# Field-Trial of Dynamic SLA in Diffserv-capable Network

Naoto Morishima[1], Akimichi Ogawa[2], Keijiro Ehara[2], Youki Kadobayashi[3]

[1] Graduate School of Information Science, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0101 Japan
[2] Graduate School of Media and Governance, Keio University, 5322 Endo, Fujisawa, Kanagawa 252-8520 Japan
[3] Cybermedia Center, Osaka University, 1-1 Yamadaoka, Suita, Osaka 565-0871 Japan

**Abstract.** As the Internet is undergoing substantial changes, there are many demands for the quality of the communication. Differentiated Services was standardized to meet such requirements. Currently, it is assumed that Diffserv is used under static Service Level Agreement (SLA). Under static SLA, the agreement between the service provider and the service subscribers are made as a result of negotiation between human agents. However, the service subscribers, especially individual users, may want dynamic SLA with which they can make their contract without human intervention and use the network resources immediately. Although many literatures addressed this point, operation of dynamic SLA under live network has not been applied. Moreover, only few experiments for dynamic SLA has been made.
In this paper, we describe our field-trial with dynamic SLA and resource reservation system at the spring retreat of WIDE Project, including its network configuration and the results of trial. Through this experiment, we attempt to reveal that how users behave with dynamic SLA, and what mechanism such system needs.

## 1 Introduction

The improvements of the Internet technology in the last decade have shifted its technical focus from the reachability to the quality of the communication. Internet Service Providers (ISPs) diversify their services and generate externality so that they can provide value added services. Today, they have the discussion on what kind of services should be offered and what kind of service management architecture the services require.

Traditional quality of service (QoS) mechanisms provide a hard bound of the resource allocation, known as deterministic guaranteed services [1]. To bring this telephone-like services to the Internet, Integrated Services (Intserv) [2] was standardized. However, there are some technical issues in this Intserv framework. First, the number of flow states the routers have to maintain is a major issue. In the Intserv architecture, all the routers on the way from the traffic source to its sink have to keep per-flow states. Therefore, the Intserv framework cannot be applied to the large scale networks such as the core backbone network in commercial ISPs. Furthermore, the Intserv does not meet an essential requirement. With the QoS guaranteed service, it is required that service subscribers does not have to specify how much bandwidth should be guaranteed in the network. In many cases, it is very hard for service subscribers to tell how much bandwidth the service needs specifically. However, still the service subscribers want to ask ISPs to guarantee the quality of services. Therefore, the QoS management requires some kind of adaptive mechanism to fulfill what the subscribers want. Expected Capacity [3] was proposed to achieve this adaptive mechanism in the Intserv framework.

Differentiated Services (Diffserv) [4] has been standardized to satisfy the requirements discussed above. The fundamental idea of the Differentiated Services is to aggregate flows which require the same quality of service and reduce state information that all routers in the core network must maintain.

Currently, it is assumed that Diffserv is used with static service level agreement (SLA). Under static SLA, the agreement between the service provider and the service subscribers (down-stream ISPs or individual users) are made as a result of negotiation between human agents.

However, it is expected that the subscribers are individual users in the smallest unit of the Diffserv-capable network, and they may want dynamic SLA with which they can make their contract without human intervention and use the network resources immediately. Nevertheless, few experiments have even been made to apply dynamic SLA to Diffserv-capable network in operation. Therefore, it has not become clear how users behave with dynamic SLA, and what mechanism such system needs.

In this experiment, we focus on dynamic SLA for the Diffserv-capable network. We have conducted this experiment for 3 days in WIDE Project retreat. In this paper, we first show Diffserv and Inter-DS-domain model, and then describe where this experiment fits in our model. We then give details of our implementation and experiment of admission control system. Finally, we show the results of this experiment and its evaluation.

## 2 Differentiated Services

Diffserv is a framework to combine multiple quality assurance mechanisms and provide statistical differentiation of services. Service level agreement(SLA) is made between subscriber and the service provider, and according to the agreement, the service provider offers various services. A service defines some characteristics of flow handling, such as flow classification or packet transmission. These characteristics are specified in quantitative or statistical indicator of throughput, delay, jitter, loss, or some other priority of access to network resources.

