Feature Articles: Research and Development Initiatives for Internet of Things Implementation

Data Stream Assist Technology Supporting IoT Services

Naoki Higo, Takuma Tsubaki, Kohjun Koshiji, Toshimitsu Tsubaki, and Takeshi Kuwahara

Abstract

NTT Network Technology Laboratories has been researching and developing data stream assist technology that can be deployed lightly and flexibly to provide a network that supports complex services using Internet of Things (IoT) devices with high functionality and performance. In this article, we introduce technologies that improve the utilization efficiency of the network and the convenience of IoT services by changing the communication protocol, communication timing, and traffic volume according to various multipurpose versions of 4K surveillance camera images.

Keywords: edge computing, IoT services, video surveillance

1. Introduction

Research and development (R&D) of the Internet of Things (IoT) has been accelerating in recent years, and it is becoming more and more common for various devices with communication modules to be connected to a communication network. With the expansion and progress of IoT, progress is being made in techniques that enable visualization of the surrounding environment such as the movement of workers in a warehouse, and remote operation and automation of work tasks, and this progress is expected to improve productivity.

Relatively simple devices such as temperature sensors and power switches are used in many IoT services that are now being provided, but other devices such as monitoring cameras, drones, robots, and VR (virtual reality) equipment with high functionality and performance are also included in the concept of IoT. Such high-functionality and high-performance devices can transmit and receive raw data relating to the state and control of recording media and raw data containing large amounts of information such as video data from devices, and it is therefore expected that services using large-volume data streams will emerge.

At NTT Network Technology Laboratories, we aim

to create a network that supports IoT services and to create and provide network technologies that support large-volume data streams and that enable an ICT (information and communication technology)-based society as a whole.

2. Characteristics and approach of IoT data streams

In this article, we define large-volume data streams transmitted from high-functionality and high-performance devices as IoT data streams. There are two key characteristics of these IoT data streams: (1) the large capacity communication occurs intermittently, and (2) different services (purposes) and qualities are required depending on the service provider or end user (**Fig. 1**).

Feature (1) means that centralizing the function that processes the IoT data stream results in extremely biased traffic. For this reason, we believe that an architecture that can minutely expand the minimum necessary functions to distribute the functions is required. In addition, because of feature (2), we believe that it is important to consider not only the sophistication and diversification of devices but also network capabilities in order to meet different services and required qualities.



Fig. 1. Features of IoT data streams.

3. Overview of data stream assist technology

NTT Network Technology Laboratories has been studying data stream assist (DSA) technology to improve the convenience of IoT services by enabling the network to assist applications and devices according to the characteristics of IoT data streams.

DSA technology flexibly distributes and arranges the network function modules that operate in lightweight software without preparing dedicated hardware, which intervenes in the processing of communication protocols in the transport and application layers. With this technology, we achieve a simple construction and improve the added value of new services by responding to various applications and required qualities for each service provider that provides IoT services and each end user who uses IoT services.

In addition, to create useful functions and use cases realized with the DSA technology, we have been striving to clarify not only the needs of communication carriers but also the cooperation between a specific service provider and our software/network functions (see section 4). For that reason, we are not only looking at creating innovative elemental technologies but also integrating existing communication protocols and data formats with commercial technology.

4. Study of specific use case (video data stream)

We have been conducting joint research and experiments with the security service provider SECOM Co., Ltd. since 2016 as part of efforts to study technologies for future IoT services. The introduction of various IoT services is anticipated in the future, and it is therefore expected that social value will be created with new IoT services that use images captured by surveillance cameras, which will have higher resolution than existing devices. However, it is also necessary to study technology related to the network that supports such new IoT services.

Currently, much of the video data of surveillance cameras is written to the recording medium inside the surveillance cameras and extracted from the cameras as necessary for analysis and/or storage purposes. Furthermore, when acquiring video via a network, it is necessary to extract multiple videos for different purposes from surveillance cameras; therefore, congestion can occur in the end users' LAN (local area network) and the communication carrier's network.

We investigated how to use images from surveillance cameras more efficiently and developed an architecture that has a function to assist video distribution such as in the edge servers of communication carriers, so that a video data stream can be duplicated on a network and used for various purposes such as analysis, monitoring, and/or storage. To effectively deliver video data to be used in different applications, it is necessary to meet three functional requirements: (1) a change in the communication protocol, (2) a change in communication timing, and (3) a change in communication volume.

