Service-specific Mobile Network Virtualization Technologies

Manabu ITO

A mobile network that provides mobile communication services becomes increasingly important as a social infrastructure. To address the problems of the existing mobile network and realize a future mobile network, the use of network virtualization technologies is a promising solution. In this paper, we introduce the fundamental technologies for applying network virtualization technologies to the mobile network.

1 Introduction

For the purpose of solving the problems of the currently used Internet and implementing the New-Generation Network that is expected to be a future social infrastructure, research and development activities on network virtualization technologies have been conducted. Network virtualization technologies are expected to be the base technologies for the New-Generation Network because the technologies will enable the provision, on a single physical network, of different network services which have their own requirements, such as transmission speed, reliability or the volume of connected user-terminals.

Furthermore, if the current situation, where mobile communication networks (hereinafter referred to as “mobile networks”) have been used widely, is taken into consideration, the New-Generation Network, in order to become a reliable social infrastructure that supports social systems whether in a normal or emergency situation, must be capable of integrating mobile networks and configuring them virtually. On the other hand, mobile-network operators have come to share the view that network virtualization technologies will be one of the keys to the realization of 5G systems, where the additional requirements such as ultra-low latency or a huge number of connections must be satisfied. Network virtualization technologies will enable the provision of ultra-low-latency services through configuring the “slices” in the neighborhood of a terminal. In addition, the reduction of the total network load will be enabled, by configuring such slices that consist of a number of necessary functional units which can conduct service-specific, instead of service-independent, controls.

In the following portions of this paper, base technologies for applying virtualization technologies to mobile networks (hereinafter referred to as “mobile-network virtualization”) will be described. First, the gateway function—indispensable for realizing mobile-network virtualization—is introduced. Next, a proof-of-concept prototype experiment, which has proved the possibility of reducing processing load in a mobile network, is introduced. Then, a scalable information sharing method among a number of gateway functions will be proposed, and its validity is shown.

2 Applying network virtualization technologies to mobile networks

2.1 Existing mobile networks and current issues

Figure 1 shows the architecture of Long Term Evolution (LTE)/Evolved Packet Core (EPC)—a widely-used mobile network. Functions of the nodes depicted in the figure are described as follows: A user-terminal (User Equipment: UE) is connected to a base station (eNodeB: eNB). Packets received/transmitted by the UE are delivered, by using the GPRS Tunneling Protocol (GTP Tunnel), to an external network or other UE. GTP tunnels are established between Serving Gateways (eNB-SGW), and between Packet-data network Gateways (SGW-PGW); those terminal nodes conduct packet-routing control and policy control. SGW

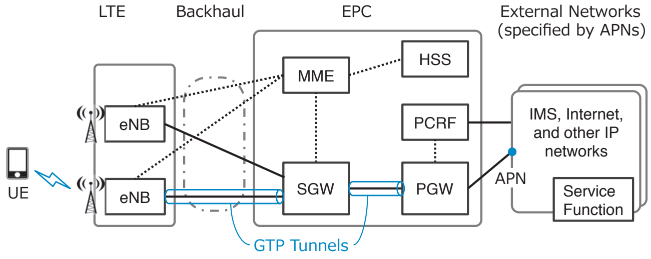

LTE
Backhaul
EPC
External Networks (specified by APNs)
HSS
MME
eNB
PCRF
IMS, Internet, and other IP networks
UE
eNB
SGW
PGW
APN
Service Function
GTP Tunnels

Fig. 1 Mobile network architecture

works as an anchor point, providing the UE with eNB-to-eNB seamless handover. PGW's functions are traffic monitoring, charging, access control and a gateway to external networks. A GTP Tunnel is established through collaboration between the Mobility Management Entry (MME) and Policy Control and Charging Rules Function (PCRF). MME, being a Control Plane (C-Plane) function, taking the responsibility of mobility management and user authentication, executes bearer (logical packet-transmission route) session control through collaboration with eNB and SGW. PCRF sets up, into PGW/SGW, the QoS (QoS class, bit rate, etc.) according to the user profile information and the media information notified by an external network (IP address, port number, protocol type, media codec, etc.).

UE, in order to establish communication with an external network (IMS, the Internet, etc.), sends to the MME a bearer-session establish-request, which carries the Access Point Name (APN)—the identifier of the external network that provides the demanded service; MME, according to the APN, identifies an appropriate SGW for establishing a GTP Tunnel between the SGW and the PGW—the connection point to the external network identified by the APN. The SGW, receiving the request from MME, establishes an eNB-to-SGW GTP Tunnel and SGW-to-PGW GTP Tunnel. Now, the UE is allowed to, through the tunnel, by using the Session Initiation Protocol (SIP)—an application-layer control protocol—or HTTP, exchange application-level session-control messages with the service function that the external network has. The service function, along with such message exchange, notifies PCRF of media information, the PCRF sends to the PGW QoS information and the media information, and the information, passing through the SGW, is sent to the MME. The MME allocates, according to the information, wireless resources to the eNB. In addition, the MME, if necessary, establishes a bearer for media or upgrades the existing bearer.

Networks in the future are required to have high-grade features such as ultra-low latency or very high efficiency, because on conventional network architectures those features are unattainable. As for ultra-low latency, which is expected to enable such new services as tactile communication, virtual offices or automatic vehicle-cruising, an end-to-end (E2E) latency of as low as 5 milliseconds is said to be required[1]—on the conventional mobile-network architectures where a PGW or a service function on an external network works as an anchor point, delays as long as several tens to several hundreds of milliseconds occur even if the subscribers belonging to the same mobile network operator are communicating. On the other hand, as for the feature of very-high efficiency, sophisticated session control is necessary—sufficiently high-speed so as to reduce the processing load required for serving a huge number of user terminals—, while in the conventional mobile networks, inflexible session controls by means of fixed parameters or procedures independent of service situations or environments are employed and cause inefficiency in the process—unnecessary processing may be done. Furthermore, on the assumption that UEs move in a mutually independent manner, inflexible session controls are conducted independently of the UE types.

