Synergistic Policy and Virtual Machine Consolidation in Cloud Data Centres

Lin Cui[1] Richard Cziva[2] Posco Tso[3] Dimitrios P. Pezaros[2]

[1]Jinan University
[2]University of Glasgow
[3]Liverpool John Moores University

INFOCOM, 2016

Outline

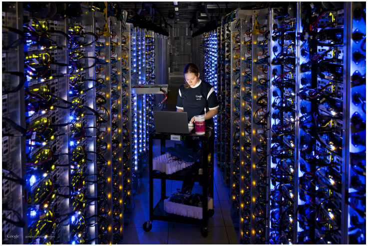
1. why this work is interesting/important
2. our work on synergistic VM and network policy management
3. our experimental results

Outline

1. why this work is interesting/important
2. our work on synergistic VM and network policy management
3. our experimental results

a closer look at inside a data centre

- Server - Loads of it. Even larger number of virtual machines.
- Network cables - Loads of it. Needed for connecting servers, network switches/routers, and middleboxes.



data centre outage is expensive[1]

- $69 Trillion would be lost per hour if every data centre in world went down at the same time.
- When Amazon.com went dark for approximately 49 minutes in January of 2013, it cost the company an estimated $4 million or more in lost sales. Another outage in August of the same year lasted only 30 minutes, but still cost the Internet giant an estimated $66,240 in lost revenue every single minute.
- Google's 5 minutes outage in 2014 was reported to have cost over $500,000 and led to 40% drop in worldwide Internet traffic.

[1] http://www.virtualhosting.com/blog/2013/outrageous-costs-data-center-downtime

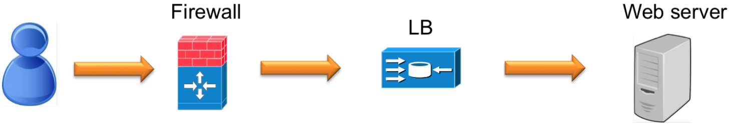
In the meantime, it was reported that the largest source of network failures in data centres stemmed from misconfiguration[2].

[2] Turner, D. et al., "NetPilot: Automating Datacenter Network Failure Mitigation." ACM SIGCOMM 12

network policy

All networks are governed by network policies, which are high-level networking objectives derived from network-wide requirements[3]:

- Internet client traffic must be checked by Firewall first and then forwarded to a LB.

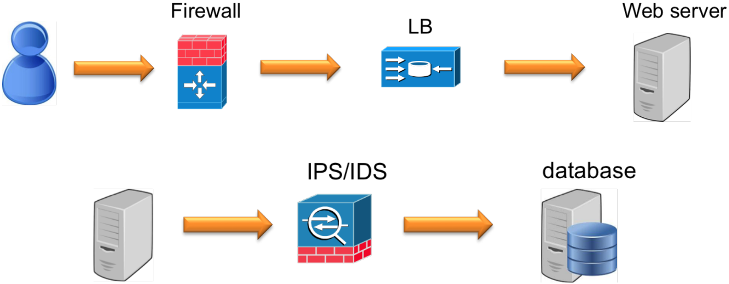


[3] "Service Function Chaining Use Cases In Data Centers", draft-ietf-sfc-dc-use-cases-04 , 2016

network policy

All networks are governed by network policies, which are high-level networking objectives derived from network-wide requirements[3]:

- Internet client traffic must be checked by Firewall first and then forwarded to a LB.
- Traffic to access database server must be checked by an IPS.



[3] "Service Function Chaining Use Cases In Data Centers", draft-ietf-sfc-dc-use-cases-04 , 2016

network policy

Network policies are implemented in middleboxes or network function boxes – such as firewalls, load balancers, SSL offloaders, web caches, and intrusion prevention boxes

how are network policies managed

In the meantime, it was reported that the largest source of network failures in data centres stemmed from misconfiguration[4].

A largely manual effort - statically insert middleboxes on to the network paths, and manually configure them.

Policy demands packets to traverse a specific set of middleboxes.

[4] Turner, D. et al., "NetPilot: Automating Datacenter Network Failure Mitigation." ACM SIGCOMM 12

how are servers managed?

cloud data centre server management

A large shared compute environment where tenants (users) come and go unpredictably - hence, unpredictable dynamic workload.

