

Guaranteed Transmission within Maximum Allowable Network Latency for Enhanced User Experience in Two-way Communication Applications

Tatsuya Fukui, Yuki Sakaue, and Katsuya Minami

Abstract

Keeping network latency below a certain level in two-way communication applications or systems enhances the user operating experience. However, the bandwidth-based shaping method commonly used in bandwidth guarantee services can generate large network latency due to burst-traffic leveling, thereby degrading the operating experience. We propose here a new shaping technique that takes required latency into account to guarantee the flow of traffic without exceeding maximum network latency.

Keywords: latency guarantee, two-way communication applications, QoE (quality of experience)

1. Introduction

The use of two-way communication applications/systems has been expanding in recent years. A typical example is a system for remotely operating heavy equipment in construction work. In this type of system, an operator sends an operating signal to a remotely located piece of equipment over the network to operate that equipment in a certain way. At this time, the operator can receive and view video taken by a camera installed on the equipment to confirm that the equipment responded as instructed. The time from operation input by the operator to confirming the operation by camera video is considered to greatly affect the user operating experience. In the case of remote operation, network communication latency is a major factor in the degradation of this experience.

Reducing this network latency to a low level improves the operating experience, but requirements in terms of allowable latency exist. In this regard, the occurrence of latency within one frame period of

camera video will not affect the display of that video. In fact, based on the authors' experience, there are times when latency even within two frame periods is not noticeable by the user, resulting in no degradation of the operating experience. In this way, there are many cases even in two-way communication applications in which low latency is not necessarily achieved but in which a certain amount of latency tolerance exists. Other examples of two-way communication applications are virtual reality, eSports, and remote desktop operation.

2. Network latency guarantee technique

In light of the above, we propose here a technique for transmitting network traffic that while not necessarily having low latency does not exceed the amount of allowable latency. Conventionally, when communication signals of the type described above are transmitted on a best-effort basis, the effects of other types of traffic might generate network latency above the level allowed, resulting in a greatly degraded operating

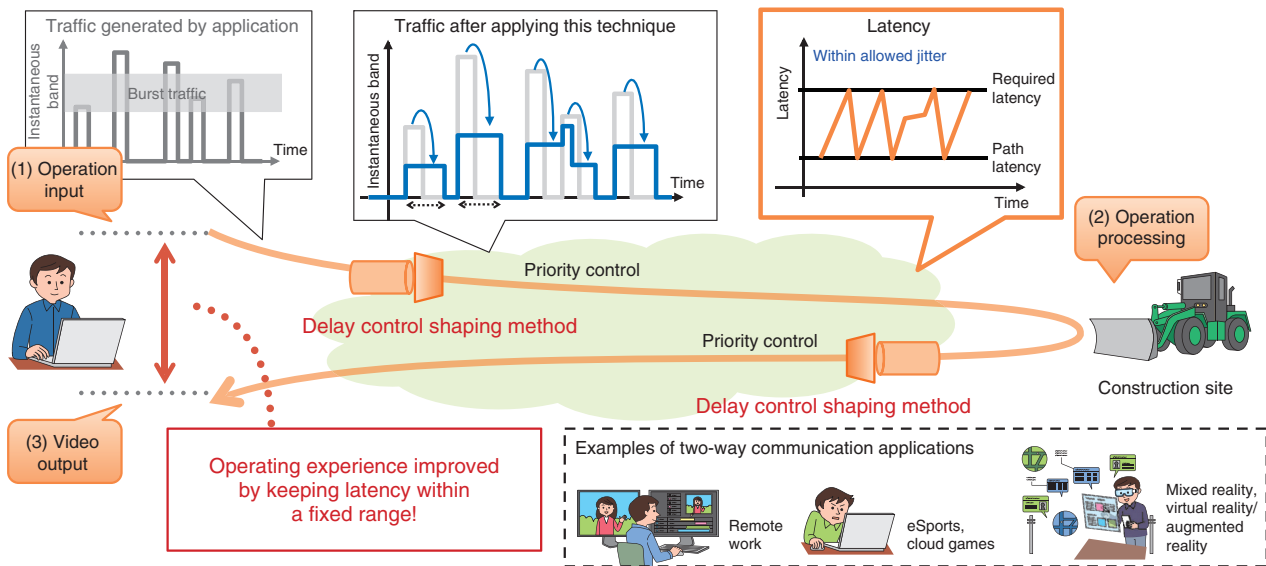


Fig. 1. Overview of network latency guarantee technique.

experience. Furthermore, while a method that treats certain types of communications as priority traffic unaffected by other types of traffic can be considered, such priority traffic may be transmitted at a level of latency lower than necessary, and since video traffic is often transmitted in bursts as mentioned above, the load on the network can be large.

If, in an effort to solve these problems, the conventional bandwidth-based shaping method used in bandwidth guarantee services is applied, large latency would occur even for a bandwidth setting larger than the average video rate, resulting in major degradation of the operating experience. This is because the burst transmission of video traffic results in a flow of traffic that instantly exceeds the set bandwidth, and shaping this traffic to fall within a certain bandwidth increases the time that signal data are stored in a buffer on network equipment. The end result is large latency.

In response to this issue, the proposed technique guarantees transmission within a certain amount of latency even for burst traffic as in video applications and suppresses the load on the network to the maximum extent possible (Fig. 1). For NTT, this technique contributes to the deployment of a competitive network infrastructure. The specific scheme behind this technique is as follows. When traffic flows into the network, the amount of arriving traffic is continu-

ously measured, and the amount of traffic to be sent into the network is calculated in real time based on a previously established amount of allowable latency and the minimally generated amount of end-to-end path latency within the network. In a relay network, prioritizing this traffic over best-effort traffic enables transmission within a certain amount of latency. In addition, shaping the traffic to the allowable latency according to the calculated amount of traffic can reduce the load on the network. We implemented a prototype of this function and evaluated its performance using a two-way communication application, and we found it to be effective in improving the user operating experience.

3. Future outlook

The application scope of the technique presented in this article is not one-way communications as in video delivery services but rather two-way communications in which processing is performed at a remote location according to an operating signal received from an operator. Going forward, we plan to perform surveys and trials of a wide range of applications and systems whose operating experiences can be improved by this technique with the aim of providing this feature as a network service.



Tatsuya Fukui

Research Engineer, Access Network Service Systems Project, NTT Access Network Service Systems Laboratories.

He received a B.E. and M.E. in applied engineering from Waseda University, Tokyo, in 2008 and 2010. He joined NTT Access Network Service Systems Laboratories in 2010, where he has been researching innovative network technologies for access applications. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE).



Katsuya Minami

Senior Research Engineer, Access Network Service Systems Project, NTT Access Network Service Systems Laboratories.

He received a B.E., M.E., and Ph.D. in information systems engineering from Osaka University in 1998, 2000, and 2003. He joined NTT in 2003. He is currently researching and developing novel access network services. He is a member of IEICE.



Yuki Sakaue

Engineer, Access Network Service Systems Project, NTT Access Network Service Systems Laboratories.

He received a B.E. and M.E. in applied physics and engineering from Hokkaido University in 2008 and 2012. He joined NTT Access Network Service Systems Laboratories in 2012, where he has been researching innovative network technologies for access applications. He is a member of IEICE.