2.2 Solutions applying service-specific controls

For solving the problems of the conventional mobile networks, mentioned in the previous subsection, a service-oriented (service-specific or adaptive) network configuration and control scheme is required. For attaining ultra-low latency, the network configuration has to be re-configured adaptively so that the physical communication distance becomes shorter—function units including anchor points must be located at a short distance from UEs.

On the other hand, for sophisticated controls, the following measures will be required: Controls of multiple UEs should be conducted in a consolidated way—conducting a number of controls at one place; and such controls that are not necessary for the concerning services should be canceled; Updating control parameters in a service-specific way according to the types of services or current service-utilization status—for example, on the occasion of providing a group-communication service to multiple UEs, the number of control processes and control messages will be reduced[2] through conducting controls for allocating a media transmitting/receiving bearer at a time to each UE, or for the terminals that are supposed not to move—such as Machine Type Communication (MTC) terminals—skipping scheduled mobile management controls will reduce processing loads[3]. In addition to the abovementioned measures for sophisticated controls in mobile networks, different methods[4]–[7] have been proposed.

2.3 Applying network virtualization technologies: key to solving problems

If independent and dedicated physical mobile networks are constructed to use on a case-by case basis for satisfying different requirements, it would lead to the rise of infrastructure construction cost for the following reasons. It would not be easy to estimate the required network facility capacity for individual mobile networks. Sophisticated control by enhancing the component function units of mobile networks—although it would be effective—, would need more time and cost for verifying that no inconsistencies exist between the existing functions because maintaining compatibility to avoid mutual conflict is required; therefore, the timely applying of sophisticated methods will be difficult, and in addition, changing controls in each function according to the service would possibly lead to the lowering of processing performance. Different from the construction of dedicated physical networks mentioned above, network virtualization technologies have attracted attention as promising technologies (network-virtualization technology refers to a technology for building logically independent networks—virtual networks—on a single physical network), for constructing dedicated mobile networks and ensuring the dynamic configuration/activation of software functions—applications executable on standardized hardware systems.

Virtualization technologies enable the realization of service-specific mobile networks on a single physical network through constructing mobile networks as virtual networks (virtual mobile networks). Such a construction method allows each mobile network to flexibly update its configuration or resource allocation. In addition, each mobile network, because it is isolated, is not required to have functions that are irrelevant to its services, nor required to maintain the compatibility of its functions with other functions (on existing virtual networks). Therefore, the construction, on a virtualization platform, of virtual mobile networks consisting of dedicated functions with EPC/IMS control matching the network’s services ensures the provision of additional network requirements, and the total efficiency of the network.

With regard to the technologies for virtualization of network resources/functions, many ideas, such as Service Defined Networking (SDN) or Network Function Virtualization (NFV) have been proposed and studied. Next in this paper, studies on the key technologies for improving the network total efficiency by applying virtualization technologies to mobile networks will be introduced.

3 Service-specific mobile-network virtualization and service-flow control technologies

3.1 Virtualized mobile-network configuration enabling service-specific control

Figure 2 shows an instance of a virtualized mobile-network configuration as follows: On a network-virtualization platform, there are constructed a virtualized mobile network (vMNW1) consisting of the functions required by standards, and a specific-service-optimized virtualized mobile network (vMNW2 and vMNW3). The vMNW2 is constructed by placing such network components as SGW, PGW and service functions that form the data communication route on a physical node neighboring the terminal. The vMNW3 is configured so that it includes the necessary C-Plane functions. Each of the functions is built by enhancing standard procedures. As for the group-communication service for example, setting-up processes of transmitting/receiving bearers to be allocated to each UE are conducted in a consolidated manner, for the purpose of reducing the processing load. Each of the virtualized networks is configured through the interconnection of virtualized servers (VM) working on IA servers via layer-2 level protocols such as Virtual eXtensible Local Area Network (VXLAN). In each VM, the functions run as software. Packets, before they are exchanged between those functions, are encapsulated by using the overlay protocol at a virtualized switch existing on the IA server. In order to make a packet sent from a UE processed in the proper virtualized network, pre-processing is required before a packet enters the IA server-group so that the packet is inspected for the identification of the services relevant to it and is encapsulated (hereinafter, referred to as “service-flow control). Such a process will be conducted either at the terminal side or at the network side; however, in the case where it is done at

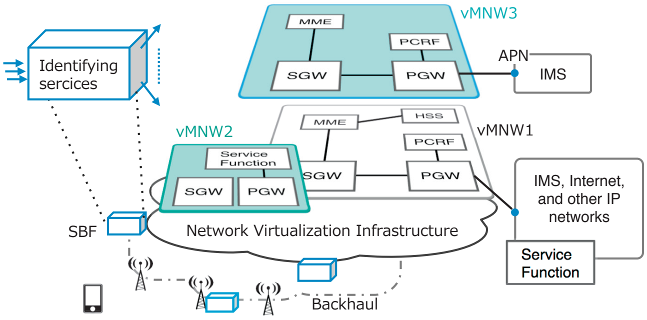


Fig. 2 Configuration of virtualized mobile-network virtualization enabling service-specific control

the terminal side, terminals must be enhanced so that they have a high-speed packet-I/O mechanism and allow their applications to handle layer-2 frames. This means that such terminal-side approach will allow only the packets sent from the enhanced terminals—not all the terminals—to be allocated to the proper virtual network. In addition, considering such a situation where MTC terminals with simplified communication features for power saving will be popular in the future, it would not be expectable that all types of terminals will be enhanced in such a way as described above. Therefore, it will be practical and desirable to implement service-flow control features at the network side.