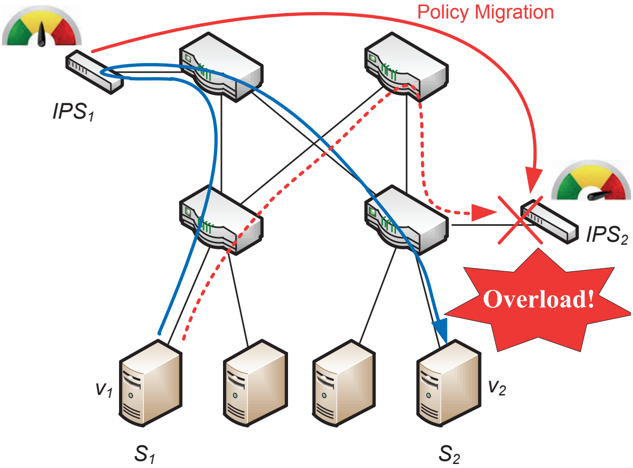
- Virtualisation (e.g. virtual machines) enables sharing and provides elasticity (i.e. dynamic scaling).
- Virtual machines can be dynamically consolidated – via live migration – to improve resource utilisation, hence the revenue ($).

joint network policy and VM management

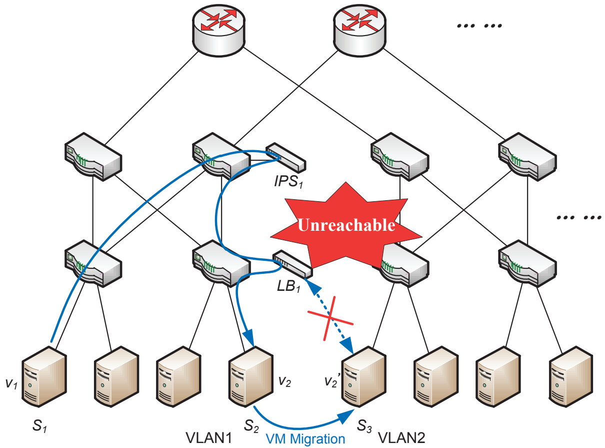
the problem

1. When a virtual machine moves, both flow state and the mapping state of IPs must be updated across the network.
   - unable to traverse a specified set of middleboxes – hence policy violation(s) – hence data centre outage
2. When a policy is deployed/migrated, locations of src/dst VMs must be considered.
   - forwarded through a longer path – hence wasting network resources – impairing performance

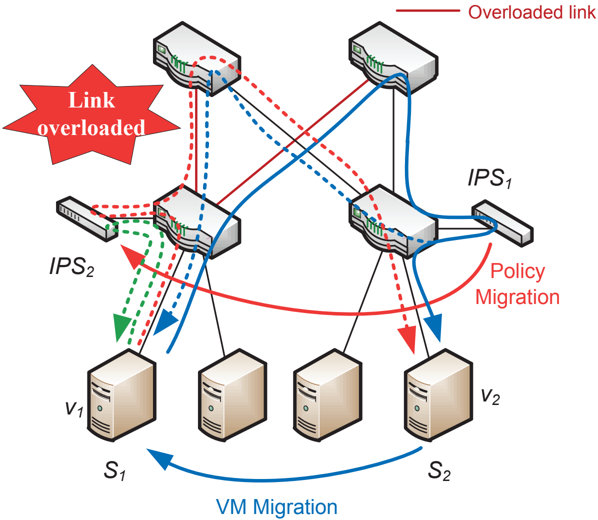
example scenario 1: MB capacity overloaded



example scenario 2: Route unreachable



example scenario 3: Performance degradation

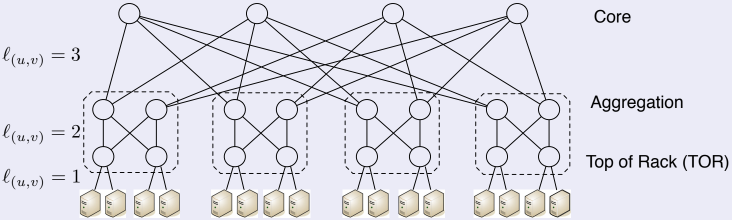


Outline

1. why this work is interesting/important
2. our work on synergistic VM and network policy management
3. our experimental results

network cost

fat-tree



network communication cost

The cost of each link in DC networks varies on the particular layer that they interconnect. High-speed core router interfaces are much more expensive (and, hence, oversubscribed) than lower-level (Top of Rack) ToR switches.

