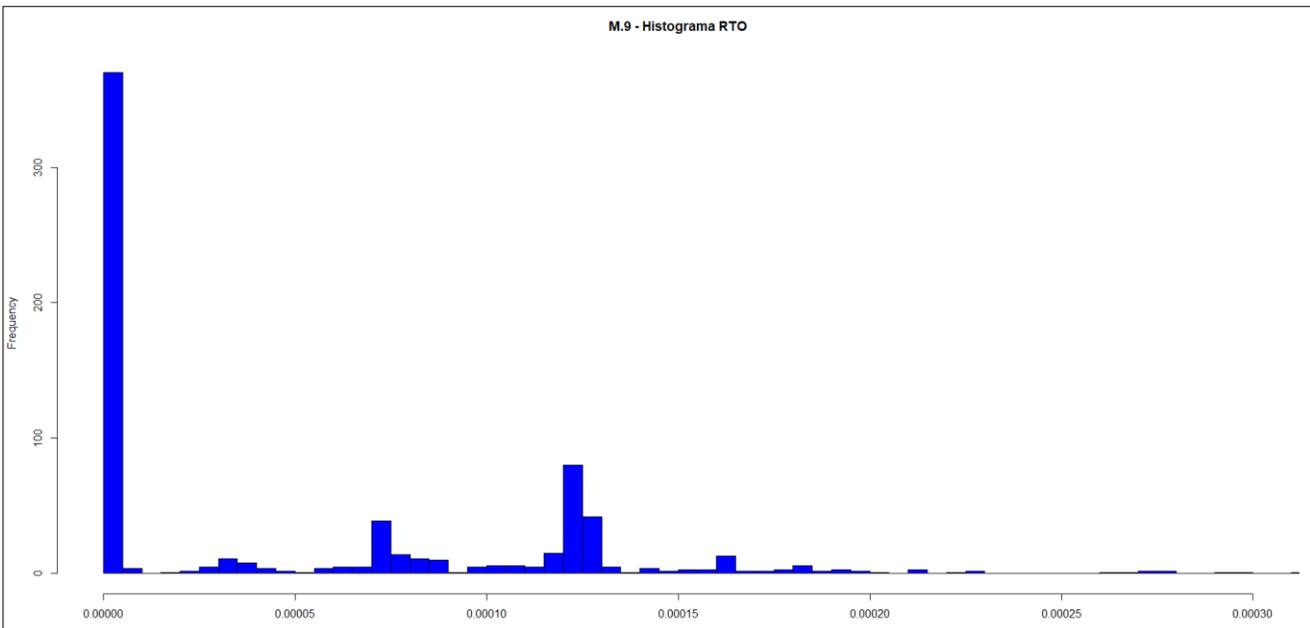
Histograma – Muestra-8:

(Longitud 100seg. - valor de Hurst=0.578)



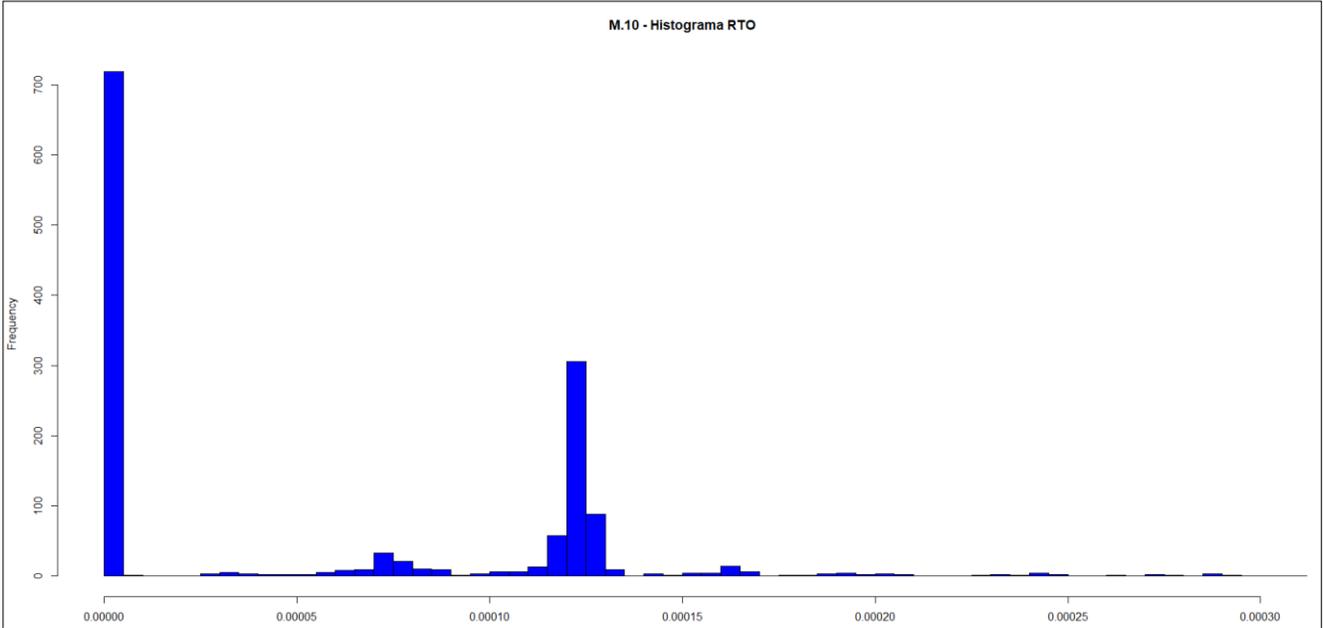
Histograma – Muestra-9:

(Longitud 100seg. - valor de Hurst=0.606)

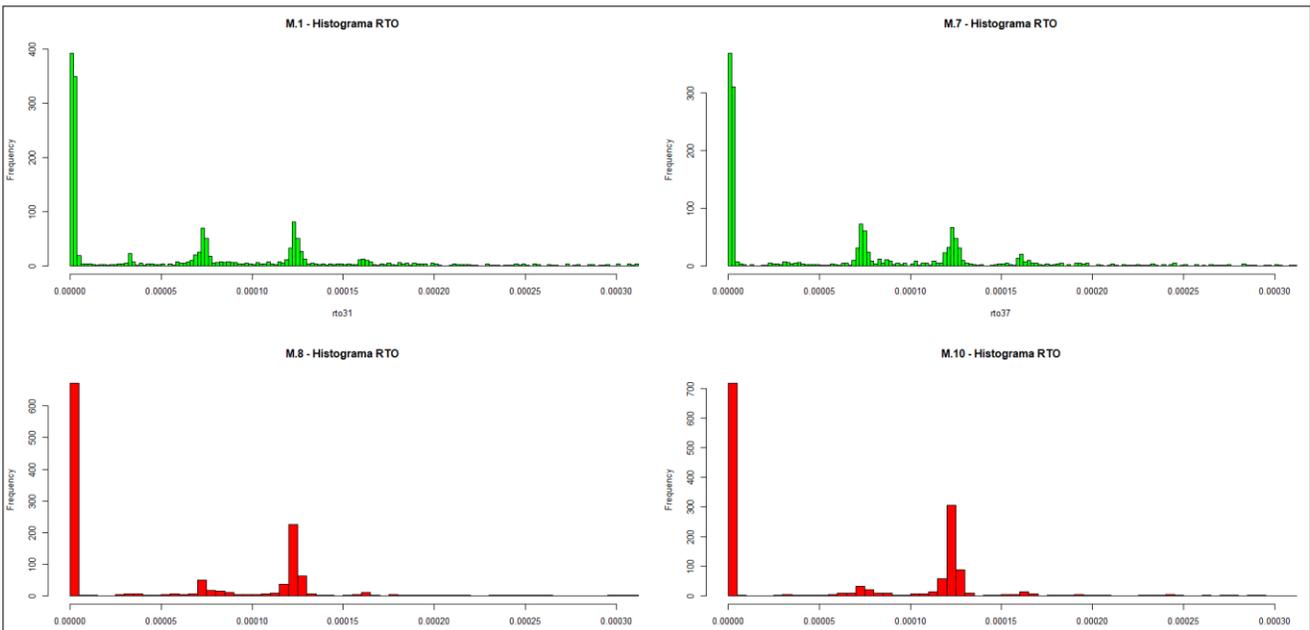


Histograma – Muestra-10:

(Longitud 100seg. - valor de Hurst=0.567)



Al analizar la distribución de los intervalos de ocurrencias, y comparar las muestras 1 y 7 (que poseen un mayor coeficiente de Hurst) con las muestras 8 y 10 (que son las que poseen un menor coeficiente de Hurst), se observa un posible patrón de conducta en las retransmisiones.



En la grafica anterior se comparan los histogramas en una misma grafica, en color verde las muestras con mayor componente autosimilar, y en roja las con menor.

Se observa una semejanza entre histogramas del mismo color. A diferencia del escenario-2, en este escenario el grado de autosimilaridad podría estar vinculado con la dispersión de las retransmisiones.

### **3.2.1.9 Interpretación de los resultados**

En el escenario-3 se superan las limitaciones presentes en el escenario-2, tales como el porcentaje de retransmisiones por muestra, tipo de datos, variación del RTT, y cantidad de conexiones TCP.

Si bien estamos frente a un escenario sin congestión, las retransmisiones son parte de la naturaleza del protocolo TCP, y este escenario no queda exento de dicho suceso. Las muestras disponen de un porcentaje de retransmisiones que no superan el 5%, y con un valor medio de 2.3%. Las retransmisiones son todos productos de vencimientos del RTO, por lo que no ocurrieron retransmisiones causadas por duplicidad de ACK.

En este punto, no quedan dudas de que la ejecución de los algoritmos de control de congestión es una condición necesaria para la existencia de autosimilaridad, y se espera que la componente autosimilar aumente a medida que la congestión se hace más evidente.

A partir del análisis de histogramas de los intervalos de ocurrencia, surge la posibilidad de que exista un patrón de conducta en la dispersión de las retransmisiones que puede estar vinculado con el grado de autosimilaridad. No obstante, dicha observación es una mera sospecha y deberá ser demostrada si el fenómeno se sigue repitiendo. Dichos resultados serán contrastados con los que se obtengan del análisis del escenario-4, y se evaluará si el supuesto es consistente.

