An Aggregation-based QoS Architecture for Network Mobility

Yaning Wang, Linghang Fan, Nadeem Akthar, Kar Ann Chew and Rahim Tafazolli

***Abstract* — In recent years, extensive studies have been undertaken in order to provide Quality of Service (QoS) for individual mobile nodes in wireless networks. With the development of Network Mobility (NEMO) technology, the provision of QoS for moving networks based on NEMO is emerging as a new research area. However, it is very difficult to directly apply existing QoS technologies to NEMO because of the unique characteristics of such networks. This paper proposes a novel two-level aggregation-based QoS architecture to guarantee QoS for moving networks based on NEMO. In this architecture, flow aggregation and dynamic SLS negotiation are applied at the two levels. Furthermore, a universal signalling protocol is being designed to meet the QoS requirement of NEMO. This signaling protocol can transfer the control information not only between the two levels in the NEMO network but also between the NEMO network and the visited network domain.**

***Index Terms*—QoS, SLS, IntServ, DiffServ, NEMO**

I. Introduction

IN recent years, the fast-developing IP mobile network technology has led to demands for fast and efficient Quality of Service (QoS) provision. On one hand, the mobility of individual nodes is supported by Mobile IPv4 [1] and Mobile IPv6 [2]. With existing QoS mechanisms, e.g., Integrated Services (IntServ) and Differentiated Services (DiffServ), many proposals [4][5] have been worked out to provide QoS for individual mobile nodes in terms of low delay, low jitter, low loss rate and high bandwidth. On the other hand, the IETF working group NEMO [3] is undertaking research on the mobility of IP moving networks, in which the whole network moves as a single unit. However, little research has been done in the area of providing QoS for Internet moving networks based on NEMO.

This paper will analyze the difficulties faced by moving networks when providing QoS, and then propose a two-level aggregation-based QoS architecture to provide QoS for NEMO network. This architecture is based on modified DiffServ mechanism. A signalling protocol is also proposed to exchange information at both levels (i.e. node-level and network-level) as well as between the NEMO network and the visited network domain.

The rest of this paper is arranged as follows. Section II introduces related work in NEMO area. Section III analyses the advantages and disadvantages of existing QoS mechanisms working on mobile nodes and moving networks. Section IV proposes the two level structure of QoS provision for NEMO network. Section V proposes the signalling protocol for SLS creation and update inside and outside the NEMO subnet. Section VI draws a conclusion for this paper.

II. Related Work On NEMO

The IETF NEMO Working Group is investigating extensions to Mobile IP to support network mobility scenarios where routers and all nodes move together as a single unit.

Although Mobile IP can handle node mobility well, it does not explicitly address the need for network mobility when an entire network, made of one or several mobile routers (MR) and attached nodes, moves as a single unit. The objective of a network mobility solution is to allow all nodes in the mobile network to be reachable via their permanent IP addresses, as well as maintain ongoing sessions when the mobile router changes its point of attachment within the Internet [6].

The basic approach to network mobility (NEMO) support is built on Mobile IPv6 with minimal extensions. Similar to Mobile IPv6, the MR of NEMO has a home agent (HA), and uses bi-directional MR-HA tunneling between the MR and HA to preserve session continuity of the nodes within its mobile network while the MR (with all the mobile nodes associated to it) moves.

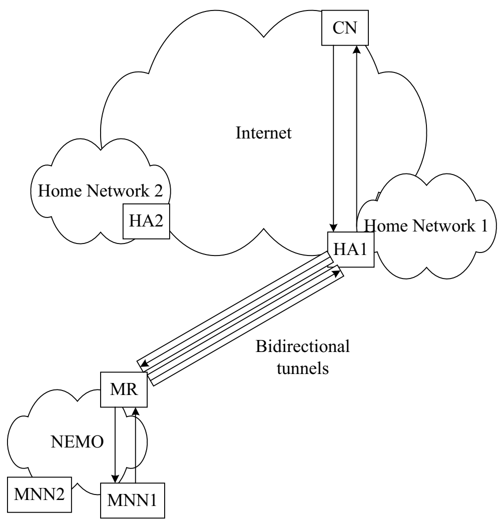

Fig. 1: Data transmission for local MNN of NEMO.