At the boundary of the Diffserv-capable network, there are ingress/egress edge nodes. This Diffserv-capable network is called DS domain. A flow entering a DS domain is classified, marked in its IP header and possibly conditioned at the ingress edge node, according to their SLA. This mark in IP header is called DS codepoint or DSCP [5]. Flows with the same DSCP are aggregated into a single behavior aggregate, and within the core of the DS domain, packets are forwarded according to the per-hop behavior associated with the DSCP.

However, the Diffserv framework mentioned above includes only intra-domain architecture, and inter-domain architecture lacks operational perspective. Namely, under current Diffserv framework, issues surrounding inter-domain architecture and dynamic SLA is not addressed at all.

Another limitation of Diffserv framework is that it does not take practical Internet operation model into account. For example, today's Internet is organized as a hierarchical chain of DS domains, rooted at major Internet exchange points. Although one DS domain must eventually interact with subscribers, such situation has not been addressed in the framework document.

Moreover, the current SLA framework in Diffserv is static one, because it makes deployment easy; decision-making about operational issues, such as bandwidth allocation plans or accounting, can be done by administrator. Such issues can be separated from the mechanisms such as router settings.

However, authors believe that such static SLA is transitional. The requirement for QoS/CoS can be roughly divided into persistent one and transient one, especially considering both ISP and subscribers' viewpoints. The former example would be a virtual leased line, and the latter example would include limited-time broadcast of live events or administrative traffic for emergent server maintenance. Thus, it's easy to imagine that QoS is shifting to the mixed environment with static SLA and dynamic SLA. In such environment, present policy framework is not sufficient, as it only supports static SLA.

We believe that the development of dynamic SLA is vital to promote widespread application of Diffserv.

## 3 Configuration of Diffserv Field-Trial

We designed, implemented and administered an operational, real Diffserv network at the year 2000 spring retreat of WIDE Project[6], which we call "WIDE camp". WIDE Project is a non-profit consortium for research and development of the Internet and related technologies. A 4-day camp for members of WIDE Project is held twice every year, once in spring and once in autumn. A time-limited operational network called "WIDE camp-net" has been built for every WIDE camp. The network is both highly operational and highly experimental; new networking technologies such as IPv6 and MPLS has been brought into operational network that most attendees rely on.

At WIDE camp-net of 2000 spring, we have done an experiment of Diffserv network with immediate reservation request from users. WIDE camp of 2000 spring was held at Isawa, Yamanashi, Japan, and there were 236 participants.

### 3.1 Field-trial overview

In order to make deployment both visible and easy, we have incorporated a number of ideas into our test-bed system.

Since our test-bed was incorporated within dual-stack, both IPv4 and IPv6-ready network, we had a number of routers, where traffic control should be applied. To eliminate operational overhead of traffic control, we isolated traffic control functions from routers by introducing ATM PVC bridging. One IPv4 router also acts as ATM PVC bridge for IPv6. In other words, all traffic has been concentrated to one router.

The QoS control to the commodity traffic of the WIDE camp-net were done at both ends of the external link. We used the COPS(Common Open Policy Service)[7] protocol for the provisioning the QoS control parameters to the routers. The COPS PDP(Policy Decision Point) that works as the provisioning server, was located within the WIDE camp-net. The COPS PEP(Policy Enhancement Point) that receive the provisioned information from the PDP, and applys the QoS configuration, was located within routers at both sides of the external link.

To avoid infinite occupation of network bandwidth, some means for promoting fair use of reservation must be implemented. We have implemented an accounting system, together with “virtual currency” that are withdrawn from each user’s account, according to usage history.

We also provided reservation tool to users, so that users can directly issue reservation requests to the Diffserv-capable network. The tool was provided for both IPv4 and IPv6.

Bandwidth reservation has been made visible through web-based front-end for bandwidth reservation, as well as remaining virtual currency in the user’s account.

### 3.2 Network Topology

In this section, we illustrate the network topology of experiment held at the WIDE camp-net. The network topology of the WIDE camp-net is shown in Fig.1. The camp-net is connected with the Internet using T1(1.5Mbps) external link. ATM was used for layer 2 protocol of the T1 external link.