## **3.2.2 CONDICION DE RED ESTRESADA: ESCENARIO-4**

### **3.2.2.1 Ámbito de realización**

La recolección de las muestras se realiza utilizando la misma topología de red descripta para el escenario-3, con la salvedad de que en este caso la red se congestiona para conseguir el escenario de interés.

Nuevamente se filtran los paquetes que ingresan a la red local desde internet y en consecuencia se analizan solo los paquetes entrantes.

Se utiliza un perfil de QoS que caracteriza con prioridad baja a las conexiones TCP y permite una salida máxima de 10Mbps, con la finalidad de que el enlace se vea saturado y, en consecuencia, se ejecute el algoritmo slow-start de TCP.

### 3.2.2.2 Calidad de servicio

En la recolección del escenario-4 se configura un perfil de QoS con baja prioridad para el tráfico TCP, permitiendo una salida máxima de 10Mbps.

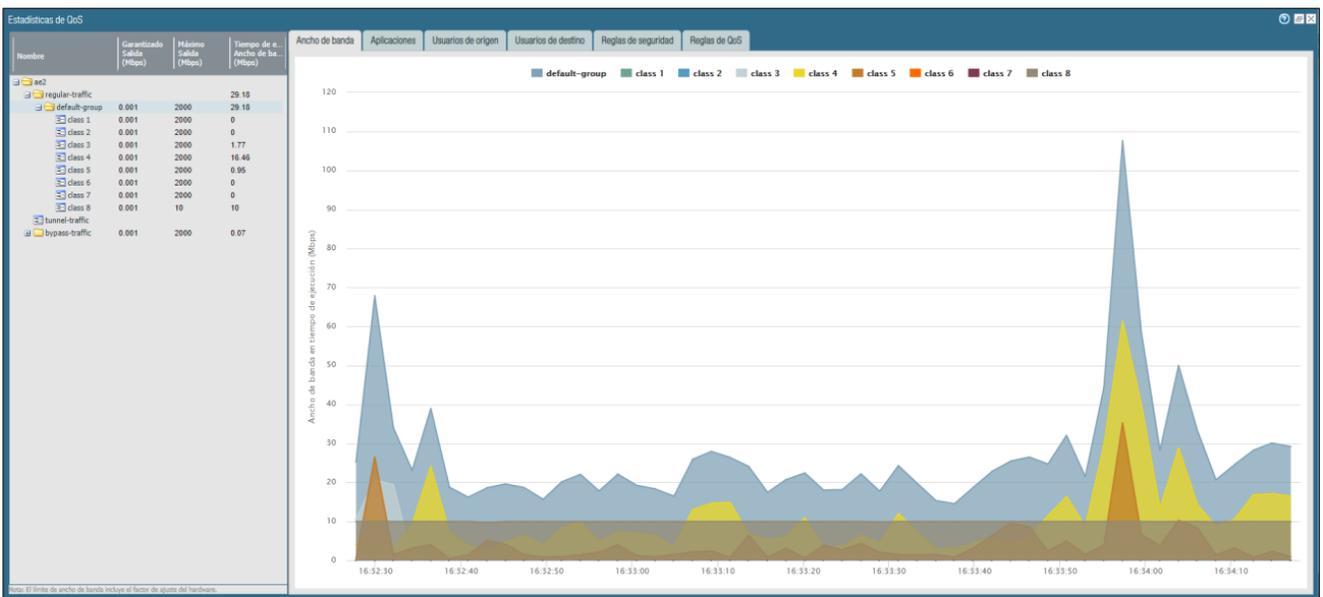
Clase	Prioridad	Máximo de salida	Salida garantizada
class8	low	10	0

Se asigna el perfil a la interface de interés y se configura la política de QoS para los cambios:

Clase	Prioridad	Máximo de salida	Salida garantizada
class8	low	10	0

Nombre	Etiquetas	IP Origen			IP Destino		Uso de reglas			Aplicación	Servicio	DSCP/ToS	Clase	Programación	
		Zona	Dirección	Usuario	Zona	Dirección	Recuento de aciertos	Último acierto	Primer acierto						
1	TCP	none	any	any	any	Internet	any	13663270	2019-10-07 10:52:53	2019-10-01 11:34:35	any	TCP	cualquiera	8	ninguno

La imagen siguiente muestra el tráfico TCP encolado con prioridad baja. Se observa que la red opera congestionada, trabajando al máximo ancho de banda asignado.

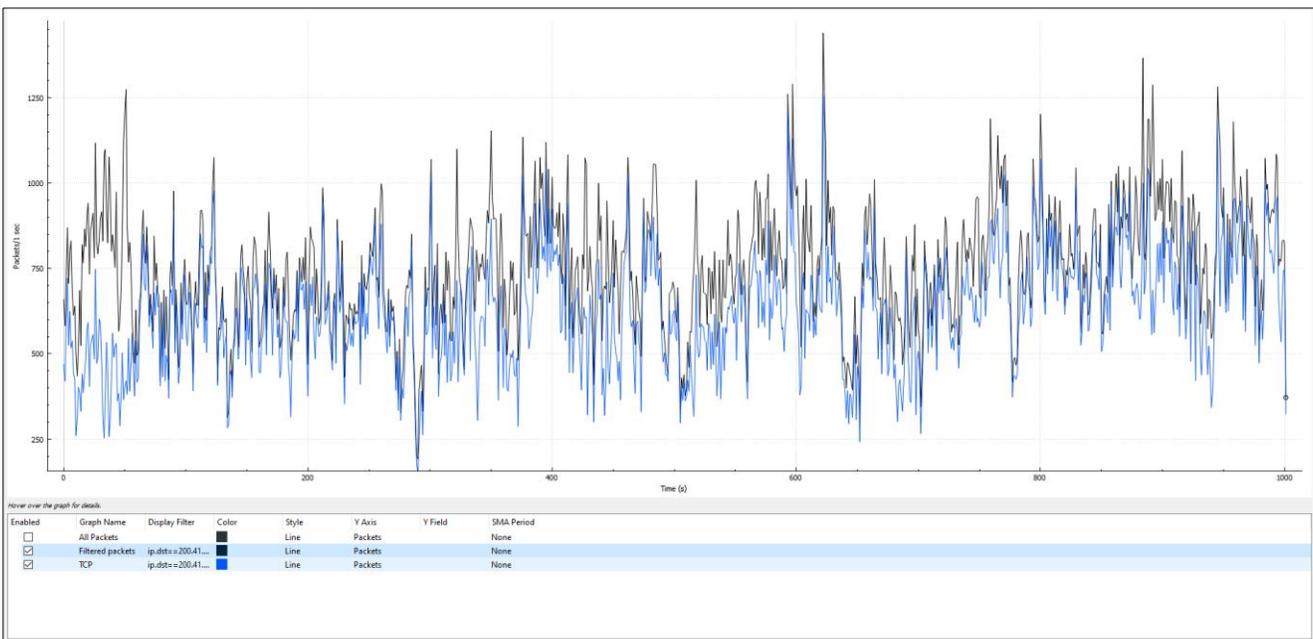


El monitoreo de la cola evidencia un escenario de red congestionada. Esta interpretación es respaldada mediante los ítems 3.2.2.3 y 3.2.2.4 que se desarrollan a continuación.

### 3.2.2.3 Análisis de tráfico

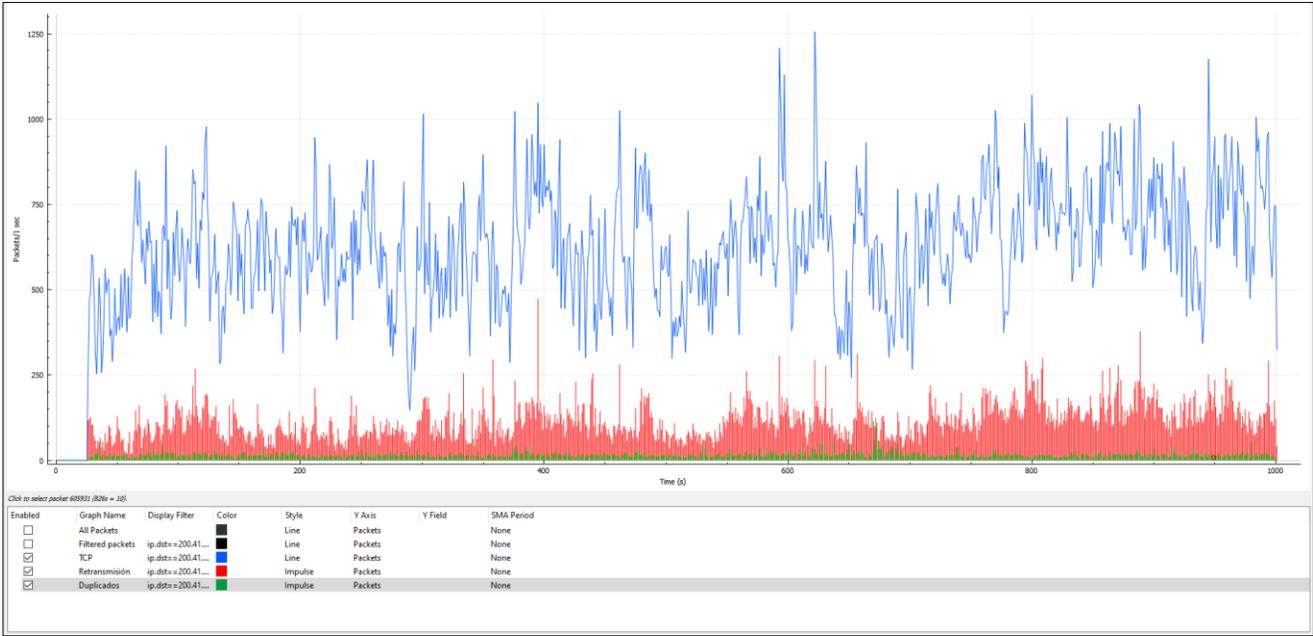
La supervisión de red detallada en el punto 3.2.1.3 es válida para todas las muestras adquiridas de la red operativa, por lo que también aplica al escenario-4.