When moving to a foreign network, the MR binds its home address (of its network interface to the access network) with a care-of address by means of binding updates to the HA in its home network (Home Network 1 in Fig. 1). The communication between MR and HA is explained as follows: When flows are transferred from MR to its HA, the MR needs to encapsulate the packets and send them in the MR-HA tunnel to the HA, which will then decapsulate them and forward them to the destinations. In the opposite direction, the HA needs to intercept all the packets intended to reach the MR and tunnel them to the MR, which will then decapsulate them and forward them to the mobile node inside its network. To recognize the mobile network nodes (MNNs) inside a NEMO network, HA needs to know the network prefixes of the NEMO subnet and set up a mapping between the MR's care-of address and the prefixes of the NEMO subnet. Therefore, when receiving packets destined to a MNN of the NEMO subnet, HA can successfully forward the packets to MR.

There are two classes of MNNs: *local nodes* (i.e. local MNN) that have the same home network with the MR, and *visiting nodes* whose home networks are other than MR's home network. Shown in Fig. 1, the *local node* MNN1 has the same home network with MR, namely Home Network 1, while the *visiting node* MNN2 has Home Network 2 as its home network and HA2 as its home agent.

The end-to-end data transmission between a *local node* MNN and its Correspondent Node (CN) is straightforward. In Fig. 1, HA1 will directly forward the flows created by MNN1 to the destination CN and the CN would also send the flows destining to MNN1 directly to HA1. In this case, all the packets need to be encapsulated once during the transmission.

If the MNN is a *visiting node*, the situation is much more complicated. Shown as Fig. 2, if MNN2 needs to communicate with its CN, according to MIPv6 and NEMO design, there would be a bidirectional tunnel between MNN2 and HA2; meanwhile, there is a bidirectional tunnel between MR and HA1. Thus, there are two tunnels that exist through the path between MR and HA1. The multiple tunnels lead to long delay and large number of signalling messages, which makes it difficult to apply existing QoS mechanisms to NEMO.

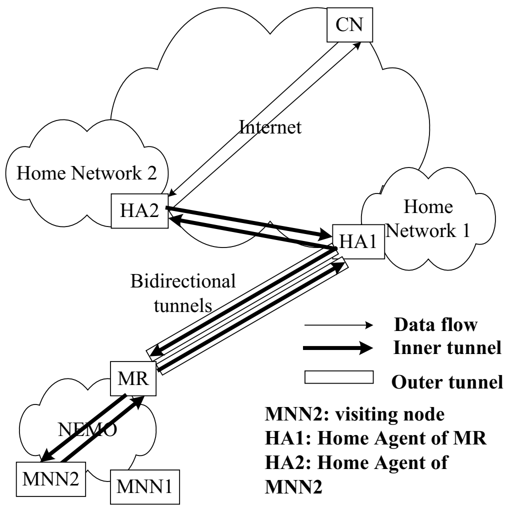

Fig. 2: Data transmission for visiting MNN of NEMO.

III. Applying existing QoS Methods to NEMO

As NEMO proposes a new mechanism to support mobility and communication for moving networks, QoS support for the NEMO networks is becoming an important research issue. In this section, the pros and cons of directly applying existing QoS mechanisms (e.g. IntServ and DiffServ) to NEMO are discussed.

A. IntServ for NEMO

To apply IntServ scheme in NEMO, we use RSVP as the signalling protocol, and the following problems arise.

1) Tunnel complexity

Because of the tunnels existing in NEMO, normal RSVP messages will go through the two ends of the tunnel without informing the routers along the tunnel. RSVP for tunnel [7] can be used to fix this problem by resending the request message to the routers between the two end points of the tunnel, but this leads to the increase of messages and delay.

2) Scalability problem

In IntServ scheme, resources are reserved for each single flow, and each intermediate router has to do the admission control. This will require significant memory and processing power to store and maintain the session information in the intermediate routers. Moreover, the complicated process can lead to long delay.

