Elastic Virtual Network Function Placement

CloudNet 2015

M. GHAZNAVI, A. KHAN, N. SHAHRIAR, KH. ALSUBHI, R. AHMED, R. BOUTABA
DAVID R. CHERITON SCHOOL OF COMPUTER SCIENCE
UNIVERSITY OF WATERLOO
Outline

Introduction
State of the Art
Problem: Elastic Virtual Network Function Placement
Solution: Simple Lazy Facility Location
Evaluation
Conclusion
Introduction

MIDDLE-BOXES
NETWORK FUNCTION VIRTUALIZATION
VNF SERVICES IN CLOUD
Middle-Boxes

“any intermediary device performing functions other than the normal, standard functions of an IP router on the datagram path between a source host and destination host” [1]

Expensive hardware
Hard to deploy
Hard to modify
Hard to scale
Provision for peak-load

Network Function Virtualization

Virtualization (Softwarization) of middle-boxes

Software middle-boxes are called Virtual Network Function (VNF)

NFV "involves the implementation of network functions in software that can run on a range of industry standard server hardware, and that can be moved to, or instantiated in, various locations in the network as required, without the need for installation of new equipment."[1]

## Network Function Virtualization

<table>
<thead>
<tr>
<th>MIDDLE-BOXES</th>
<th>VIRTUAL NETWORK FUNCTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Expensive</strong> hardware</td>
<td><strong>Low-cost</strong> software</td>
</tr>
<tr>
<td><strong>Hard</strong> to deploy</td>
<td><strong>Easy</strong> to deploy</td>
</tr>
<tr>
<td><strong>Hard</strong> to modify</td>
<td><strong>Easy</strong> to modify</td>
</tr>
<tr>
<td><strong>Hard</strong> to scale</td>
<td><strong>Easy</strong> to scale</td>
</tr>
<tr>
<td>Provision for peak-load</td>
<td>Scale resources on demand</td>
</tr>
</tbody>
</table>
VNF Services in Cloud

Offered by cloud providers
- IBM Bluemix
- Microsoft Azure
- Amazon EC2

Services
- Riverbed STEELHEAD WAN optimizer [1]
- McAfee Next Generation firewall [2]
- Virtual LoadMaster load balancer [3]

VNF Services in Cloud

WHAT CLOUD PROVIDER SHOULD SUPPORT

Pay per use
- Clients pay only for real used resources

Elasticity
- Scale resources on demand
  - Upon arrival or departure of service request
  - Variation of workload of admitted service request

CHALLENGES OF CLOUD PROVIDER

Minimizing Costs:
- Trade-off between Host & Bandwidth Resources

Elasticity
- Which mechanisms to apply
- Elasticity benefit vs. its overhead
VNF Services in Cloud

Where to place VNF instances?

Which request must be assigned to which VNF instance?
VNF Services in Cloud

A solution can be
- $v_1$ serves the first and second service traffics
- $v_2$ serves the third and forth service traffics
State of the Art

COMPARISON OF STATE OF THE ART
## Comparison of State of the Art

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Elastic Virtual Network Function Placement (EVNFP)</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Elasticity in Cloud [1, 2, 3]</td>
<td>✔</td>
<td>✕</td>
<td>✔</td>
</tr>
<tr>
<td>Dynamic VM Placement [2, 4]</td>
<td>✔</td>
<td>✕</td>
<td>✔</td>
</tr>
<tr>
<td>Network Aware VM Placement [5, 6, 7]</td>
<td>✔</td>
<td>✔</td>
<td>✕</td>
</tr>
<tr>
<td>Virtual DPI Placement [8]</td>
<td>✔</td>
<td>✔</td>
<td>✕</td>
</tr>
</tbody>
</table>

Problem: Elastic Virtual Network Function Placement (EVNFP)

SCOPE AND ASSUMPTIONS
CONFLICTING OBJECTIVES
ELASTICITY MECHANISMS AND OVERHEAD
Scope and Assumptions

SCOPE

Single cloud provider
Single data-center
Centralized optimization

ASSUMPTIONS

One VNF instance-type
Multi-tenancy
Elasticity Mechanisms
  ◦ Horizontal Scaling
  ◦ Migration of VNF instances
  ◦ Reassignment of workload
Conflicting Objectives

Minimizing the bandwidth cost, and

Minimizing the number of installed VNFs
Conflicting Objectives

Minimizing the bandwidth cost:

- **12** Unit of Bandwidth over **12** Links
- **4** VNF instances

\[
\begin{align*}
&\text{Minimizing the bandwidth cost:} \\
&\quad \text{◦ 12 Unit of Bandwidth over 12 Links} \\
&\quad \text{◦ 4 VNF instances}
\end{align*}
\]
Conflicting Objectives

Minimizing the number of installed VNFs
- 1 VNF instance
- 34 Unit of Bandwidth over 20 Links
Elasticity Mechanisms and Overhead

MECHANISMS

Horizontal Scaling of VNF instance
  ◦ Installing a new VNF instance
  ◦ Removing an existing VNF instance

Migration of a VNF instance

Reassignment of workload to another VNF instance

OVERHEAD

Migration overhead

Reassignment overhead
Elasticity Mechanisms and Overhead

1. Initial Placement
   - src₁
   - src₂
   - src₃
   - v
   - trg₁
   - trg₂
   - trg₃

2. v* Installation and Reassignment
   - src₁
   - src₂
   - src₃
   - v*
   - trg₁
   - trg₂
   - trg₃

