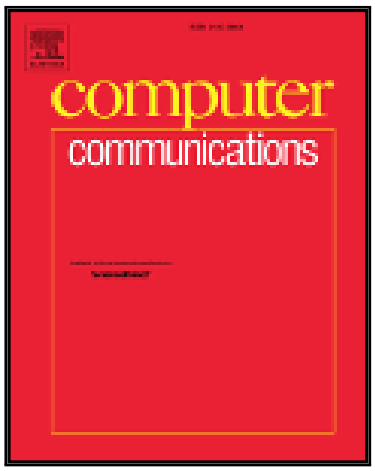

Hierarchical Load Balancing as a Service for federated Cloud networks

Anna Levin, Dean Lorenz, Giovanni Merlino, Alfonso Panarello, Antonio Puliafito, Giuseppe Tricomi

PII: S0140-3664(18)30316-5
DOI: 10.1016/j.comcom.2018.07.031
Reference: COMCOM 5750

To appear in: *Computer Communications*

Received date: 3 April 2018
Revised date: 19 July 2018
Accepted date: 24 July 2018

Please cite this article as: Anna Levin, Dean Lorenz, Giovanni Merlino, Alfonso Panarello, Antonio Puliafito, Giuseppe Tricomi, Hierarchical Load Balancing as a Service for federated Cloud networks, *Computer Communications* (2018), doi: 10.1016/j.comcom.2018.07.031

Hierarchical Load Balancing as a Service for federated Cloud networks

Anna Levin[a], Dean Lorenz[a], Giovanni Merlino[b], Alfonso Panarello[b], Antonio Puliafito[b], Giuseppe Tricomi[b]

[a] *IBM Haifa Research Labs, Haifa, Israel*
[b] *Università degli Studi di Messina, Messina, Italy*

Abstract

The Software Defined Networking (SDN) paradigm, according to the most popular definition, proposes the ambitious vision of making network infrastructure (e.g., routers and switches) fully programmable. This approach introduces suitable levels of abstraction, in order to adaptnetwork infrastructure functions at runtime through powerful and expressive APIs. In this context, the concept of network virtualisation is of particular importance, namely the idea to create virtual partitions of the physical network infrastructure. When virtualization is applied in this domain, it allows several controller instances and their applications to populate and manage the assigned partitions. Even more useful are the aforementioned concepts when dealing with wide-area networks of cloud-hosting datacenters, especially when trying to provide, e.g., cloud-agnostic and transparent QoS, i.e., cloud-bursting. Scientific research is tackling these new trends following two approaches: multi-cloud and federated cloud. In this work we will pursue the latter, because it leaves the end-users (application owners or companies) free to focus their efforts on application-related activities. Load balancing is among the best practices to distribute user workloads fairly and dynamically among all the nodes in a scaling group, either in a data-center or across clouds. This strategy becomes even more relevant in a scenario featuring

Email addresses: lanna@il.ibm.com (Anna Levin), dean@il.ibm.com (Dean Lorenz), gmerlino@unime.it (Giovanni Merlino), apanarello@unime.it (Alfonso Panarello), apuliafito@unime.it (Antonio Puliafito), gtricomi@unime.it (Giuseppe Tricomi)