3) Limit of wireless bandwidth

In the NEMO architecture, all the reservation messages must transit the link between MR and the access network. With higher number of MNNs, the limited bandwidth of the wireless link will be occupied, creating serious network congestion. RSVP aggregation [8] can be used between MR and its HA to alleviate this problem. However, only a limited number of flows can be aggregated, and it is actually not an IntServ solution but using DiffServ in the aggregated segment of the path.

B. DiffServ for NEMO

Since a NEMO subnet is a group of nodes moving together, it is efficient and simple to regard the subnet as a DS domain. As the node sending and receiving all control and data packets for other nodes, MR is suitable to act as the administrator of the subnet. In this way, dynamic SLAs could be negotiated between MR and the administrator of the visited domain.

A typical way to negotiate dynamic SLA is proposed in [9]. In this mechanism, all the SLA negotiations and updates are triggered by the request of MNN. This per-flow-triggered DiffServ mechanism can cause the same problems as IntServ.

1) Scalability Problem

Like IntServ, the dynamic SLA negotiation is also triggered by a new flow of the MNN. Many flows may be created by the MNNs in a NEMO subnet in a short interval, so the SLA will be updated at a very high rate. This imposes high demand on memory and processing power.

2) Limit of wireless bandwidth

Because of the frequent negotiation and update of SLA, the number of signalling messages is so large that the limited wireless link would be occupied by the signals. Thus, congestion may happen in the wireless link.

To resolve the problems mentioned above, an architecture that can alleviate the scalability problem and a new signalling protocol that can fast and efficiently exchange information are proposed in the following sections.

IV. A New Architecture Providing QoS for NEMO

To improve the scalability and reduce the number of signalling messages, a two-level aggregation-based QoS architecture is proposed to provide QoS for NEMO. The first level is *node-level* on individual MNNs. Each MNN collects the QoS requirements of all flows it created and aggregates them into one request to MR to ask for resources for all flows. The second level is *network-level* lying at the MR. MR collects the QoS requests from its member MNNs, aggregates them into a single SLS request for the entire NEMO subnet, and sends the request to the resource manager of the visited domain. The architecture is shown as Fig. 3.

A. Node-Level Aggregation

In the NEMO network, each MNN can perform aggregation for different flows created by itself and send a single QoS request to the MR to ask for an aggregate network resource.

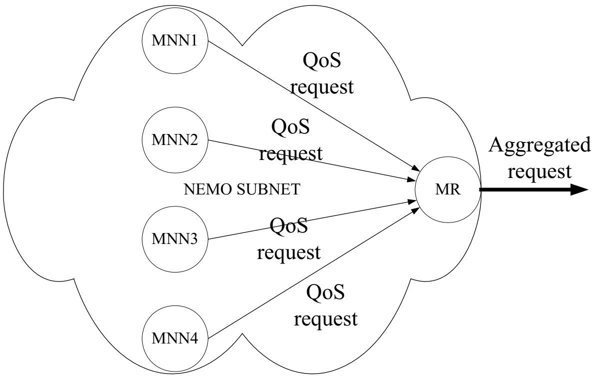

Fig. 3: Two-level aggregation-based QoS architecture.

Each request sent by a MNN is not for a single flow but for all the existing flows of the MNN. The main purpose of this approach is not only to improve the scalability but also to utilize the network resources more efficiently.

With node-level QoS aggregation, the requests can be created on a per-MNN basis instead of on per-flow basis. MNN can request resources once for several flows or applications and then distribute the resources to these flows.

B. Network-Level Aggregation

Besides the node-level aggregation, at the higher level of the proposed architecture, MR plays the aggregating part for the whole NEMO subnet. MR should collect the QoS information of all member MNNs inside the subnet and aggregate different QoS requests into one combined request, as shown in Fig. 3. In this way, the number of QoS requests will be reduced remarkably and the scalability of QoS protocol will be improved.

As both MR and MNN provide QoS for flow aggregates, a universal signalling protocol will be designed to carry the requests from both the node level and the network level.

V. Signalling Protocol Design

Based on the proposed QoS architecture, a universal signalling protocol is proposed to exchange Service Level Specification (SLS) [12] between MNNs and MR and between MR and the visited domain.