For the purpose of conducting service-flow controls on the network side, it is necessary to analyze packets down to their payload level (hereinafter referred to as "packet inspection"), because "5-tuple" (source/destination IP address, source/destination port number, and protocol type) will not be sufficient for the network to know what service is requested. It is difficult to, only by using 5-tuple, precisely sort out signaling messages for bearer-session control in an EPC, SIP or HTTP messages, to the level compatible with the granularity of services. Such situations can be described, as follows: A signaling message transmitted by a UE (S1 AP message[8]), encapsulated by means of the Stream Control Transmission Protocol (SCTP)[9], is sent to its MME; however, the alternation of MME according to the terminal type or APN (external network identifier) will never occur, but the destination MME is determined according to the terminal's International Mobile Subscriber Identity (IMSI) used for identifying terminals. In addition, SIP messages (or HTTP messages) transmitted by the UE, independently of the demanded service, are sent to a fixed SIP server (or webRTC server). Therefore, it is required, by means of packet inspection, to identify a UE or APN written in the bearer-session control message and the demanded service written in the service-session control-message, and encapsulate and transmit those to the proper destinations. Hereinafter, the function for inspecting messages down to the granularity of services is referred to as the Service Binding Function (SBF).

3.2 Efficiency technologies for service-flow control

When SBF is implemented, inspecting each packet by using character-string matching will cause a large overhead (processing load) that will not only cancel out the benefits brought by the employment of network virtualization technologies but increase the risk of consuming a large volume of resources. Hence, for the purpose of improving efficiency, such a method is proposed as picking up the packets that seem necessary to inspect by using a combination of 5-tuple filtering. While, as for the terminals for dedicated use—only used for specific services or in fixed environments—pre-setting 5-tuple information to SBF would be enough, however, in such a case where terminals will be used in a virtual network, occasionally (dynamically) for some certain services or in some certain environments, SBF has to be capable of dynamically accepting the 5-tuple information and character-string-matching rules depending on the situation. The service-flow control method that is proposed in the following will be effectively applicable to such a situation.

Figure 3 shows the outline of the scheme; vMNW3 is pre-configured; SBF consists of a 5-tuple flow-identification function and a service-identification function; and the virtual NW management mechanism is enhanced to have the capability of updating the flow entry and the rule information stored in the SBF in collaboration with the network side. Procedures are described as follows, using a case where a UE attempts to start using a group communication service: Step 1: a service-request transmitted by the UE (SIP INVTTE), via the flow identification function, reaches the SIP server (P-CSCF1) on vMNW1—it is assumed that the flow entry of the flow identification function is pre-set so that a SIP message, if it is sent from a specific source IP address, bypasses service identification to be sent to vMNW1 using the encapsulated transmission method. Step 2: P-CSCF1 processes the SIP message; Step 3: P-CSCF1, while processing the message, notifies the virtual NW management function of the service identifier (also UE information, etc.) included in the SIP-header; Step 4: the virtual NW management function, according to the information notified by P-CSCF1, makes a copy of the state information (UE connection status) held by

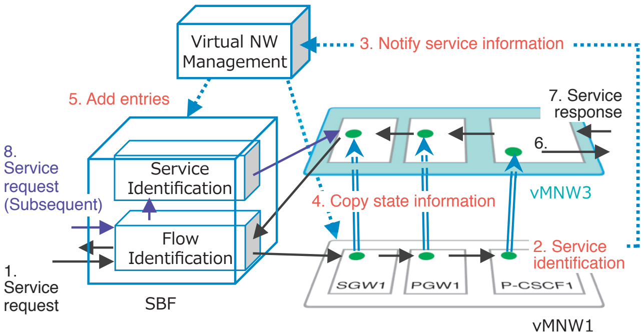


Fig. 3 Service-flow control efficiency method

a vMNW1 function, and places it in the corresponding function in vMNW3; Step 5: at the same time, the virtual NW management function adds a flow entry on the flow identification function so that SIP messages from a specific UE are transmitted to the service identification function, where, as for the state information copying method, redundant server technology, the relocation method (specified by standards), or a method applying hand-over-messages will be applicable; Step 6: the SIP message is processed by vMNW3's functions; Step 7: SBF, on receiving the reply messages from the Step 6 process, updates the parameters written in the SIP message if necessary—if a function of vMNW3 is based on a different IP system from that of vMNW1, it is required to rewrite the parameters written in the header of SIP, IP or GTP. Step 8: The subsequent request SIP message from the UE—in normal situations, several pairs of request/reply messages are exchanged—, sent by the flow identification function to the service identification function, processed there for service identification by means of character-string matching, is sent, according to the result of service identification by encapsulation transmission, to vMNW3; and then the setting of all the media transmit/receive bearers to be allocated to each of the UEs belonging to the same group is carried out in a consolidated manner.

