Towards a Network Aware VM Migration: Evaluating the Cost of VM Migration In SDN-Based Cloud Computing Network

Sachin Shetty
Associate Professor
Department of Electrical and Computer Engineering
Tennessee State University
Nashville, TN

June 16, 2016
Motivation

- Migrating VMs is a good defense mechanism but comes at a network performance cost
  - Consume nearly the entire bandwidth
  - Affects performance of competing applications
- Cloud applications deadlines must be met [Kohavi et. al. SIGKDD, 2007]
  - 100 ms additional latency on Amazon EC2 = 1% drop in sales
  - 500 ms increase in Google search time = 20% drop in revenues
- VM Migration process is unreliable and only 87% of migrations successfully finish [Tomas et. al., TSP, 2015]
The Cloud Network – Resource Sharing

**Amazon Data Center in California**

**Google Data Center in California**

**Amazon Data Center in Virginia**

**Google Data Center in California**

**Internet User in Virginia**

- VM Migration Of VMs 1, 2, 3
- Real Time Online data back up
- HD Content Streaming
- Web browsing

**Cloud Network of networks**

**Core Network**

**Regional Network**

**Internet User in Virginia**

- Web browsing

**Workshop on Science of Security through Software-Defined Networking, Illinois Institute of Technology**

**June 16-17, 2016**
Research Challenge

- Need for a framework to quantify the cost of VM Migration
- Need for a mechanism that intelligently allocates network resources during VM Migration
- Ensure a fair share of resources during VM Migrations
- Minimize degradation of network performance
The Problem

- Live VM Migrations are expensive
  - Nearly consumes entire bandwidth.
  - Impacts performance of competing flows

- Vendors
  - Recommend separate network for VM mobility – costly

- Know the cost of VM Migration
- Fair share of resources during VM Migration
  - Migration policies, Network topology, Data access, QoS policies
  - Cost estimation model \( N = M \frac{1 - (R/L)^{n+1}}{1 - (R/L)} \) – simulated
Related Efforts

VM Migration

Migration Mechanisms
- Migration Downtime
- Migration Total Time

Migration Policies

Network Topology
- Data Access
- Hotspots
- E2E QoS Policies
Related Work - Remedy

- Builds on pre-copy migration technique [V. Mann, Networking 2012]
  - Used by VMware vMotion and KVM Live Migration
  - Given the Memory size $M$ of a VM in MB,
  - Page dirty rate $R$ of a VM in MB/s,
  - Bandwidth of the link used for migration $L$ in MB/s,
  - Switchover goal time $T$ and
  - Minimum required progress amount $X$ in MBs,
Related Work - Remedy

- Remedy cost Model is described as:
  - The number of pre-copy cycles:  \( n = \min \left( \left\lfloor \log_{R/L} \frac{T.L}{M} \right\rfloor, \left\lfloor \log_{R/L} \frac{X.R}{M(L-R)} \right\rfloor \right) \)
  
- Total traffic generated by the migration:  \( N = M \frac{1-(R/L)^{n+1}}{1-(R/L)} \)

- Total migration time:  \( Tm = N / L \)

- Time spent on a stop-copy transfer  \( W(L) = \frac{M}{L} \left( \frac{R}{L} \right)^n \)
  - Downtime experienced
Goals and Objectives

- **Goal:** Compute cost of VM Migration and ensure successful VM Migration without degrading performance of network applications in a SDN based cloud computing network

- **Objectives:**
  - Empirically evaluate Remedy cost estimation model on GENI testbed
  - Openflow Controller to enforce QoS and fair share of resources during VM migration
  - Ensure VM Migration completes successfully
  - Improve the performance of competing flows
Methodology

Controller

Data Center

Core
Methodology

- Design and implement an OpenFlow controller
  - Monitors network utilization
  - Calculates the cost of VM Migration on all network paths
    - Cost of VM migration, Network topology and Network traffic
    - Optimal bandwidth for VM Migration flows
    - Enforce end-to-end QoS policies to reserve optimal Bandwidth

