A Novel Approach for Service Function Chain (SFC)Mapping with Multiple SFC instances in a Fog-To-Cloud Computing System

A.Zamani

Supervised by: Dr. Sharifian

Amirkabir University of Technology

ICSPIS Conference, December 2018

Outline

Introduction

- IoT
  - interconnects billions or even trillions of diverse devices
  - generate a massive amount of data
  - should be transmitted to the cloud for computing

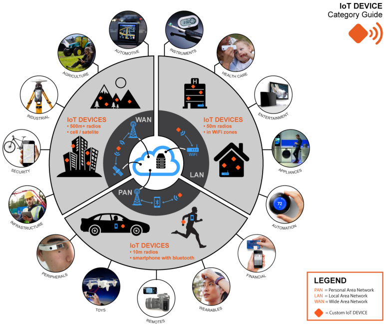


Figure: IoT devices

Introduction

- Cloud computing
  - cloud offers various benefits such as scalability and elasticity
  - consolidation and centralization lead to many network hops
  - results in high latencies and high bandwidth consumption

Figure: Cloud computing

Introduction

- Healthcare

Figure: Healthcare

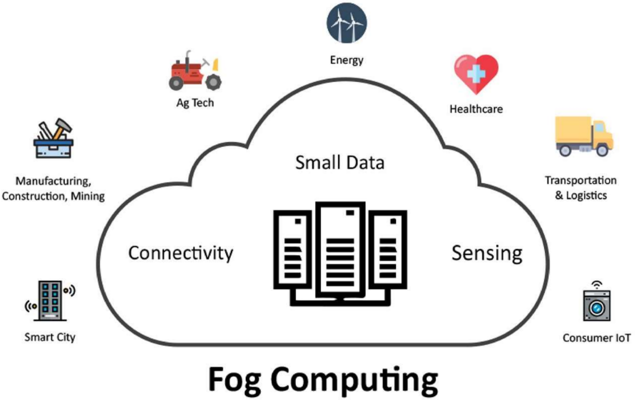
Introduction

- Augmented reality

Figure: Augmented reality

Introduction

- Fog computing
  - offers distributed edge cloud close to the Things



- Fog-to-Cloud computing system
  - fog and cloud work together
  - provide computing, storage, and application servicesin the IoT domain
  - complex management of such a network of distributed fogs

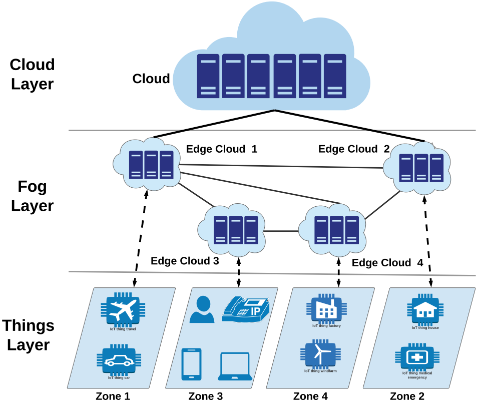


Figure: Fog-to-Cloud computing system

- Software Defined Network(SDN)
  - SDN separates the control and data planes

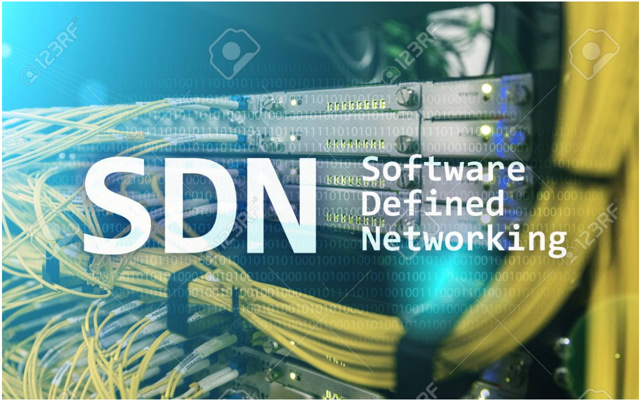


Figure: Software Defined Network

- network function virtualization(NFV)
  - NFV reshapes dedicated hardware functionality
  - software modules named virtual network functions (VNFs)
  - agile and scalable service placement
  - reducing Capital Expenditure (CAPEX) and Operation Expense (OPEX)

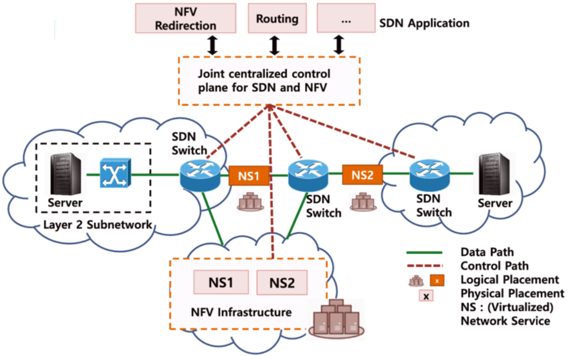


Figure: SDN based network with support for NFV

- Service Function Chaining(SFC)
  - specific set of VNFs
  - joint VNF placement and traffic routing are called SFC mapping

