Chatty Tenants and the Cloud Network Sharing Problem

H. Ballani, E. Jang, T. Karagiannis, C. Kim,
D. Gunawardena, G. O'Shea

Presented by Evguenia (Elmi) Eflov
CS 856 Fall 2015

Outline

Background

- Network bandwidth is a resource shared by all tenants
- Intra-tenant traffic control addressed by existing network sharing policies
- Inter-tenant traffic not managed by existing sharing policies
  - Which communication partner should dictate network allocations?

Network Sharing Requirements

- Associate VMs with minimum bandwidth guarantees
- Ensure high network utilization
- Divide network resources in proportion to tenant payments

Problem Statement

- Providing minimum bandwidth guarantees
- Bounding maximum network impact
- Doing so in presence of inter-tenant traffic

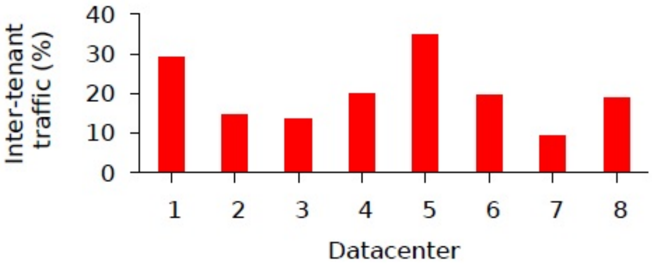
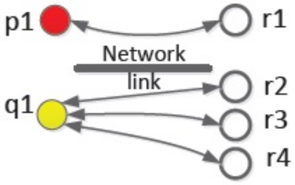


Figure 1: Inter-tenant traffic, as a % of the datacenter's total traffic, for eight production datacenters.

Bandwidth Allocation Approaches

- Most existing approaches lead to unfair allocation when applied to inter-tenant traffic
- PS-L and PS-P assign bandwidth according to weights associated with each tenant



(a)

Allocation	P	Q	R
per-flow	250	750	1000
per-src (→)	500	500	1000
per-src (←)	250	750	1000
PS-L	333	666	1000
PS-P	250	750	1000

(b)

Figure 2: Inter-tenant traffic: Tenants *P*, *Q* and *R* have one ($p1$), one ($q1$) and four VMs respectively.

Outline

Contributions

- Definition of payment proportionality that ensures robust network sharing in presence of inter-tenant traffic
- Bandwidth allocation policy to match defined proportionality
- Relaxed bandwidth guarantee semantics to improve multiplexing for a provider
- VM placement algorithm to satisfy the guarantees using max-flow network formulation

Proportionality Redefined

- Upper Bound Proportionality - maximum bandwidth for each tenant and each VM defined by their payment
- The bound applies to both inter- and intra-tenant traffic

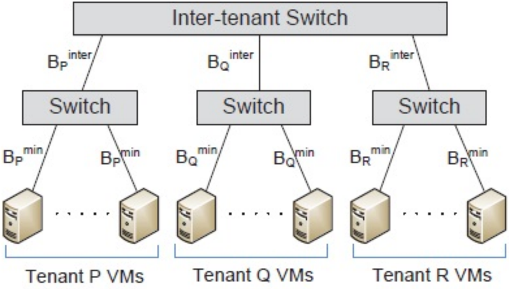


Figure 4: Hierarchical hose model gives per-VM minimum bandwidth for intra-tenant traffic and per-tenant minimum for inter-tenant traffic.

Proposed Allocation Model: Hose

- Flow allocation according to the hose model
- Such allocation facilitates satisfying min-guarantee requirement
- $B_P{}^{min}$ - minimum guarantee for intra-tenant traffic, per VM
- $B_P{}^{inter}$ - for inter-tenant traffic, defined for the tenant

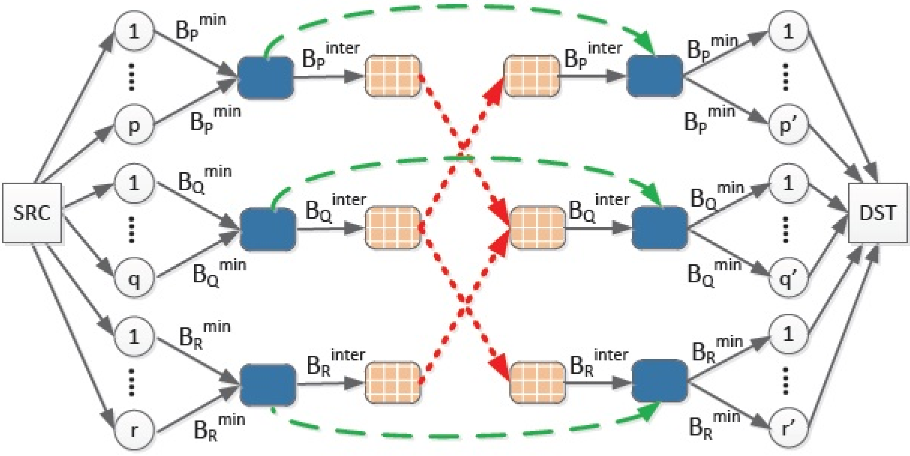
Implementation

- **Hadrian** - network sharing framework
  - **VM placement** - manager that performs admission control and VM allocation
  - **Hose-compliant bandwidth allocation**

