

Caracterización y Análisis del Desempeño de la Red de Datos de la Pontificia Universidad Javeriana

Jairo Sebastian Rodríguez Quiroga, 22-05-2017, jaroq80@gmail.com

Resumen— En este artículo se dan a conocer los resultados alcanzados en el desarrollo de la caracterización y análisis del desempeño de la Pontificia Universidad Javeriana, mostrando como es el comportamiento de la red, los cuellos de botella existentes que impactan la percepción de servicio de la red, como los elementos de CORE se ajusta a un modelo de colas y como estas se comportan en un modelo de red de colas abiertas. Con la motivación de dar a conocer el desempeño de la red de CORE de la Universidad, por medio de un acercamiento teórico y observando cómo se puede llegar a evaluar la red, sin la necesidad de realizar acciones intrusivas que puedan impactar la red. Con los resultados obtenidos de evaluación de desempeño, se puede llegar a realizar acciones sobre esta red o simplemente tener como referencia para realizar ingeniería sobre esta.

Palabras Clave— Colas, Distribución de Probabilidad, Red de Colas, Cuellos de Botella, Caracterización.

I. INTRODUCCIÓN

Para mantener y administrar una red de datos, es indispensable manejar una calidad de servicio óptima para el uso de esta. Por tal razón es que, a las redes de comunicaciones, se les debe realizar una evaluación de desempeño que asegure su correcta operación, por medio de métricas y del conocimiento de sus patrones de funcionamiento. La evaluación de Para mantener y administrar una red de datos, es indispensable manejar una calidad de servicio óptima para el uso de esta. Por tal razón es que, a las redes de comunicaciones, se les debe realizar una evaluación de desempeño que asegure su correcta operación, por medio de métricas y del conocimiento de sus patrones de funcionamiento. La evaluación del desempeño de una red, es la cuantificación del servicio entregado por esta.

Este trabajo de grado, consistió en realizar la caracterización del tráfico que cursa en un instante de tiempo, por los nodos de CORE de la red de datos de la Pontificia Universidad Javeriana, analizando su desempeño y así poder detectar el principal punto débil de la red.

La red de datos de la Pontificia Universidad Javeriana, a nivel de los equipos de CORE, es muy simple ya que se cuenta con 2 enrutadores que soportan toda la red. Estos equipos se comunican entre sí, son el eje principal de los accesos a la red de toda la universidad y son el medio con el cual se tiene acceso a Internet.

En el desarrollo del trabajo realizado se debió identificar la cantidad de tráfico interno y externo que cursa sobre la red, para reconocer la causa de congestión sobre los nodos de CORE que fueron objeto del estudio.

Al realizar la caracterización del tamaño de paquetes y el tiempo entre paquetes, se logró analizar su comportamiento en un instante de tiempo, donde se presenta el mayor tráfico de datos sobre la red. Observando cómo al ser elementos de CORE, estos dos componentes tienen un comportamiento que refleja las fuentes del tráfico transportado sobre estos elementos.

Haciendo uso y aplicando la teoría de colas para el desarrollo del trabajo planteado, más adelante en este documento, se expondrá lo encontrado de forma detallada y para cada uno de los elementos estudiados para esta red. Adicionalmente a esto, se darán a conocer los detalles de cómo interactúan entre sí estas colas, al formar una red de colas. El desempeño de una red, es la cuantificación del servicio entregado por esta.

Una red de datos que exige tanta demanda de servicio como en la Pontificia Universidad Javeriana, se debe monitorear su desempeño y conocer el tipo de tráfico que está cursando sobre esta. Cada semestre, esta red abarca una cantidad de usuarios superior a la del semestre inmediatamente anterior, lo que le implica que es de gran importancia poder detectar los puntos más débiles de la red (cuellos de botella), que llevan a la congestión o hasta el colapso de la red, al no ser corregida en el momento adecuado. Estas correcciones no se han realizado de la mejor forma o conscientemente, ya que el departamento de DTI de la Universidad no ha realizado análisis profundos de la red, ya sea por falta de recursos en tiempo o personal, estas razones han motivado la realización de este trabajo.

Este trabajo de grado, consistió en realizar la caracterización del tráfico que cursa en un instante de tiempo, por los nodos de CORE de la red de datos de la Pontificia Universidad Javeriana, analizando su desempeño y así poder detectar el principal punto débil de la red.

Al realizar la caracterización del tamaño de paquetes y el tiempo entre paquetes, se logró analizar su comportamiento en un instante de tiempo, donde se presenta el mayor tráfico de datos sobre la red. Observando cómo al ser elementos de CORE, estos dos componentes tienen un comportamiento que

refleja las fuentes del tráfico transportado sobre estos elementos.

Haciendo uso y aplicando la teoría de colas para el desarrollo del trabajo planteado, más adelante en este documento, se expondrá lo encontrado de forma detallada y para cada uno de los elementos estudiados para esta red. Adicionalmente a esto, se darán a conocer los detalles de cómo interactúan entre sí estas colas, al formar una red de colas.

II. METODOLOGÍA

Para entender el trabajo desarrollado y los resultados alcanzados en el trabajo propuesto, se debe partir que la topología real de la red no constaba de 5 equipos como se informó al momento de proponer el trabajo de grado. Realmente la red se basa en dos nodos principales, que soportan toda la red y forman tres áreas en la red:

- El *CORE A* que está compuesto por tres equipos, que a nivel lógico se representan como un solo equipo. Estos equipos son Ingeniería, Hospital y Ático. Este equipo del *CORE A* se encuentra conectado con varios equipos de red, que forman parte de la distribución en esta parte de la red.
- El *CORE B* está compuesto por dos equipos, que al igual que en el *CORE A*, a nivel lógico se representan como un solo equipo. Estos dos equipos son Central y Biblioteca. Siguiendo lo descrito del *CORE A*, este también se encuentra conectado con los equipos de distribución y se comunica con el *CORE Alcatel*.
- El *CORE Alcatel* es la antigua red *CORE* de la Universidad, en la cual solo se encuentra en funcionamiento con un solo equipo de red, el cual se encarga de comunicar el Datacenter y proveer de la salida a Internet de la Universidad.