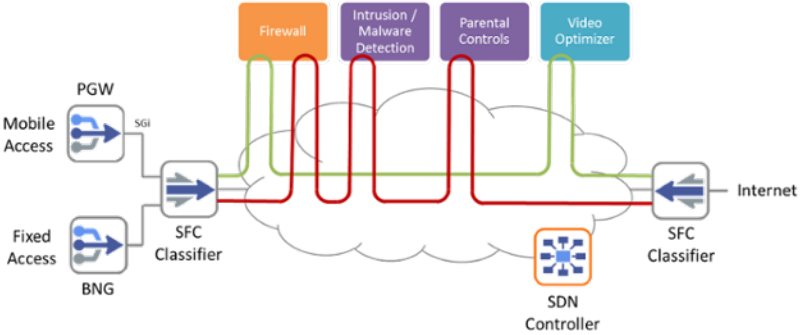


Figure: Service Function Chaining

Author	Year	Mapping	Solution	Objective
Draxler	2017	VNFs mapping	heuristic and exact	total data rate and total latency
Huin	2018	Placement	exact	total latency
Masri	2017	Select optimal fog or cloud	exact	total latency
Fan	2017	Task scheduling in a fog-to-cloud computing system	heuristic	maximizing the profits of fog service provider
Gupta	2017	VNFs mapping	ILP column -generation based model	bandwidth consumption

Table: Related Work

Related Work

- do not consider SFC mapping in the F2C architecture which is critical in the IoT domain
- do not consider the Service Level Agreement(SLA) as a constraint, while in our formulation it is considered
- we propose an ILP solution that provides the exact solution for SFC mapping in the F2C architecture
- in order to minimize total latency of network
  - the number of instances for each SFC and replicas for each VNF are considered

System Model

- Problem statement
  - network topology
  - capacity and latency of links
  - computing resources at the fogs and cloud nodes
  - traffic flows between two pairs of fog or fogs and cloud requiring a specific SFC
  - users SLA
  - instances number
  - placement of VNFs
  - corresponding traffic routing
  - users assignment to the SFC instances
  - minimize overall latency of network

System Model

- Service Function Chaining (SFC)

$$[SFC\ c] \quad f_{\sigma_1(c)} \rightarrow f_{\sigma_2(c)} \rightarrow \cdots \rightarrow f_{\sigma_{n_c}(c)} \tag{1}$$

- generate all configurations of each SFC $c \rightarrow \hat{\Gamma}_c$
- aggregation of all configuration of SFCs $\rightarrow \hat{\Gamma}$
- Each configuration ($\hat{\gamma}$) of SFC c is characterized by:
  - Location of VNFs $\rightarrow a_{vi}^{\hat{\gamma}}$
  - Connectivity of located VNFs $\rightarrow b_{i\ell}^{\hat{\gamma}}$
  - User assignment $\rightarrow \delta_{sd}^{\hat{\gamma}}$

Configuration

$$f_{\sigma_1(c_1)} \rightarrow f_{\sigma_2(c_1)} \rightarrow f_{\sigma_3(c_1)} \quad (f_{\sigma_1(c_1)} = f_1, \quad f_{\sigma_2(c_1)=f_5}, \quad f_{\sigma_3(c_1)} = f_8)$$

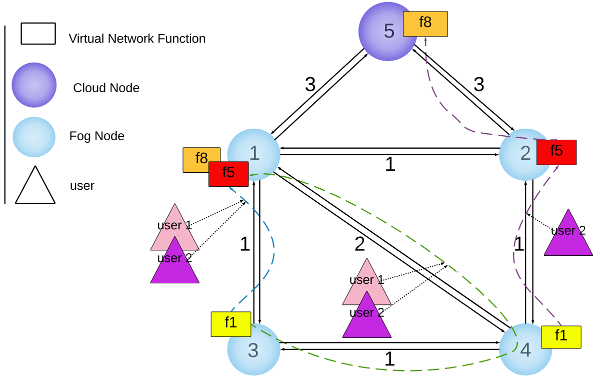


Figure: Three configurations

Objective function

$$min. \sum_{c\in C} \sum_{\hat{\gamma}\in\hat{\Gamma}_c} (\sum_{sd\in SD} \sum_{\ell\in L} \sum_{i=1}^{n_c-1} b_{i\ell}^{\hat{\gamma}} \delta_{sd}^{\hat{\gamma}} delay_\ell) z_{\hat{\gamma}} + \sum_{c\in C} \sum_{\ell\in L} \sum_{sd\in SD} delay_\ell (y_\ell^{f_1(c),sd} + y_\ell^{f_{n_c}(c),sd}) \quad (2)$$

