Q∗: Energy and delay-efficient dynamic queue management in TCP/IP virtualized data centers

Enzo Baccarelli, Paola G. Vinueza Naranjo, Mohammad Shojafar, Michele Scarpiniti

PII: S0140-3664(16)30689-2
DOI: 10.1016/j.comcom.2016.12.010
Reference: COMCOM 5425

To appear in: *Computer Communications*

Received date: 21 July 2016
Revised date: 24 October 2016
Accepted date: 15 December 2016

Please cite this article as: Enzo Baccarelli, Paola G. Vinueza Naranjo, Mohammad Shojafar, Michele Scarpiniti, Q∗: Energy and delay-efficient dynamic queue management in TCP/IP virtualized data centers, *Computer Communications* (2016), doi: 10.1016/j.comcom.2016.12.010

Q^*: Energy and delay-efficient dynamic queue management in TCP/IP virtualized data centers

Enzo Baccarelli[a], Paola G. Vinueza Naranjo[a], Mohammad Shojafar[a], Michele Scarpiniti[a,*]

[a]*Department of Information Engineering, Electronic and Telecommunication, Sapienza University of Rome, via Eudossiana 18, 00184, Rome, Italy*

Abstract

The emerging utilization of Software-as-a-Service (SaaS) Fog computing centers as an Internet virtual computing commodity is raising concerns over the energy consumptions of networked data centers for the support of delay-sensitive applications. In addition to the energy consumed by the servers, the energy wasted by the network devices that support TCP/IP reliable inter-Virtual Machines (VMs) connections is becoming a significant challenge. In this paper, we propose and develop a framework for the joint characterization and optimization of TCP/IP SaaS Fog data centers that utilize a bank of queues for increasing the fraction of the admitted workload. Our goal is two-fold: (i) we maximize the average workload admitted by the data center; and, (ii) we minimize the resulting networking-plus-computing average energy consumption. For this purpose, we exploit the Lyapunov stochastic optimization approach, in order to design and analyze an optimal (yet practical) online joint resource management framework, which dynamically performs: (i) admission control; (ii) dispatching of the admitted workload; (iii) flow control of the inter-VM TCP/IP connections; (iv) queue control; (v) up/down scaling of the processing frequencies of the instantiated VMs; and, (vi) adaptive joint consolidation of both physical servers and TCP/IP connections. The salient features of the resulting scheduler (e.g., the Q^* scheduler) are that: (i) it admits distributed and scalable implementation; (ii) it provides deterministic bounds on the instantaneous queue backlogs; (iii) it avoids queue overflow phenomena; and, (iv) it effectively tracks the (possibly unpredictable) time-fluctuations of the input workload, in order to perform joint resource consolidation without requiring any *a priori* information and/or forecast of the input workload. Actual energy and delay performances of the proposed scheduler are numerically evaluated and compared against the corresponding ones of some competing and state-of-the-art schedulers, under: (i) Fast - Giga - 10Giga Ethernet switching technologies; (ii) various settings of the reconfiguration-consolidation costs; and, (iii) synthetic, as well as real-world workloads. The experimental results support the conclusion that the proposed scheduler can achieve over 30 percent energy savings.

Keywords: TCP/IP virtualized data centers, energy-vs.-delay tradeoff, dynamic resource reconfiguration-consolidation, SaaS cloud, Fog computing.

*Corresponding author. Tel.: +390644585869.

Email addresses: enzo.baccarelli@diet.uniroma1.it (Enzo Baccarelli), paola.vinueza@uniroma1.it (Paola G. Vinueza Naranjo), shojafar@diet.uniroma1.it (Mohammad Shojafar), michele.scarpiniti@uniroma1.it (Michele Scarpiniti)