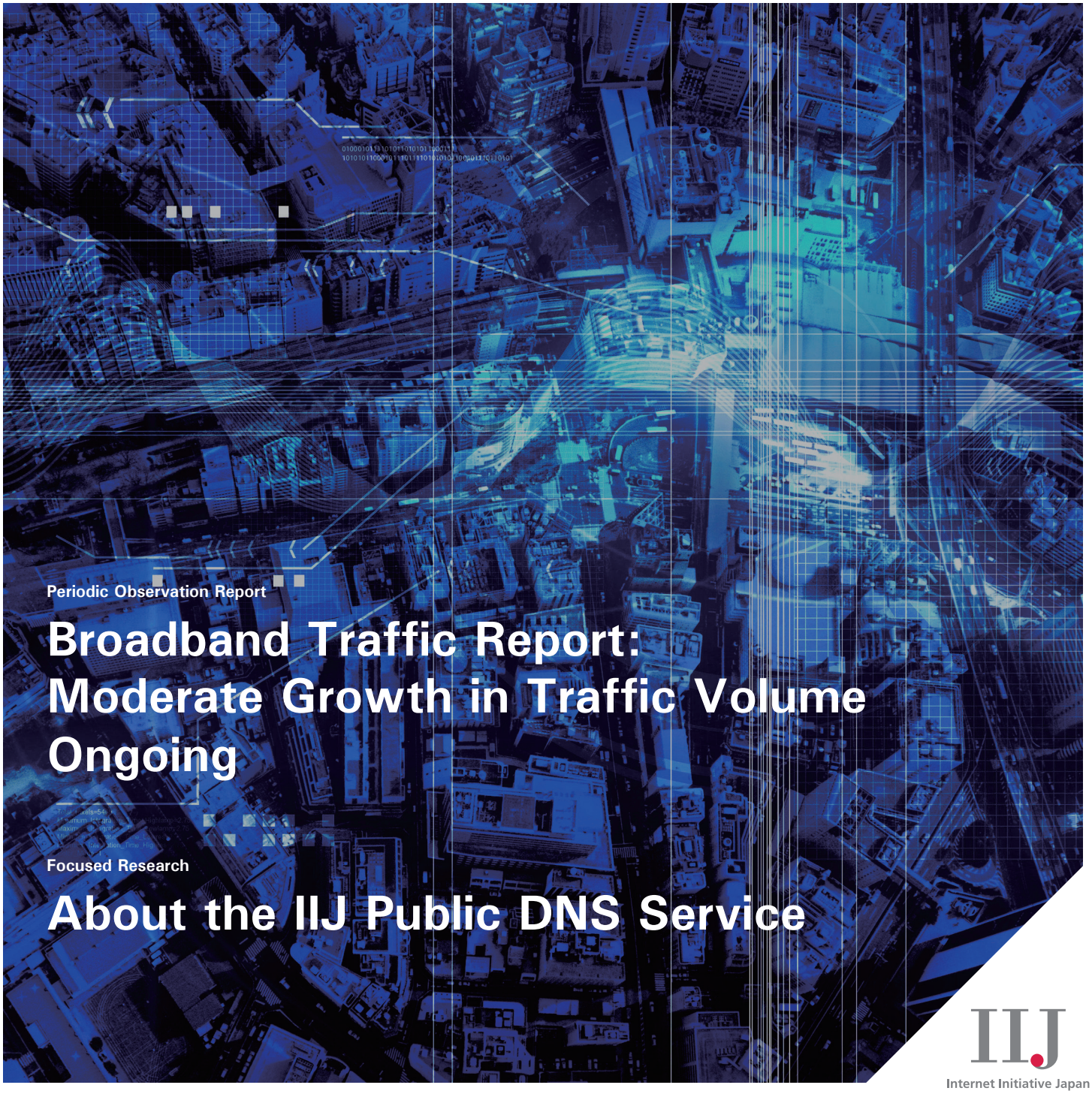
# IIR

Internet
Infrastructure
Review

Nov.2019
Vol. 44



Periodic Observation Report

## Broadband Traffic Report: Moderate Growth in Traffic Volume Ongoing

Focused Research

## About the IIJ Public DNS Service

IIJ
Internet Initiative Japan

# Internet Infrastructure Review

November 2019 Vol.44


Executive Summary ····· 3

1. Periodic Observation Report ····· 4
1.1 Overview ····· 4
1.2 About the Data ····· 4
1.3 Users' Daily Usage ····· 5
1.4 Usage by Port ····· 8
1.5 Conclusion ····· 9