- Emulate data center networks by utilizing computing and networking resources on GENI infrastructures
Experiment Setup

- Reserving resources: Data center, Core, OF controller
  - 8 kvm hosts and 7 ovs switches
  - NFS server for shared storage
Experiment Setup

- **Virtualization Manager**
  - Libvirt toolkit
  - Ubuntu 12.04 Desktop VM, 2048 MB Memory on H1
  - Openvswitch bridged networking
  - Manage, Clone, Live-Migrate VMs across the data center

- **VM Workload and Network Flows**
  - RUBiS Workload
    - Auction site prototype modeled after eBay.com
    - 50 concurrent clients: average page dirty rate of 500 pages/second
  - Ping flows
  - iperf flows
Global Environment for Network Innovations (GENI)

- GENI is a nationwide suite of infrastructure for “at scale” experiments in networking, distributed systems, security, and novel applications.
- GENI provides compute resources that can be connected in experimenter specified Layer 2 topologies.

Evaluation in GENI Testbed

Virtualization Manager
- Manage, Clone and Live-migrate VMs

Migration Traffic
- Ping Traffic
- Iperf Traffic
- RUBIS Traffic

DATA CENTER NETWORK

OpenFlow Controller
- Resource Monitor
- Resource Allocator

NFS Server
- Shared Storage

Evaluation Diagram:
- Sw1
- Sw2
- Sw3
- Sw4
- Sw5
- Sw6
- Sw7
- H1
- H2
- H3
- H4
- H5
- H6
- H7
- H8

- I/O task on guest VM
- Ping Server
- Iperf Server
- RUBIS Server
- RUBIS App Server
- Page Modification Benchmark

- Ping Traffic
- Iperf Traffic
- RUBIS Traffic

Workshop on Science of Security through Software-Defined Networking, Illinois Institute of Technology
June 16-17, 2016
Experiment: Resource Manager

- **Bandwidth Monitor**
  - Discover all links in the network
  - Poll for port and flow statistics from all switches
  - Discover heaviest flows and links
  - Discover available bandwidth in the network

- **Memory Monitor**
  - Virtualization manager: VM memory size, page dirty rate
Experiment: QoS Enforcer

- OVS capabilities: create 2 queues on each port of OpenFlow switch
  - Egress traffic shaping

- Cost of Migration Estimation – Migrate H1 to H8
  - User specified VM migration deadline (25 seconds)
  - VM’s page dirty rate (500MB/s)
  - Memory size (2048 MB),
  - User specified minimum progress amount (100MB)
  - Switchover goal time (100ms),
  - Minimum bandwidth used by the migration (80MB/s).
Experiment: QoS Enforcer

- Bandwidth of dedicated queue = estimated minimal value.
- OF rule to Enqueue migration traffic to dedicated queue.
- Predicted Migration Times
## Results

- **Actual Migration Times and Average Ping Latencies**

<table>
<thead>
<tr>
<th>Reserved Bandwidth (MB/s)</th>
<th>Migration Elapsed Time (s)</th>
<th>Average Ping Latency (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No reservation</td>
<td>19.465</td>
<td>19.059</td>
</tr>
<tr>
<td>80</td>
<td>19.576</td>
<td>10.066</td>
</tr>
<tr>
<td>90</td>
<td>19.436</td>
<td>5.865</td>
</tr>
<tr>
<td>100</td>
<td>19.116</td>
<td>3.31</td>
</tr>
<tr>
<td>120</td>
<td>18.896</td>
<td>2.4844</td>
</tr>
</tbody>
</table>

- Increasing bandwidth reservation has minimal effect on migration time, but heavily improves ping flow performance (47.18%)
Results

- Impact of QoS enforcer on iperf flow
  - H3 IP Traffic

- Incoming traffic throttled in the absence of QoS policy
- Enforcing QoS policies in terms of bandwidth reservation relieves the network of possible overloads during migration
Results – Evaluating Remedy

- Which parameters are more critical?
  - Page dirty rate $R$,
  - Available bandwidth $L$,
  - User specified parameters progress amount $X$,