4 Enhancement of service-flow control functions (SBF) for practical usage

4.1 Rule-Sharing among multiple SBFs and scalability

In a real situation, multiple SBFs are deployed in such a way that they are geographically distributed; and furthermore, multiple deployment of SBF is necessary for load distribution. SBFs, while working between a UE and the virtualization platform, are desirably located as close as possible to the UE so that anchor functions should be available in the neighborhood of the UE. Hence, attaching SBF to eNB seems to be effective; however, equipping every eNB with SBF would not be practical from the standpoint of cost—capital expenditure/operating expenditure (CAPEX/OPEX)—, because it is highly predicted that more base stations will be required as the reduction of service cell size goes forward for the purpose of improving base-station capacity. Then, a more practical way would be equipping back-halls or some macro-eNBs with SBF functions. In any of the situations described above where multiple SBFs are physically located in a distributed way, the sharing of necessary rules will be indispensable for the purpose of transmitting packets to a proper virtualized network wherever the terminal moves. However, because the number of the necessary rules is huge and, furthermore, situations will be dynamically changeable, synchronization of all the rules among every SBF is unfeasible from the standpoint of keeping scalability. Therefore, for solving the problem, the application of centralized-control technologies[10], which are supporting SDN technologies as their base, will be effective. Each SBF, each time it receives an unknown packet (having no rule information, so the demanded service is un-identifiable), transmits the unknown packet to the controller, and the controller, conducting packet inspection (the controller holds every rule), installs necessary rules into the SBF; in such a way, even if the UE is moving, necessary rules will be properly transmitted to the corresponding SBF.

On the contrary, as for the controller, another problem will arise: the controller, having all the rules—the total volume of which is huge—, consumes some certain amount of time for packet inspection[11]; as a result, in such a situation where a certain volume of unknown packets, at a unit time interval, flows into the controller, packet-processing at the controller becomes a bottleneck, causing large transmission latency to the virtualized network. Although load-distributing among some controllers might be effective, the growth of management cost will cause another problem in the case of multiple controllers because, if a huge volume of rules is required to be updated, every controller has to keep consistency among rules[12]. Hence, in order to prevent the number of controllers from increasing, the key to the problem will be preventing the occurrence of the concentration of too many unknown to-be-inspected packets on a controller.

4.2 Redirection-based rule-sharing method

Here, the redirection-based rule-sharing method—aiming at the improvement of the scalability in the rule-sharing method using the centralized control method mentioned in the previous subsection—is introduced. An SBF in this method redirects an unknown packet not to the controller but to the SBF accommodating the base station that had provided the UE with services, and tries to obtain necessary information from the SBF, expecting that the SBF, conducting packet-inspection, identifying service, returns the rules for the service. Because every SBF existing on the transmission route through which the unknown packet travels saves the rules in cache, the rules are shared among such SBFs.

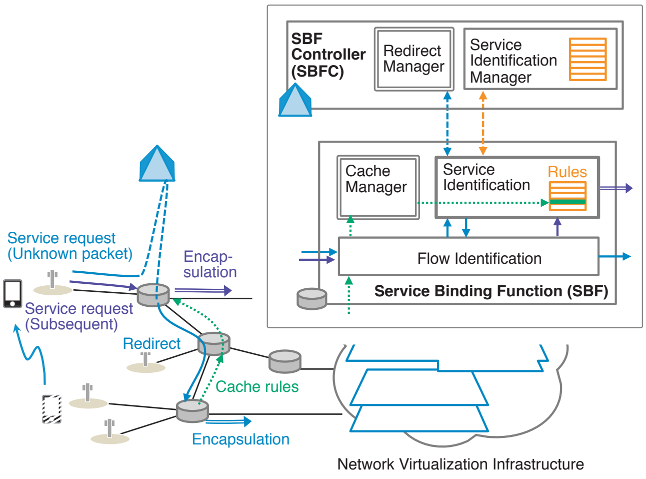

SBF Controller (SBFC)
Redirect Manager
Service Identification Manager
Cache Manager
Service Identification
Rules
Flow Identification
Service Binding Function (SBF)
Service request (Unknown packet)
Encap-sulation
Service request (Subsequent)
Redirect
Cache rules
Encapsulation
Network Virtualization Infrastructure

Fig. 4 Redirection-based rule-sharing method: outline

Figure 4 shows how the system using the proposed method works. SBFs are functionally-enhanced to enable such redirection and rule-caching. An SBF, on the receipt of an unknown packet, in order to identify the redirection destination, transfers a part of the packet (the part holding the information for identifying the source-UE) to the controller (hereinafter, referred to as SBF Controller, SBFC), and buffers the packet body. The SBF, when failing to obtain the redirection destination, transfers the buffered packet body to the SBFC to request packet inspection of the unknown packet, or, when successfully obtaining the redirection SBF from the SBFC, transfers the buffered packet to the redirection destination SBF. The redirection destination SBF, through applying the rules it holds, inspects the transferred packet. The SBF, when successfully identifying service, encapsulates the packet and transfers it to the proper virtualized network, and at the same time returns the rules—usable for identifying service—toward the redirection origin SBF, and when failing to identify, transfers the packet to the SBFC. The cache management function, which has been added to SBFs, records the I/F information on the occasion of unknown packet reception so that rule caching by hop-by-hop transfer is enabled. The SBF that received the rules, caches the rules and transfers them to the registered I/F.

The SBFC is equipped with the redirect manager function for holding the mapping information of the UE and the redirection destination SBF. The SBFC, when receiving a packet having no other information than UE-identification information, returns the redirection destination SBF corresponding to the UE. The SBFC, when having no mapping information available, returns a request of the whole packet to the SBF. The SBFC, when receiving the whole packet, conducts service identification by packet inspection and returns the corresponding rules to the SBF.