3 virtual circuits were configured over the T1 external link. These VCs were used for the following purpose:

1. connection to the IPv4 Internet
2. connection to the IPv6 Internet via Keio University at Fujisawa
3. connection to the IPv6 Internet via NAIST, Nara

Each edge node described in Fig.1 were configured as PEP. PDP was located within the WIDE camp-net. In this experiment, the PEPs with the T1 external link was considered as a DS(Diffserv) domain.

The layer 2 configuration at the external gateway of WIDE camp-net is shown in Fig.2. The external T1 line is connected to the ATM switch. The ATM switch

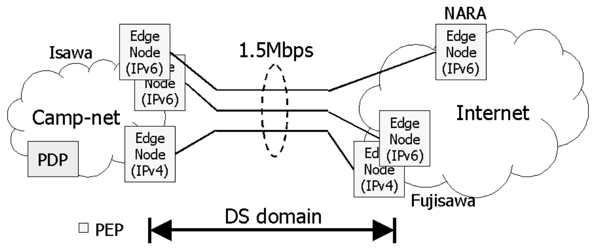


**Fig. 1.** Network Topology of WIDE camp-net

is connected to a PC router. Connectivity to the IPv4 Internet is provided within the WIDE camp-net via the PC router. The ATM switch is also connected to two IPv6 routers. The VC for both IPv6 network is connected to the IPv6 Internet via the PC router. For traffic queueing, ALTQ[8] is used within the PC router. For the traffic from WIDE camp-net to the Internet, Diffserv AF style marking was done at the input queue of the PC router. In the output queue for the traffic from WIDE camp-net to the Internet, HFSC[9] with RIO[10] was used for scheduling and queueing.

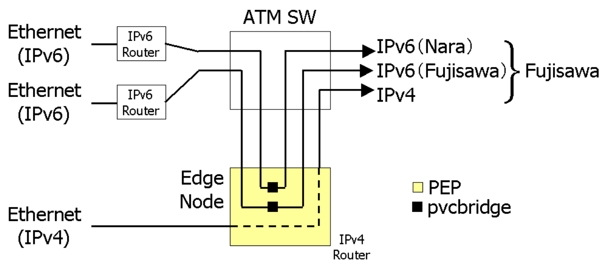


**Fig. 2.** Layer 2 Configuration of Gateway of Isawa

Each PC router in Isawa and Fujisawa was configured as a PEP. The PDP was located within the WIDE camp-net. The TCP connection between the PDP and PEP were kept alive during the experiment.

### 3.3 Administrative Model

In this section, we will show the administrative model of the experiment. In this experiment, the reservation services were provided to users within the WIDE camp-net.

For reservation, we divided the bandwidth of the external link into 19 blocks (Fig.3). 18 blocks were used for reservation by users, and 1 block was used for the default traffic. Each 18 blocks consists of 64kbps. In this experiment, each user request results to reservation of one 64kbps block. Since the number of blocks for reservation is 18, the number of available reservation is limited to 18. However, during non-congested times, the available bandwidth will be shared by the existing flows.

When a user reserves a 64kbps block, the reserved flow will be marked blue as in Diffserv AF manner. When the traffic of the reserved flow exceeds 64kbps, the reserved flow will be marked yellow. When the traffic of the reserved flow exceeds 64kbps extremely, the reserved flow will be marked red. The red marked packet will be treated as same as the default traffic.

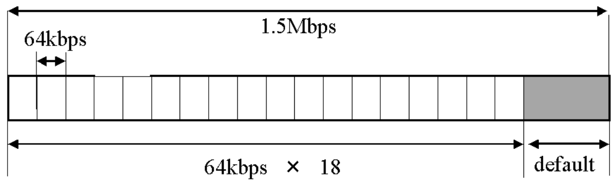


**Fig. 3.** Bandwidth Division of 1.5Mbps External Link

The reservation from the users is shown in Fig.4. The reservation from the users are done in the following steps. 1) First, a user sends a request to the nearby PEP using TCP. 2) The PEP that received the request from the user sends a COPS REQ(Request) message to its PDP. The HANDLE for the COPS REQ message is created by the PEP. 3) If the PDP decides that the request from the PEP is acceptable, the PDP sends a COPS DEC(Decision) message to the connected PEPs. The HANDLE for the COPS DEC message is created by the PDP. 4) PEPs configures the configuration given by the COPS DEC message from the PDP, and reports the result to the PDP using the COPS RPT message. 5) After receiving RPT messages from all the PEPs, the PDP sends a COPS DEC message to the PEP using the HANDLE generated with the COPS REQ message at 2). 6) The user is informed about the result of reservation by the PEP.