SLS is a set of parameters and their values that together define the service offered to a traffic stream by a DiffServ domain. A SLS can be regarded as the technical version of an SLA. An SLS encompasses the specification of how traffic meeting certain conditions and arriving from a peer network on a certain interface will be treated. In the proposed QoS architecture, SLS is introduced to carry QoS information for the traffic aggregates. This proposed signalling protocol will focus on SLS negotiation, which is the technical part of SLA.

From the view point of NEMO network, the SLS can be created at the node level, namely between MNN and MR, or at the network level, namely between MR and the visited network domain. In the former situation, MR acts as the provider and MNN acts as the customer; in the latter situation, the administrator of the visited network is the provider, while MR is customer. In order to keep the simplicity and efficiency of the architecture, we design a signalling protocol based on the abstract concepts of provider and customer, working at both of the aggregation levels.

In this section, the SLS classes and parameters for NEMO QoS are first proposed, and then signalling procedures are specified.

A. SLS Classes

SLS should include several classes in which different services are provided. Some different mechanisms have been proposed, e.g., UMTS QoS classes [11]. The different Internet services can be categorized into classes shown in Table 1.

Expedited forwarding (EF), also called premium services, is used for the real-time and interactive services such as VoIP. Assured forwarding is also used but not all the four classes defined in [10] are included. Since DiffServ AF does not define a clear range of the classes, it is one of the future tasks to define

TABLE I SLS Classes

Traffic class	Characteristics	Example
EF	Jitter sensitive, strictly real-time, conversational pattern	VoIP
AF Class 3	Real-time and Highly interactive	Video streaming
AF Class 2	Not real-time but time dependent, interactive	Online reservation
AF Class 1	Low time dependent, not interactive	Web browsing
Best Effort	Background traffic, no interactivity, not real-time	FTP service

the limits of QoS parameters for each AF class. All the services that are time dependent and interactive are classified into AF classes, while time independent services or the so-called background services are served by best-effort class, which does not ensure quality of service, as its name indicates.

B. SLS Parameters

Each SLS class is identified by several QoS parameters. The exact values of the parameters need to be negotiated between the two network entities. Based on DiffServ, the QoS parameters included in SLS are *maximum end-to-end delay, maximum end-to-end jitter, packet-loss rate and bandwidth*. In the negotiation process, MNN should map the individual flow parameters to the parameter values of node-level aggregates while the MR should map the node-level aggregate requests to the network-level request.

C. Signalling Procedures and Flows

An entire suite of procedures (i.e. SLS creation, update and termination) has been specified in the signalling protocol. These procedures are described as follows.

1) SLS creation

The creation of SLS starts after the handoff process. If a NEMO subnet handoffs to a visited network domain, a SLS should be negotiated between the MR and the administrator of the domain. Similarly, if a MNN joins a NEMO subnet, a SLS should be negotiated between the MNN and the MR. After detecting a new customer, the provider sends an advertisement to the customer's Care-of Address (CoA) to announce the available resources that can be guaranteed to the customer. Receiving the advertisement, the customer decides the resources it wants based on the existing or future traffic created and sends out an aggregate QoS request to the provider. The provider checks the request, and if it is conformant to the available guarantee, the provider will send an acknowledgment to accept the request to finish the SLS negotiation, otherwise a denial will be sent out and the creation process should repeat. The procedure is shown clearly in Fig. 4.