In the proposed method, an SBF, according to the mapping information that the SBFC holds, redirects an unknown packet to another SBF. However, a case would occur where the desired rules do not exist in the redirection destination SBF. For instance, the rules, although they should be held, might be deleted by a time-out, or the UE might request a different service from the service requested previously. In such a situation, the SBF, having no means for identifying service, transfers a packet to the SBFC.

In the method described above, the redirection destination SBF places the rules used for service identification on the transmission route in either case where the service identification is accomplished using the rules it holds, or where such identification is done through transferring the unknown packet to the SBFC. Hence, the proposed method will contribute to reduction of unknown-packet transfer volume to the SBFC as a whole, because the volume of the rules cached on each of the SBFs that exist on the transmission route grows in any case—regardless of the availability of the proper rules on the redirect destination SBF.

5 Evaluations by prototype-implementation and simulations

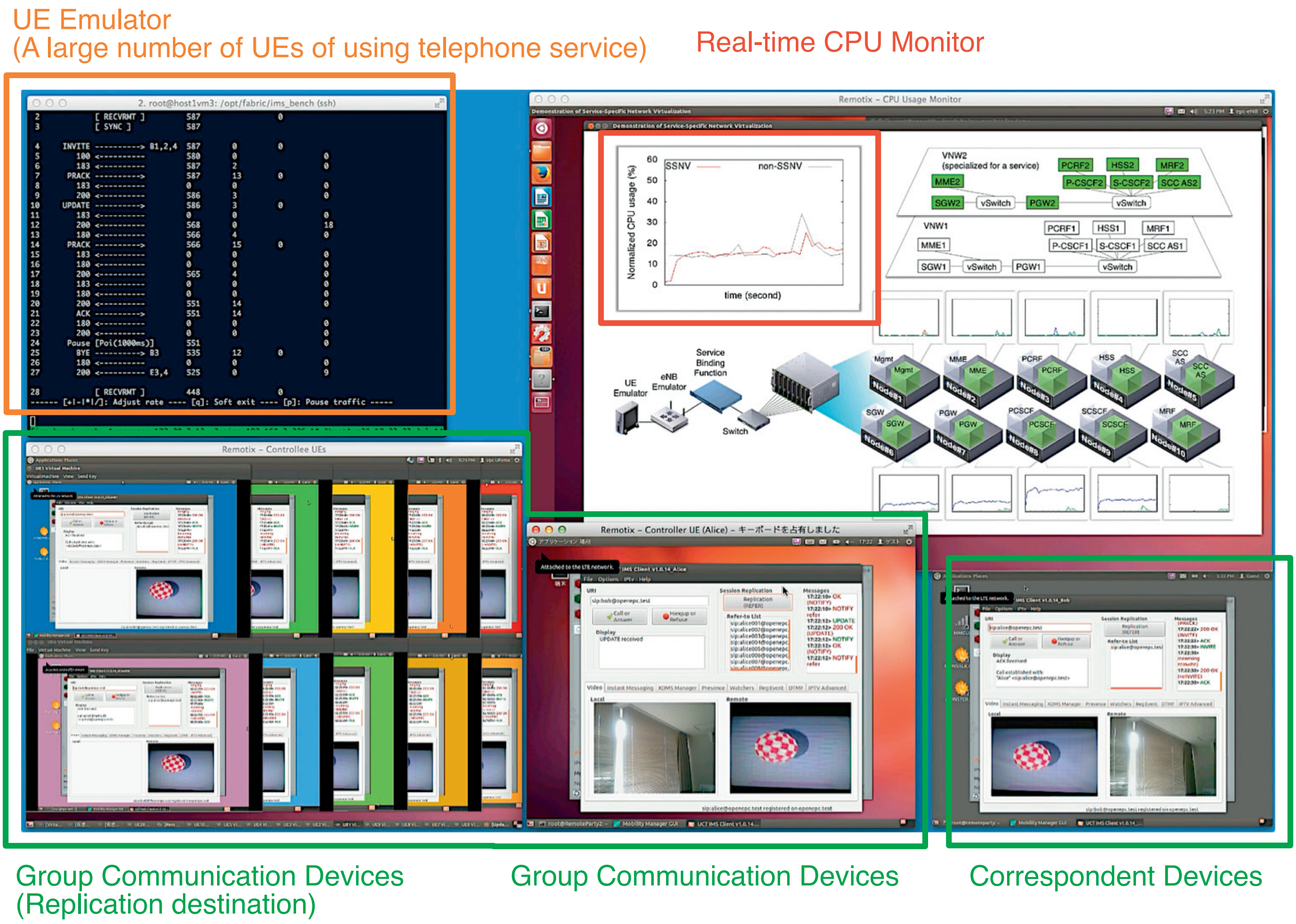
In the following Subsection **5.1**, evaluations conducted on a prototype implementation are introduced. They proved that on the service-specific mobile network virtualization, overall network efficiency is attainable. In Subsection **5.2**, simulations on the effects of the redirection-based rule-sharing method expected to solve the problems in the real situations (implementation problems) are introduced, which validated the method.

