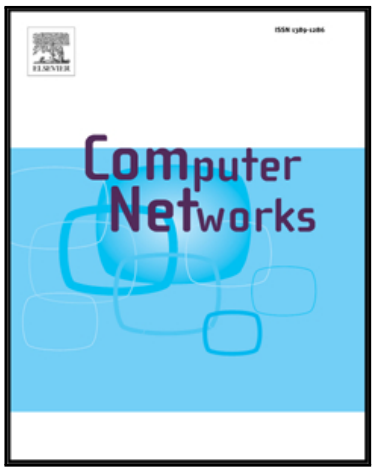

Multi-Gbps HTTP Traffic Analysis in Commodity Hardware Based on Local Knowledge of TCP Streams

Carlos Vega, Paula Roquero, Javier Aracil

PII: S1389-1286(17)30001-4
DOI: 10.1016/j.comnet.2017.01.001
Reference: COMPNW 6082

To appear in: *Computer Networks*

Received date: 2 September 2016
Revised date: 9 December 2016
Accepted date: 1 January 2017

Please cite this article as: Carlos Vega, Paula Roquero, Javier Aracil, Multi-Gbps HTTP Traffic Analysis in Commodity Hardware Based on Local Knowledge of TCP Streams, *Computer Networks* (2017), doi: 10.1016/j.comnet.2017.01.001

Multi-Gbps HTTP Traffic Analysis in Commodity Hardware Based on Local Knowledge of TCP Streams

Carlos Vega[a], Paula Roquero[a], Javier Aracil[a]

[a] *Departamento de Tecnología Electrónica y de las Comunicaciones, Escuela Politécnica Superior, Universidad Autónoma de Madrid. C/Francisco Tomás y Valiente 11 (28049) Madrid*

Abstract

In this paper we propose and implement novel techniques for performance evaluation of web traffic (response time, response code, etc.), with no reassembly of the underlying TCP connection, which severely restricts the traffic analysis throughput. Furthermore, our proposed software for HTTP traffic analysis runs in standard hardware, which is very cost-effective. Besides, we present sub-TCP connection load balancing techniques that significantly increase throughput at the expense of losing very few HTTP transactions. Such techniques provide performance evaluation statistics which are indistinguishable from the single-threaded alternative with full TCP connection reassembly.

Keywords: HTTP, Traffic Analysis, High Speed Analysis

1. Introduction

Large organizations such as banks, etc. make an increasing share of their business through the Internet [1]. Typically, HTTP is the protocol of choice to deliver services to the end-user, thanks to the widespread deployment of web clients in all kinds of mobile and desktop devices. Therefore, measuring the Quality of Service (QoS) provided by web portals [2] becomes of strategic importance. The same applies to other application protocols (VoIP, SIP, RTP, RTCP) [3] but we focus on HTTP due to its widespread usage. Such QoS evaluation is normally based on response time statistics (from HTTP query to reply) and also on the analysis of response codes for the detection of anomalous behaviour in the monitored web services. For example, an HTTP error 500 indicates an internal server error, which must be taken care of.

The dissection and analysis of HTTP traffic can also be performed for cybersecurity purposes. However, the latter analysis is very fine-grain because security threats try to masquerade themselves among normal HTTP traffic.

Email addresses: carlosgonzalo.vega@predoc.uam.es (Carlos Vega), paula.roquero@uam.es (Paula Roquero), javier.aracil@uam.es (Javier Aracil)

Preprint submitted to Computer Networks *Received: date / Accepted: date*