A continuación se observa el ancho de banda consumido para al momento de la recolección de las muestras, donde se grafica en color negro todo el tráfico que circula por la interface y en azul el tráfico TCP. Se hace evidente la falta de priorización del tráfico TCP mediante la configuración de QoS, ya que las gráficas son ligeramente diferentes.



### 3.2.2.4 Comportamiento de TCP

En la gráfica siguiente se observan en azul todos los paquetes TCP que ingresan a la LAN, en rojo los paquetes retransmitidos y en verde los paquetes duplicados. En este caso no ocurren 3 ACK en simultáneo asociados al mismo segmento, por lo que no hay retrasmisiones por duplicidad de ACK y el algoritmo FR/FR no tiene razón de ser en el escenario-4.



Los paquetes retransmitidos representan el 15.3% de la totalidad de los paquetes entrantes a la LAN, para el escenario de estudio.

Time	Source	Destination	Protocol	Length	Info
23 0.020968	216.58.202.35		TCP	100	[TCP Retransmission] 443 → 604
109 0.141392	74.125.107.170		TCP	1462	[TCP Retransmission] 443 → 388
111 0.145092	185.28.222.11		TCP	1514	[TCP Retransmission] 443 → 551
162 0.198215	192.16.59.1		TCP	1514	[TCP Retransmission] 443 → 553
190 0.225581	185.28.222.11		TCP	1514	[TCP Retransmission] 443 → 551
237 0.300095	74.119.119.132		TCP	102	[TCP Retransmission] 443 → 550
241 0.306635	74.125.107.170		TCP	1462	[TCP Retransmission] 443 → 388
266 0.323591	200.124.123.65		TCP	1514	[TCP Retransmission] 443 → 468
299 0.368115	74.125.107.170		TCP	1462	[TCP Retransmission] 443 → 388

Frame 23: 100 bytes on wire (800 bits), 100 bytes captured (800 bits)  
 Ethernet II, Src: JuniperN\_e9:8f:2f (ec:3e:f7:e9:8f:2f), Dst: PaloAlto\_00:0a:01 (00:1b:17:00:0a:01)  
 Internet Protocol Version 4, Src: 216.58.202.35, Dst: 200.41.39.18  
 Transmission Control Protocol, Src Port: 443, Dst Port: 60443, Seq: 198, Ack: 1, Len: 46

```

000  00 1b 17 00 0a 01 ec 3e f7 e9 8f 2f 08 00 45 00  .....> .../..E.
010  00 56 93 33 00 00 7c 06 19 d5 d8 3a ca 23 c8 29  -V.3..| . ...:#. )
020  27 12 01 bb ec 1b 50 72 ef 2c 83 4e 3e d0 50 18  '....Pr ., .N> .P.
030  00 f6 4f fe 00 00 17 03 03 00 29 00 00 00 00 00  --0-... ..).....
    
```

cap4in.pcap Packets: 755821 · Displayed: 115516 (15.3%)

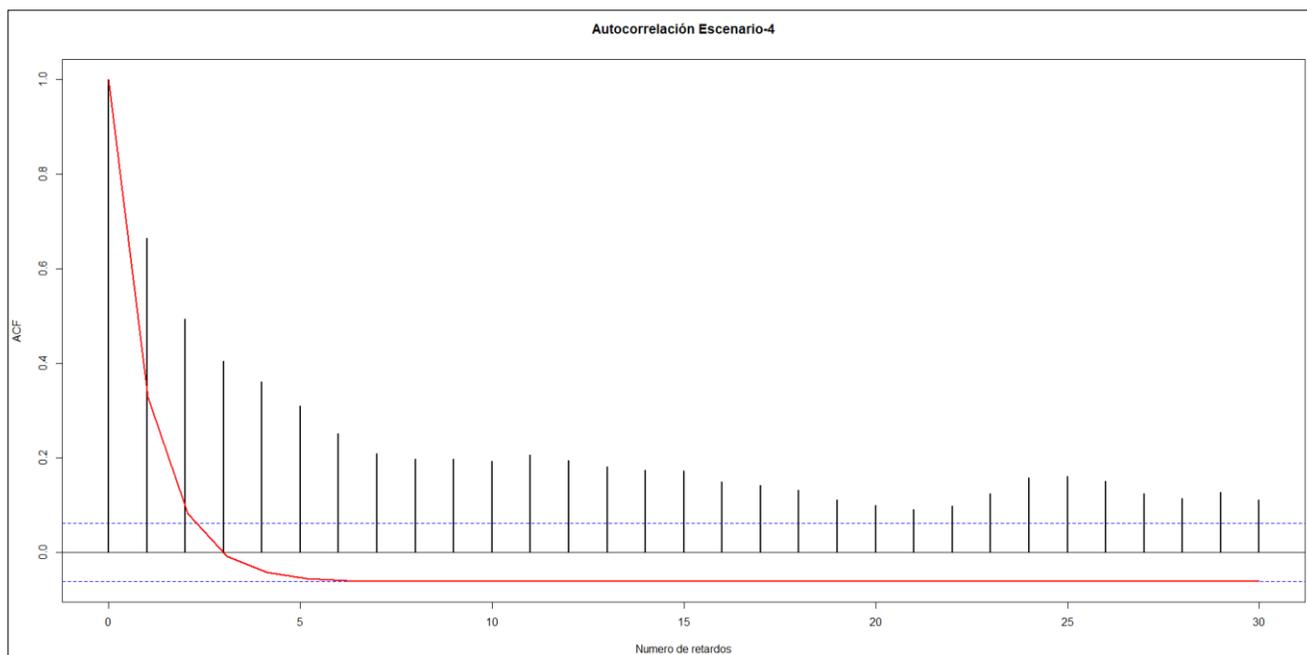
### 3.2.2.5 Recolección de las muestras

Se procesan los paquetes capturados y se exportan en formato texto plano. Luego se contabilizan los paquetes por segundos.

En el vector *muestra4* se almacena la contabilización de los paquetes capturados en cada segundo. Luego se secciona el vector *muestra4* en los vectores *m41*, *m42*, *m43*, *m44*, *m45*, *m46*, *m47*, *m48* y *m49*, *m410* y se obtienen las diez muestras a analizar.

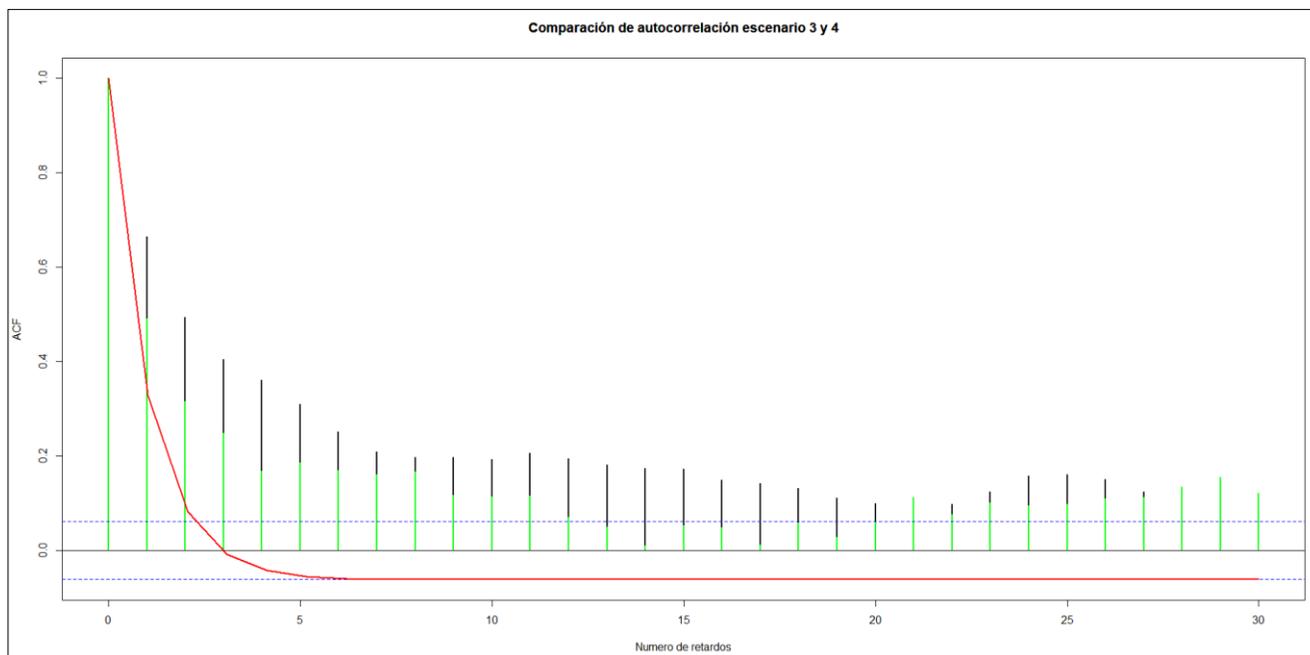
### 3.2.2.6 Autocorrelación

Nuevamente se utiliza la función *ACF* de RGui, para realizar el cálculo de autocorrelación y se la compara con una curva representativa de un decaimiento exponencial:



El escenario en cuestión se corresponde con una red operativa estresada, y a partir del correlograma podemos observar una autocorrelación lenta, y en consecuencia una dependencia a largo plazo entre las muestras.

Si comparamos la autocorrelación del escenario-4 (red estresada) con el escenario-3 (red sin congestión), se visualiza el siguiente resultado:



Donde la grafica color negro corresponde al escenario-4, la grafica en color verde corresponde al escenario-3 y la grafica roja es representativa de un decaimiento exponencial. Queda a la vista que la red estresada presenta una autocorrelación con decaimiento más lento en el tiempo que el escenario-3, y en consecuencia mayor dependencia entre las muestras.

### 3.2.2.7 Autosimilaridad

Calculamos el coeficiente de Hurst que nos permitirá concluir si las muestras dadas contienen una componente autosimilar y en qué medida está presente.