### 3.4 User Tool for Reservation

At year 2000 spring WIDE camp, attendees were considered as users of reservation system. We distributed 2000WU (Wide Unit, a virtual currency) to every attendee, for reservation purposes. Reservation of 64kbps for 1 minute could be obtained by paying 10WU.

We created a tool for requesting reservation from users. The tool was called "PEPe". PEPe establishes a TCP connection to the PEP when sending a reser-

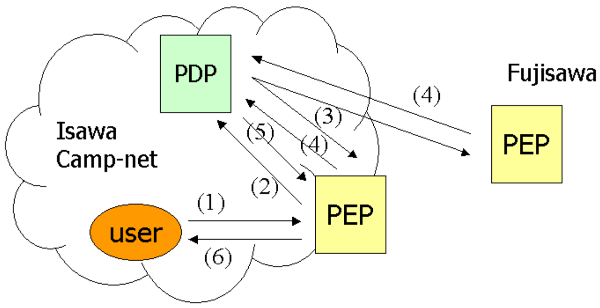


**Fig. 4.** Overview of Reservation from users

vation request. The request message sent by PEPe is shown in Fig.5. The user id and password was used to identify users. Flow information is included within the PEPe message.

| 0 | 8 | 16 | 24 |
|---|---|---|---|
| Length | | Username Len | Password Len |
| Username (Variable Length) | | | |
| Password (Variable Length) | | | |
| Service | | | |
| Bandwidth | | | |
| Start Time | | | |
| Time To Live | | | |
| Protocol | Address Family | Reserved | |
| Source Port | | Destination Port | |
| Source IP Address | | | |
| Source IP Netmask | | | |
| Destination IP Address | | | |
| Destination IP Netmask | | | |

**Fig. 5.** PEPe message

After every transaction for the request is processed within PEP and PDP, the PEPe receives a report from the PEP. The PEPe prints out the report from the PEP, and terminates the connection with the PEP.

# 4 Evaluation

In this section, we discuss evaluation of the experiment.

For evaluation, we created artificial congestion within the external link. The network topology for evaluation is shown in Fig.6.

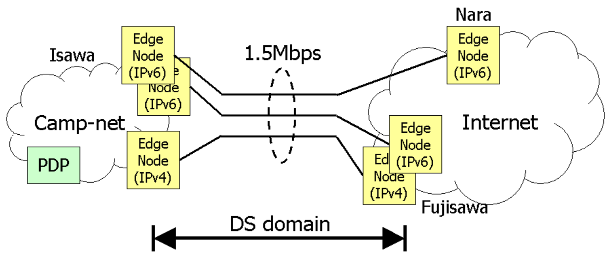


**Fig. 6.** Network topology for evaluation

First, we used netperf[11] with the options shown in Fig.7. Netperf was used

netperf –H isawa–gw –l 100000 – – –s 64K –S 64K –m 1460

**Fig. 7.** options for netperf

in the following pattern, 1) from out side to inside of WIDE camp-net, 2) from inside to outside of WIDE camp-net, and 3) from both sides of WIDE camp-net. Next we emulated congestion by sending UDP traffic from outside to inside of WIDE camp-net. The traffic caused by sending UDP was CBR(Constant Bit Rate) traffic.

The start time, end time and type of artificial congestion is shown in Table.1.

The normal traffic at the external link is shown in Fig.8. Fig.8 shows traffic from outside to inside of the WIDE camp-net. Fig.9 shows traffic from outside to inside of the widecamp network with netperf. Since the queue length in the gateway was large, there were no packet loss within the external link. Fig.10 shows traffic from outside to inside of the WIDE camp-net with UDP traffic. The network was consumed most, when there was artificial congestion by sending UDP traffic. There were packet loss while UDP traffic was sent.

The number of reservation requests from the users are shown in Fig.11. The reservation request is sent most during heavy congestion at the external link.

