Abstract - This paper presents the design principles of Dynamic Service Negotiation Protocol (DSNP). DSNP is a protocol to negotiate the SLS (Service Level Specification) in IP layer. It can be used for service negotiation from host to network, network to host, and network to network. The automated negotiation makes service negotiation efficient in terms of time, cost, and correctness. The dynamic negotiation not only allows users to adapt their needs dynamically, but also let providers better utilize the network. DSNP can be used in both wireline and wireless networks. It is, however, particularly useful in mobile environment. To demonstrate the usefulness of DSNP, a reference wireless QoS architecture based on DiffServ is presented. Exemplary applications and experimental results are illustrated.

1. INTRODUCTION

Today, many different wireless systems exist, ranging from PANs (Personal Area Networks), wireless LANs (Local Area Networks) to outdoor cellular systems. They are typically not compatible with each other, making it difficult to roam from one system to another. PANs, Wireless LANs, and cellular wireless systems are also being developed and are evolving independently (e.g., from 3G to 4G, 802.11 to 802.11a/802.11b, HIPERLAN to HIPER-LAN II, etc.). Although ITU IMT-2000 has been trying to unify third-generation (3G) wireless systems, incompatible systems are expected to co-exist in the future. No wireless technology has emerged as a common and long-term universal solution.

IP (Internet Protocol), which is already a universal network-layer protocol for wireline packet networks, is becoming a promising universal network-layer protocol over all wireless systems. An IP device, with multiple radio interfaces or software radio, could roam between different wireless systems if they all support IP as a common network layer. Unlike today’s radio systems that continue to depend heavily on proprietary technologies, IP provides a globally successful open infrastructure for services and applications. Such an all-IP wireless and wireline network could also make wireless networks more robust, scalable, and cost-effective.

A key challenge in realizing the all-IP wireless networks is how to guarantee Quality of Service (QoS). To guarantee QoS on the Internet, Int-Serv (Integrated Services) [1] and Diff-Serv (Differentiated Services) [2], which differ in the technique for resource provisioning and the granularity of service differentiation, have been proposed. Both approaches, however, are limited to stationary hosts and cannot be applied to the mobile environment directly.

Today the service level agreement (SLA) is usually agreed, either verbally or in writing, by both the user and the service provider when a user signs up for a service. The service provider stores it in some repository and uses it to condition the traffic sending to/from the user. To change the SLA, normally a user has to contact and negotiate with the authority of the service provider, which then manually changes it. Usually service provider allows this kind of re-negotiation or changes only in a large time scale, for instance, once per year.

It is expected that a user will use different terminals with different capabilities in different situation. For example, PC may be used at home or inside an office. While driving, small handset might be more suitable. PDA or laptop will be used when traveling. They are different not only in size, but also in processing and communication capabilities. Different applications will also be used in different terminals. They generally require different bandwidth for network transmission.

As stated above, mobility and the diversity in transmission media (Bluetooth, 802.11, cellular, etc.) and user terminals (PC, laptop, PDA, etc.) create a very dynamic environment. It is hardly for a service provider to plan uniform resource in all networks and to envision fixed bandwidth requirement from all users. Users are also hardly to project what they really need due to mobility and the difference in terminal they may carry. In addition, users may roam to other service providers and still wish to enjoy the same QoS as they had in the home provider. It is desirable that there is a way to negotiate the service dynamically. This dynamic service negotiation should be automated and should be able to do in a small time scale in opposition to today’s manual negotiation in a large time scale. A user should be able to negotiate with both home and visiting service providers dynamically. Similarly, the service provider can also negotiate with a user depending on the resource available. A service provider may advertise unused resource if the resource is underutilized. On the other hand, a service provider can negotiate with users to lower their service grade if the network is over-provisioned. While the user is roaming, the home and visiting providers should also be able to negotiate with each other to decide the service which can be offered to the user. There should be a standard protocol which can be used for service negotiation for multi-vendors and multi-service providers.