5.1 Example of efficiency improvement using service-specific mobile network virtualization: proof by prototype implementation

The effectiveness of Service-Specific Mobile Network Virtualization Method for reduction of the processing load of the whole network, was determined through observing its CPU utilization when group communication service is used, on the prototype on which the method was implemented (Fig. 5). The configuration of the prototype is shown below. The operation verification and the performance evaluation of the method were conducted using a virtualization platform with ten IA-servers connected. Two types of virtualized networks were built by using nine

IA servers. Functions of EPC/IMS were implemented as software packages on VMs which were interconnected using VXLAN technology—an overlay technology. On one of the two virtualized networks mentioned earlier, functions required by the standards were working (VNW1 shown at the upper right of the Fig. 5). On another network, functions enhanced for supporting group communication services were working (VNW2 shown at the upper right of Fig. 5). On the remaining IA-server out of the ten, a VM where a virtualized network management function was working (shown as Mgmt in Fig. 5) was placed. SBFs, working as service-flow control functions, were implemented as a software package on a physical node connected to another physical node emulating eNB at its network-side (up-stream). The physical node emulating multiple UEs was connected to EPC/IMS through eNB and SBF—the EPC/IMS or the eNB was enhanced to have the necessary functions, by using OpenEPC software[13]. SBFs, Mgmt's and service functions (SCC AS or MRF shown in Fig. 5) were newly developed for working as software. Furthermore, a UTC IMS Client[14]—open-source IMS client was enhanced so that it works as the UE using group communication services. IMS Bench SIPp[15]—an open-source emulator—was used for emulating multiple UEs using general service (telephone service).

On the prototype experiment configuration described above, group communication service was activated, under the situation where telephone service was under way at a call request rate of 20 cps (calls per second), and the CPU utilization rate was measured for SBFs and tens of IA servers. Figure 6 shows the measured CPU utilization rates—normalized so that the total of the CPU utilization for each node is between 0 to 100 percent—for comparing the case (non-SSNV) where the service-specific virtual network (VNW2) was not used with the case (SSNV) where it was used. The results show that CPU utilization reduction of approximately 25 percent was obtained—at the same time, it reveals that the sophisticated SBF contributes to suppressing the overall overhead, even if additional overhead might occur in each of the virtualized network, the virtual network management function, or SBFs.



```
2            [ RECVRMT ]          587                0
3            [ SYNC ]             587

4     INVITE ----------> B1,2,4   587      0         0
5        100 <----------          580      0                   0
6        183 <----------          587      2                   0
7      PRACK ---------->          587      13        0
8        183 <----------          0        0                   0
9        200 <----------          586      3                   0
10    UPDATE ---------->          586      3         0
11       183 <----------          0        0                   0
12       200 <----------          568      0                   18
13       180 <----------          566      4                   0
14     PRACK ---------->          566      15        0
15       183 <----------          0        0                   0
16       180 <----------          0        0                   0
17       200 <----------          565      4                   0
18       183 <----------          0        0                   0
19       180 <----------          0        0                   0
20       200 <----------          551      14                  0
21       ACK ---------->          551      14
22       180 <----------          0        0                   0
23       200 <----------          0        0                   0
24     Pause [Poi(1000ms)]        551                          0
25       BYE ----------> B3       535      12        0
26       180 <----------          0        0                   0
27       200 <---------- E3,4     525      0                   9

28           [ RECVRMT ]          448                0
------ [+|-|*|/]: Adjust rate ---- [q]: Soft exit ---- [p]: Pause traffic -----
```

Fig. 5 A Display of Prototype Demonstration

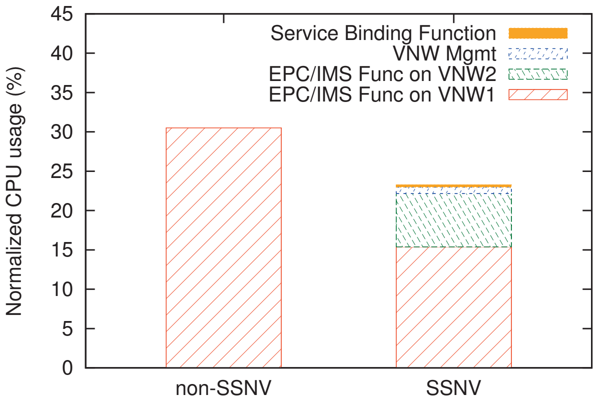


Fig. 6 CPU-utilization reduction through Service-Specific Mobile Network Virtualization

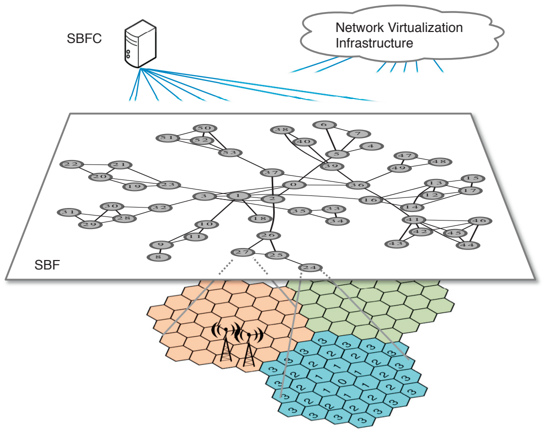


Fig. 7 Configuration of simulation system of redirection-based rule-sharing method

5.2 Effectiveness of redirection-based rule-sharing method

The redirection-based rule-sharing method was proved to have the potentiality of reducing the volume of unknown-packet transfer, by conducting simulations using an event-driven queuing system. Simulations were conducted through assuming such a network configuration as shown in Fig. 7. Several SBFs, interconnected with each other, are located before a network virtualization platform. The SBFs are configured according to a Transit-Stub type topology, which is generated by applying GT-ITM (a topology-generation tool)[16]. eNBs are connected to each of the SBFs; the number of connected eNBs is determined by the I/F capacity of the SBF and the capacity of the eNB. In this simulation, the I/F capacity of the SBF and the number of eNBs are constant (corresponding to a metropolitan area), and then the number of required SBFs is alternated—corresponding to the situation where eNB capacity will grow in future. Each of SBFs is connected to the SBFC. The simulation is conducted in the following way:

INVITE—service-request in SIP—was generated according to the Poisson distribution. An INVITE is processed differently according to its history—whether it is for the same service as the previous service or whether it was transmitted after crossing an SBF domain boundary. The probability of the former event—demoted as P_{same},—is used as a variable parameter for the simulation. The event probability of the latter case was determined using the fluid flow model[17] for emulating a terminal movement. The following shows how the event probability of SBF domain boundary crossing is estimated. As shown in Fig. 7, each SBF domain is composed of n-eNB cells. Hence, the average number per unit time of occurrences of such an event that a terminal crosses the SBF domain boundary, R_d, is expressed,

$$R_d = \frac{v \times L_d}{\pi \times A_d} \cong \frac{v \times (12n+6)l}{\pi \times \{3n(n+1)+1\} \times 2.6l^2} \qquad (1)$$

here,

v: mean terminal-movement rate (m / s)

L_d: SBF domain outer-perimeter (m)

A_d: SBF domain area (m^2)

l: cell radius (m)

A terminal sends a message after crossing a SBF domain boundary at the probability expressed by the equation (2).

$$P_d = \frac{R_d \times N_{UE}}{\lambda} \qquad (2)$$

here,

N_{UE}: the number of terminals;

λ : the sum of the number of service-request messages per unit time transmitted from all the UEs.

As for the time required for service identification by packet inspection, in the pre-experiment, measuring the time consumed for pattern matching using the open-source intrusion detection system, snort 2.9.7.0[18], on 10,000 packets (each packet, with a size of 1,330 bytes, encapsulated by a GTP header) with alternating packet-inspection rules, it was revealed that the service identification time per packet, T_s, is expresses as follows:

$$T_s = T_{msg} + T_d = T_{msg} + 8.0813 \times 10^{-6} \times \left(N_{rules}\right)^{0.4335} \qquad (3)$$

here,

T_{msg}: packet processing time

T_d: time for character string matching

N_{rules}: number of rules

Applying the parameters listed in Table 1, the simulation was conducted. Figure 8 shows the results of the measurement of the arrival rate of unknown packets while varying the number of SBFs; in Fig. 8, P_{same}, is the same by its

Table 1 Simulation Parameters

Parameters		Values				
No. of UEs	N_{UE}	500,000				
No. of eNBs	N_{eNB}	5,000				
Radius of a cell [m]	l	500				
Avg. velocity of UE [m/s]	v	1 (walk)				
No. of rings	n	8	5	4	3	2
No. of SBFs	N_{SBF}	27	54	81	135	270
SBF domain crossing rate [10^{-3}/s]	R_d	0.12	0.18	0.22	0.39	0.63
Sum of the service request rate [/s]	λ	1000				
Probability of the same service as last time	P_{same}	0, 0.2, 0.4, 0.6				
Probability of initial service	–	1				
No. of rules at SBF Controller	–	1,000,000				
No. of rules at SBFs in initial state	–	10,0000				
Link delay [s]	–	0.001				
Time of flow identification at SBF	–	0.000015				
Time of packet processing	T_{msg}	0.0006				
Simulation time [s]	T_{sim}	10000				

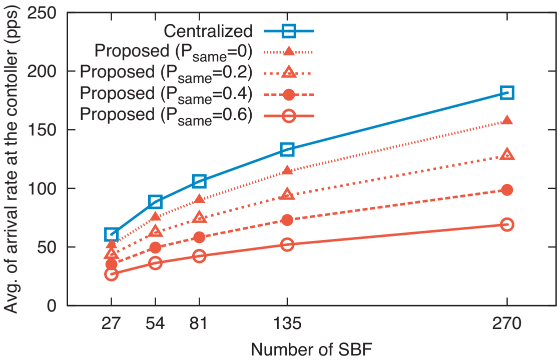


Fig. 8 Unknown packet arrival rate to SBFC

definition as the probability that a redirection destination SBF has the desired rules. As shown in Fig. 8, the reduction of the unknown-packet arrival rate to SBFC was attained by the redirection mechanism. Furthermore, even in such a case where all the redirected unknown packets are transferred to the controller (meaning that such a probability is zero that redirection destination SBFs have desired rules), the arrival rate of unknown packets to the controller was reduced. Such reduction was attained through the mechanism where, as a result of the occurrences of redirections, the SBFs on the route saved the rules in cache, and the probability that an SBF succeeds to identify the service for an unknown packet was increased.

5.3 Challenges for future

It has been introduced, so far, that network virtualization technologies can improve the performance of a mobile network, the mobile network has several virtualized networks—each virtualized network is selectable according to the demanded service and controlled in a service-specific way, and such a configuration, contributing to the reduction of the total network-processing load, enables efficiency of the processing of the mobile network. For further efficiency improvement of mobile networks, one of the challenges will be to enable such a dynamic construction/reconfiguration of virtualized mobile networks that enables prompt/quick allocation or restoration on the concerned virtualized network.

In addition, a method using the element technologies for SDN has been introduced as an efficient rule-sharing method among gateway systems (SBFs). The method is to be evaluated for its efficiency and performance through implementation experiments. However, prior to the implementation experiments, the method must be improved so that its easy implementation on a market-available flow-based switch is ensured.