The number of reservation request errors (that is, failure to satisfy reservation requests) are shown in Fig.12. The reservation request error is reported most during heavy congestion at the external link. Most of the errors during

**Table 1.** Artificial Congestion

| Start time | End time | type |
|---|---|---|
| 3/15 18:14:50 | 3/15 18:19:05 | 2 |
| 3/15 23:02:57 | 3/16 00:02:16 | 3 |
| 3/16 12:33:37 | 3/16 13:39:28 | 1 |
| 3/16 13:42:54 | 3/16 13:49:03 | 1 |
| 3/16 13:54:22 | 3/16 15:06:01 | 3 |
| 3/16 16:19:47 | 3/16 19:39:47 | UDP |
| 3/16 20:55:27 | 3/16 23:49:24 | UDP |

heavy congestion were caused by overflow of reservations. In this experiment, we arranged 18 entries of 64Kbps bandwidth for reservation. When the 18 entries were full, the reservation request from the user resulted in error.

This shows that the attendees feel the network bandwidth is worth reserving and want additional value to connectivity during heavy congestion. As a result, requests for bandwidth were mostly made during this period of time. On the contrary, during the term of no congestion, reservations were not requested by users.

This also shows that the request is made more frequently with the progress of WIDE camp. This is because the experimental was first exposure to dynamic SLA system for most attendees, therefore at first they had some hesitation to reserve bandwidth. However, once they have learned effectiveness of bandwidth reservation system, they frequently used it. This particular observation indicates effectiveness of bandwidth reservation system under certain circumstances.

## 5 Conclusion

While Diffserv has been standardized to meet growing demands from diverse Internet applications, currently supported service model, i.e., static SLA, is rather limited. The authors believe that dynamic SLA will be used in near-term future.

In order to achieve widespread use of dynamic SLA, it is necessary to understand user behavior under dynamic SLA-enhanced Diffserv Internet. We have designed, implemented and operated live Diffserv-capable Internet with specific focus on dynamic SLA. During 4 days of WIDE retreat, most attendees were able to use immediate bandwidth reservation system that we have developed.

Since this was first exposure to bandwidth reservation system for most attendees, they had some reluctance to use this system until they face severe congestion, where they really need it. We enforced every attendee to learn and use this new tool by creating artificial congestion. After they have learned effectiveness of bandwidth reservation system, they eagerly used it. This particular observation indicates effectiveness of bandwidth reservation system under certain circumstances.

Thus, static SLA is useful only until dynamic SLA is widespread among several ISPs and subscribers are aware of its effectiveness. Our observation through field-trial at WIDE retreat confirms this argument. We believe that successive work on this topic will further support and amplify this discussion.

## References


1. D.D.Clark, S.Shenker, and L.Zhang. Supporting real-time applications in an integrated services packet network: Architecture and mechanism. In *ACM Computer Communication Review*, volume 22, October 1992.
2. R. Braden, D. Clark, and S. Shenker. *Integrated Services in the Internet Architecture: an Overview*, June 1994. RFC 1633.
3. D.D.Clark. A model for cost allocation and pricing in the internet. *MIT Workshop on Internet Economics*, March 1995.
4. S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss. *An Architecture for Differentiated Service*, December 1998. RFC 2475.
5. K. Nichols, S. Blake, F. Baker, and D. Black. *Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers*, December 1998. RFC 2474.
6. WIDE Project. http://www.wide.ad.jp/.
7. J. Boyle, R. Cohen, D. Durham, S. Herzog, R. Rajan, and A. Sastry. *The COPS (Common Open Policy Service) Protocol*, January 2000. RFC 2748.
8. K.Cho. A framework for alternate queueing: Towards traffic management by pc-unix based routers. In *USENIX*, 1999.
9. I.Stoica, H.Zhang, and T.S.Eugene Ng. A hierarchical fair service curve algorithm for link-sharing. In *SIGCOMM '97*, 1997. Real-Time and Priority Service.
10. D.D.Clark and W.Fang. Explicit allocation of best effort packet delivery service. *the IEEE ACM Transactions on Networking*, 6(4):362–373, August 1998.
11. Hewlett packard Company. Netperf: A network performance benchmark revision 2.1. Technical report, Hewlett-packard Company, 1995.