DSNP (Dynamic Service Negotiation Protocol) is a protocol proposed to dynamically negotiate the SLS (Service Level Specification) [3] in IP layer. A SLS is a set of parameters and their values which together define the service offered to a traffic stream by a DS (differentiated services) domain. DSNP can be used for service negotiation from host to network, network to host, and network to network. The automated negotiation makes service negotiation efficient in terms of time, cost, and correctness, etc.

Section 2 briefly reviews DSNP. Section 3 describes the DSNP messages. To demonstrate the usefulness of DSNP, Section 4 presents a reference QoS architecture. Section 5 discusses the traffic conditioning in wireless DiffServ. Section 6 illustrates some exemplary applications. Section 7 presents the tested and experimental results. Section 8 summarizes the papers.
2. DSNP OVERVIEW

DSNP can be used in both wireline and wireless networks. It is, however, particularly useful in mobile environment. For example, a mobile user may roam to a new service provider which does not have any contract with either the mobile user or its old service provider. The inter-domain negotiation might be necessary in order to get network service. Even though the old and new providers have certain level of service agreement, the new service provider may still need to negotiate with the old service provider. When roaming inside the same domain, the following are motivation to support dynamic intra-domain negotiation:

1. A user may carry a different terminal at different time to access the network. The capabilities of these terminals may be different, thus the network resource requirements may be different too.
2. A user may roam to networks with different physical and link layers, for instance, from Bluetooth to IEEE 802.11 to IS-95 (or other outdoor cellular systems). The available resource typically are different in different type of networks.
3. Due to mobility, the provisioning of network resource may not be accurate for actual demand. For example, a special event in a city may gather many unexpected network users.

Dynamic service negotiation not only allows users to adapt their needs dynamically, but also let the service provider better utilize the network.

Service negotiation may involve human. If so, some applications may be needed. The user or the service provider may also pre-define and store their policy so the negotiation can be done without human interactions. In either case, DSNP is a protocol for hosts and networks to negotiate SLS in IP layer. DSNP can be used in any architecture frameworks. It is independent of network architecture, and how resource reservation or provisioning is done. DSNP, however, is particularly useful in Diffserv. Diffserv is built on the concept of classifying packets and keeping per-customer state at the network edge and letting the core deal with aggregates of traffic. In operation, routers use DS byte to differentiate traffic flows belonging to different service classes. The edge routers perform conditioning functions to keep traffic “in profile” with the TCS (Traffic Conditioning Specification).

In order to condition the traffic properly, the edge router needs to know the QoS profile of a user. The changes in SLS should be known by the necessary edge routers (ERs). In mobile environment, there should be an efficient way to distribute mobile’s QoS profile to possible ERs. It is also desirable to reduce QoS-related signaling messages for every handoff so fast handoff can be achieved.

Next section lists the DSNP messages. A QoS architecture then will be presented to demonstrate how DSNP can be used for SLS negotiation and traffic distribution so mobile users can perform fast handoff while also can guarantee the same level of QoS.

3. DSNP MESSAGES

This section explains the various messages used in DSNP. An host to network negotiation scenario is assumed. A host acts as the DSNP client and a service provider’s QoS authority acts as the DSNP server.

**SLS_LIST_REQUEST:** This message is sent by a DSNP client to the DSNP server, to request for a list of SLSs offered by the DSNP server. A DSNP client may send this message when it has just booted up and does not have SLS parameters.

**SLS_LIST_RESPONSE:** This message is sent by the DSNP server in response to the SLS_LIST_REQUEST message. This message lists all the SLSs that are provided by the DSNP server. The cost and the time of availability for each service may also be included in the list.

