2015 International Symposium on Nonlinear Theory and its Applications
NOLTA2015, Kowloon, Hong Kong, China, December 1-4, 2015

Optimization of Computational Resource Allocation for Virtualized Mobile Core Network

Kohei Hayashi†, Hiroyuki Yasuda†, Yozo Shoji‡ and Mikio Hasegawa†

†Graduate School of Engineering, Tokyo University of Science
6-3-1 Niijuku, Katsushikaku, Tokyo, 125-8585 Japan
‡National Institute of Information and Technology
4-2-1 Nukui-Kitamachi, Koganei, Tokyo, 184-8795 Japan
Email: †{hayashi, hirobacon}@haselab.ee.kagu.tus.ac.jp, hasegawa@ee.kagu.tus.ac.jp,
‡shoji@nict.go.jp

Abstract– With significant increase of mobile traffics and mobile services, signaling processing load is making big impact to the mobile communication network. In order to deal with the increasing signaling loads, service-specific network virtualization has been developed. By network function virtualization, it is possible to improve efficiency and fault tolerance by flexibly and dynamically assigning the signaling loads to appropriate servers. However, since there are the processing overheads in virtualization, it is an important issue to reduce the signaling processing delay caused by the virtualization. In this paper, we minimize the processing delay by optimizing the assignment of the signaling functions to multiple servers. Based on the actual signaling, we formulate an optimization problem to minimize the communication delay cost. Our computer simulations show that the proposed scheme could minimize the processing delay.

1. Introduction

All-IP mobile communication network has been developed and standardized. Recently, various wireless communication services, such as voice calls, short messages and various IP-based applications, are used on various types of smart mobile phones. By significant increase of such mobile services, signaling processing load to establish such mobile communication sessions is becoming a big impact to the mobile core networks [1]. In the mobile core networks, the Evolved Packet Core (EPC) and the IP multimedia subsystem (IMS) work to establish various wireless communication services. With increase of mobile service requests, processing loads to the EPC and the IMS is becoming serious issue.

In order to solve this issue, network virtualization technology is introduced to the mobile core networks [2] [3]. The functions of the EPC and the IMS have been implemented as a virtual function, which is software running on a number of physical servers.

In the mobile core network virtualization, although flexibility of assignment of the signaling load becomes very high, the processing speed of the virtualized servers is slower than physical dedicated servers. It is important to reduce such overhead of virtualization.

In this paper, we minimize the virtualization overhead by optimizing assignment of the computational resource of software-defined EPC/IMS. We formulate the optimization problem based on the problem definition in ref. [4]. By solving the optimization problem, we show the possibility of minimization of virtualization overhead.

2. The Architecture of Mobile Core Network

Figure 1 illustrates the mobile communication network architecture assumed for this study, which is composed of EPC and IMS. This network provides mobile communication services such as calling and messaging for user equipments (UEs). EPC bases on IP protocols, as standardized by IETF and OMA, enhanced by the 3GPP specifications and can coordinate several wireless access systems such as WLAN, WMAN and 3GPP2 radio access systems not only LTE [5]. The Core Network provides the handover and the control of Quality of Service (QoS). These are not dependent on a specific wireless access system and realized by a common mechanism. IMS is composed of multiple functions that manage and control voice service, messaging service, the databases of subscriber information and the network gateways.

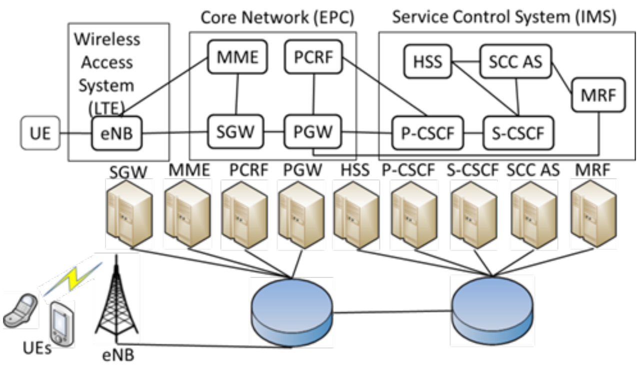


Figure 1: Architecture of mobile core network.

EPC is composed of a mobility management entity (MME), a serving gateway (SGW), a packet data network gateway (PGW), and policy and charging rules function (PCRF). The IMS is composed of two types of call session control functions (CSCFs), home subscriber server (HSS), and application servers (ASs), service centralization and continuity AS (SCC AS), and media resource function (MRF). Figure. 2 shows the initial attach procedure of UE. A user needs to register with the network to receive services that require registration. This registration is described as network attachment. When the power is turned on, the UE registers itself with the EPC (this is the ‘‘attach’’ procedure). In this procedure, the EPC performs location registration of the UE, assigns it an IP address, and establishes a default bearer for exchanging SIP messages between the UE and IMS. After completing Attach, the UE exchanges SIP messages (REGISTER and its response) with the IMS and register itself.