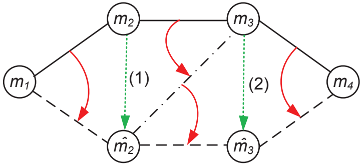
Hence, we define the *Communication Cost* of all traffic from VM v_i to v_j as

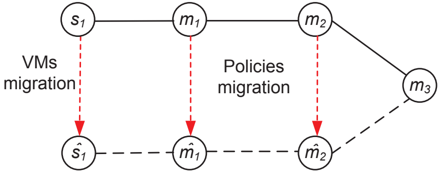
$$
\begin{aligned}
C(v_i, v_j) &= \sum_{p_k \in P(v_i,v_j)} f_k.rate \sum_{L_s \in R_k(v_i,v_j)} c_s \\
&= \sum_{p_k \in P(v_i,v_j)} \{C_k(v_i, p_k.in) \\
&\quad + \sum_{j=1}^{p_k.len-1} C_k(p_k.list[j], p_k.list[j+1]) \\
&\quad + C_k(p_k.out, v_j)\}
\end{aligned}
\tag{1}
$$

Given the set of VMs $\mathbb{V}$, servers $\mathbb{S}$, policies $\mathbb{P}$ and MBs $\mathbb{M}$, we need to find an allocation A that minimizes the total communication cost:

$$\begin{aligned} \min \quad & \sum_{v_i \in \mathbb{V}} \sum_{v_j \in \mathbb{V}} C(v_i, v_j) \\ s.t. \quad & A(v_i) \neq \emptyset \ \&\& \ |A(v_i)| = 1, \forall v_i \in \mathbb{V} \\ & p_k \text{ is satisfied}, \forall \ p_k \in \mathbb{P} \\ & \sum\nolimits_{v_i \in A(s_j)} r_i \preceq h_j, \forall s_j \in \mathbb{S} \\ & \sum\nolimits_{p_k \in A(m_i)} f_k.rate \leq m_i.capacity, \forall m_i \in \mathbb{M} \end{aligned} \tag{2}$$

policy migration among middleboxes

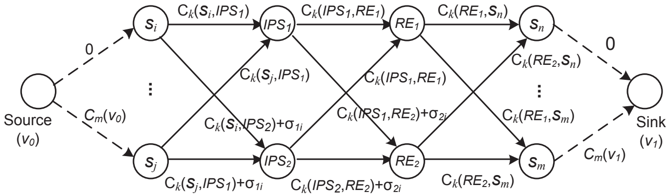




Their network cost reduction, i.e. utility, is equal. MBs and VMs can be migrated separately.

two-phase based algorithm

- Phase I: policy migration
  - Build Cost Network
  - Find the shortest path
  - Vote for candidate servers



two-phase based algorithm

- Phase II: VM migration
  - build preference lists for both VMs and servers
  - find the most "stably matched" servers for VMs
  - use stable matching to solve preference conflicts between them

Sync can be converged and output a stable matching results within O(VS), where V and S are the numbers of input VM and server respectively.

Outline

1. why this work is interesting/important
2. our work on synergistic VM and network policy management
3. our experimental results

simulation

We have evaluated the performance of Sync scheme over a simulated fat-tree DC topology with k = 14 (i.e., 931 nodes, including 686 servers and 245 switches) in ns-3.

- each policy flow is configured to traverse 1- 3 middleboxes.
- comparing against S-CORE[a].

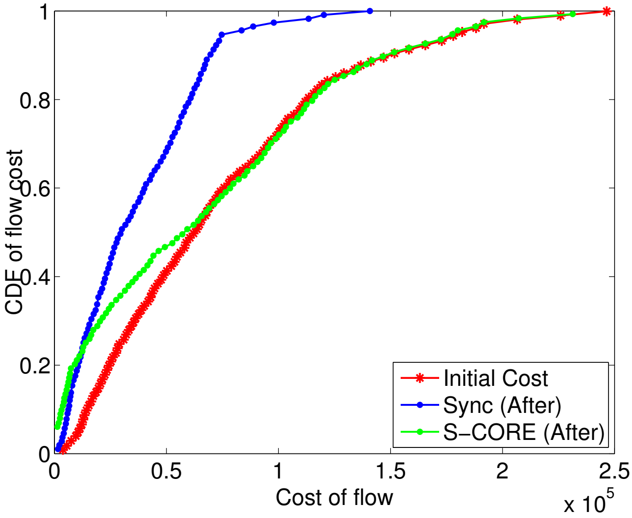
[a] Tso, F. P., Oikonomou, K., Kavvadia, E., & Pezaros, D. P. Scalable traffic-aware virtual machine management for cloud data centers. IEEE ICDCS, 2014