**SLS_NEGO_REQUEST:** This message is usually sent by an DSNP client to the DSNP server, to request for a particular SLS. The requested SLS could either be customized or one of those listed in the SLS_LIST_RESPONSE message. This message is used for both requesting a new SLS as well as updating an existing one. The DSNP server can also send this message to the hosts under some circumstances. For example, if network resources become scarce, the DSNP server sends this message to the hosts that have a SLS with the DSNP server requesting them to update their existing SLS to suit the current network conditions. The DSNP server could offer cost incentives to the hosts that accept the suggested SLS. Similarly, when there are unused resources available, the DSNP server could offer them at a lower price to the DSNP clients. It could do an advertisement by sending out SLS_NEGO_REQUEST messages with the available SLSs and the cost. Also, if the DSNP server wants to forcefully terminate a SLS of an DSNP client due to some reason, it sends a SLS_NEGO_REQUEST message to the DSNP client with appropriate fields set to ZERO.

**SLS_NEGO_RESPONSE:** This message is sent in response to the SLS_NEGO_REQUEST. This message indicates whether the requested SLS is accepted or not. If the requested SLS is not accepted, then the reason for not accepting is also provided. For example, if the DSNP server does not accept the SLS of an DSNP client due to lack of resources, it sends back a response indicating a reject along with the maximum SLS that could be supported.

**SLS_STAT_REQUEST:** This message is sent by a DSNP client to the DSNP server asking for a feedback on the statistics of the actual usage of each SLS. The DSNP server collects the necessary information and sends it to the requested DSNP client.

The same messages could also be used in an inter-domain negotiation between two service providers. For instance, one service provider requests for some service (and hence acts as the DSNP client) and the other provides the requested service (and hence acts the DSNP server). The nature of interaction remains the same. For detailed DSNP packet formats, please refer to [4].

4. REFERENCE QOS ARCHITECTURE

Although DSNP is independent of QoS architecture, this paper uses the ITSUMO QoS architecture as a reference architecture. The ITSUMO1(Internet Technologies Supporting Universal Mobile Operation) [5] is a research project that focuses on the design of next generation wireless IP networks. It envisions an end-to-end wireless/wireline IP platform for supporting real-time and non-real-time multimedia services in the future. Its goal is to use IP and next generation wireless technologies to design a wireless platform that allows mobile users to access all services on a next generation Internet.

---

1ITSUMO is a trademark of Telcordia Technologies, Inc. It means “all the time, anytime” in Japanese.
In the ITSUMO QoS architecture, there is at least one logical global server and several local nodes in each administrative domain, which is essentially a DS domain. The server is referred to as the QoS Global Server (or QGS), and local nodes are referred to as QoS Local Nodes (or QLN). The architecture is depicted in Fig. 1. In addition to other necessary components in the system such as DHCP [6]/DRCP [7] server and AAA (authentication, authorization, and accounting) server, there are three major QoS components:

MS (mobile station): MS is the device that allows users to communicate, and also provides means of interaction between users and the networks. Traffic is generated/received by MS and may be queued in the MS while waiting for transmission/reception.

QGS (QoS global server): As shown in Fig. 1, there is one logical QGS in each administrative domain. The QGS has the global information of the resource available in the whole domain. Essentially, it is a bandwidth broker (BB) with extension for wireless networks. The MS interacts with the QGS when the MS requests certain degree of QoS in the domain. The QGS is the entity for QoS negotiation and signaling between MS and the network control system, i.e. it is in control plane, as that of MGC (Media Gateway Controller) in MEGACO [8]. The QGS decides what services are available for each MS and sends the decision to related QLNs. Thus, the QGS is an intelligent entity for decision making. It is similar to PDP (Policy Decision Point) in Policy Framework [9].