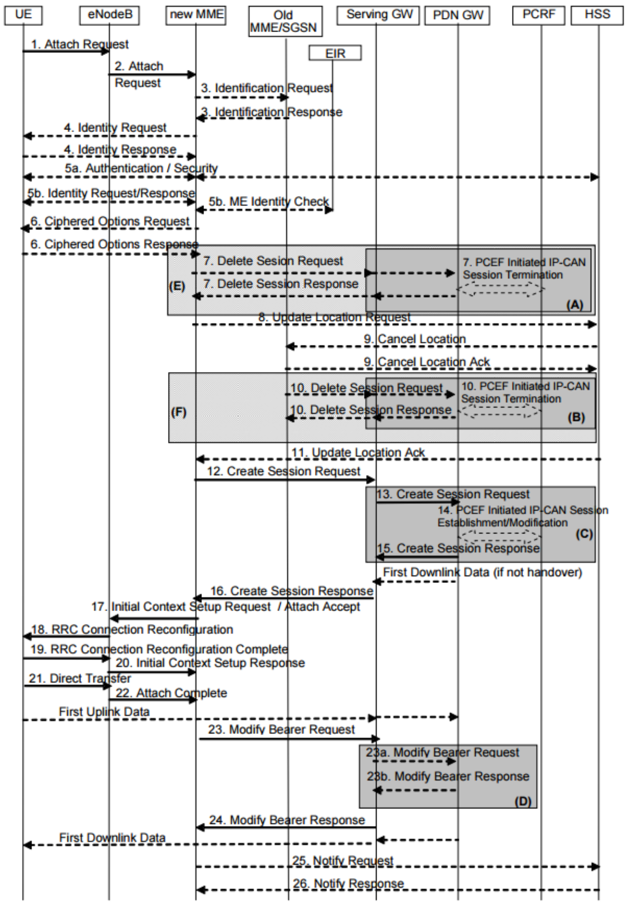


Figure 2: Attach procedure of User Equipment.

3. Virtualization of Mobile Core Network

Communications congestion caused by the spreads of mobile devices, it is necessary for the core network to minimize the impact on services of the user even when a failure occurs. In recent years, server virtualization and network virtualization such as Software-Defined Network has been studied to expand and apply to the network of the telecommunications operator [6] [7]. While control signals and traffic in the network increase, it is necessary to prepare properly the computational resources required to service requests of the users in the core network. Applying server virtualization and network virtualization to core networks enables a flexible allocation of server resources in response to the service request quantity. Also even if there is a failure in a part of the server to assign other server newly, by alternative processing of a new service request, it is possible to service operation that does not continue the fault effects [8].

4. Combined Optimization of Assignment for Virtualized Functions on the Virtual Infrastructure

For Virtualized Mobile Core Network, it is necessary to operate effectively to assess the overall processing delay. Virtualized EPC/IMS network is composed of functions in IMS/EPC as virtual machines (VMs) on physical servers [2].

In this paper, we formulate a combinatorial optimization problem by defining the state variable including the virtual function assignment to minimize the delay of session establishment. We define the state of the function assignment by one state variable x_{lip} , which becomes 1 when the virtual function i runs on the physical server p specified by the service l , otherwise x_{lip} becomes 0. The communication sequence of specific service (e.g., voice call, messaging or multi-device service) provided by EPC/IMS is given. We define s_{lki} as the generation of a signaling message at the virtual function, which is 1 when the function i send the signaling message at the k-th in the sequence of the service l, otherwise 0. P is the number of physical servers in the system, I is the total number of virtual functions, L is the number of services provided by EPC/IMS, K_l is the number of processing in service l .