4.1 Functional requirement: communication protocol change

There are two reasons for changing the communication protocol. The first is to adapt to cases where the required quality differs among multiple destinations that want to distribute videos. For destinations where videos will be recorded, real-time access is not required, so we send video data streams using protocols with high data reachability. In contrast, for destinations where videos will be viewed in real time, we send video data streams using a protocol that emphasizes real-time functionality.

The second reason is to adapt to cases where different protocols can be used for the destination and for the sender that delivers the video. There are situations in which the types of protocols that can be used by the receiver are limited due to the fact that there are few protocols supported by certain kinds of software such as the media player operating on a device, and also because users often prefer to use widely popular protocols such as the Hypertext Transfer Protocol (HTTP). With surveillance cameras, it is difficult to adapt to new protocols because the replacement frequency of such cameras is very low. The protocol used also depends on the compatibility between the protocol and the hardware.

4.2 Functional requirement: communication timing change

There are two reasons for changing the communication timing. The first is the need to send video data streams according to the state of the network to the destination where the video data stream will be distributed. When the network is congested, video data streams that are not required for real-time capabilities can be temporarily held back, making it possible to reduce network congestion and avoid retransmission requests due to video streaming/buffering problems.

The second reason is to use the latest video at a high frame rate backwards in time. It is assumed that videos stored in the destination intended for recording are stored at a lower frame rate in order to reduce the storage capacity. However, to quickly confirm the content in videos when accidents or incidents occur, it is desirable to have high frame rate videos. Therefore, the communication carrier temporarily stores video data streams at a high frame rate and delivers the latest video to the destination as necessary, so that the high frame rate video delivery can be performed without increasing the end user's storage capacity.

4.3 Functional requirement: communication volume change

A change in the communication volume may be required depending on the bandwidth of the network to the destination and/or the image processing capacity at the destination. When radio transmission capabilities are included in the network, the bandwidth is often narrowed, and congestion occurs when the sender tries to flow traffic at a higher bandwidth. Also, under the assumption that image analysis is performed by artificial intelligence, which has advanced remarkably in recent years, the number of video frames that can be analyzed per unit time varies depending on the analysis time. For example, if it takes more than 100 ms to process one video frame, image analysis cannot be done at more than 10 frames per second (fps); therefore, no more video frames are needed. Furthermore, even when people watch video in monitoring applications, image delivery exceeding 30 fps is unnecessary.

These three functional requirements must be achieved without transcoding video data at the network side. Many in-market devices or software that enable changes in communication protocols, communication timing, and communication volume have high capacity that enables connection of many surveillance cameras and/or viewing devices, as well as various functions such as object recognition in videos. However, such equipment is expensive, as it requires dedicated hardware or many computational resources. In addition, there are cases where transcoding is done, and changes are made to the image quality and the captured content. Furthermore, as the conventional system architecture is not designed considering dispersed and/or divided functions, the system needs to be deployed on a physical or virtual machine. For these reasons, the computational resources required to operate unnecessary functions, and the scalability of handling an increasing number of IoT devices such as surveillance cameras are issues that need to be addressed.

In contrast to the existing technology as described above, the proposed technology enables us to change the communication protocol, communication timing, and communication volume without decoding the information, by replicating a video data stream and appropriately distributing it to multiple destinations through implementation of a finely decentralized network function.



Fig. 2. Application example of network and service cooperatively using RTP.

5. Video DSA technology

This section explains the functions used to satisfy the above requirements and the system that manages these functions.

We have assumed the use of the Real-time Transport Protocol (RTP) as an appropriate layer to construct the functions for the three functional requirements without video transcoding. Most cameras support RTP, which is suitable for transmitting a real-time stream, and the Real-time Streaming Protocol (RTSP), which controls the stream.

A usage example of an RTP header used cooperatively between a network and a service is shown in **Fig. 2**. The RTP header contains information on the timestamp, sequence number, media codec, and other details, but it does not have information on the video content. However, we can detect or obtain the end of the video frame and the time the video frame was generated by using the information in the header. Furthermore, an RTP extension header enables us to add more specific information such as a flag indicating the motion of some objects in the video. Such information can be used to judge whether or not to stream the video, thereby reducing the amount of traffic.