Para el desarrollo de este trabajo, no se tuvieron en cuenta los demás elementos de red que no hacían parte del *CORE*, como lo son los equipos de distribución y de acceso. Se tomó la red de la siguiente forma:

- *CORE A* -Distribución, equipo Arboleda.
- Enlaces *CORE A* - *CORE B*.
- *CORE B* – Distribución, equipo Biblioteca.
- Enlaces *CORE B* – Alcatel, equipo Datacenter (salida a Internet).

Para cada nube representada en la Fig 1, se modelizaron 4 colas a excepción del *CORE Alcatel*, el cual únicamente se modelizaron dos colas.

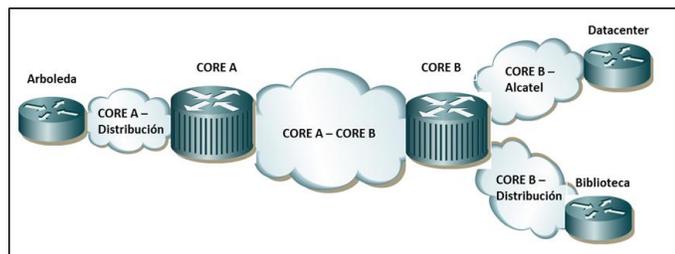


Fig. 1. Modelo de red analizado

A. Identificación tráfico interno y externo de la Red

Para la identificación del tráfico interno y externo, se realizó con ayuda del gestor de red IMC, que permitió extraer la información necesaria para alcanzar parcialmente el objetivo planteado para este punto.

B. Caracterización de tráfico

En la realización de la caracterización de tráfico, se hicieron capturas para las cuatro partes de la red, en las que se dividió esta. Estas capturas se realizaron en horario de alto tráfico que según datos tomados del gestor de los equipos IMC, debía ser un día entre semana en el periodo de 11:00 a 12:00. Por el volumen de información capturada. Las capturas oscilaron entre 10 minutos y 30 minutos. La caracterización del tráfico se realiza con ayuda del software de Matlab, el cual permite realizar el ajuste a distribuciones de manera de prueba y error. Para realizar la caracterización de tráfico se realizó usando una metodología, que consta de realizar una captura de datos de los equipos por medio del sniffer Wireshark, procesar la información con Matlab y realizando la modelización, con ayuda de una herramienta dentro de Matlab que se llama Distribution Fitting Tool. Esta aplicación permite realizar el ajuste a una distribución, para los datos capturados de la forma de prueba y error. Es decir, que los datos se pudieron ajustar a varias distribuciones de probabilidad y observar cuál se ajustaba mejor, a los datos capturados. Para cada una de las partes de la red, se obtiene la distribución estadística de tiempo entre paquetes y tamaño de paquetes. Se hace de manera unificada, sin discriminar por interfaz. Las distribuciones del tamaño del paquete se cortan abruptamente en 1500 Bytes, ya que se debe tener en cuenta que la MTU (Unidad Máxima de Transferencia) de Ethernet es de 1500 Bytes. Las muestras obtenidas que no se tienen en cuenta en la caracterización, se trabajan como fuentes de valores fijos.

C. Análisis de colas

Se realiza el análisis de colas para las cuatro partes de la red, únicamente para el tráfico TCP y modelizando 4 colas por parte, a excepción del *CORE B* – Alcatel, el cual sólo mantiene activa una interfaz, la cual se modeliza con 2 colas. Para cada cola se realizó el modelo a una cola tipo G/G/1, esto debido a que las tasas de arribo y de servicio, son procesos generales. Para cada una de las colas G/G/1 se realizó el cálculo de:

- L_q : Número medio de usuarios en la cola.
- L : Número medio de usuarios en el sistema.
- W_q : Tiempo medio de espera en la cola.
- W : Tiempo medio de espera en el sistema.
- C_a : Coeficiente de variación de la tasa de llegadas.
- C_s : Coeficiente de variación de la tasa de servicio.
- ρ : Utilización del servidor.

Por medio de las siguientes ecuaciones:

$$\rho = \frac{\lambda}{\mu}$$

λ : Tasa de llegadas

μ : Tasa de servicio

$$C_a = \frac{\sigma_a}{\lambda}$$

σ_a : Desviación estándar de tasa de llegadas

$$C_s = \frac{\sigma_s}{\mu}$$

σ_s : Desviación estándar de tasa de llegadas

$$L_q = \left[\frac{C_a^2 + C_s^2}{2} \right] \times \frac{\rho^2}{1 - \rho}$$

$$W_q = \frac{L_q}{\lambda}$$

$$W = W_q + \frac{1}{\mu}$$

$$L = \lambda W$$

D. Análisis de cuellos de botella

Para el análisis de cuellos de botella, se debió realizar primero el análisis de colas para cada nodo. Esto con el objetivo de obtener la información necesaria de cada cola y poder encontrar el eslabón más “débil” en la red. Con esto, lo que se toma es el tiempo promedio total en el sistema (para cada cola), de la siguiente manera:

$$R = \sum_{k=1}^n V_k R_k$$

V_k : Número de visitas de un paquete en la cola k.

R_k : Tiempo de respuesta medio de la cola k.