Primero calculamos el coeficiente de Hurst para la muestra-4, de longitud mil (1000) segundos:

```
> hurstexp(muestra4)
Simple R/S Hurst estimation: 0.7313868
Corrected R over S Hurst exponent: 0.8355433
Empirical Hurst exponent: 0.7976361
Corrected empirical Hurst exponent: 0.7706869
Theoretical Hurst exponent: 0.5388682
```

Coefficiente de Hurst para la muestra en el intervalo 1-100 segundos:

```
> hurstexp(m41)
Simple R/S Hurst estimation: 0.7409994
Corrected R over S Hurst exponent: 0.9298582
Empirical Hurst exponent: 0.6438704
Corrected empirical Hurst exponent: 0.6660849
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 101-200 segundos:

```
> hurstexp(m42)
Simple R/S Hurst estimation: 0.6293973
Corrected R over S Hurst exponent: 0.7142604
Empirical Hurst exponent: 0.6287995
Corrected empirical Hurst exponent: 0.6523029
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 201-300 segundos:

```
> hurstexp(m43)
Simple R/S Hurst estimation: 0.6942275
Corrected R over S Hurst exponent: 0.8382684
Empirical Hurst exponent: 0.653167
Corrected empirical Hurst exponent: 0.6746205
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 301-400 segundos:

```
> hurstexp(m44)
Simple R/S Hurst estimation: 0.6840416
Corrected R over S Hurst exponent: 0.8516127
Empirical Hurst exponent: 0.6943111
Corrected empirical Hurst exponent: 0.7126868
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 401-500 segundos:

```
> hurstexp(m45)
Simple R/S Hurst estimation: 0.5420304
Corrected R over S Hurst exponent: 0.5423664
Empirical Hurst exponent: 0.6140039
Corrected empirical Hurst exponent: 0.638842
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 501-600 segundos:

```
> hurstexp(m46)
Simple R/S Hurst estimation: 0.7343195
Corrected R over S Hurst exponent: 0.9084794
Empirical Hurst exponent: 0.6621486
Corrected empirical Hurst exponent: 0.6828906
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 601-700 segundos:

```
> hurstexp(m47)
Simple R/S Hurst estimation: 0.6969171
Corrected R over S Hurst exponent: 0.8795716
Empirical Hurst exponent: 0.6862853
Corrected empirical Hurst exponent: 0.7052259
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 701-800 segundos:

```
> hurstexp(m48)
Simple R/S Hurst estimation: 0.7216854
Corrected R over S Hurst exponent: 0.9020983
Empirical Hurst exponent: 0.6868647
Corrected empirical Hurst exponent: 0.705764
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 801-900 segundos:

```
> hurstexp(m49)
Simple R/S Hurst estimation: 0.7254853
Corrected R over S Hurst exponent: 0.9266958
Empirical Hurst exponent: 0.6177777
Corrected empirical Hurst exponent: 0.6422686
Theoretical Hurst exponent: 0.5229461
```

Coefficiente de Hurst para la muestra en el intervalo 901-1000 segundos:

```
> hurstexp(m410)
Simple R/S Hurst estimation: 0.6241954
Corrected R over S Hurst exponent: 0.7003049
Empirical Hurst exponent: 0.6432282
Corrected empirical Hurst exponent: 0.6654962
Theoretical Hurst exponent: 0.5229461
```

Para el caso de estudio, nuevamente el enfoque utilizado es el análisis R/S.

Se observa que para el escenario-4 el coeficiente de Hurst muestra un valor que oscilan entre  $H=0.542$  y  $H=0.74$ , con un valor medio de  $H=0.678$ . También se calculo el valor del coeficiente para la muestra de 1000 segundos, que dio como resultado  $H_2=0.731$  y que muestra un valor superior al coeficiente de Hurst de la *muestra3* que dio como resultado  $H_1=0.669$ .

### 3.2.2.8 Caracterización de las muestras

En línea con los escenarios anteriores, nuevamente el alcance del presente ítem se acota siguiente desarrollo:

- ✓ Tipos y cantidad de retrasmisiones por cada muestra.

- ✓ Histograma de la dispersión de las retransmisiones.
- ✓ Efecto de las retransmisiones por cada muestra.

### Tipos y cantidad de retransmisiones por cada muestra

A continuación, la primera tabla representa una comparación entre la totalidad de los paquetes transmitidos contra las retransmisiones presentes en cada muestra.

MUESTRAS	HURST	LARGO DE LA MUESTRA	RETRANSMISIONES	RETRANSMISIONES/MUESTRAS (%)
<b>GENERAL</b>	0.731	754628	115354	15.28%
<b>1</b>	<b>0.74</b>	<b>74494</b>	<b>8514</b>	<b>11.42%</b>
<b>2</b>	0.629	67171	9914	14.75%
<b>3</b>	0.694	64506	8312	12.88%
<b>4</b>	0.684	77658	11736	15.11%
<b>5</b>	<b>0.542</b>	<b>76210</b>	<b>11552</b>	<b>15.15%</b>
<b>6</b>	<b>0.734</b>	<b>74948</b>	<b>10658</b>	<b>14.22%</b>
<b>7</b>	0.696	71720	10737	14.97%
<b>8</b>	0.721	76688	13429	17.51%
<b>9</b>	0.725	85988	16134	21.03%
<b>10</b>	<b>0.623</b>	<b>85245</b>	<b>14368</b>	<b>16.85%</b>

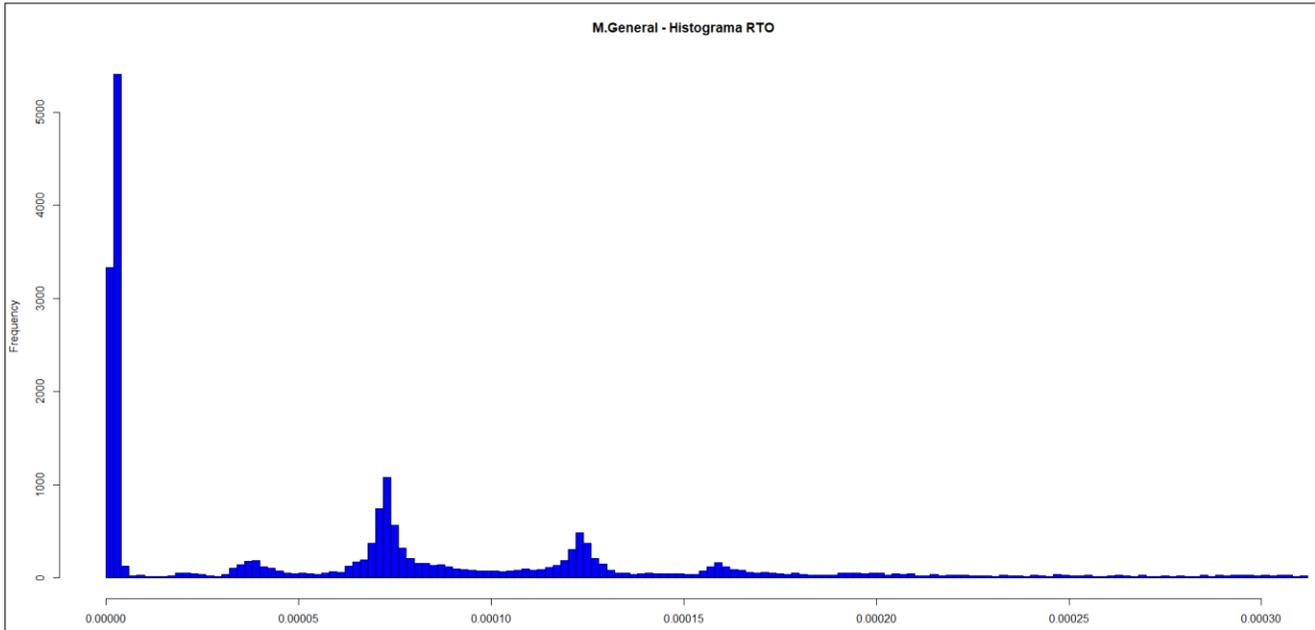
El este escenario, al igual que en el escenario-3, todas las retransmisiones son causadas por vencimiento del RTO. Al analizar la tabla anterior, se resalta en color verde las dos muestras con mayor autosimilaridad y en color rojo las dos muestras con menor autosimilaridad. No se observa relación entre el grado de autosimilaridad con la cantidad de retransmisiones presentes en las muestras.

### Histograma de la dispersión de las retransmisiones

Se procede a graficar el histograma haciendo uso de RGui. En función de los resultados se contrastará el supuesto del escenario-3, con la finalidad de comprobar si dicho comportamiento puede ser un patrón de conducta relacionado con el aumento de la autosimilaridad.

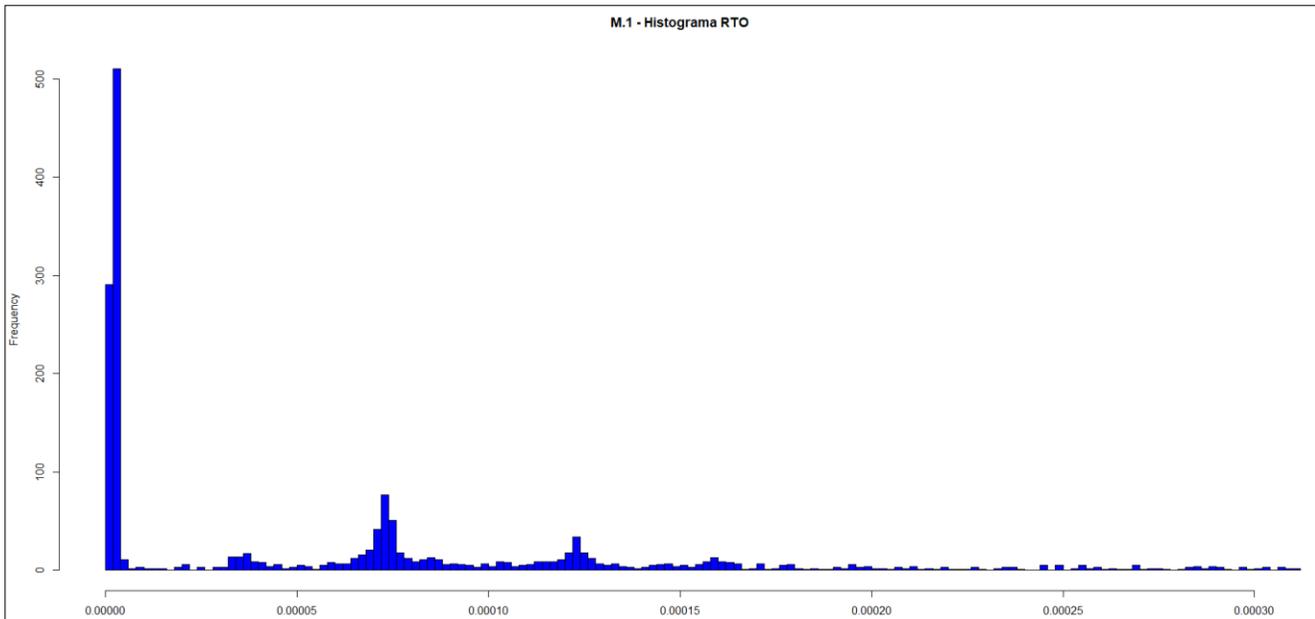
El histograma - Muestra general:

(Longitud 1000seg. - valor de Hurst= 0.731)



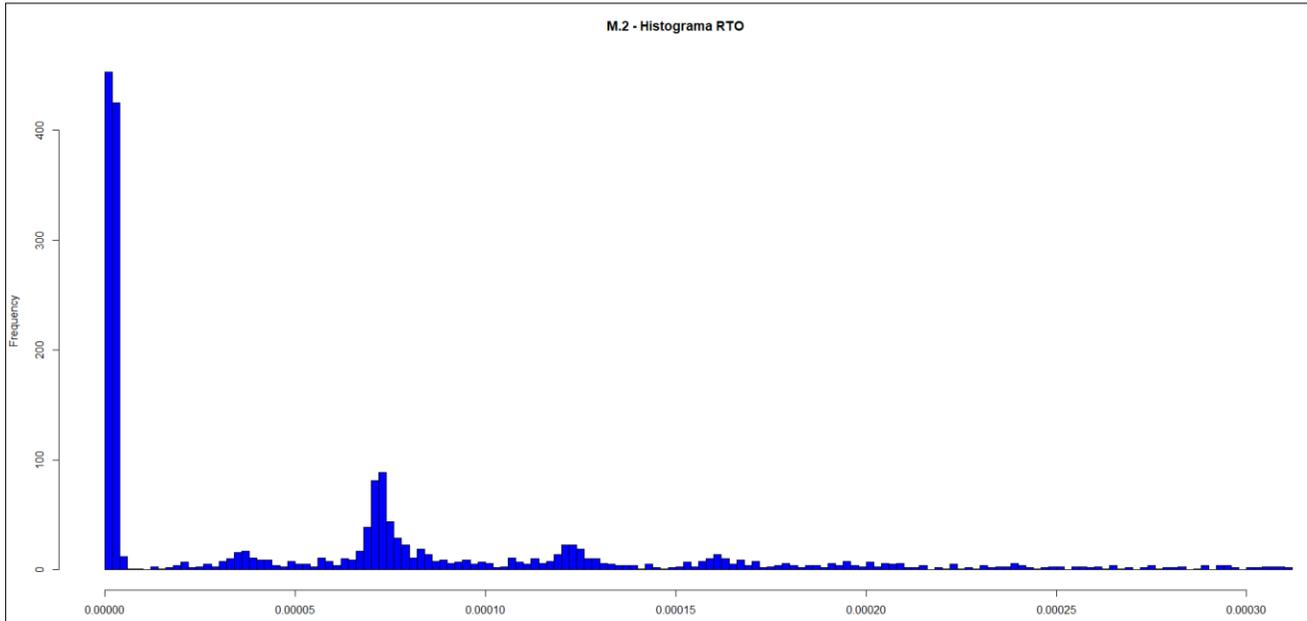
Histograma – Muestra-1:

(Longitud 100seg. - valor de Hurst= 0.74)



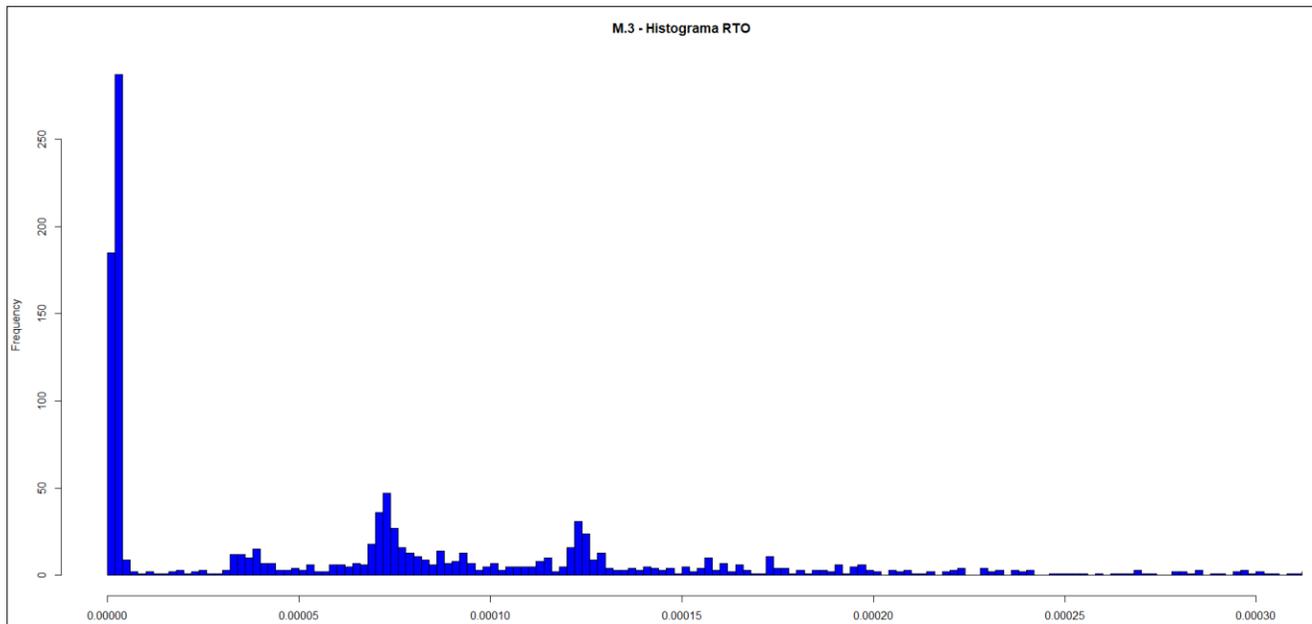
Histograma – Muestra-2:

(Longitud 100seg. - valor de Hurst= 0.629)



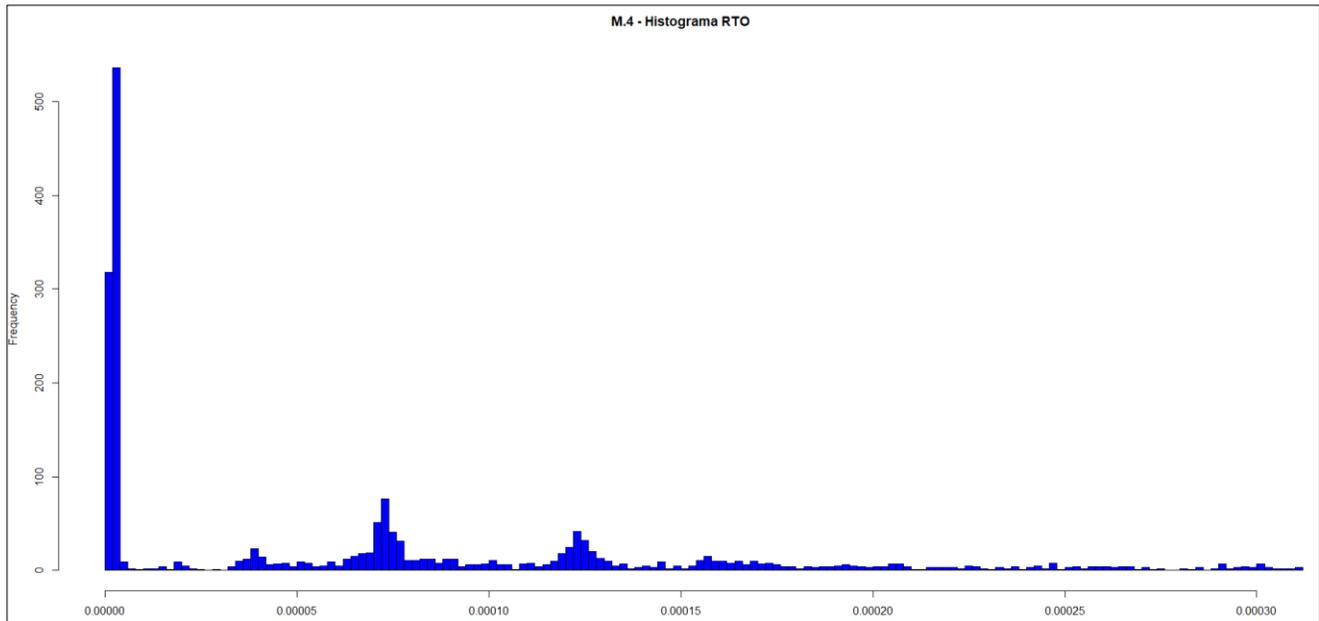
Histograma – Muestra-3:

(Longitud 100seg. - valor de Hurst= 0.694)