testbed

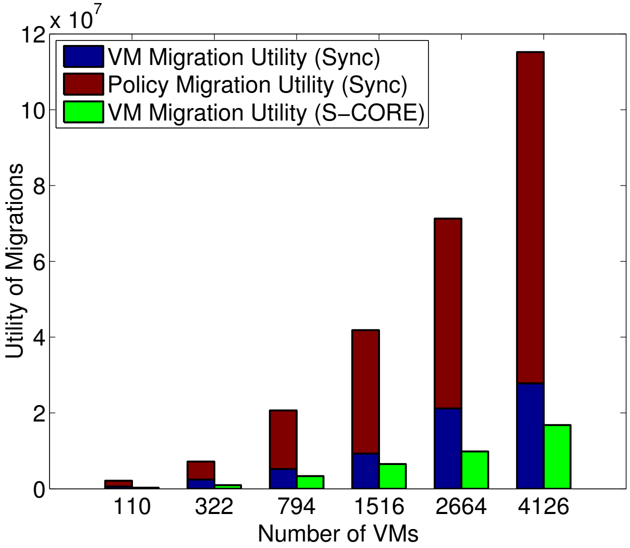
A centralised controller is implemented on top of Ryu SDN controller to collect all flow statistics and running time.

- A CentOS 6 host with Intel 2.1GHz CPU and 4GB RAM.
- Flow statistics are collected from all software switches (Open vSwitch 2.3.1).

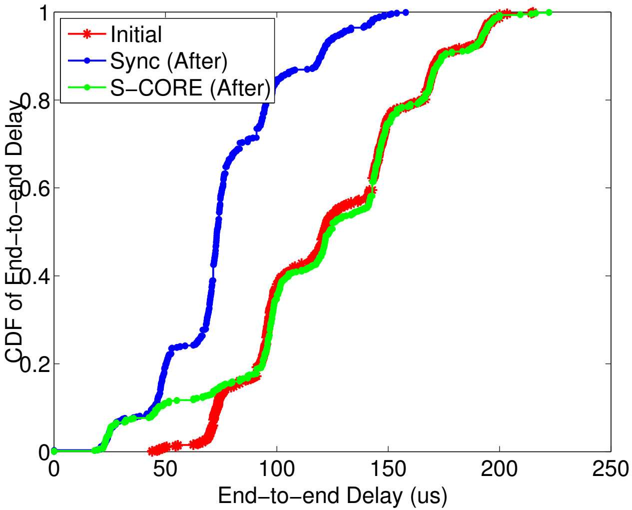
simulation – cost of flows for $k = 14$



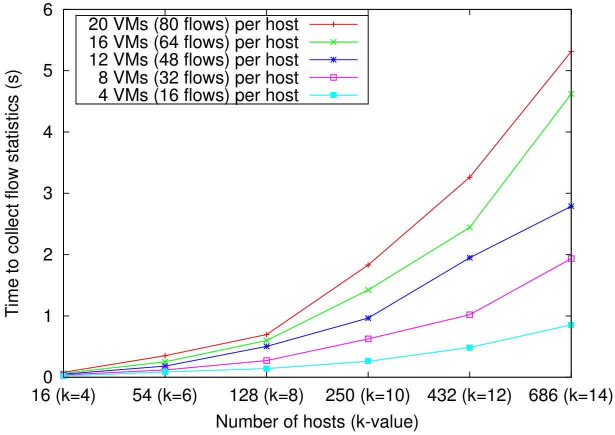
simulation – migration utility



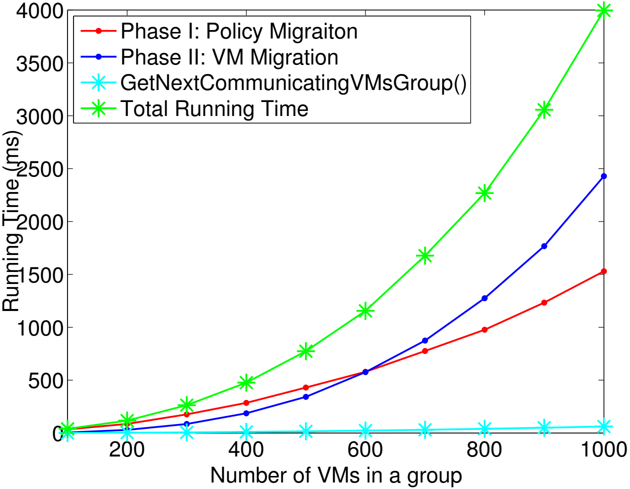
simulation – end to end relay



testbed – time spent to collect flow statistics



testbed – time spent for Sync migration algorithm



Take-away

- Middleboxes (network policies) and virtual machines management have been treated as independent problems.
- Data centre network policy management is challenging due to multidimensional dynamism.
- Data centre network needs synergistic management - exemplified by our joint network policy and virtual machine consolidation.