2. Focused Research ····· 10
2.1 Introduction ····· 10
2.2 What is DoT/DoH ····· 10
2.2.1 DNS and Privacy ····· 10
2.2.2 DNS Transport Encryption ····· 10
2.2.3 Transport Encryption and DNSSEC ····· 11
2.3 IIJ Public DNS Service and DoT/DoH ····· 12
2.3.1 The TCP barrier ····· 12
2.3.2 The TLS Barriers ····· 12
2.3.3 The HTTP barrier ····· 13
2.3.4 Have we overcome the barriers? ····· 14
2.4 Public DNS and DoT/DoH ····· 14

## Executive Summary

In the afternoon of August 23, Amazon Web Services (AWS) experienced an approximately 6-hour outage in the Tokyo region. The incident caused major global cloud services to go offline, and this impacted systems and services at a great many companies, making those companies' services unavailable to general users. While the Internet is the ultimate distributed network, the rise of cloud computing means that data processing is becoming increasingly centralized. That an incident at a single company (AWS in this case) could have such a far-reaching impact is a stark reminder of this centralization on the Internet. Then on August 29, a number of ISPs reported Internet faults thought to stem from the sharp increase in traffic caused by a Microsoft Windows Update. This was also a notable event in that the activities of a single company had a major impact on the Internet infrastructure.

The IIR introduces the wide range of technology that IIJ researches and develops, comprising periodic observation reports that provide an outline of various data IIJ obtains through the daily operation of services, as well as focused research examining specific areas of technology.

Our periodic observation report for this issue, in Chapter 1, is our usual rundown of broadband traffic. This report has been part of the IIR every year since 2009, and we are proud to present this valuable data tracking trends in Internet traffic over more than 10 years. This year, although we observed increases in both fixed broadband and mobile services similar to what we saw last year, the results indicate that usage volumes at the individual user level have not changed much over the past few years.

Our focused research report in Chapter 2 looks at DNS, one important part of the Internet's foundations. Since May this year, IIJ has been providing the IIJ Public DNS Service, which performs name resolution via DNS over TLS (DoT) and DNS over HTTPS (DoH). The service is available to anyone, not just IIJ users. The report explains technical aspects of DoT/DoH and how they differ from DNSSEC and then discusses the IIJ Public DNS Service implementation and the customizations we made.

Through activities such as these, IIJ strives to improve and develop its services on a daily basis while maintaining the stability of the Internet. We will continue to provide a variety of services and solutions designed to serve the full needs of the infrastructure that underpins companies' business endeavors.

**Junichi Shimagami**

Mr. Shimagami is a Senior Executive Officer and the CTO of IIJ. His interest in the Internet led to him joining IIJ in September 1996. After engaging in the design and construction of the A-Bone Asia region network spearheaded by IIJ, as well as IIJ's backbone network, he was put in charge of IIJ network services. Since 2015, he has been responsible for network, cloud, and security technology across the board as CTO. In April 2017, he became chairman of the Telecom Services Association of Japan MVNO Council..

# Broadband Traffic Report: Moderate Growth in Traffic Volume Ongoing

## 1.1 Overview

In this report, we analyze traffic over the broadband access services operated by IIJ and present the results each yearr[*1][*2][*3][*4][*5][*6][*7][*8][*9][*10]. Here, we again report on changes in traffic trends over the past year, based on daily user traffic and usage by port.

Figure 1 shows the overall average monthly traffic trends for IIJ's fixed broadband services and mobile services. IN/OUT indicates the direction from the ISP perspective. IN represents uploads from users, and OUT represents user downloads. Because we cannot disclose specific traffic numbers, we have normalized the data, setting the latest OUT observation in each dataset to 1.

Since the previous edition of this report, the broadband data have included IPv6 IPoE traffic. The thin line labeled "broadband-IPoE" excludes IPv6 IPoE traffic. IPv6 traffic on IIJ's broadband services comprises both IPoE and PPPoE traffic[*11], but IPoE traffic does not pass directly through IIJ's network as we use Internet Multifeed Co.'s transix service for IPoE, and IPoE is therefore excluded from the analysis that follows here. As of June 2019, IPoE accounted for 19% of IN and 14% of OUT broadband traffic overall. Respectively, these are year-on-year increases of 7 and 6 points, so use of IPoE is rising.