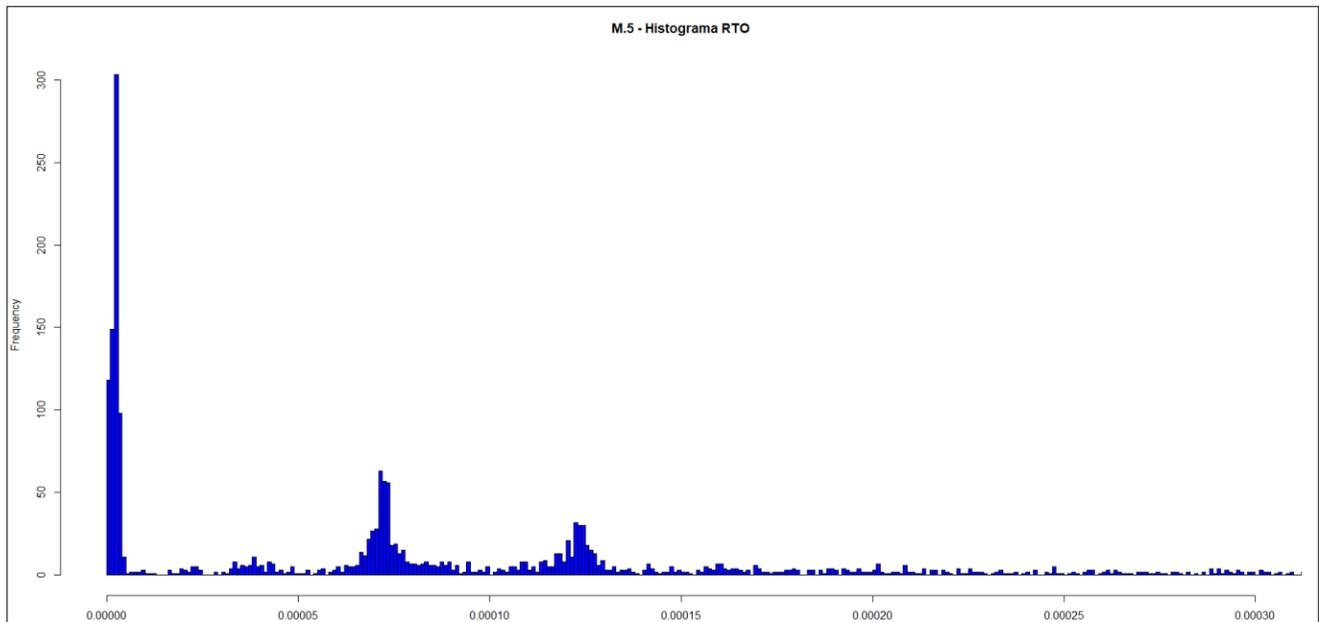
Histograma – Muestra-4:

(Longitud 100seg. - valor de Hurst= 0.684)



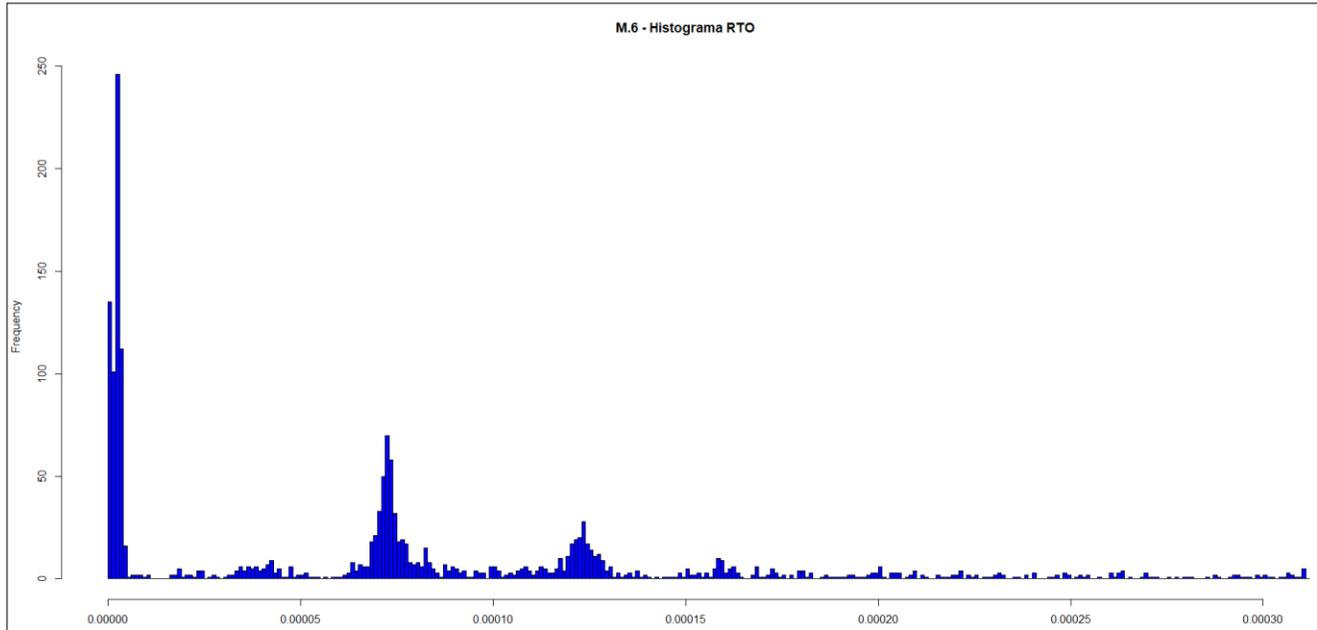
Histograma – Muestra-5:

(Longitud 100seg. - valor de Hurst= 0.542)



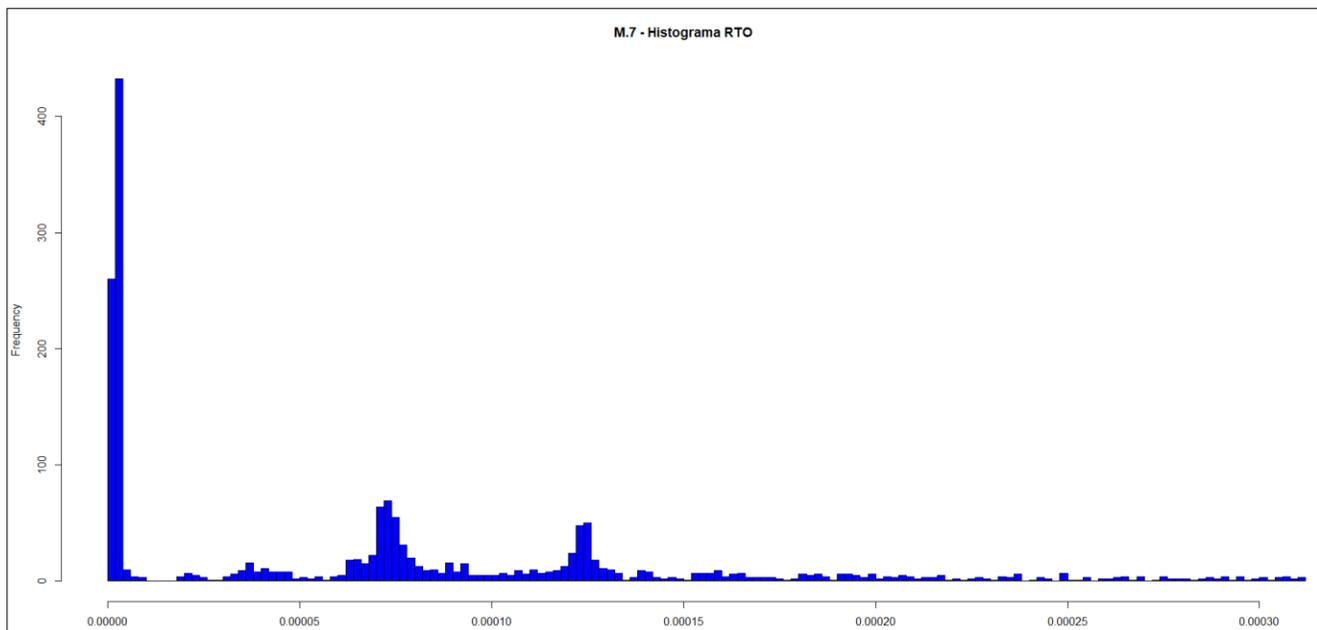
Histograma – Muestra-6:

(Longitud 100seg. - valor de Hurst= 0.734)



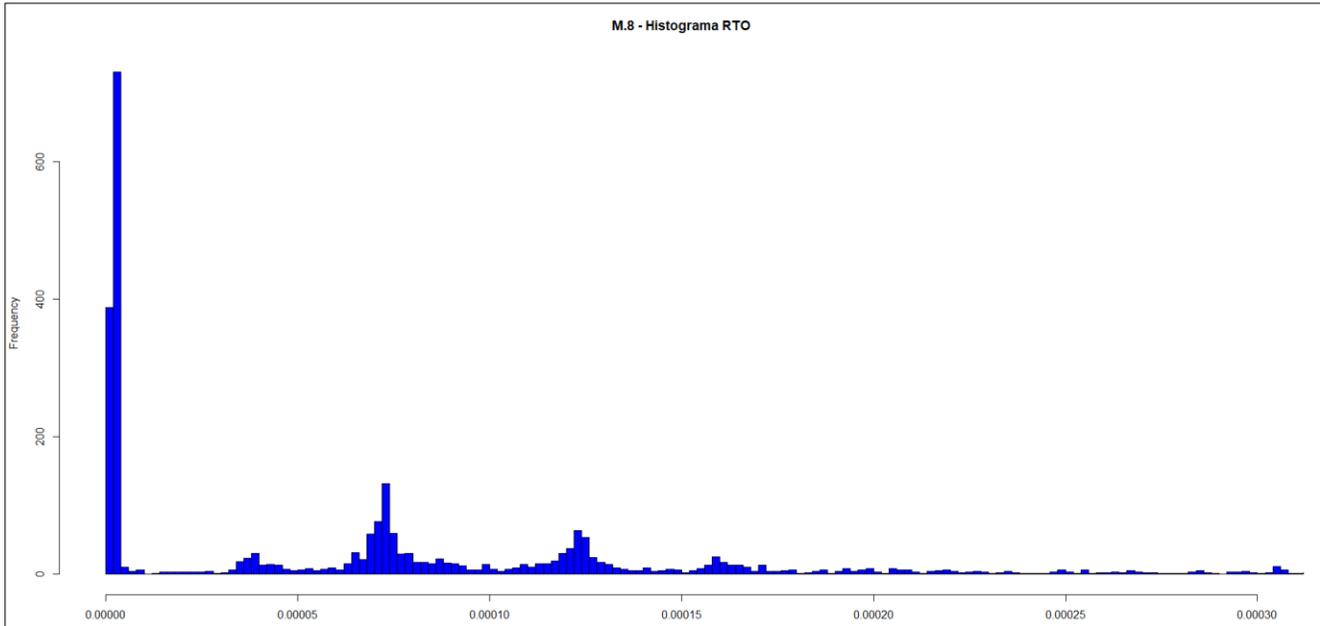
Histograma – Muestra-7:

(Longitud 100seg. - valor de Hurst= 0.696)



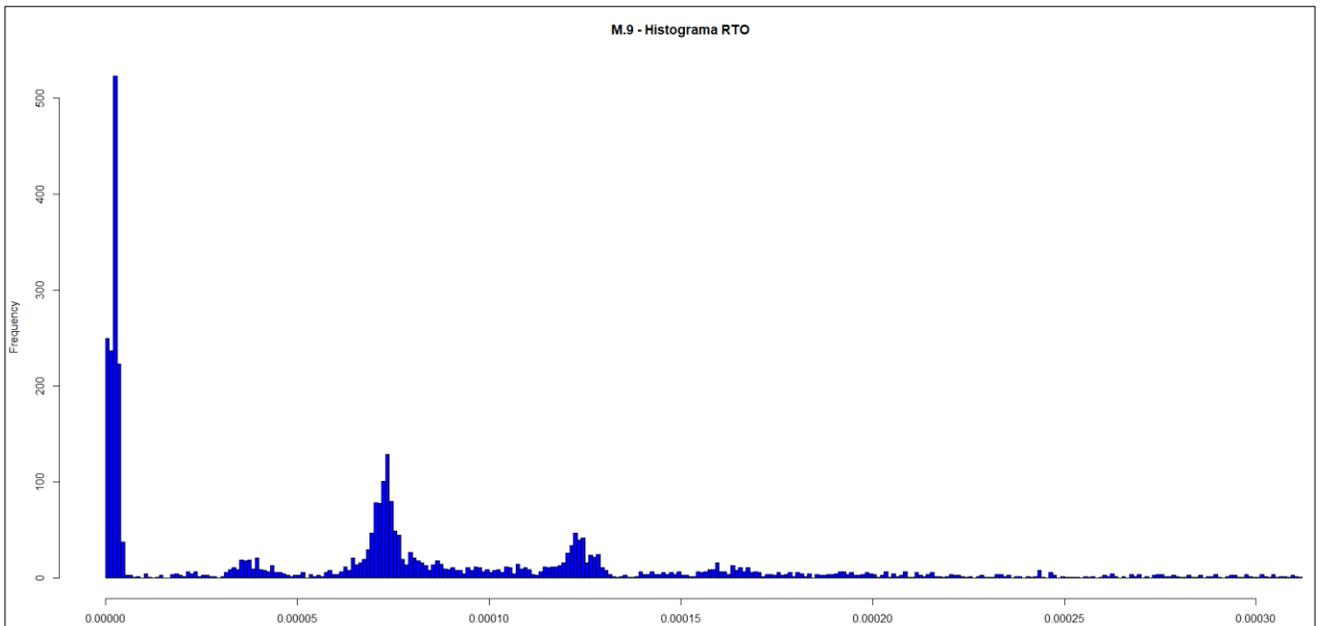
Histograma – Muestra-8:

(Longitud 100seg. - valor de Hurst= 0.721)



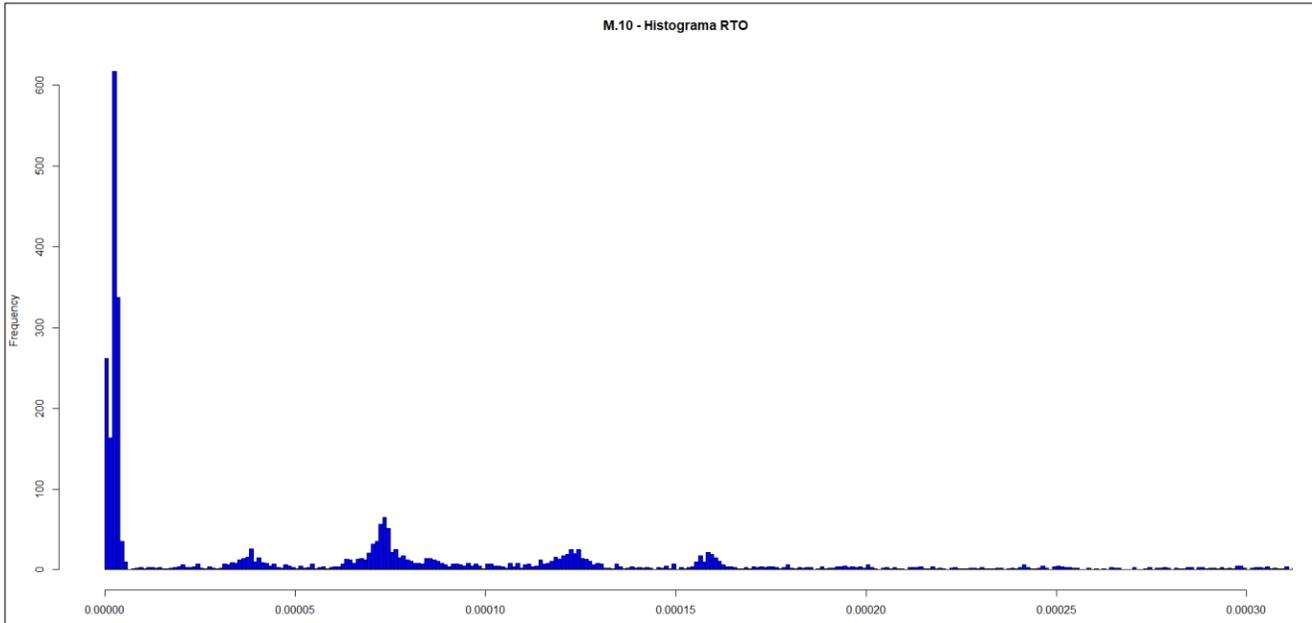
Histograma – Muestra-9:

(Longitud 100seg. - valor de Hurst= 0.725)



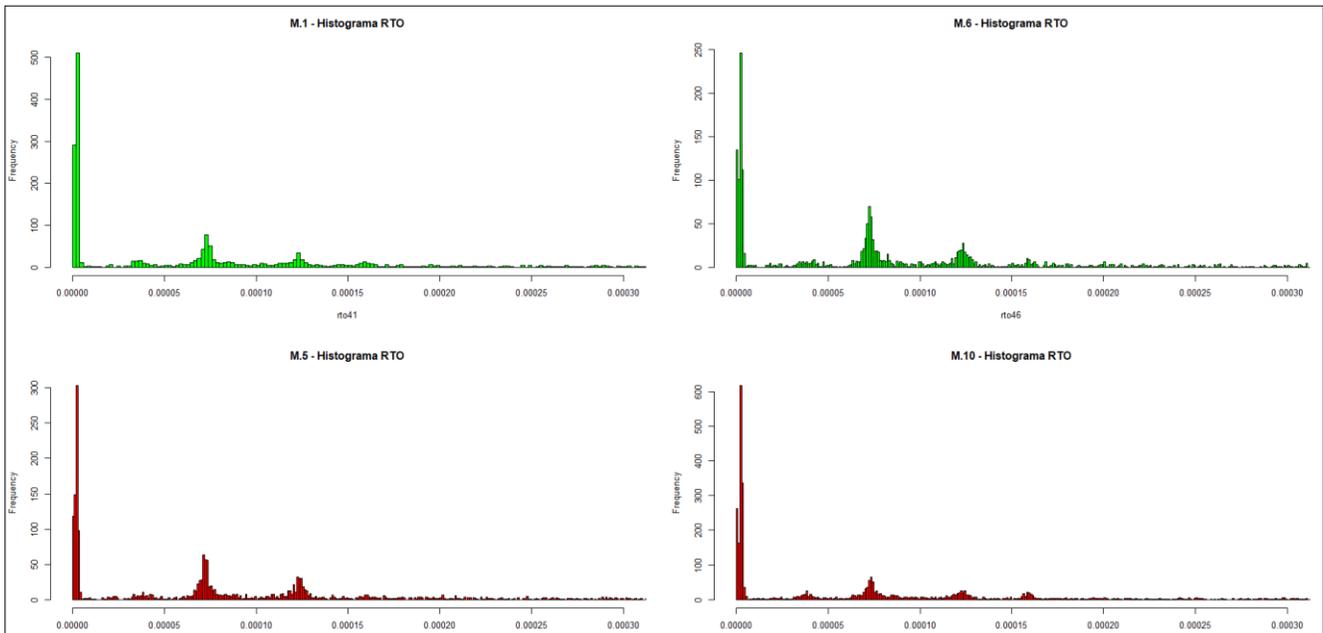
Histograma – Muestra-10:

(Longitud 100seg. - valor de Hurst= 0.623)



Al analizar la distribución de los intervalos de ocurrencias, no se observan comportamientos que conduzcan a sospechar de la existencia de un posible patrón en las retransmisiones.

A continuación se comparan los histogramas con mayor y menor componente autosimilar. En color verde las muestras con mayor autosimilaridad, y en roja las con menor.



Las graficas son semejantes, independientemente del grado de autosimilaridad presente en la muestra. El escenario no acompaña las interpretaciones realizadas en el escenario anterior.

Se concluye que no hay evidencia que vincule la autosimilaridad con el momento en que ocurren las retransmisiones.

### **3.2.2.9 Interpretación de los resultados**

El escenario-4 representa el tráfico de datos en una red estresada. Las muestras disponen de un porcentaje alto de retransmisiones que promedian en un 15.2%. Las retransmisiones son todos productos de vencimientos del RTO por lo que, incluso bajo condición de congestión, no ocurrieron retransmisiones causadas por el envío de 3 ACK en simultáneo, y en efecto no se ejecuta el algoritmo FR/FR.

Nuevamente podemos confirmar que la autosimilaridad tiene una fuerte relación con la congestión. Se observa que los coeficientes de Hurst calculados para las muestras del presente escenario son, en general, superiores a los obtenidos en el escenario 2 y 3. Esto nos permite validar que, cuando el porcentaje de retransmisiones tiene un aumento considerable, los valores del coeficiente de Hurst tienden a aumentar.