VM Placement

- Consistent with bandwidth requirements and tenant communication patterns
- Requires knowledge of tenant communication partners
- Greedy placement that attempts to minimize traffic higher in the switch tree

Flow Network to Compute Max-Flow



Bandwidth Allocation

- Rate for the flow is determined by minimum rate of switches on the path
- First packet header embeds B_P^{min}, N and flow weight
- Hypervisors at both ends compute weight from minimum allocation
- Weight is used by switches to compute rate
- Extended design - switches maintain per-tenant state to adjust rates with arrival of new flows from the same tenants

Outline

Experimental Setup

- Testbed and large-scale simulation
- VM Placement algorithms
  - Greedy
  - Dependency-aware
  - Hadrian - mimumum bandwidth and dependency aware
- Bandwidth allocation policies
  - Per-flow
  - Per-source
  - PS-L sharing
  - Hose-compliant
  - Reservations

VM Requests

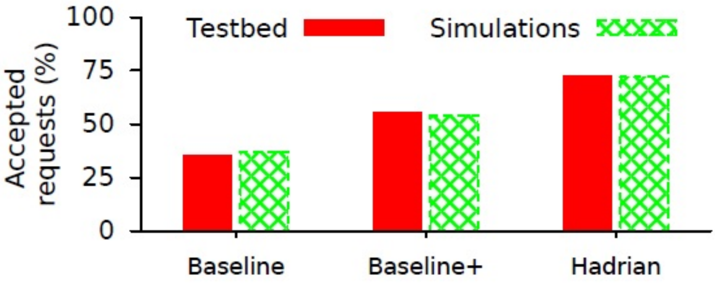


Figure 6: Accepted requests in testbed and simulator.

VM Placement Benefit

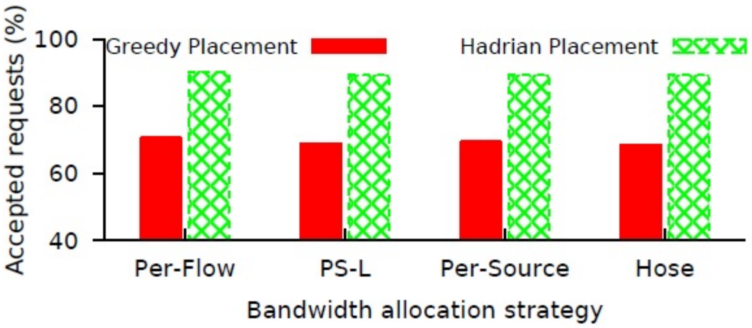
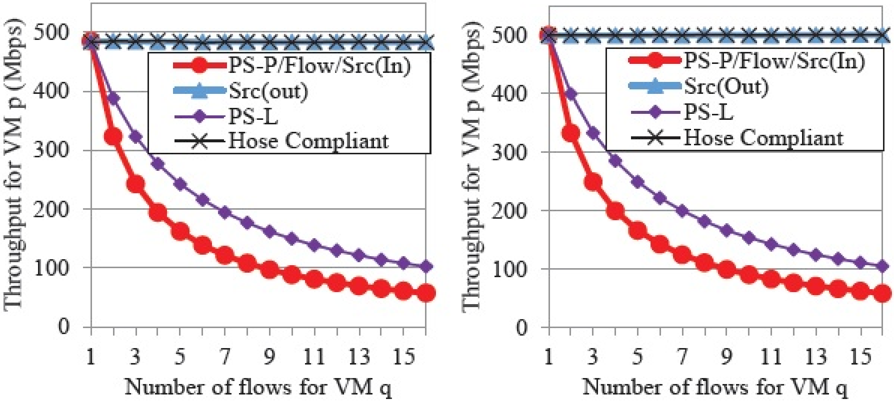


Figure 9: With non-aggressive tenants, Hadrian's placement provides most of the gains.

Throughput for Tenants



(a) Testbed (b) Simulator

Summary

- Robust yet proportional network sharing
- VM placement according to communication requirements
- Bandwidth allocation for both inter- and intra-tenant requests