The research and development activities described in this article have been conducted focusing on sophisticated processing. However, as a hugely growing number of terminals are connected, an increase in processing load will arise as a serious issue, and furthermore, an increase in the volume of information held on the network will become another issue. Hence, for realizing a highly-flexible network in the situations described above, technologies for efficiently managing huge information must be established.

6 Conclusion

How to apply network virtualization technologies, which are the base technologies for the New-Generation Network, to mobile networks was introduced. Network virtualization technologies are expected to enable the provision of "service-communication integrated mobile communication"—in its true sense—through the implementation of service-specific configuration/control. Furthermore, network virtualization technologies are expected, through the building of a virtualized network environment covering multiple mobile networks of different mobile network operators, to realize the stable provision of public critical communication and a 100-percent area coverage for objects, and also enable local governments or corporations to provide efficient network services fine-tuned to the daily lives of local residents. The latter—corporations (which are not mobile network operators) providing mobile network services—seems to have been partially realized through the introduction of the mobile virtual network operator (MNVO) scheme. However, what is currently provided is just network connectivity. There still remain high barriers preventing organizations, which are not mobile network operators, from conducting network controls depending on services—one such obstacle is that the knowledge accumulated through network operation is not open to other operators. For lowering such barriers, public institutions such as NICT should develop and provide a validation environment for the purpose of promoting collaboration and sharing of operating practices among the organizations. NICT is ready to take such a role.

Acknowledgments

The author acknowledges that the research described in this paper would not have been achieved without ample advice and suggestions from the following people and others: Yoshinori KITATSUJI, Group Leader, Mobile Network Group, KDDI R&D Laboratories, Inc.; Hideo MIYAHARA, Professor, Osaka University; Masayuki MURATA, Professor, Osaka University; Makoto IMASE, Vice President, Network Research Headquarters, NICT; Nozomu NISHINAGA, Director of New Generation Network Laboratory; Yozo SHOJI, Research Manager, New Generation Network Laboratory; Kiyohide NAKAUCHI, Senior Researcher, New Generation Network Laboratory. The author expresses here its thanks to all these people.

References

1 P. Agyapong, M. Iwamura, D. Staehle, W. Kiess, and A. Benjebbour, “Design considerations for a 5G network architecture,” IEEE Communications Magazine, Vol.52, No.11, pp.65–75, 2014.

2 M. Ito, K. Nakauchi, Y. Shoji, N. Nishinaga, and Y. Kitatsuji, “Service-Specific Network Virtualization to Reduce Signaling Processing Loads in EPC/IMS,” IEEE Access, Vol.2, pp.1076–1084, 2014.

3 T. Taleb, A. Kunz, “Machine type communications in 3GPP networks: potential, challenges, and solutions,” IEEE Communications Magazine, Vol.50, No.3, pp.178–184, March 2012.

4 I. Widjaja, P. Bosch, and H. La Roche, “Comparison of MME signaling loads for long-term-evolution architectures,” in Proc. IEEE 70th Veh. Technol. Conf. Fall (VTC-Fall) , pp.1–5, Sept. 2009.

5 I. Sato, A. Bouabdallah, and X. Lagrange, “Improving LTE/EPC signaling for sporadic data with a control-plane based transmission procedure,” in Proc. 14th Int. Symp. Wireless Personal Multimedia Commun. (WPMC) , pp.1–5, Oct. 2011.

6 Y. Hong, C. Huang, and J. Yan, “Mitigating SIP overload using a control-theoretic approach,” in Proc. IEEE Global Telecommun. Conf. (GLOBECOM) , pp.1–5, Dec. 2010.

7 H. Tang, “An adaptive paging area selection scheme,” in Proc. 3rd Int. Conf. Commun. Netw. Chin. , pp.1106–1110, Aug. 2008.

8 S1 Application Protocol (S1AP) Release 11, document 3GPP TS 36.413 v11.5.0, 2013.

9 Stream Control Transmission Protocol, IETF Standard RFC 4960, 2007.

10 M. Casado et al., “Rethinking Enterprise Network Control,” IEEE/ACM Transactions on Networking, Vol.17, No.4, pp.1270–1283, 2009.

11 A. Tongaonkar, S. Vasudevan, and R. Sekar, “Fast packet classification for Snort by native compilation of rules,” in Proc. 22nd Conference on Large Installation System Administration, pp.159–165, Nov. 2008.

12 R. Ahmed and R. Boutaba, “Design considerations for managing wide area software defined networks,” IEEE Communications Magazine, Vol.52, No.7, pp.116–123, 2014.

13 (Dec. 18, 2012). OpenEPC Rel. 4. [Online]. Available: http://www.openepc.net/_docs/OpenEPC-Whitepaper_nov2012.pdf

14 (Dec. 17, 2013). UCT IMS Client. [Online]. Available: http://sourceforge.net/projects/uctimsclient.berlios/

15 (Nov. 5, 2013). IMS Bench SIPp. [Online]. Available: http://sipp.sourceforge.net/ims_bench/

16 E. W. Zegura, K. L. Calvert, and S. Bhattacharjee, “How to model an internetwork,” in Proc. IEEE INFOCOM ’96. Conference on Computer Communications, Vol.2, pp.594–602, 1996.

17 W. Wang and I. F. Akyildiz, “Intersystem location update and paging schemes for multitier wireless networks,” in Proc. 6th Annual International Conference on Mobile Computing and Networking - MobiCom ’00, pp.99–109, Aug. 2000.

18 (Feb. 9, 2015). Snort 2.9.7.0. [Online]. Available: https://snort.org

Manabu ITO

Research Expert, New Generation Network Laboratory, Network Research Headquarters
Mobile Network Architecture, Network Virtualization