E. Análisis red de colas

Para la realización del modelo de una red de colas, de la red en estudio, se realiza la conexión de las colas y se definen los caminos que podría llegar a tomar un paquete, según la lógica en la que opera la red. Como la red de CORE de la Universidad, está conformada por dos enrutadores, el modelo de red de colas para esta red, se plantea interconectando las colas según el sentido y el volumen de tráfico que manejan. Para este modelo, se hizo uso de los datos históricos de los enrutadores, con ayuda del gestor de red IMC y así definir el sentido del tráfico, por donde se envía la información en la red. La red de colas modelada, tiene cinco procesos externos, que inyectan tráfico y cinco salidas por donde el tráfico es entregado a los usuarios finales. En la Fig. 2 se ilustra cómo fue modelada la red de colas:

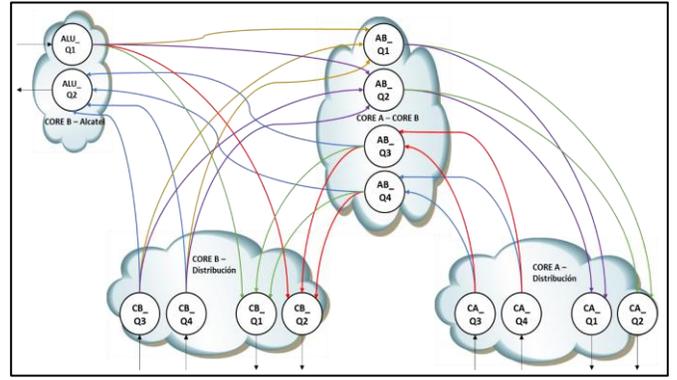


Fig. 2. Modelo de Red de Colas, de la red de datos de la Universidad

Se realiza el modelo de red de colas, basado en redes de Jackson abiertas. En este modelo se asume que todos los paquetes salen de la red y hay ingreso de paquetes desde el exterior de la red. El modelo aplicado es una red de Jackson Abierta, donde se realizó el cálculo de los siguientes parámetros de la red:

$$\lambda_j = r_j + \sum_{i=1}^n P_{ij} \lambda_i; \text{ tasa de arribo de la cola } j$$

r_j : Tasa de arribo externa.

P_{ij} : Probabilidad de pasar de la cola j a la cola i.

λ_i : Tasa de arribo de la cola i.

n : Número de colas en la red.

$\rho_j = \frac{\lambda_j}{\mu_j}$; utilización de la cola j

μ_j : Tasa de servicio de la cola

$E[Q_j] = \frac{\rho_j}{1 - \rho_j}$; tamaño medio de la cola j

$E[D_j] = \frac{E[Q_j]}{\lambda_j}$; tiempo medio en la cola j

$E[D] = \frac{\sum_{j=1}^n E[Q_j]}{\sum_{j=1}^n r_j}$; tiempo total medio en la red

$P(Q_j = 0) = \prod_{j=1}^n (1 - \rho_j)$; prob. de que la cola este vacía

III. RESULTADOS

Los resultados obtenidos se dan a conocer, en el orden que se mencionó la metodología usada.

A. Tráfico interno y Externo de la red.

La gran mayoría del tráfico sobre esta parte de la red, está relacionado con el tráfico interno de la universidad. Este tráfico

según lo indicado por el departamento DTI de la Universidad, es normal y corresponde a equipos de seguridad. Se observa que el tráfico externo aumenta a partir de las 6:00 a.m. y disminuye a las 10:00 p.m. El tráfico interno mantiene un típico por encima de 120 GB por cada hora y que al igual que el tráfico externo, este aumenta en el periodo entre las 6:00 a.m. y las 10:00 p.m.

En la Fig. 3, se muestra cómo se comporta el tráfico en la universidad durante un día. Se recuerda que esta información únicamente corresponde al CORE A. La información se muestra en volumen de datos:

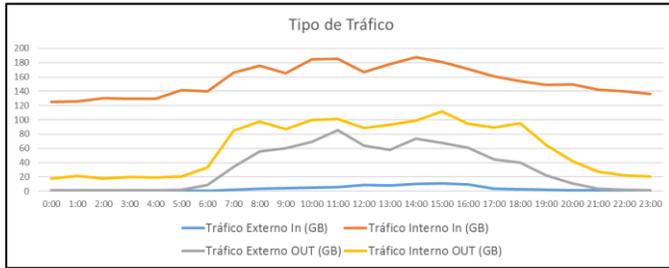


Fig. 3. Tráfico interno y externo durante el día

B. Caracterización de tráfico

La caracterización hecha para cada parte de la red, evidencia que las distribuciones de probabilidad obtenidas para el tamaño de paquetes y tiempo entre paquetes, son la suma de varias distribuciones. Las distribuciones obtenidas son:

- CORE A – CORE B

Tiempo entre paquetes, se tiene una distribución estadística compuesta por la suma de cuatro distribuciones, una distribución Gamma, una distribución Weibull y dos distribuciones Normales:

$$X \sim G(\alpha, \lambda) + W(\alpha, \lambda) + N_1(\mu_1, \sigma_1^2) + N_2(\mu_2, \sigma_2^2)$$

$$G(\alpha, \lambda), \text{ con } \alpha = 12.5 \text{ y } \lambda = 2.9e-06$$

$$W(\alpha, \lambda), \text{ con } \alpha = 1.2e-4 \text{ y } \lambda = 1.9$$

$$N_1(\mu_1, \sigma_1^2), \text{ con } \mu_1 = 9.998e-05 \text{ y } \sigma_1^2 = 7.8e-11$$

$$N_2(\mu_2, \sigma_2^2), \text{ con } \mu_2 = 2.02e-4 \text{ y } \sigma_2^2 = 1.99e-10$$

Tamaño de paquetes, es la suma de una distribución Exponencial, con dos distribuciones Normales:

$$X \sim Ex(\lambda) + N_1(\mu_1, \sigma_1^2) + N_2(\mu_2, \sigma_2^2)$$

$$Ex(\lambda), \text{ con } \lambda = 616$$

$$N_1(\mu_1, \sigma_1^2), \text{ con } \mu_1 = 174.052 \text{ y } \sigma_1^2 = 1766.06$$

$$N_2(\mu_2, \sigma_2^2), \text{ con } \mu_2 = 1006.25 \text{ y } \sigma_2^2 = 58013.5$$