When virtual functions communicate between different physical servers, we define the coefficients of the commutation delay by β_{pq} ; the delay time T_1 can be formulated as

$$T_1 = \sum_{l=1}^{L} \sum_{k=1}^{K_l} \sum_{i=1}^{I} \sum_{j=1}^{I} \sum_{p=1}^{P} \sum_{q=1}^{P} \beta_{pq} s_{lki} s_{l(k+1)j} x_{lip} x_{ljq}, \quad (1)$$

when virtual functions running on the same physical server p communicate signaling messages, we formulate the communication delay T_2 using the coefficient γ_p as

$$T_2 = \sum_{l=1}^{L} \sum_{k=1}^{K_l} \sum_{i=1}^{I} \sum_{j=1}^{I} \sum_{p=1}^{P} \gamma_p s_{lki} s_{l(k+1)j} x_{lip} x_{ljp}, \quad (2)$$

using δ_{lk} as the coefficients of the processing delay at the k-th process in the sequence of service l and processing time T_3 is

$$T_3 = \sum_{l=1}^{L}\sum_{k=1}^{K_l}\sum_{i=1}^{I}\sum_{p=1}^{P}\delta_{lk} s_{lki} x_{lip}, \qquad (3)$$

each physical server can have multiple virtual machines and one virtual machine must be allocated to one physical server

$$\sum_{p=1}^{P} x_{lip} = 1, i \in I, l \in L. \qquad (4)$$

By using the equation (1)-(4) we formulate a combinatorial optimization problem to optimize the total delay of providing mobile communication services as follows,

$$\begin{aligned} &\underset{\mathbf{x}}{\text{Minimize}} \ f(\mathbf{x}) = T_1 + T_2 + T_3 \\ &\text{subject to} \ \sum_{p=1}^{P} x_{lip} = 1, i \in I, l \in L. \end{aligned} \qquad (5)$$

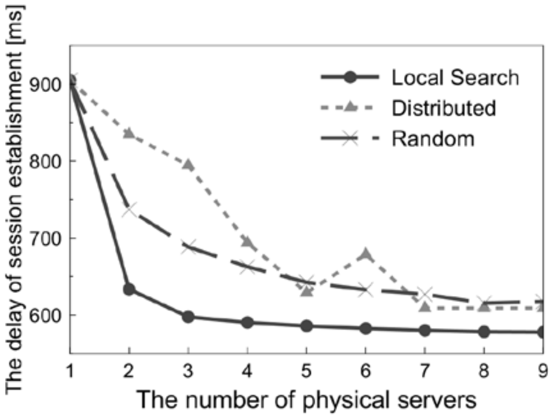
In order to minimize the total delay of providing mobile communication services, we have to find the optimal state variable x_{lip} which minimize the sum of T_1, T_2 and T_3.

5. Simulations and Evaluation

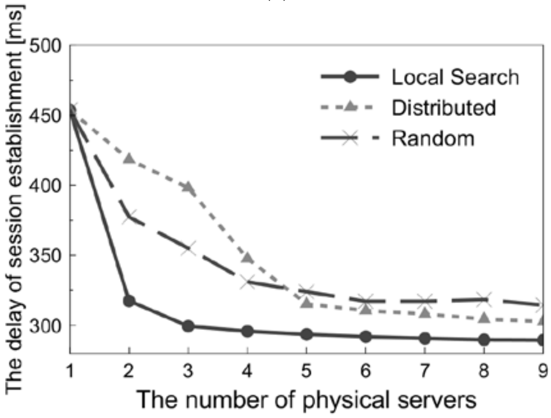
We simulated the delay of session establishment, when the network has nine kinds of functions on multiple physical servers. We used the sequences of multi-device service and voice call service as examples.

Based on actual measurements, the parameters are used as follows. The delay of communication between servers coefficient β_{pq} is 0.7ms, the coefficient γ_p is set to 0.3ms using a virtual switch within one physical server. Also in this case, it is assumed that all processing in a virtual machine takes one by 2.0ms. We evaluate the total delay of the assignment using local search to the optimization problem as compared with the random allocation and the distributed allocation. In this paper, we use 2-opt exchanges as local search method.

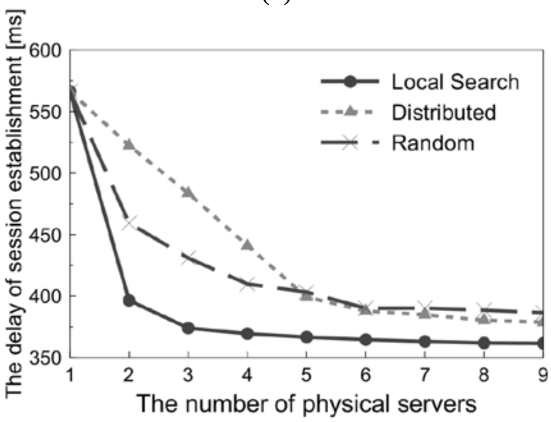
Figure 3 (a) – (d) shows the overall delay to provide the service for each user in virtualized EPC/IMS toward the four kinds of ratio of voice call and multi device service. The ratio of voice call and multi device service in (a) is 1:1, (b) is 9:1, (c) is 8:2 and (d) is 3:7. The multi device service needs more signaling loads than voice call service. The results show that the optimization of the resource allocation for virtualized functions can reduce the total delay for the providing services when the signaling load is big.