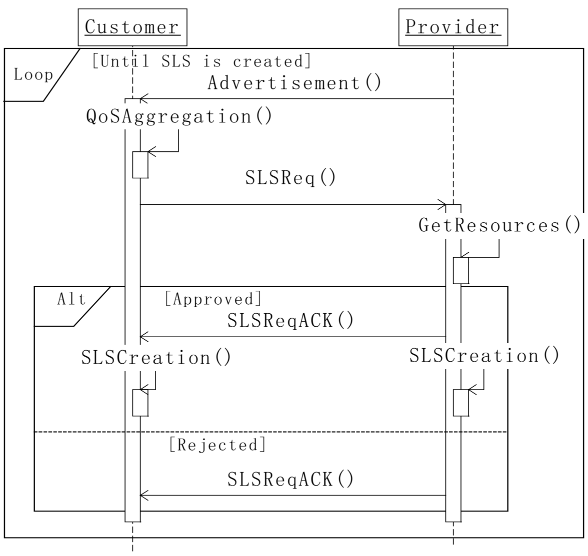

Fig. 4. Signal flow of SLS creation

The SLS creation and negotiation is optional and not mandatory for a customer. If the customer does not want to set up a QoS association with the provider, it can simply ignore the advertisement. The SLS agreement does not have to be established only when there is an existing traffic. A customer can negotiate the SLS with the provider in advance of the possible traffic and the requested resources can be invoked and released in following SLS update procedure.

2) SLS update

In order to provide resources to the customer efficiently at any time, the SLS should be updated dynamically with the change of resources requested by the customer.

The update process can be triggered by changes of the MNN membership, e.g. a new visiting MNN, or changes of the QoS requirements, e.g. the creation of a new flow. These events can cause too frequent SLS updates. To control the frequency of updates and optimize the network performance, factors, such as timers, thresholds and priorities, will be considered in the algorithm design to trigger aggregated update.

The update procedure is shown as Fig. 5. The customer sends an UPDATE message to the provider, noticing the new values of the SLS. If the request is approved, the provider will re-configure the SLS parameters and allocate consequent

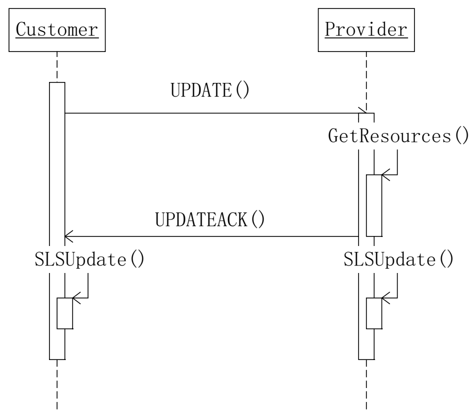

Fig. 5. Signal flow of SLS update

resources to the customer. Finally, the provider returns an ACK message to confirm the update. If the request is rejected, the UPDATEACK message will be returned with error information.

If we consider the frequency-controlled mechanism to be realized, an example of update can be shown in Fig. 6. Both MNN and MR maintain a frequency control scheme based on the algorithms determining the timer periods and thresholds as well as priorities. So MNN and MR will not send SLS update request to their providers as soon as a triggering event happen (e.g. the first new flow and the SLS creation request from the new member MNN). When the second new flow is created, the condition of frequency control schemes is reached, and then MNN sends the UPDATE message to its provider (i.e. MR).

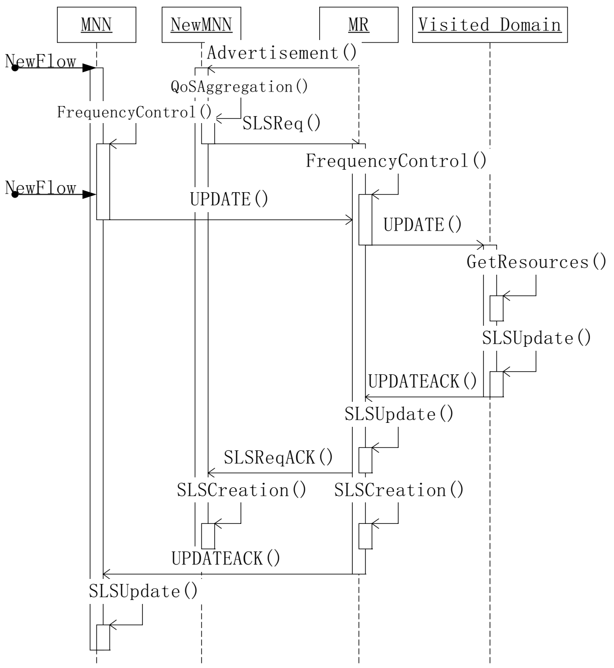

Fig. 6: An example of frequency-controlled SLS update

Receiving the UPDATE message, the condition of the MR frequency control schemes is reached, so MR creates an UPDATE message to the administrator of the visited network domain. If receiving a positive UPDATEACK message, MR distributes the reserved resources to its two member MNNs and finishes the update process.

