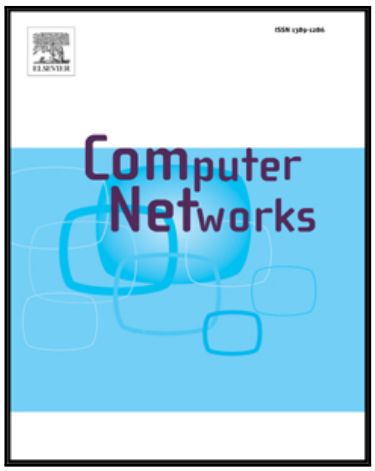

Network and Server Resource Management Strategies for Data Centre Infrastructures: A Survey

Fung Po Tso, Simon Jouet, Dimitrios P. Pezaros

PII: S1389-1286(16)30229-8
DOI: 10.1016/j.comnet.2016.07.002
Reference: COMPNW 5956

To appear in: *Computer Networks*

Received date: 8 March 2016
Revised date: 28 June 2016
Accepted date: 4 July 2016

Please cite this article as: Fung Po Tso, Simon Jouet, Dimitrios P. Pezaros, Network and Server Resource Management Strategies for Data Centre Infrastructures: A Survey, *Computer Networks* (2016), doi: 10.1016/j.comnet.2016.07.002

Network and Server Resource Management Strategies for Data Centre Infrastructures: A Survey

Fung Po Tso[a,*], Simon Jouet[b], Dimitrios P. Pezaros[b]

[a] *Department of Computer Science, Liverpool John Moores University, L3 3AF, UK*
[b] *School of Computing Science, University of Glasgow, G12 8RZ, UK*

Abstract

The advent of virtualisation and the increasing demand for outsourced, elastic compute charged on a pay-as-you-use basis has stimulated the development of large-scale Cloud Data Centres (DCs) housing tens of thousands of computer clusters. Of the significant capital outlay required for building and operating such infrastructures, server and network equipment account for 45% and 15% of the total cost, respectively, making resource utilisation efficiency paramount in order to increase the operators' Return-on-Investment (RoI).

In this paper, we present an extensive survey on the management of server and network resources over virtualised Cloud DC infrastructures, highlighting key concepts and results, and critically discussing their limitations and implications for future research opportunities. We highlight the need for and benefits of adaptive resource provisioning that alleviates reliance on static utilisation prediction models and exploits direct measurement of resource utilisation on servers and network nodes. Coupling such distributed measurement with logically-centralised Software Defined Networking (SDN) principles, we subsequently discuss the challenges and opportunities for converged resource management over converged ICT environments, through unifying control loops to globally orchestrate adaptive and load-sensitive resource provisioning.

Keywords: Cloud Data Centre, Virtualisation Management, Network

*Corresponding author
Email address: p.tso@ljmu.ac.uk (Fung Po Tso)