- CORE A – Distribución

Tiempo entre paquetes, es la suma de una distribución Normal y una distribución Weibull:

$$X \sim N(\mu, \sigma^2) + W(\alpha, \lambda)$$

$$N(\mu, \sigma^2), \text{ con } \mu = 1.001e-4 \text{ y } \sigma^2 = 9.24e-11$$

$$W(\alpha, \lambda), \text{ con } \alpha = 1.25e-4 \text{ y } \lambda = 1.925$$

Tamaño de paquetes, es la suma de una distribución Normal, una distribución Rayleigh y una distribución Uniforme:

$$X \sim N(\mu, \sigma^2) + R(\theta) + U(a, b)$$

$$N(\mu, \sigma^2), \text{ con } \mu = 159.687 \text{ y } \sigma^2 = 3056.68$$

$$R(\theta), \text{ con } \theta = 541.865$$

$$U(a, b), \text{ con } a=1010 \text{ y } b=1500$$

- CORE B – Distribución

Tiempo entre paquetes, es la suma de dos distribuciones Normales, una distribución Weibull y una Rayleigh:

$$X \sim R(\theta) + N_1(\mu_1, \sigma_1^2) + N_2(\mu_2, \sigma_2^2) + W(\alpha, \lambda)$$

$$R(\theta), \text{ con } \theta = 2.94e-05$$

$$N_1(\mu_1, \sigma_1^2), \text{ con } \mu_1 = 1.03e-4 \text{ y } \sigma_1^2 = 1.91e-10$$

$$N_2(\mu_2, \sigma_2^2), \text{ con } \mu_2 = 1.001e-4 \text{ y } \sigma_2^2 = 3.9e-11$$

$$W(\alpha, \lambda), \text{ con } \alpha = 2.09e-4 \text{ y } \lambda = 9.64342$$

Tamaño de paquetes, es la suma de una distribución Normal y una distribución Weibull:

$$X \sim N(\mu, \sigma^2) + W(\alpha, \lambda)$$

$$N(\mu, \sigma^2), \text{ con } \mu = 136.34 \text{ y } \sigma^2 = 1112.32$$

$$W(\alpha, \lambda), \text{ con } \alpha = 295.242 \text{ y } \lambda = 1.19522$$

- CORE B – Alcatel

Tiempo entre paquete, es la suma de una distribución Rayleigh y 5 distribuciones Normales:

$$X \sim N_1(\mu_1, \sigma_1^2) + R(\theta) + N_2(\mu_2, \sigma_2^2) + N_3(\mu_3, \sigma_3^2) + N_4(\mu_4, \sigma_4^2) + N_5(\mu_5, \sigma_5^2)$$

$$N_1(\mu_1, \sigma_1^2), \text{ con } \mu_1 = 3.56e-5 \text{ y } \sigma_1^2 = 1.23e-10$$

$$R(\theta), \text{ con } \theta = 4.42e-5$$

$$N_2(\mu_2, \sigma_2^2), \text{ con } \mu_2 = 6.86e-5 \text{ y } \sigma_2^2 = 3.88e-11$$

$$N_3(\mu_3, \sigma_3^2), \text{ con } \mu_3 = 1.01e-4 \text{ y } \sigma_3^2 = 5.49e-11$$

$$N_4(\mu_4, \sigma_4^2), \text{ con } \mu_4 = 2.61e-5 \text{ y } \sigma_4^2 = 5.64e-12$$

$$N_5(\mu_5, \sigma_5^2), \text{ con } \mu_5 = 4.53e-5 \text{ y } \sigma_5^2 = 1.9e-11$$

Tamaño de paquetes, es la suma de una distribución Normal Logarítmica y una Normal:

$$X \sim N(\mu, \sigma^2) + LogN(\mu, \sigma^2)$$

$$N(\mu, \sigma^2), \text{ con } \mu = 147.7 \text{ y } \sigma^2 = 190,04$$

$$LogN(\mu, \sigma^2), \text{ con } \mu = 4,2 \text{ y } \sigma^2 = 154,6$$

Al realizar la caracterización de tráfico sobre la red, se logra observar que la mayor parte del tráfico que cursa sobre la distribución de este nodo, está relacionada con TCP y seguido por el protocolo de encapsulamiento GRE, el protocolo UDP es de muy poco tráfico en relación con los dos primeros. Aunque

no se tenía como meta la medición del protocolo GRE, se incluyó por tener gran cantidad del tráfico cursado sobre la red. El protocolo GRE (Protocolo Genérico de Encapsulamiento), es un protocolo creado por Cisco Systems, el cual permite establecer túneles a través de Internet, y encapsular varios protocolos hasta de capa 3, entre los que están OSPF, RIP, ARP y MPLS. En la Fig. 4, se puede observar cómo se reparte el tráfico de estos tres protocolos, donde se puede evidenciar la gran mayoría del protocolo TCP versus los demás.

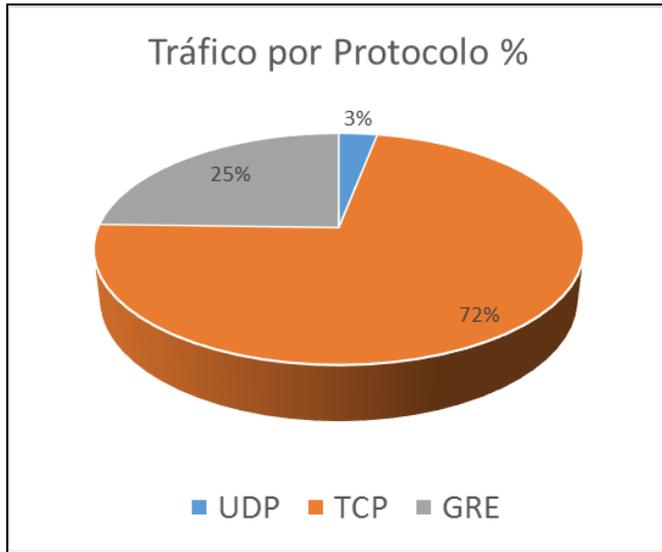


Fig. 4. Tráfico por protocolo

Como se observa en la Fig. 5, el tráfico durante el día, los protocolos de UDP y GRE, aumentan cuando inicia la jornada de estudio en la universidad. Se observa un crecimiento en el tráfico TCP en el mismo horario, pero este último maneja un tráfico mínimo de 144 GB durante el día, en el que se relaciona con lo indicado por el departamento DTI de la Universidad, en que es normal y corresponde a equipos de seguridad y plataformas de servicio de la Universidad.