RTP is designed as the protocol above UDP (User Datagram Protocol), but RTSP interleave enables us to utilize TCP (Transmission Control Protocol) as the protocol under RTP. Therefore, we can make the most use of the features of RTP and change the under-layer protocols, the timing of communication, and the amount of traffic without decoding the video. Thus, RTP can be said to be the layer where the network and the service cooperate and create value for each other.

The difference between the conventional approach and our proposed approach in the situation where a video stream is conveyed with changes in the communication protocol, communication timing, and communication volume corresponding to the destination is shown in **Fig. 3**.

In the conventional approach, video streams transmitted from surveillance cameras are decoded and re-encoded at the intermediary entity and then conveyed to each device (Fig. 3(a)). While the existing technology needs to transcode video in the intermediary



(a) Existing system



(b) Proposed system

Fig. 3. Comparison of existing and proposed system architectures.

entity, our technology does not do so. Instead of transcoding, our technology changes the communication protocol, communication timing, and communication volume corresponding to the destination by handling RTP packets of one video stream; ensuring *distribution* of flow, which duplicates RTP packets and changes their under-layer protocols [1], applying a *filter*, which drops some RTP packets to adjust the frame rate, and implementing *reservation* of packets to store the video stream as RTP packets, but not as video containers (Fig. 3(b)).

An example of the operation flow of the DSA technology is shown in **Fig. 4**. These functional modules are designed to rapidly process RTP packets. In addition, because this technology handles each RTP packet, if packets are encrypted as SRTP (Secure RTP), the distribution function and the filter function in Fig. 4 do not need to have encrypted keys. These functional modules are constructed to enable pluggability, and they connect to each other like a pipeline.



Fig. 4. Workflow of data stream assist.

This enables us to add new functions and to scale out with ease.

There are some advantages in processing packetby-packet if we can use the semantics of the data format loaded on the RTP packet payload, for example, protocol conversion of JPEG over RTP from RTSP to HTTP or ROS (Robot Operating System), and clipping a specific area from an image.

6. Field trial

To provide our modules dynamically, we have adopted Docker [2] as the environment in which the processes run and Kubernetes [3] as the orchestrator of containers in the current prototype system. In the field trial that began in December 2017 [4, 5], we clustered the physical machines at some central offices (COs) by using Kubernetes and deployed containers dynamically. An overview of the trial is depicted in **Fig. 5**. We confirmed through the trial using 4K quality surveillance cameras that sufficient availability of our proposed approach was achieved by deploying functional modules based on containers, and that utilizing COs (or edge-computing platforms) was effective in terms of low latency and distributed processing.

When the target video data stream is registered by the user of the camera on our system, containers are deployed based on declarations about physical machines and regions, and a process that manages media information regarding the data stream takes place. Then, in correspondence to requests from



NGN: Next Generation Network

PC: personal computer VPN: virtual private network

Fig. 5. Network configuration of demonstration experiment.

devices such as a viewing terminal, a function process in a related container boots up and duplicates the data stream. The duplicated data streams are transmitted to other functions, and the communication protocol, communication timing, and communication volume are changed. The cooperative operation of these functions is achieved with the Advanced Message Queuing Protocol (AMQP)^{*1}.

A series of actions, registration of the video data stream, container deployment, and the booting process in response to requests are shown in **Fig. 6**. The location of functions is provisioned by a stream management function.

We demonstrated the prototype system and discussed new use cases for this technology in a project of ATII (APAC Telecom Innovation Initiative)^{*2} [6]. We also confirmed that it was possible to achieve smart buildings by integrating the technology with existing systems.

7. Future prospects

IoT data streams do not only contain video. A stream of raw data used for actuation and sensing can be used to control robots and drones. At the NTT R&D Forum held in February 2018, the concept of maintaining low latency by utilizing multiple routes for raw data obtained from actuation and sensing of haptic devices enabling *touching* of objects in virtual space was exhibited. In the future, when high-performance devices become more widely available, it is expected that services accompanying new IoT data streams will emerge.