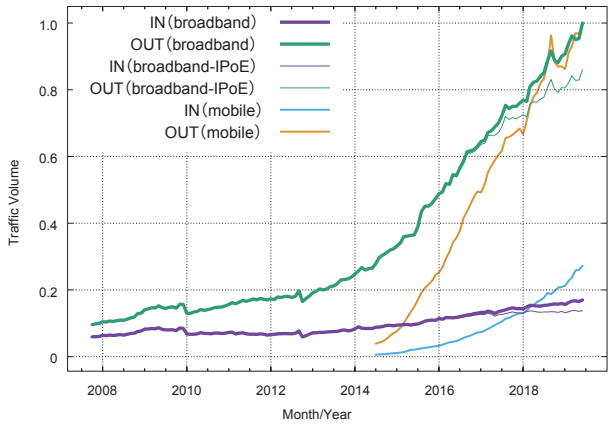



**Figure 1: Monthly Broadband and Mobile Traffic over Time**

Growth in both broadband and mobile traffic has risen, with some ups and downs, over the past two years or so. These fluctuations in broadband and mobile have been mostly synchronous with each other, suggesting that the underlying factors are the same.

Over the past year, broadband IN traffic increased 12% and broadband OUT traffic increased 19%, virtually the same growth rates as the year-earlier figures of 12% and 20%. Growth in mobile traffic has slowed, IN from 69% a year earlier to 60% this year, and OUT from 36% to 22%. And the total volume of mobile traffic remains an order of magnitude lower than broadband traffic.

## 1.2 About the Data

As with previous reports, for broadband traffic, our analysis uses data sampled using Sampled NetFlow from the routers that accommodate the fiber-optic and DSL broadband customers of our personal and enterprise broadband access services. For mobile traffic, we use access gateway billing information to determine usage volumes for personal and enterprise mobile services, and we use Sampled NetFlow data from the routers used to accommodate these services to determine the ports used.

Because traffic trends differ between weekdays and weekends, we analyze traffic in one-week chunks. In this report, we look at data for the week of May 27 through June 2, 2019, and compare those data with data for the week of May 28 through June 3, 2018, which we analyzed in the previous edition of this report.

*1 Kenjiro Cho. Broadband Traffic Report: Download Growth Slows for a Second Year Running. Vol.40. pp4-9. September 2018.
*2 Kenjiro Cho. Broadband Traffic Report: Traffic Growth Slows to a Degree. Internet Infrastructure Review. Vol.36. pp4-9. September 2017.
*3 Kenjiro Cho. Broadband Traffic Report: Traffic Growth is Accelerating. Internet Infrastructure Review. Vol.32. pp28-33. August 2016.
*4 Kenjiro Cho. Broadband Traffic Report: Comparing Broadband and Mobile Traffic. Internet Infrastructure Review. Vol.28. pp28-33. August 2015.
*5 Kenjiro Cho. Broadband Traffic Report: Traffic Volumes Rise Steadily Over the Past Year, and HTTPS Use Expands. Internet Infrastructure Review. Vol.24. pp28-33. August 2014.
*6 Kenjiro Cho. Broadband Traffic Report: The Impact of Criminalization of Illegal Downloads was Limited. Internet Infrastructure Review. Vol.20. pp32-37. August 2013.
*7 Kenjiro Cho. Broadband Traffic Report: Traffic Trends over the Past Year. Internet Infrastructure Review. Vol.16. pp33-37. August 2012.
*8 Kenjiro Cho. Broadband Traffic Report: Examining the Impact of the Earthquake on Traffic on a Macro Level. Internet Infrastructure Review. Vol.12. pp25-30. August 2011.
*9 Kenjiro Cho. Broadband Traffic Report: Traffic Shifting away from P2P File Sharing to Web Services. Vol.8. pp25-30. August 2010.
*10 Kenjiro Cho. Broadband Traffic Report: Increasing Traffic for General Users. Internet Infrastructure Review. Vol.4. pp18-23. August 2009.
*11 Akimichi Ogawa. Appendix A.3 "IPv6 PPPoE and IPv6 IPoE" in Professional IPv6 (in Japanese). Lambda Note. July 2018.