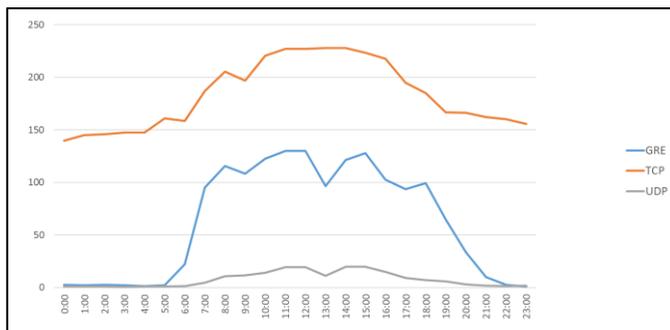


Fig. 5. Volumen de datos por protocolo durante el día

En la Fig. 6, se observa como la mayor parte del tráfico (74%) corresponde a aplicaciones http y https, se observa un 24% del total de tráfico para aplicaciones no conocidas sobre TCP. Se presentan muchos más protocolos, pero al tener muy poca presencia en los datos analizados, no se incluyen en el gráfico.

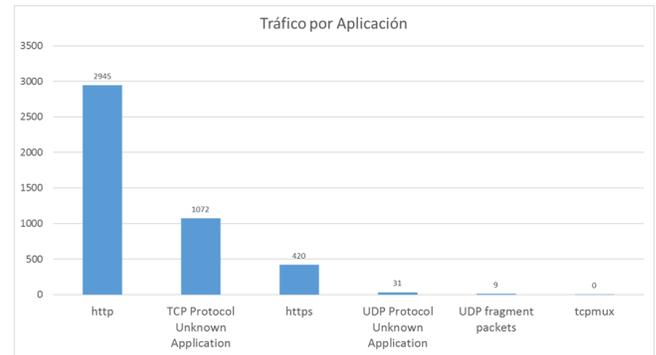


Fig. 6. Tráfico por Aplicación

C. Modelo de colas

En el modelo de colas se realiza asumiendo varias fuentes que suministran las distribuciones para la tasa de arribo y la tasa de servicio para cada una de las colas. Estas fuentes no son únicamente fuentes de alguna distribución, sino que también se tienen fuentes de valores fijos.

El modelo realizado de colas, arroja a que las interfaces de los equipos de CORE son del tipo G/G/1 (es decir, llegadas y tiempo de servicio generales, un servidor con buffer infinito), este tipo de cola se ajusta al modelo, por la razón que los procesos de llegada y servicio, son dados por procesos no exponenciales o poissonianos.

En la Tabla 1, se dan a conocer las tasas de servicio y de arribo en cada cola. así como su descripción.

En la Tabla 2, se muestran los resultados del modelo de cola para cada una de estas. Estos valores reflejan cómo se comportan las colas a nivel individual, en base de las capturas realizadas sobre los equipos. Es de aclarar que el modelo se aplica tomando paquetes como si fueran usuarios y las tasas están dadas en paquetes/segundos.

En ambas tablas se puede ver como la utilización en cada cola es inferior al 50%.

TABLA 1
DESCRIPCIÓN DE CADA COLA

No. Cola	Cola	Descripción	λ	μ
1	CA_Q1	Cola que va del CORE A hacia Distribución	335630	935370
2	CA_Q2	Cola que va del CORE A hacia Distribución	331980	969430
3	CA_Q3	Cola que va de Distribución al CORE A	400430	1219500
4	CA_Q4	Cola que va de Distribución al CORE A	830,872 9	18939000
5	AB_Q1	Cola que va del CORE B al CORE A	867460	1854700
6	AB_Q2	Cola que va del CORE B al CORE A	1328600	8659300
7	AB_Q3	Cola que va del CORE A al CORE B	348320	882890
8	AB_Q4	Cola que va del CORE A al CORE B	454200	918190

9	CB_Q1	Cola que va del CORE B hacia Distribución	551060	1124400
10	CB_Q2	Cola que va del CORE B hacia Distribución	56595	19984000
11	CB_Q3	Cola que va de Distribución al CORE B	51675	19139000
12	CB_Q4	Cola que va de Distribución al CORE B	103200	1804100
13	ALU_Q1	Cola que va del CORE Alcatel a CORE B	15442	58661
14	ALU_Q2	Cola que va del CORE B a CORE Alcatel	164140	1616100

Para μ y λ , están dadas como paquetes/segundo

TABLA 2
DESCRIPCIÓN DE CADA COLA

No. Cola	ρ	L_q	W_q	W	L
1	3,59e-01	2,21	6,58e-06	7,65e-06	2,57
2	3,42e-01	5,22	1,57e-05	1,68e-05	5,57
3	3,28e-01	2,59	6,46e-06	7,28e-06	2,91
4	4,39e-05	8,49e-10	1,02e-12	5,28e-08	4,39e-05
5	4,68e-01	1,23	1,42e-06	1,96e-06	1,70
6	1,53e-01	1,85e+02	1,39e-04	1,40e-04	1,85e+02
7	3,95e-01	7,92	2,27e-05	2,39e-05	8,31
8	4,95e-01	1,25e+01	2,75e-05	2,86e-05	1,30e+01
9	4,90e-01	7,83	1,42e-05	1,51e-05	8,32
10	2,80e-03	1,01	1,79e-05	1,79e-05	1,01
11	2,70e-03	1,73e-01	3,35e-06	3,40e-06	1,76e-01
12	5,72e-02	5,86e-01	5,68e-06	6,23e-06	6,43e-01
13	2,63e-01	9,41	6,10e-04	6,27e-04	9,68
14	1,02e-01	4,71e-01	2,87e-06	3,49e-06	5,72e-01

ρ : utilización, L_q : paquetes en la cola, W_q : tiempo medio en la cola (segundos), W : tiempo medio en el sistema (segundos), L : paquetes en el sistema

D. Cuellos de botella

Un cuello de botella es el nodo o elemento dentro de una red, que hace que el tiempo de respuesta de toda la red se maximice. En palabras más simples, es el elemento más débil de toda la red.