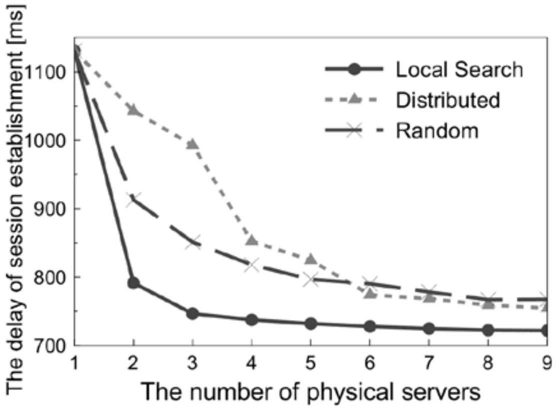
(a)



(b)



(c)



(d)

Figure 3: The delay of providing the mobile communications services per user. (a) The ratio of voice call to multi-device service is 1:1. (b) The ratio is 9:1. (c) The ratio is 8:2. (d) The ratio is 3:7.

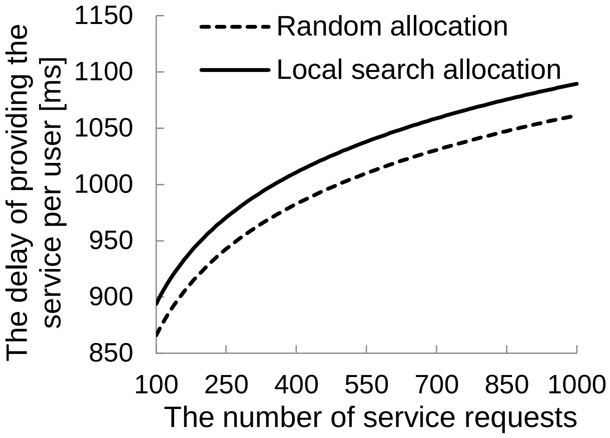


Figure 4: The total delay of providing the mobile communication services per user at the increasing the number of service requests.

Figure 4 shows the relationship between the delay and the number of service requests. The delay at the optimized allocation is smaller than the delay at the random allocation.

6. Conclusion

To reduce the load of processing signaling messages in a mobile communication network, an approach that creates several virtual networks that are composed of network functions specialized for particular services have been proposed in the study of [2]. In this paper, we study a method that optimizes computational resource allocation for virtual machines in that virtualized Mobile Core Network. By taking a virtual network specialized for a multi-device service and a call service as examples, we conducted a sample numerical analysis where network virtualization overheads are taken into account. In addition, we have shown that the proposed mechanism could reduce the load under our assumption. In the future, we will confirm the effect of load reduction by simulation and a demonstration experiment.

References

[1] G. Brown, "White paper: The evolution of the signaling challenge in 3G & 4G networks," Heavy Reading, New York, NY, USA, Tech. Rep., pp. 1-14, 2012.

[2] M. Ito, K. Nakauchi, Y. Shoji, N. Nishinaga and Y. Kitatsuji, "Service-Specific Network Virtualization to Reduce Signaling Processing Loads in EPC/IMS," 2014.

[3] Kostas Pentikousis, Yan Wang and Weihua Hu, "Mobileflow: Toward software- defined mobile networks," 2013.

[4] M. Ito, K. Nakauchi, Y. Shoji and Y. Kitatsuji, "Mechanism of Reducing Signaling Processing Load in EPC/IMS Using Service-Specific Network Virtualization," 2014.

[5] 3GPP TS23.402 V12.7.0, "Universal Mobile Telecommunications System (UMTS); LTE; Architecture enhancements for non-3GPP accesses," 2015.

[6] A. Bradai, K. Singh, T. Ahmed and T. Rasheed, "Cellular software defined networking: a framework," 2015.

[7] J. Costa-Requena, "SDN integration in LTE mobile backhaul networks," Proceedings of International Conference on Information Networking, pp. 264-269, 2014.

[8] A. Basta, W. Kellerer, M. Hoffmann, K. Hoffmann and E. D. Schmidt, "A Virtual SDN-Enabled LTE EPC Architecture: A Case Study for S-/P-Gateways Functions," 2013 IEEE Software Defined Networks for Future Networks and Services, pp. 1-7, 2013.