Las autocorrelación calculada evidencia la estrecha relación que existe entre la autosimilaridad y la dependencia entre las muestras.

A partir del análisis de histogramas de la dispersión, quedan desestimada la supuesta existencia de un patrón que relacione la dispersión de las retransmisiones con la autosimilaridad de la muestra.

## 4. CONCLUSIÓN

---

En línea con los objetivos planteados, la tesis se focalizó en estudiar el protocolo de control de transmisión (TCP) y su influencia en el comportamiento autosimilar del tráfico en redes convergentes.

Se tomaron muestras de tráfico, tanto de un ambiente de red controlado como en uno operativo, y se observó la relación entre el mecanismo de control de congestión de TCP con el grado de autosimilaridad presente.

Se observó en el escenario-1 que, ante la ausencia de retransmisiones, hay un comportamiento antipersistente en los datos, donde una tendencia de aumento en el pasado es acompañada de una tendencia a disminuir en el futuro, y viceversa.

Al introducir congestión en el escenario-2, se generan retransmisiones y aparace el fenómeno de autosimilaridad en el tráfico, que se hace evidente en alguna de las muestras y desaparece en otras. En este escenario las retransmisiones representan un 1.5% de la totalidad de los datos, y resulta un coeficiente de Hurst medio de  $H=0.633$ .

El escenario-3, cuya muestra es tomada de una red operativa, representa un esquema sin congestión. Debido a la naturaleza del protocolo TCP, este escenario no queda exento de las retransmisiones, y se observa que el 2.3% en promedio de los datos corresponden a retransmisiones. En este caso resulta un coeficiente de Hurst medio de  $H=0.629$ .

Por último se trabajó en el escenario-4, cuya muestra es tomada de una red operativa estresada. En esta ocasión las retransmisiones representan el 15% de la totalidad del tráfico, y resultó un coeficiente de Hurst medio de  $H=0.678$ .

Los resultados evidencian que la congestión es una condición necesaria para la existencia de un comportamiento autosimilar en el tráfico de datos.

Se observa que hay una tendencia a aumentar la componente autosimilar cuando las retransmisiones se incrementan considerablemente.

Cuando los porcentajes de retransmisión aumentan en pequeñas proporciones, los valores de autosimilaridad no parecen tener relación con dicho incremento. Esto nos permite afirmar que la relación entre la congestión y la autosimilaridad no es lineal, y posiblemente influyan otros factores que nada tienen que ver con la congestión.

El desarrollo también profundiza en el estudio de la dispersión de las retransmisiones, con la finalidad de buscar un patrón que relacione los intervalos de ocurrencia de las retransmisiones con los cambios en la componente autosimilar. El análisis no refleja relación entre la dispersión de las retransmisiones con la autosimilaridad.

Se recomienda a futuro, para dar seguimiento al tema y en relación con las pruebas realizadas, estudiar el tipo de tráfico que circula por la red y su relación con la congestión. El análisis propuesto puede abrir nuevas líneas de investigación con altas posibilidades de éxito en esta búsqueda constante de la causa raíz del comportamiento autosimilar en las redes multiservicio.

Un obstáculo que se presenta y dificulta analizar el tipo de tráfico que circula por la red es la encriptación, que en cierta forma enmascara y oculta información fundamental para concluir determinados ensayos.

## 5. GLOSARIO

---

**ACF:** del inglés Autocorrelation Function , en español función de autocorrelación.

**ACK:** del inglés acknowledgement, en español acuse de recibo.

**ARMIA:** del inglés Autoregressive Integrated Moving Average, en español modelo autoregresivo integrado de promedio móvil.

**BSD:** del inglés Berkeley Software Distribution, en español distribución de software Berkeley.

**BUFFER:** espacio en un disco o en un instrumento digital reservado para el almacenamiento.

**ECN:** del inglés Explicit Congestion Notification, en español notificación explícita de congestión.

**FIFO:** del inglés First In First Out, en español lo primero que entra es lo primero que sale.

**FR/FR:** del inglés Fast Retransmit / Fast Recovery, en español retransmisión rápida / recuperación rápida.

**FTP:** del inglés File Transfer Protocol, en español protocolo de transferencia de fichero.

**HTTP:** del inglés Hypertext Transfer Protocol, en español protocolo de transferencia de hipertexto.

**IDD:** Idénticamente Distribuido.

**IP:** del inglés Internet Protocol, en español protocolo de internet.

**LAN:** del inglés Local Area Network, en español redes de área amplia.

**LRD:** del inglés Long Range Dependent, en español dependencia de rango largo.

**MPEG:** del inglés Moving Picture Experts Group y es el nombre de un grupo de estándares de codificación de audio y video.

**MTU:** del inglés Maximum Transmission Unit, en español unidad máxima de transferencia.

**OSI:** del inglés Open Systems Interconnection y es un modelo de referencia para los protocolos de comunicaciones.

**PC:** del inglés personal computer, en español computadora personal.

**PRACMA:** del inglés Practical Numerical Math Functions y es una librería del software R.

**QoS:** del inglés Quality of Service, en español calidad de servicio.

**RED:** del inglés Random Early Detect y es un mecanismo de gestión de colas activo.

**RFC:** del inglés Request For Comments y es un conjunto de documentos que sirve de referencia para la comunidad de internet.

**RGui:** software que sirve de plataforma para programar en R.

**RTO:** del inglés Retransmission Time Out, en español temporizador de retransmisión.

**RTT:** del inglés Round Trip Time, en español tiempo de ida y vuelta.

**SLL:** del inglés Secure Sockets Layer, en español capa de puertos seguros.

**SMTP:** del inglés Simple Mail Transfer Protocol, en español protocolo para transferencia simple de correo.

**SNMP:** del inglés Simple Network Management Protocol, en español protocolo simple de administración de red.

**SRD:** del inglés Short Range Dependent, en español dependencia de rango corto.

**SSS:** del inglés Strict Sense Stationarity y significa estacionario en sentido estricto.

**TAIL-DROP:** es un algoritmo simple de gestión de colas.

**TCP/IP:** conjunto de protocolos que permiten la comunicación en una red.

**TCP:** del inglés Transmission Control Protocol, en español protocolo de control de transmisión.

**TELNET:** del inglés Telecommunication Network y es un protocolo de red que permite el acceso a otras máquinas para manejarla remotamente.

**TIC:** Tecnología de la Información y las Comunicaciones.

**UDP:** del inglés User Datagram Protocol y es un protocolo de nivel de transporte.

**UNIX:** es un sistema operativo.

**VBR:** del inglés Variable Bit Rate, en español tasa de bit variable.

**WIMAX:** del inglés Worldwide Interoperability for Microwave Access y es una norma de transmisión de datos que utiliza ondas de radio.

**WSS:** del inglés Wide Sense Stationarity y significa estacionario en sentido amplio.

## 6. BIBLIOGRAFÍA

---

- [1] M. Santiago, «La invención del telégrafo,» *Redhistoria*, 2012.
- [2] <https://ethw.org/>, «Antonio Meucci,» *Engineering and Technology History Wiki*, 2016.
- [3] <https://es.wikipedia.org/>, «Alexander Graham Bell,» *Wikipedia*, 2020.
- [4] A. Figueiras, Una panorámica de las telecomunicaciones, Madrid: Perarson, 2002.
- [5] R. C. Abad, Introducción a la simulación y a las teorías de colas, España: Netbiblo, 2002.
- [6] Leland W. E., Taqqu M. S., Willinger W. y Wilson D. V., «On the Self-Similar Nature of Ethernet Traffic,» Bellcore, Morristown, NJ, USA, 1994.
- [7] R. D. R., «Introducción a los modelos de tráfico para las redes de banda ancha,» Departamento de matemática aplicada i telemática, Universidad Politécnica de Catalunya, España, 1998.
- [8] M. M. A. A., «Introducción al tráfico autosimilar en redes de comunicaciones,» Universidad Distrial, Colombia, 2001.
- [9] P. S. C., «Análisis del comportamiento autosimilar del tráfico ethernet en las redes de datos,» Grupo de investigación y desarrollo en TIC, Universidad Tecnológica Nacional, Facultad Regional de Mendoza, Mendoza, Argentina, 2010.
- [10] S. R. J. R., «Estudio del tráfico autosimilar orientado a la simulación mediante la utilización de Wavelests u Onditas,» Universidad Nacional de La Plata, La Plata, Argentina, 2010.
- [11] R. D., «Desarrollo de un modelo generador de tráfico autosimilar para la evolución de los algoritmos de asignación de ancho de banda en WIMAX,» Universidad Nacional de La Plata, La Plata, Argentina, 2010.
- [12] Leon A. P., Piedrahita E. M., Salcedo O., «Aplicaciones del modelo On/Off al tráfico agregado en las redes de comunicaciones,» *Tecnura*, vol. 15, pp. 129-147, 2011.
- [13] Astaiza-Hoyos E., Bermúdez-Orozco H., Méndez-Suárez D., «Evaluación de desempeño de un modelo autosimilar para tráfico en redes 802.11,» *Tecno Lógicas*, vol. 31, pp. 13-36, 2013.
- [14] Pérez S., Facchini H.'Bisaro L., «Estudio sobre la distribución de tráfico autosimilar en

redes ethernet,» Grupo UTN de investigación y desarrollo TICs, Argentina, 2012.

- [15] G. J. A., «Impacto de la dependencia de rango largo en el QoS del tráfico de video,» Universidad Nacional de Colombia, Colombia, 2017.
- [16] Strelkovskaya I., Solovskaya I., Makoganiuk A., «Predicting characteristics of self-similar traffic,» International Conference on Information and Telecommunication Technologies and Radio Electronics, Ucrania, 2018.
- [17] Z. Dídimo, «Señales Aleatorias,» Universidad Nacional de Río Cuarto, Río Cuarto, Córdoba, Argentina, 2008.
- [18] M. H., «Redes de información,» Universidad Nacional de Río Cuarto, Río Cuarto, Córdoba, Argentina, 2010.
- [19] S. William, Comunicaciones y redes de computadores, Pearson, 2000.
- [20] <https://www.itu.int/>, «Redes de la próxima generación,» *Unión Internacional de Telecomunicaciones*, 2020.