Al analizar una red de datos, se pueden encontrar múltiples elementos que la limitan. Estos elementos, son los cuellos de botella presentes en la red. Para resolver un cuello de botella o mejor, cambiar la condición de un elemento, para que deje de ser el eslabón más débil dentro de la red, no se debe necesariamente atacarse directamente este elemento, puede presentarse por consecuencia de otro elemento, que altera el comportamiento de este. El resolver un cuello de botella, hará que surja dentro de la red otro u otros cuellos de botellas, por lo cual antes de resolver un cuello de botella, se debe analizar y

estimar que tanto va a aliviar la red y que tanto se impactarán los demás elementos.

Para la red en estudio, se realiza el cálculo para cada cola, asumiendo (número de visitas de un paquete en la cola), debido a que en condiciones normales un paquete no debe pasar dos veces sobre una misma interfaz. Así el tiempo total de respuesta es:

$$R = \sum_{k=1}^n V_k R_k = 1,015e-3$$

Ahora para determinar el cuello de botella de la red, se toma el valor que maximiza R , el cual es la cola ALU_Q1 que corresponde a la cola que va del CORE Alcatel y entra al CORE B. Este tráfico proviene del Datacenter y de Internet con destino a la red de la Universidad. Aunque se tomó una muestra parcial del total de tráfico que realmente transita sobre esta interfaz, aun así, se puede observar como es el de mayor tiempo en el sistema y por ende el cuello de botella en la red.

Adicional, se tiene una segunda cola que está por encima de la mayoría de las demás colas. Esta cola es la AB_Q2, que es la cola que va del CORE B al CORE A, el tráfico que cursa sobre esta cola en su mayoría corresponde a contenidos que provienen de internet. Ya que como se dijo anteriormente, la forma como está desplegada la red, nos lleva a que la única salida y entrada de información hacia Internet, sea por estas interfaces.

En la Tabla 3, se puede observar el total del análisis de cuellos de botella presentes en la red. Se resaltan los valores de las colas que se tienen como cuellos de botella.

TABLA 3
CUELLOS DE BOTELLA

Cola	V_k	R_k
CA_Q1	1	9,78e-5
CA_Q2	1	4,65e-5
CA_Q3	1	7,28e-6
CA_Q4	1	5,28e-8
AB_Q1	1	1,96e-6
AB_Q2	1	1,40e-4
AB_Q3	1	2,39e-5
AB_Q4	1	2,86e-5
CB_Q1	1	1,22e-5
CB_Q2	1	1,79e-5
CB_Q3	1	3,40e-6
CB_Q4	1	6,23e-6
ALU_Q1	1	6,27e-4
ALU_Q2	1	3,49e-6

$$R = \sum_{k=1}^n V_k R_k = 1,015e-03$$

E. Modelo de red de colas

Se realiza el cálculo de una red de Jackson abierta, tomando como entrada los valores obtenidos de tasa de servicio y tasa de arribo en el modelo de colas. Se toma esta red de colas como la conexión de las colas como se definió anteriormente en la Figura 4. Inicialmente se toman estas colas tipo M/M/1.

Para realizar el modelo se debe armar un arreglo de probabilidades, en el que la suma de cada fila no puede ser mayor a 1. Cada cola Q_j , tiene una probabilidad P_{ij} de pasar a la cola Q_i . Los cálculos de esta red nos arrojan los siguientes valores:

Tiempo total medio en la red:

$$E[D] = 3,4867 \mu s$$

Probabilidad de que la red este vacía:

$$P(Q_j = 0) = 0,1754$$

Los valores de cada cola calculados se pueden ver en la Tabla 4.

En la red modelada se puede observar que como se determinó en el análisis de cuellos de botella, la cola que más alto tiene el tiempo medio en la cola es ALU_Q1, su utilización es de 26% con los datos modelados, aunque como se dijo la salida a Internet se encuentra con utilización al 100%, con los datos que se lograron capturar no se alcanza a llegar a esta.

Además, si se compara el total de los datos obtenidos, con los datos modelados para cada cola, se evidencia que las colas intermedias su tasa de arribo es menor. Esto se debe a que, en el desarrollo del trabajo, no se modelaron todas las interfaces y al no ser modeladas en su totalidad se hace evidente esta diferencia.

TABLA 4
RESULTADOS OBTENIDOS EN CÁLCULO DE LA RED DE COLAS

Cola	j	ρ_j	$E[Q_j]$	$E[D_j], (\mu s)$
CA_Q1	1	0,103100	0,1150	1,1347
CA_Q2	2	0,065500	0,0701	1,1343
CA_Q3	3	0,328400	0,4889	1,2209
CA_Q4	4	0,000044	0,00004	0,0528
AB_Q1	5	0,069000	0,0742	0,5792
AB_Q2	6	0,004100	0,0041	0,116
AB_Q3	7	0,297200	0,4229	1,6117
AB_Q4	8	0,151200	0,1781	1,2831
CB_Q1	9	0,131600	0,1516	1,0389
CB_Q2	10	0,006800	0,0068	0,0504
CB_Q3	11	0,002700	0,0027	0,0524
CB_Q4	12	0,057200	0,0607	0,5879
ALU_Q1	13	0,263200	0,3573	23,138
ALU_Q2	14	0,057000	0,0605	0,6562

Analizando el tiempo total medio en la red, que es un valor muy bajo a pesar de que hablamos de interfaces de 10 Gbps en

su mayoría que se encuentran en la red, es debido a lo anteriormente mencionado.