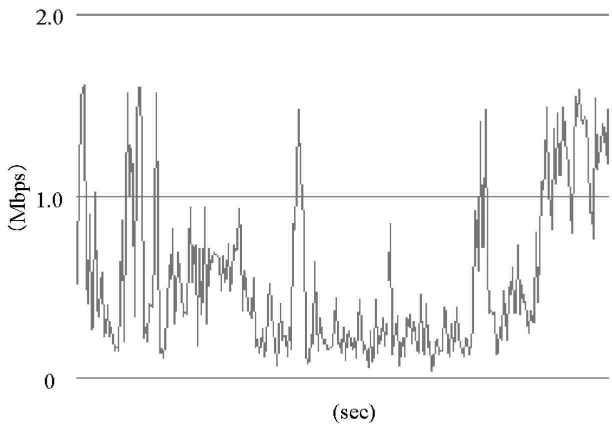


**Fig. 8.** Normal Traffic

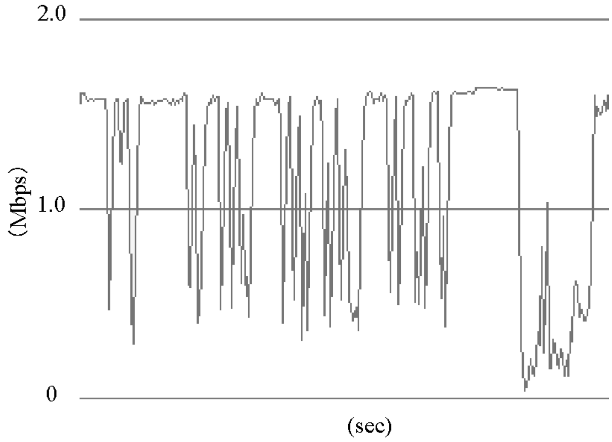


**Fig. 9.** Traffic with netperf

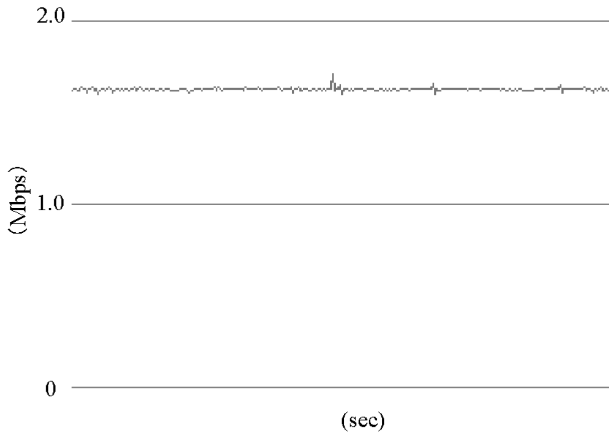


**Fig. 10.** Traffic with UDP

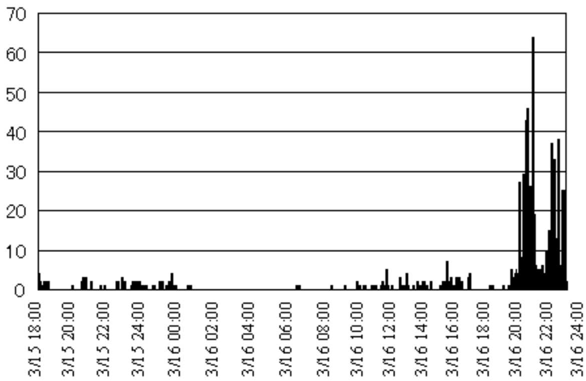


**Fig. 11.** The number of reservation requests

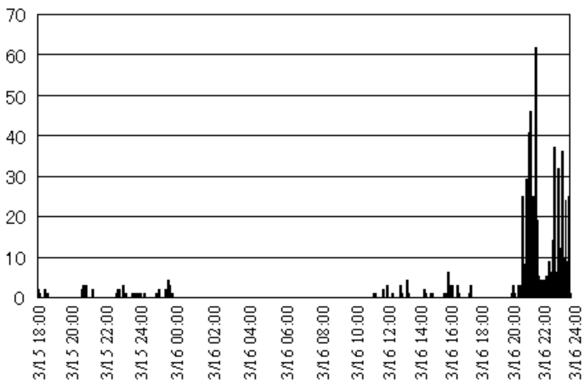


**Fig. 12.** The number of reservation request errors