QLN (QoS local node): QLN is the edge router residing in the boundary of the DS domain. Fig. 1 depicts that QLN could be part of edge router, or could reside in a component inside RAN (radio access network) such as BS (base station). QLN is like a wireless bandwidth broker which retains the local information of the resource in the local domain. However QLN does not interact with MS directly for QoS negotiation and signaling. In stead, this local information is provided to the QGS. QLN then performs traffic conditioning according to the instruction from the QGS. Therefore, it functions like PEP (Policy Enforcement Point) in Policy Framework [9]. In contrast to QGS, QLN handles the actual traffic thus it is in transport plane, similar to MG (Media Gateway) in MEGACO [8].

The QGSs in domain 1 and domain 2 may contact with each other directly, or through a public QGS which may attach to the Global IP Network in Fig. 1. Since QGS retains the global information of the whole administrative domain, dynamic service negotiation can be achieved easily and efficiently. The MS only needs to negotiate with the QGS, which makes the decision based on the up-to-date global information. Once it is done, the QGS will instruct the related QLNs how to condition the MS’s traffic. Therefore, the MS can move freely inside the domain. How does the QGS allocate resource and reach the decision can be done in many different ways which have been proposed in literature, for example [10]. By separating control and transport, the architecture is also flexible, easy to add new services, easy to integrate with other network components, and easy to interoperate with legacy networks.

5. DISTRIBUTION OF QOS PROFILE AND TRAFFIC CONDITIONING

In wired network, it is easy to locate a user, therefore it is easy to locate the edge router that will need to condition the traffic for a specific user. In wireless networks, however, users can roam anywhere. Thus potentially all edge routers will need to know the QoS profile of a users. One straightforward solution is to let all edge routers in the domain maintain the QoS profiles of all users. Each time when service negotiation is done, the new SLS is broadcast to all edge routers. It is, however, inefficient to maintain the same copy of data in all edge routers. It causes unnecessary broadcast too. If the number of edge routers in the domain is huge, the data are replicated unnecessarily in many places. The database in each edge router will be huge too if there are many users in the domain. In addition, once a MS moves or changes its SLS, the same transaction for updating the database must be performed for all edge routers. Many edge routers may never have traffic coming from or going to the MS.

It is desirable to maintain the QoS profiles of all users in the domain only in a central repository. Edge routers only keep the necessary data without maintaining the QoS profiles of all users in the domain. Referring to the QoS architecture presented in Sec-
tion 4, the QGS of the domain retains the database of all users. Each time when the service negotiation is done, the QGS sends the new SLS of the user to potential QLNs only. The potential QLNs may be the neighboring QLNs of current serving QLN. Each time when the MS moves, the QGS can select the new set of potential QLNs, as that the QGS maintains the global information and is the decision maker. Therefore, the QoS profile of the MS can be distributed to the new QLN before the MS moves. Since everything is done by the network, the MS does not need to perform QoS-related signaling once the initial negotiation is done. It reduces the amount of QoS signaling messages, conserves MS’s energy, and achieves fast handoff.

6. EXEMPLARY APPLICATIONS

6.1. Initial QoS Negotiation

When a MS is powered up, first it may need to perform registration and configuration with DHCP/DRCP to get an IP address. Before the MS sends actual traffic, it initiates the QoS signaling with the QGS if there is no service agreement/specification or the MS wants to renegotiate it. The QGS may need to consult with AAA or other servers if necessary. Based on the interaction with other servers, the global information in QGS, service agreement, and other information such as mobility pattern, etc., the QGS will either admit or reject the QoS request. If the request is accepted, the SLS for the MS will be multicast to the potential QLNs. As discussed in Section 5, the potential set of QLNs may include current serving QLN and its neighboring QLNs.

Fig. 2 shows an example in that the MS uses DHCP to get an IP address for the subnet in the initial set-up. It then makes a QoS request for the list of available SLS. Based on the response, the MS negotiates with the QGS for the SLSs it wants. The QGS consults with AAA server, then makes the decision. Assuming that the QGS decides to offers certain kind of services to the MS. It sends the decision to the related QLNs so they can condition the MS’s traffic accordingly. COPS [11] or SNMP [12] can be used for the communication between QGS and QLNs. After receiving the SLS_NEGO_RESPONSE, the MS can send its actual traffic in Fig. 2. QLNi indicates the potential set of QLNs.