La probabilidad de que la red se encuentre vacía, es muy alta para ser el horario en que se realizaron las capturas, un 17% es una probabilidad muy alta.

Para realizar un mejor modelo de cómo se comporta esta red de colas, se realizan dos simulaciones de redes de colas, la primera trabajando las colas como una cola tipo M/M/1 y la segunda como colas tipo G/G/1. Esto con razón de determinar que los datos calculados y simulados para la red de colas M/M/1 sean consistentes y al realizar el cambio en la simulación por colas G/G/1, sea una simulación válida.

En la Tabla 5, se puede observar los datos obtenidos para la primera simulación, la cual se toma realizando una simulación con datos de entrada los valores obtenidos para la tasa de arribo y la tasa de servicio, del modelo realizado para cada una de las colas.

Tiempo total medio en la red:

$$E[D] = 15,728 \mu s$$

Probabilidad de que la red este vacía:

$$P(Q_j = 0) = 0,1754$$

En la simulación se tiene que los valores obtenidos para $E[Q_j]$, son muy diferentes a los calculados, pero en cambio

los datos obtenidos para $E[D_j]$, tienden a ser más parecidos y la diferencia es mínima.

En cuanto al tiempo medio total en la red, se aproxima más a un valor en condiciones reales de la red. Realizando un cálculo aproximado, un paquete de 1514 Bytes que se procese por una interfaz de 10 Gbps, demora aproximadamente un tiempo de 1,2 μs , asumiendo que el paquete pase por tres interfaces y sea el único paquete en la red, su tiempo será de 3,6 μs . Es el valor obtenido en los cálculos realizados, pero si le adicionamos la variable de carga real sobre cada nodo, este tiempo aumentará.

TABLA 5
RESULTADOS OBTENIDOS EN SIMULACIÓN DE LA RED DE COLAS M/M/1

Cola	j	ρ_j	$E[Q_j]$	$E[D_j], (\mu s)$
CA_Q1	1	0,103116	0,6313	1,1347
CA_Q2	2	0,065548	0,5723	1,1343
CA_Q3	3	0,328356	1,0455	1,2222
CA_Q4	4	0,000044	0,50025	0,0528
AB_Q1	5	0,069049	0,5473	0,5792
AB_Q2	6	0,004054	0,5033	0,116
AB_Q3	7	0,297234	0,9075	1,6112
AB_Q4	8	0,151207	0,6813	1,2833
CB_Q1	9	0,131638	0,6335	1,0389
CB_Q2	10	0,006757	0,5053	0,0504
CB_Q3	11	0,002700	0,5023	0,0524
CB_Q4	12	0,057203	0,5653	0,5879
ALU_Q1	13	0,263241	0,8374	23,132
ALU_Q2	14	0,057046	0,5573	0,6562

En la Tabla 6, se dan a conocer los valores obtenidos en la simulación de la red de colas, tomando cada cola como una G/G/1. Para esta simulación se, realiza la simulación de todas las colas y luego se realiza la conexión de estas, tal y como se trabajaron con las anteriores simulaciones.

TABLA 6
RESULTADOS OBTENIDOS EN SIMULACIÓN DE LA RED DE COLAS G/G/1

Cola	j	ρ_j	$E[Q_j]$	$E[D_j], (\mu s)$
CA_Q1	1	0,042304841	0,532994924	0,1175
CA_Q2	2	0,548605147	0,512562814	0,7107
CA_Q3	3	0,033872324	0,512562814	0,3509
CA_Q4	4	0,000020800	0,502512563	0,0252
AB_Q1	5	0,007478636	<u>0,502512563</u>	0,3463
AB_Q2	6	0,023712337	0,502512563	0,1800
AB_Q3	7	0,040439806	0,512562814	0,3422
AB_Q4	8	0,017119283	0,502512563	0,2755
CB_Q1	9	0,030502254	0,522613065	0,1983
CB_Q2	10	0,000505115	0,502512563	0,0121
CB_Q3	11	0,001386992	0,502512563	0,0526
CB_Q4	12	0,005473189	0,502512563	0,5378
ALU_Q1	13	0,020047957	0,502512563	1,4546
ALU_Q2	14	0,066247341	0,512562814	0,8374

Tiempo total medio en la red:

$$E[D] = 47,05 \mu s$$

Probabilidad de que la red este vacía:

$$P(Q_j = 0) = 0,3361$$

Al realizar el modelo con colas G/G/1, los valores son muy diferentes a los calculados y simulados con colas M/M/1. Los valores de $E[Q_j]$ son similares a los simulados con colas M/M/1, pero $E[D_j]$ varía mucho. Hay algo que resaltar y es que la cola ALU_Q1, siempre tiende a comportarse de la misma forma en los tres casos, es que siempre va a ser la de mayor tiempo en el sistema. El tiempo medio en cada cola G/G/1, en la simulación, tienden a ser de valores muy inferiores que para las colas M/M/1, por la razón que el proceso exponencial en la cola M/M/1 de la tasa de servicio, hace que la máxima probabilidad sea para la tasa de servicio más bajas y esta probabilidad decrece exponencialmente entre más altas sean estas tasas, en cambio en la cola G/G/1, la probabilidad no decrece exponencialmente, sino que puede aumentar, en valores de tasas de servicio más altas. Esta es la principal razón por la cual los valores tienden a ser tan diferentes entre ambas simulaciones.

Cuando se realiza una simulación del modelo con colas M/M/1, se tiene que el comportamiento de las tasas de servicio y de arriba sean de un comportamiento determinado, por así decirlo. En cambio, al realizar el modelo con colas G/G/1, se tienen múltiples fuentes que cambian el comportamiento de la red de colas. Esto nos lleva a entender que el comportamiento

de las colas G/G/1 modelizadas en la red, son muy dependientes de la cantidad de paquetes simulados. Se realiza con el mayor posible y la simulación arroja el mejor resultado expuesto en la Tabla 6.