Results are aggregated by subscription for broadband traffic, and by phone number for mobile traffic as some subscriptions cover multiple phone numbers. The usage volume for each broadband user was obtained by matching the IP address assigned to users with the IP addresses observed. We gathered statistical information by sampling packets using NetFlow. Sampling rates were set between 1/8,192 and 1/16,382, taking into account router performance and load. We estimated overall usage volumes by multiplying observed volumes with the reciprocal of the sampling rate.

IIJ provides both fiber-optic and DSL broadband services, but fiber-optic access now accounts for the vast majority of use. Of users observed in 2019, 98% were using fiber-optic connections and accounted for 99% of overall broadband traffic volume.

## 1.3 Users' Daily Usage

First, we examine daily usage volumes for broadband and mobile users from several angles. Daily usage indicates the average daily usage calculated from a week's worth of data for each user.

Starting with this edition, we use daily usage data only on services provided to individuals. Enterprise usage varies widely and is readily influenced by the usage patterns of a subset of users, such that the overall distribution is clearly distorted. Individual usage, by contrast, shows a smooth distribution that remains stable. So we determined that using only the individual data to ascertain usage patterns would yield more generally applicable and easily interpretable conclusions. Note that because of the difficulty of distinguishing between individual and enterprise usage, the analysis of usage by port in the next section does include enterprise data.

Figure 2 and Figure 3 show the average daily usage distributions (probability density functions) for broadband and mobile users. Each compares data for 2018 and 2019 split into IN (upload) and OUT (download), with user traffic volume plotted along the X-axis and user frequency along the Y-axis. The X-axis shows volumes between 10KB (104) and 100GB (1011) using a logarithmic scale. Most users fall within the 100GB (1011) range, with a few exceptions.

The IN and OUT broadband traffic distributions are close to a log-normal distribution, which looks like a normal distribution on a semi-log plot. A linear plot would show a long-tailed distribution, with the peak close to the left and a slow gradual decrease toward the right.

The OUT distribution is further to the right than the IN distribution, indicating that download volume is more than an order of magnitude larger than upload volume. The peaks of both the IN and OUT distributions for 2019 are slightly further to the right than the peaks of the 2018 distributions, indicating that overall user traffic volumes are increasing. But that rightward shift in the distribution in 2019 was smaller than it had been in the past few years.

The peak of the OUT distribution, which appears toward the right in the plot, has steadily been moving rightward over

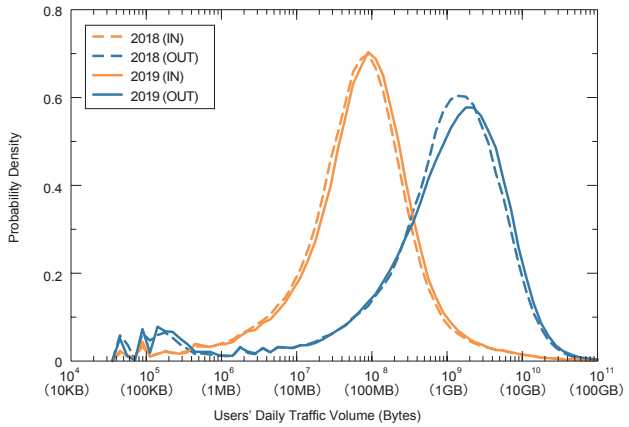



Figure 2: Daily Broadband User Traffic Volume Distribution Comparison of 2018 and 2019

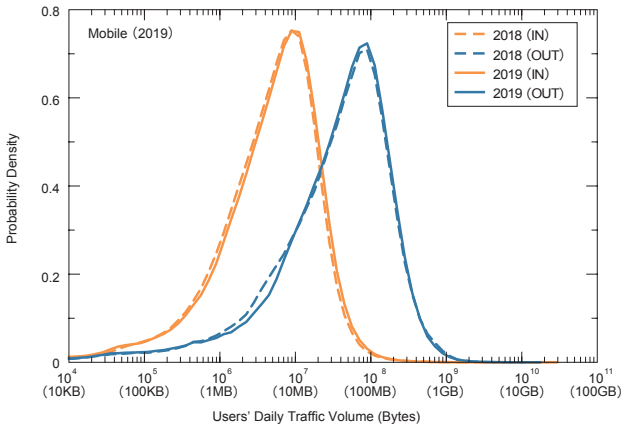



Figure 3: Daily Mobile User Traffic Volume Distribution Comparison of 2018 and 2019

the past few years, but heavy-user usage levels have not increased much, and as a result, the distribution is becoming less symmetric. The IN distribution on the left, meanwhile, is generally symmetric and closer to a log-normal distribution.