- Conditions under which the model works/fails
  - $L \geq 1\text{Gbps}$
  - $R \leq 3000\ \text{pages/seconds}$
  - Higher $R$, Lower $L$ : Model fails to perform
    - Irrespective of VM memory size
### Varying Page dirty rate (R), (1 GB VM)

#### Migration deadline = 330 s, L=0.1Gbps

<table>
<thead>
<tr>
<th>R (Pages/s)</th>
<th>Remedy Bandwidth (Mbps)</th>
<th>Our Bandwidth (Mbps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>30</td>
<td>33</td>
</tr>
<tr>
<td>250</td>
<td>38</td>
<td>41</td>
</tr>
<tr>
<td>500</td>
<td>45</td>
<td>51</td>
</tr>
<tr>
<td>1000</td>
<td>60</td>
<td>70</td>
</tr>
<tr>
<td>2000</td>
<td>90</td>
<td>103</td>
</tr>
</tbody>
</table>

#### Migration deadline = 80, L=1Gbps

<table>
<thead>
<tr>
<th>R (Pages/s)</th>
<th>Predicted Bandwidth (Mbps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>112</td>
</tr>
<tr>
<td>250</td>
<td>120</td>
</tr>
<tr>
<td>500</td>
<td>128</td>
</tr>
<tr>
<td>1000</td>
<td>144</td>
</tr>
<tr>
<td>2000</td>
<td>160</td>
</tr>
</tbody>
</table>

#### Migration deadline = 8 s, L=10Gbps

<table>
<thead>
<tr>
<th>R (Pages/s)</th>
<th>Predicted Bandwidth (Mbps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>1144 same</td>
</tr>
<tr>
<td>250</td>
<td>1144 same</td>
</tr>
<tr>
<td>500</td>
<td>1144 same</td>
</tr>
<tr>
<td>1000</td>
<td>1144 same</td>
</tr>
<tr>
<td>2000</td>
<td>1144 same</td>
</tr>
</tbody>
</table>

#### Migration deadline = 80, L=15Gbps, M=2GB

![Graph showing migration times](image)

---

Workshop on Science of Security through Software-Defined Networking, Illinois Institute of Technology
June 16-17, 2016
Available bandwidth, $L$

- Dictates range of user migration deadline

- For 20 MB/s link, and $R=2000$ pages/s
  - You can only reserve 11-18 MB/s
  - Allowable deadline 200-800 s

- For 2000 MB/s link, and all values of $R$
  - You can reserve 0-2000 MB/s
  - Allowable deadline 2-23 s
User specified progress amount, X

- Corrects discrepancies in the model caused by:
  - Insufficient bandwidth, L
  - Higher page dirty rates, R

X=50MB, M=64MB, T=0.1s  
X=250MB, M=64MB, T=0.1s
### User specified progress amount, $X$

- **Range of $X$ under which the model works/fails ($R=10,000$ pages/s)**

<table>
<thead>
<tr>
<th>Memory size (MB/s)</th>
<th>Range of $X$ To select MB/s</th>
<th>Range of $X$ NOT To select MB/s</th>
<th>Range of $X$ To select MB/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>64</td>
<td>0-30</td>
<td>30-500</td>
<td>&gt; 500</td>
</tr>
<tr>
<td>512</td>
<td>0-100</td>
<td>100-1250</td>
<td>&gt; 1250</td>
</tr>
<tr>
<td>1024</td>
<td>0-250</td>
<td>250-2000</td>
<td>&gt; 2000</td>
</tr>
<tr>
<td>2048</td>
<td>0-500</td>
<td>500-5000</td>
<td>&gt; 5000</td>
</tr>
</tbody>
</table>
Ongoing and Future work

- **Significance of cost estimation model parameters**
  - Modify parameters to suit types of applications, optimize model

- **Performance of model on a wide area network**
  - Variable link bandwidths – QoS, OF
  - Migration across Multiple Data Centers
    - GENI - USA
    - iMinds - Europe

- **Deploy the model to compute cost of VM Migration based Moving Target Defense technique in the cloud computing network**
REFERENCES


Thank you!