6.2. Client Moves within the Same Domain

When the MS is roaming inside the same administrative domain, i.e., the domain serving by the same QGS, the MS only needs to get a new IP address if changing subnet. It does not need to have any QoS signaling since the decision made by the QGS has been sent to all potential QLNs. As discussed in Section 5, the set of potential QLNs may be changed dynamically while the MS is moving. Thus the MS can transmit/receive traffic without initiating new QoS signaling while it is moving within the same administrative domain. Fig. 3 is an example flow for moving within the same domain but the subnet is changed.

6.3. Client Renegotiates SLS within the Same Domain

Once the MS is up and the QoS negotiation is done, the MS is free to move within the same domain without any QoS signaling. However, the MS may want to change the SLS and renegotiate with the network for the service level. Fig. 4 plots an example flow for this purpose. It is similar to Fig. 2 except that the MS has the IP address and the list of SLS already.

6.4. Client Moves into a New Domain

When the MS moves to a new administrative domain, it must initiate the QoS signaling with the new QGS. The new QGS may consult with the new AAA server, old AAA server, and the old QGS to decide whether it should admit or reject the QoS request. After that, it is similar to what described above.

7. TESTBED AND EXPERIMENTS

The ITSUMO testbed consists of a number of laptops, base stations, and routers. The radio layer transport is carried over an
IEEE 802.11 complaint WaveLAN system. Each MS also equips with a digital camera. The indoor testbed is composed of three distinct wireless subnets. The OS running is Linux.

An initial implementation of the protocol has been developed to perform experiments. Fig. 5 shows the experimental result of a real-time video conferencing application. Initially, both MS and CH are in the same subnet. The MS negotiates with the QGS by using DSNP for a video rate of 200 Kbps. Once the request is admitted, the MS roams away from the CH. The MS changes subnet twice therefore there are two IP handoffs which are performed by Mobile IP. The MS then negotiates with the QGS by DSNP again for a video rate of 200 Kbps. After it is accepted, the MS roams back to the same subnet of the CH. There are two IP handoffs on the way back too. Fig. 5 indicates that there are two major different video rates: 200 Kbps and 100 Kbps. It also indicates that the similar rate can be maintained after handoffs. During handoff, however, the video rate is dropped then reached a peak point. This is because packets are stopped when Mobile IP performs IP handoff. Excess packets pass through QLN immediately after the IP handoff is done.

Instead of real application, Fig. 6 shows the result from three traffic generators. Initially, only the TCP best-effort traffic is sending, which retains almost all of the 1.2 Mbps bandwidth. However, once QoS type I starts with 600 Kbps, the bandwidth used for best-effort traffic is dropped. The bandwidth for best-effort drops again when QoS type II stops, best-effort traffic regains the bandwidth. Around 800 sec. in the time axis, the QoS type I negotiates with the QGS by DSNP and drops its bit rate to 400 Kbps. Therefore, the best-effort traffic gets more. After QoS type I closes the session, the best-effort traffic returns to its original bit rate.

8. SUMMARY

This paper presents the motivation of DSNP and its design principles. Dynamic service negotiation not only allows users to adapt their needs dynamically, but also let the service provider better utilize the network. Although DSNP is a generic protocol for SLS negotiation, the ITSUMO QoS architecture, which can be conceived as a wireless extension to Diffserv and bandwidth broker, is presented to illustrate the usefulness and applications of DSNP. Experimental results show that ITSUMO QoS architecture and DSNP achieve certain important QoS objectives in mobile environment.

9. ACKNOWLEDGMENTS

The authors wish to acknowledge the contributions of other members of the ITSUMO team from Telcordia Technologies, Inc. and Toshiba America Research Incorporated.

REFERENCES