Constrants

$$\sum_{\hat{\gamma}\in\hat{\Gamma}_c} \leq l_c \quad c \in C \tag{3}$$

$$\sum_{v\in V} x_{vf} \leq R_f \quad f \in F \tag{4}$$

$$\sum_{c\in C}\sum_{\hat{\gamma}\in\hat{\Gamma}_c}\sum_{sd\in SD} D^c_{sd}\delta^{\hat{\gamma}}_{sd} * \\ (\sum_{f\in F}\sum_{i=1}^{n_c} T^c_{fi} n^{CORE}_f a^{\hat{\gamma}}_{vi}) z_{\hat{\gamma}} \leq n^{FOGCORE}_v \quad v \in V_{FOG} \tag{5}$$

Constrants

$$\sum_{c\in C}\sum_{\hat{\gamma}\in\hat{\Gamma}_c}\sum_{sd\in SD} D^c_{sd}\delta^{\hat{\gamma}}_{sd}*$$

$$(\sum_{f\in F}\sum_{i=1}^{n_c} T^c_{fi} n_f^{CORE} a^{\hat{\gamma}}_{vi}) z_{\hat{\gamma}} \leq n_v^{CLOUDCORE} \quad v \in V_{CLOUD} \tag{6}$$

$$\sum_{c\in C}\sum_{\hat{\gamma}\in\hat{\Gamma}_c}\sum_{sd\in SD} D^c_{sd}*$$

$$(y_\ell^{f_1(c),sd} + y_\ell^{f_{n_c}(c),sd} + \sum_{\hat{\gamma}\in\hat{\Gamma}_c} \delta^{\hat{\gamma}}_{sd} z_{\hat{\gamma}} \sum_{i=1}^{n_c b-1 b^{\hat{\gamma}}_{i\ell}}) \leq CAP_\ell \tag{7}$$

$$\sum_{\hat{\gamma}\in\hat{\Gamma}_c}\sum_{\ell\in L} delay_{\ell} * (y_{\ell}^{f_1(c),sd} + y_{\ell}^{f_{n_c(c)},sd} + \tag{8}$$

$$\sum_{\hat{\gamma}\in\hat{\gamma}_c} \delta_{sd}^{\hat{\gamma}} z_{\hat{\gamma}} \sum_{i=1}^{n_c-1} b_{i\ell}^{\hat{\gamma}} \leq SLA_{sd}^{c} \quad c \in C, sd \in SD$$

$$\sum_{\hat{\gamma}\in\hat{\Gamma}_c} \delta_{sd}^{\hat{\gamma}} z_{\hat{\gamma}} = 1 \quad c \in C, sd \in SD : D_{sd}^{c} > 0 \tag{9}$$

Numeric Results

- 10 node pairs
- each VNF uses one core to run
- Each traffic flow is 1Gbps

Service Chain	Chained VNFs
Online Gaming	NAT-FW-VOC-WOC-IDPS

Numeric Results

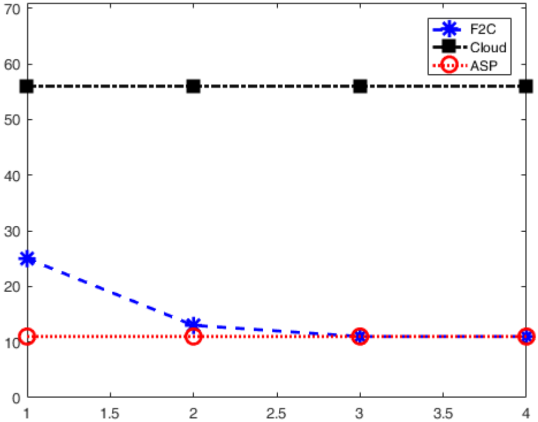


Figure: The overall end-to-end latency of network vs. number of instances.

Conclution

- multi Service Function Chain (SFC) mapping
- multiple SFC instances
- fog-to-cloud architecture
- formulate our problem as an Integer Linear Program
- minimize the overall end-to-end latency
- our model can achieve the least overall end-to-end latency when the number of instances increases

References

[7] Drxler, S., H. Karl, and Z.. Mann. Joint Optimization of Scaling and Placement of Virtual Network Services. in 2017 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID). 2017.

[8] Huin, N., B. Jaumard, and F. Giroire, Optimal Network Service Chain Provisioning. IEEE/ACM Transactions on Networking, 2018. 26(3): p. 1320-1333.

[9] Masri, W., et al. Minimizing delay in IoT systems through collaborative fog-to-fog (F2F) communication. in 2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN). 2017

[10] Fan, J., et al. Deadline-Aware Task Scheduling in a Tiered IoT.Infrastructure. in GLOBECOM 2017 - 2017 IEEE Global Communications Conference. 2017.

[11] Gupta, A., et al. Service Chain (SC) Mapping with Multiple SC Instances in a Wide Area Network. in GLOBECOM 2017 - 2017 IEEE Global Communications Conference. 2017.

Thanks for your attention.