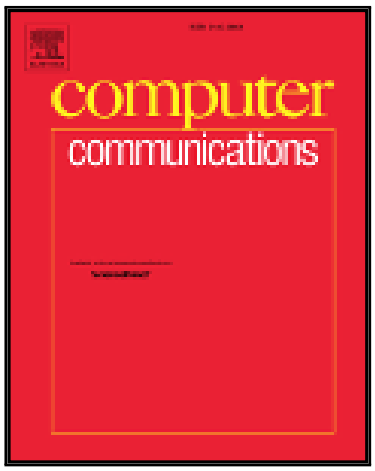

PSPAT: software packet scheduling at hardware speed

Luigi Rizzo, Paolo Valente, Giuseppe Lettieri, Vincenzo Maffione

PII: S0140-3664(17)30080-4
DOI: 10.1016/j.comcom.2018.02.018
Reference: COMCOM 5656

To appear in: *Computer Communications*

Received date: 18 January 2017
Revised date: 26 November 2017
Accepted date: 19 February 2018

Please cite this article as: Luigi Rizzo, Paolo Valente, Giuseppe Lettieri, Vincenzo Maffione, PSPAT: software packet scheduling at hardware speed, *Computer Communications* (2018), doi: 10.1016/j.comcom.2018.02.018

PSPAT: software packet scheduling at hardware speed

Luigi Rizzo[a], Paolo Valente[c], Giuseppe Lettieri[b], Vincenzo Maffione[b]

[a]*Dipartimento di Ingegneria dell'Informazione, Università di Pisa; now at Google*
[b]*Dipartimento di Ingegneria dell'Informazione, Università di Pisa*
[c]*Dipartimento di Science Fisiche, Informatiche e Matematiche, Università degli Studi di Modena e Reggio Emilia*

Abstract

Tenants in a cloud environment run services, such as Virtual Network Function instantiations, that may legitimately generate millions of packets per second. The hosting platform, hence, needs robust packet scheduling mechanisms that support these rates and, at the same time, provide isolation and dependable service guarantees under all load conditions.

Current hardware or software packet scheduling solutions fail to meet all these requirements, most commonly lacking on either performance or guarantees.

In this paper we propose an architecture, called PSPAT, to build efficient *and robust* software packet schedulers suitable to high speed, highly concurrent environments. PSPAT decouples clients, scheduler and device driver through lock-free mailboxes, thus removing lock contention, providing opportunities to parallelise operation, and achieving high and dependable performance even under overload.

We describe the operation of our system, discuss implementation and system issues, provide analytical bounds on the service guarantees of PSPAT, and validate the behaviour of its Linux implementation even at high link utilization, comparing it with current hardware and software solutions. Our prototype can make over 28 million scheduling decisions per second, and keep latency low, even with tens of concurrent clients running on a multi-core, multi-socket system.

Email address: rizzo@iet.unipi.it (Luigi Rizzo)