3. Migration of v
   - src₁
   - v
   - trg₁
   - src₃
   - trg₃

4. Removing v
   - src₁
   - v*
   - trg₁

- v VNF instance
- srcᵢ Source of service traffic i
- trgᵢ Target of service traffic i
- Service traffic increase
- Service traffic decrease
Solution: **Simple Lazy Facility Location** *(SLFL)*

**IDEA**

SLFL: SIMPLE LAZY FACILITY LOCATION
Idea

Arrival and departure of a request, or workload variation alter the locality

SLFL locally optimizes the placement of VNF instances in a greedy manner
SLFL: Simple Lazy Facility Location

UPON REQUEST ARRIVAL OR WORKLOAD INCREASE

Installation potential
- Installing a VNF instance
- Set of reassignments
  - The difference of operational cost before and after installing the VNF instance and reassignments

Migration potential
- Migration of a VNF instance
  - The difference of operational cost before and after migration of the VNF instance

UPON REQUEST DEPARTURE OR WORKLOAD DECREASE

Removing potential
- Removing a VNF instance
- Set of reassignments
  - The difference of operational cost before and after removing the VNF instance and reassignments

Emigration potential
- Migration of a VNF instance
  - The difference of operational cost before and after migration of the VNF instance
SLFL: Simple Lazy Facility Location

**UPON REQUEST ARRIVAL OR WORKLOAD INCREASE**

Apply the best action among:
- Installing a VNF instance
  - Considering the installation potential
- **Migrating a VNF instance**
  - Considering the migration potential of the VNF instance
- Assign to one of existing VNFs
  - Considering bandwidth cost

**UPON REQUEST DEPARTURE OR WORKLOAD DECREASE**

Apply the best action among:
- Removing a VNF instance
  - Considering the installation potential
- **Migrating a VNF instance**
  - Considering the emigration potential of the VNF instance
Evaluation

EXPERIMENTAL SETUP AND OBJECTIVES
ACCEPTANCE RATIO AND OPERATIONAL COST
RESOURCE UTILIZATION
Experimental Setup and Objectives

EXPERIMENTAL SETUP

Network
- Fat-tree of 99 nodes
- 54 hosts with 8 Core CPU
- 1 GB full bisection bandwidth

VNF
- Bro IDS [2]: 80 Mbps, 1 vCPU, 1GB of memory

Requests
- 20,000 requests
- Arrival: Poisson distribution
- Duration: Exponential distribution

OBJECTIVES

Evaluating
- The acceptance ratio
- Operational cost
  - Balancing bandwidth and host resource costs
- Resource Utilization
  - Balancing bandwidth and host resource utilization?

Comparison to
- Random Placement
- First-Fit Placement
Acceptance Ratio and Operational Cost

**ACCEPTANCE RATIO**

<table>
<thead>
<tr>
<th>SLFL</th>
<th>Random</th>
<th>Total Operational Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLFL</td>
<td>Random</td>
<td>Total Operational Cost</td>
</tr>
<tr>
<td>SLFL</td>
<td>Random</td>
<td>Total Operational Cost</td>
</tr>
</tbody>
</table>

SLFL accepts $\sim 2 \times$ workload vs basic algorithms
- SLFL: 97% acceptance ratio
- Random: 48% acceptance ratio
- FirstFit: 45% acceptance ratio

SLFL accepts $\sim 2 \times$ workload with less cost
- 9% operational cost less than Random
- 4% operational cost less than FirstFit
Resource Utilization

**BANDWIDTH RESOURCE UTILIZATION**

- Random
- SLFL
- FirstFit

**HOST RESOURCE UTILIZATION**

- Random
- SLFL
- FirstFit

82% Utilization of bandwidth resources
91% Utilization of host resources
Conclusion

SUMMARY
Summary

Elastic Virtual Network Function Problem
- Bandwidth and host resources cost trade-off
- Elasticity Overhead

Simple Lazy Facility Location
- Balancing the bandwidth and host resource cost trade-off
- Carefully selecting the correct elasticity mechanisms
- Optimizing the elasticity overhead
- Accepting $\sim2 \times$ workload vs basic algorithms
Acceptance Ratio and Resource Utilization

![Acceptance Ratio Graph]

![Host Resource Util. Graph]

![Bandwidth Resource Util. Graph]

![VNFS Util. Graph]
Operational Cost

- Total Operational Cost
- Bandwidth Resource Cost
- Host Resource Cost
- Elasticity Overhead Cost
Assumptions-Horizontal Scaling

Why horizontal scaling and ignoring vertical scaling

- On the fly vertical resource scaling is not supported in most cases
- Might require system reboot
  - SLA violation
## Assumptions-One VNF instance-type

<table>
<thead>
<tr>
<th></th>
<th>Scenario</th>
<th>One small flavor</th>
<th>Multiple flavors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Resource Consumption</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Host Res.</td>
<td>~</td>
<td>- Worse</td>
<td>+ Better</td>
</tr>
<tr>
<td>Bandwidth Res.</td>
<td>~</td>
<td>+ Better</td>
<td>- Worse</td>
</tr>
<tr>
<td><strong>Elasticity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Installation</td>
<td>In a same machine</td>
<td>+ Better</td>
<td>- Worse</td>
</tr>
<tr>
<td>Removal</td>
<td>In a same machine</td>
<td>+ Better</td>
<td>- Worse</td>
</tr>
<tr>
<td>Migration overhead</td>
<td>~</td>
<td>+ Better</td>
<td>- Worse</td>
</tr>
<tr>
<td>Reassign. overhead</td>
<td>~</td>
<td>= Equal</td>
<td>= Equal</td>
</tr>
</tbody>
</table>