In our R&D, we aim to realize new IoT services in collaboration with service providers by creating new

^{*1} AMQP: An application layer protocol of message-oriented middleware.

^{*2} ATII: An R&D initiative to realize virtual infrastructure technology by improving communication services in the APAC (Asia-Pacific) region and creating new network services after 2020.



API: application programming interface

Fig. 6. Container deployment and process activation.

technologies, improving existing technologies and infrastructures, and integrating them.

References

- N. Higo, R. Ishibashi, Y. Sueda, Y. Takita, Y. Hamaishi, and Y. Kaieda, "A Proposal for the Distributary Function to Assist Multi-purpose Usages of Data-streams," IEICE Tech. Rep., Vol. 117, No. 33, pp. 5–10, 2017.
- [2] Docker, https://www.docker.com/
- [3] Kubernetes, https://kubernetes.io/
- [4] NTT press release issued on December 19, 2017 (in Japanese). http://www.ntt.co.jp/news2017/1712/171219a.html

- [5] Focus on the News, "SECOM x NTT Collaboration Achievement Began Field Trials of Network Assist Function Converting Communication Data According to Purpose," NTT Technical Journal, Vol. 30, No. 5, pp. 55–57, 2018 (in Japanese).
- [6] T. Kuwahara, I. Hitoshi, and K. Suzuki, "Global Collaboration Initiatives Revolutionizing Research and Development of Network Technologies," NTT Technical Review, Vol. 16, No. 6, 2018. https://www.ntt-review.jp/archive/ntttechnical.php?contents= ntr201806fa7.html

Trademark notes

All brand names, product names, and company names that appear in this article are trademarks or registered trademarks of their respective owners.



Naoki Higo

Researcher, Innovative Network Technology and Collaboration Group, Network Architecture Design and Promotion Project, NTT Network Technology Laboratories. He received a B.S. and M.S. in control engi-

He received a B.S. and M.S. in control engineering and statistics from Tokyo Institute of Technology in 2011 and 2013. He joined NTT Network Technology Laboratories in 2013, where he conducted research on softwaredefined networking/network functions virtualization and network architecture for future IoT. His current research interest is multi-access edge computing. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE).

Takuma Tsubaki

Researcher, Innovative Network Technology and Collaboration Group, Network Architecture Design and Promotion Project, NTT Network Technology Laboratories.

He received a B.E. and M.E. in electrical engineering from Yokohama National University in 2012 and 2014. He joined NTT in 2014 and worked on network reliability evaluation and network design for disaster resilience. He is currently researching multi-access edge computing. He is a member of IEICE.



Toshimitsu Tsubaki

Senior Research Engineer, Innovative Network Technology and Collaboration Group, Network Architecture Design and Promotion Project, NTT Network Technology Laboratories.

NTT Network Technology Laboratories. He received a B.S. and M.S. in geophysics from Tohoku University, Miyagi, in 1994 and 1996. He joined NTT in 1996 and was involved in R&D of medium access control for broadband wireless access systems, RFID (radio frequency identifier) tagging systems, and non-contact high-speed transmission systems. His current research interests are network assistance technologies for IoT services. He is a member of IEICE.



Senior Research Engineer, Supervisor, Group Leader, Innovative Network Technology and Collaboration Group, Network Architecture Design and Promotion Project, NTT Network Technology Laboratories.

He received a B.E. and M.E. from Waseda University, Tokyo, in 1995 and 1997. Since joining NTT in 1997, he has been engaged in R&D of ATM (asynchronous transfer mode) networks, IP-VPNs (Internet protocol virtual private networks), cloud systems, network security, and edge computing technologies. He was in charge of telecom carrier relations regarding network R&D while based in Washington, D.C., USA, and led the establishment of ATII while in hyle in hyle in the source position. He is a member of IEICE.



Kohjun Koshiji

Research Engineer, Innovative Network Technology and Collaboration Group, Network Architecture Design and Promotion Project, NTT Network Technology Laboratories

NTT Network Technology Laboratories. He received a B.E. and M.E. in management science from Tokyo University of Science in 2007 and 2009. He joined NTT in 2009 and worked on reliability evaluation/design/management of telecommunication networks. He is currently involved in multi-access edge computing. He is a member of IEICE.