The data for mobile traffic indicate that usage volumes are significantly lower than for broadband. And limits on mobile data usage mean that heavy users, which fall on the right-hand side of the distribution, account for only a small proportion of the total, so the distribution is asymmetric. There are also no extremely heavy users. The variability in each user's daily usage volume is higher for mobile than for broadband owing to there being users who only use mobile data when out of the home/office as well as limits on mobile data. Hence, the daily average for a week's worth of data shows less variability between users than the data for individual days. Plotting the distributions for individual days in the same way results in slightly lower peaks and correspondingly higher tails on both sides, but the basic shape and modal values of the distribution remain largely unchanged. The difference in the mobile distributions versus last year is also minimal.

Table 1 shows trends in the mean and median daily traffic values for broadband users as well as the mode (the most frequent value, which represents the peak of the distribution).

**Table 1: Trends in Mean and Mode of Broadband Users' Daily Traffic Volume**

| | IN (MB/day) | | | OUT (MB/day) | | |
|---|---|---|---|---|---|---|
| Year | Mean | Median | Mode | Mean | Median | Mode |
| 2007 | 436 | 5 | 5 | 718 | 59 | 56 |
| 2008 | 490 | 6 | 6 | 807 | 75 | 79 |
| 2009 | 561 | 6 | 6 | 973 | 91 | 100 |
| 2010 | 442 | 7 | 7 | 878 | 111 | 126 |
| 2011 | 398 | 9 | 9 | 931 | 144 | 200 |
| 2012 | 364 | 11 | 13 | 945 | 176 | 251 |
| 2013 | 320 | 13 | 16 | 928 | 208 | 355 |
| 2014 | 348 | 21 | 28 | 1124 | 311 | 501 |
| 2015 | 351 | 32 | 45 | 1399 | 443 | 708 |
| 2016 | 361 | 48 | 63 | 1808 | 726 | 1000 |
| 2017 | 391 | 63 | 79 | 2285 | 900 | 1259 |
| 2018 | 428 | 66 | 79 | 2664 | 1083 | 1585 |
| 2019 | 479 | 75 | 89 | 2986 | 1187 | 1995 |

**Table 2: Trends in Mean and Mode of Mobile Users' Daily Traffic Volume**

| | IN (MB/day) | | | OUT (MB/day) | | |
|---|---|---|---|---|---|---|
| Year | Mean | Median | Mode | Mean | Median | Mode |
| 2015 | 6.2 | 3.2 | 4.5 | 49.2 | 23.5 | 44.7 |
| 2016 | 7.6 | 4.1 | 7.1 | 66.5 | 32.7 | 63.1 |
| 2017 | 9.3 | 4.9 | 7.9 | 79.9 | 41.2 | 79.4 |
| 2018 | 10.5 | 5.4 | 8.9 | 83.8 | 44.3 | 79.4 |
| 2019 | 11.2 | 5.9 | 8.9 | 84.9 | 46.4 | 79.4 |