3) SLS termination

When the NEMO subnet leaves the visited network domain or the MNN leaves the NEMO subnet, the reserved resources should be released by terminating the relevant SLSs. In order to utilize the network resources efficiently, the leaving NEMO subnet or MNN will eliminate the SLS information and send a SLS termination message to its provider when they detect a coming handoff. As shown in Fig.7, upon receiving the message, the provider should remove its SLS information and release the resources. As a back-up scheme, the provider will end the SLS without hearing the customer for a long time.

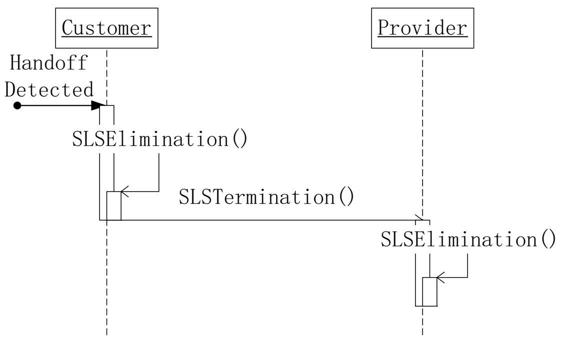

Fig. 7. Signal flow of SLS termination

VI. Conclusions

This paper explored the QoS provision for moving networks. The problems that appear when existing QoS mechanisms are applied in NEMO network were discussed. A novel two-level aggregation-based QoS architecture was proposed as a possible solution to provide QoS for NEMO network. QoS aggregation and SLS negotiation are introduced in both node-level and network-level in order to efficiently manage the QoS. A framework of signalling protocol was specified, not only between the two levels inside the NEMO network, but also between the NEMO network and the visited domain, which paved the way to future performance evaluation.

In the future, the signalling protocol needs to be implemented in more detail, and the performance of the proposed QoS architecture and its related signaling protocol will be evaluated via analytical and simulation means.

Acknowledgment

Authors would like to thank Dr. Dan He for valuable comments and advice. We also thank colleagues in the Ambient Networks project for many ideas and creative discussion.

References

[1] C. Perkins, "IP Mobility Support for IPv4," *IETF RFC 3344*, August 2002

[2] D. Johnson, C. Perkins, J. Arkko, "Mobility Support in IPv6," *IETF RFC 3775*, June 2004

[3] Vijay Devarapalli, et al, "Network Mobility (NEMO) Basic Support Protocol," *draft-ietf-nemo-basic-support-02.txt, IETF draft*, December 2003

[4] Bongkyo Moon, Aghvami, H., "Diffserv extensions for QoS provisioning in IP mobility environments," *Wireless Communications, IEEE,* Volume: 10, Issue: 5, Oct 2003, Pages:38 – 44

[5] Shou-Chih Lo, et al, "Architecture for Mobility and QoS Support in All-IP Wireless Networks," *IEEE Journal on Selected Areas in Communications*, Vol. 22, No. 4, May 2004

[6] Hong-Yon Lach, "Network Mobility in Beyond-3G Systems", *Communications Magazine*, IEEE, Volume 41, Issue: 7, pp.52-57, July 2003

[7] A. Terzis, J. Krawczyk, J. Wroclawski, and L. Zhang, "RSVP Operation Over IP Tunnels," *IETF RFC2746*, Jan. 2000.

[8] F. Baker, C. Iturralde, F. L. Faucheur, and B. Davie, "Aggregation of RSVP for IPv4 and IPv6 Reservations," *IETF RFC 3175*, Sept. 2001.

[9] Xipeng Xiao, et al., "Internet QoS: A big Picture," *IEEE Network*, March/April 1999, pp. 8-18

[10] J. Heinanen, et al., "Assured Forwarding PHB Group," *IETF RFC 2597*, June 1999

[11] "Quality of Service (QoS) concept and architecture," *3GPP Technical Specification 23.107, v6.1.0*

[12] D. Grossman, "New Terminology and Clarifications for Diffserv," *IETF RFC 3260,* Apr. 2002.