IV. CONCLUSIONES

La caracterización y análisis de desempeño realizado a la red de datos de la Pontificia Universidad Javeriana, da cuenta de varios aspectos que tal vez no se tienen presentes en el saber general de los administradores y usuarios de la red. Estos aspectos nos llevan a concluir que:

1. La red de datos la Pontificia Universidad Javeriana, es una red en la que el tráfico que cursa sobre esta en su mayoría es interno. Tráfico que se relaciona con equipos de seguridad y de servicios, que ofrece la institución dentro de las instalaciones.

2. Esta red tiene su mayor demanda en los días hábiles de lunes a viernes en horario de 11 de la mañana a 12 del mediodía, donde la mayor parte del tráfico que se atiende en capa de transporte es TCP y cuando inicia la jornada académica, aumenta considerablemente este tráfico.

3. La presencia del protocolo de encapsulamiento GRE es de gran relevancia dentro de la red, se presenta en una mayor cantidad que UDP, aumenta y únicamente está presente en la red durante la jornada académica de la Universidad. El tráfico de este protocolo se evidenció en gran cantidad en las capturas realizadas y en el gestor IMC.

4. Los protocolos de capa de aplicación que más demandan tráfico dentro de la red corresponden a HTTP y HTTPS.

5. Los nodos de CORE no tienen una distribución definida para los paquetes que arriban (tamaño y tiempo entre llegadas), se comportan como la suma de varias fuentes, cada una con su propia distribución.

6. Al realizar un análisis teórico de la red permitió encontrar el modelo de colas que mejor se ajusta a la red. El modelo de colas aplicado sirvió para dar cuenta de la demanda real de la red de CORE. El modelo realizado sirve para realizar un trabajo posterior, donde se necesite relacionar estos elementos y su comportamiento.

7. Es una red con tendencia a tener una ocupación considerable. Con el modelo de colas realizado se pudo observar que, si bien no se encuentra en un valor cercano a un estado de congestión, la ocupación se puede presentar hasta valores cercanos del 50%.

8. Dentro del análisis de colas, el cuello de botella presente en la red es la salida a Internet contratada. Se pudo determinar, gracias al modelo de colas realizado, que la ocupación es más alta en esta conexión y es el elemento que más impacta el desempeño del servicio para un usuario final.

9. El modelo de red de colas aplicado a la red CORE de la Universidad, se modelizó como una red de Jackson abierta, la cual por las características de las colas de cada equipo y por el número de colas incluidas en el modelo, no se obtiene un modelo que aplique en su totalidad a la red de la Universidad, como se pretendía al iniciar este trabajo. Los valores obtenidos al realizar cálculos y simulaciones, lleva a que el modelo de redes de Jackson, como fue aplicado no se ajuste completamente a esta red.

Para una futura continuación de este trabajo se recomienda:

- Mejorar el modelo de colas, con la totalidad de las interfaces que llegan al CORE. Esto permitirá establecer mejor los modelos realizados.
- La red de colas que se debe trabajar, debe ser abierta como se trabajó en este documento e incluir la totalidad de las colas de todas las interfaces, esto con el ánimo de que las fuentes estén completas y el modelo se ajuste mejor a la realidad.
- Usar en lo posible un equipo de cómputo con una mejor capacidad de procesamiento y una interfaz 10 Gigabit Ethernet.

BIBLIOGRAFÍA

- [1]Alberto Dainotti, A. P. (2006). A Packet-level Characterization of Network Traffic. Napoli, Italia: 11th International Workshop on Computer-Aided Modeling, Analysis and Design of Communication Links and Networks.
- [2]Boudec, J.-Y. L. (2015). Performance Evaluation of Computer and Communication Systems. EPFL.
- [3]Haitao Wang, L. S. (2013). Study On Network Measurement Technologies And Performance Evaluation Methods. Jiangsu, China: Cross Strait Quad-Regional Radio Science and Wireless Technology Conference (CSQRWC).
- [4]Juan Li, Y. Q. (2011). A New Method of Network Bottleneck Diagnosis. Beijing: 12th IFIP/IEEE International Symposium on Integrated Network Management (IM 2011) and Workshops.
- [5]V. Bolotin, J. Coombs-Reyes, D. Heyman, Y. Levy and D. Liu. IP Traffic Characterization for Planning and Control. AT&T Labs, Middletown, NJ U.S.A.
- [6]Zukerman, M. (2000). Introduction to Queueing Theory and Stochastic Teletraffic Models. City University of Hong Kong: EE Department.
- [7]Xiaoqiao (George) Meng, Starsky H.Y. Wong, Yuan Yuanz, Songwu Lu (2.004). Characterizing Flows in Large Wireless Data Networks. University of California & University of Maryland. MobiCom '04 Proceedings of the 10th annual international conference on Mobile computing and networking.
- [8]Thomas Bonald, M. F. (2011). Network Performance Analysis. London, UK: ISTE.
- [9]U. Narayan Bhat (2008). An Introduction to Queueing Theory - Modeling and Analysis in Applications. Dallas, TX, U.S.A. Birkhäuser

AUTOR

Jairo Sebastian Rodríguez Quiroga, nacido en Bogotá D.C. Colombia, en 1986. Con título profesional de Ingeniero Electrónico de la Universidad Manuela Beltrán, Bogotá D.C. Colombia, en 2010. Desde el año 2009 a 2010, fue Analista Técnico en INDRA Colombia Ltda., realizando labores de monitoreo de servicios de red celular para el proyecto de Movistar Colombia. Desde el año 2010 a 2014, se desempeñó como Técnico de Centro Gestión, en Colombia Telecomunicaciones S.A., realizando labores de supervisión, monitoreo y gestión de fallas de la red móvil celular. A partir del año 2014 se desempeña como Profesional de Centro Gestión y Monitoreo.

Sus temas de interés se encuentran relacionados con telecomunicaciones, redes y Bases de Datos.