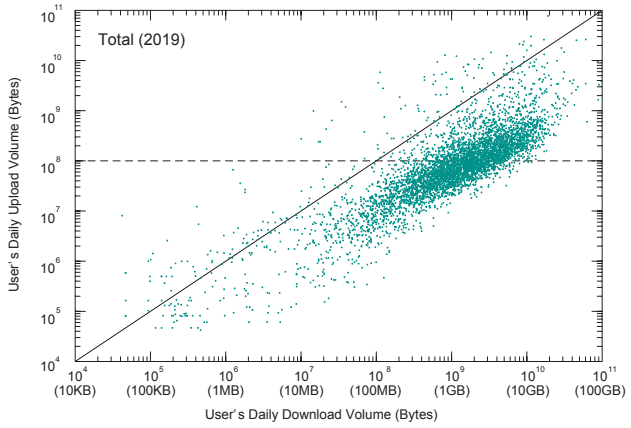



**Figure 4: IN/OUT Usage for Each Broadband User**

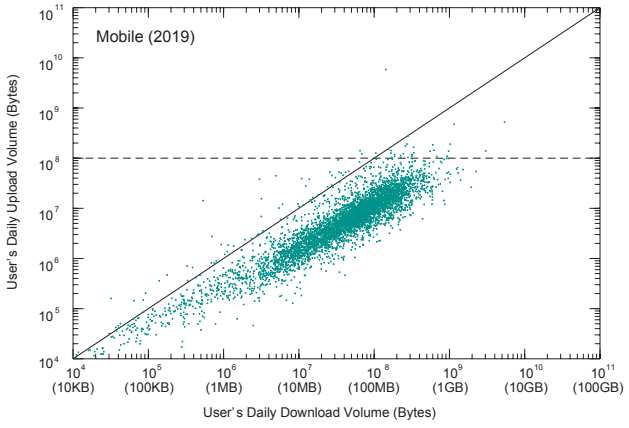



**Figure 5: IN/OUT Usage for Each Mobile User**

The peak was slightly off from the center of the distribution, so the distribution was adjusted to bring the mode toward the center. Comparing the values for 2018 and 2019, the IN mode rose from 79MB to 89MB and the OUT mode rose from 1,585MB to 1,995MB, translating into a growth factor of 1.3 for both IN and OUT. Meanwhile, because the means are influenced by heavy users (on the right-hand side of the distribution), they were significantly higher than the corresponding modes, with the IN mean being 479MB and the OUT mean being 2,986MB in 2019. The 2018 means were 428MB and 2,664MB, respectively.

For mobile traffic (Table 2), the mean and modal values are close owing to the lack of heavy users. In 2019, the IN mode was 9MB and the OUT mode was 79MB, while the means were IN 11MB and OUT 85MB. The modes for both IN and OUT traffic were identical to those for the previous year. The modes were unchanged but the means and medians increased, which reflects a slight decrease in the proportion of light users, corresponding to the part of the distribution to the left of the peak in Figure 2.

Figure 4 and Figure 5 plot per-user IN/OUT usage volumes for random samples of 5,000 users. The X-axis shows OUT (download volume) and the Y-axis shows IN (upload volume), with both using a logarithmic scale. Users with identical IN/OUT values fall on the diagonal.

The cluster spread out below and parallel to the diagonal in each of these plots represents typical users with download volumes an order of magnitude higher than upload volumes.

For broadband traffic, there was previously a clearly recognizable cluster of heavy users spread out thinly about the upper right of the diagonal, but this is now no longer discernible. Variability between users in terms of usage levels and IN/OUT ratios is wide, indicating that there is a diverse range of usage styles. There is almost no difference between these plots and those for 2018, too.

For mobile traffic, the pattern of OUT being an order of magnitude larger also applies, but usage volumes are lower than for broadband, and there is less variability between IN and OUT.

Figure 6 and Figure 7 show the complementary cumulative distribution of users' daily traffic volume. On these log-log plots, the Y-axis values represent the proportion of users with daily usage levels greater than the corresponding X-axis values. These plots are an effective way of examining the distribution of heavy users. The linear-like decline toward the right-hand side of the plots indicates that the distributions are long-tailed and close to a power-law distribution. Heavy users appear to be distributed statistically and do not appear to constitute a separate, special class of user.

On mobile, heavy users appear to be distributed according to a power-law for OUT traffic, but the linear-like slope breaks down somewhat for IN traffic, with a large proportion of users uploading large volumes of data.

Traffic is heavily skewed across users, such that a small proportion of users accounts for the majority of overall traffic volume. For example, the top 10% of broadband users

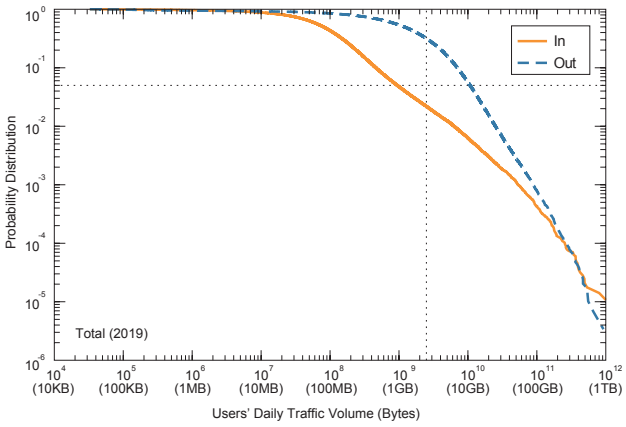



Figure 6: Complementary Cumulative Distribution of Broadband Users' Daily Traffic Volume

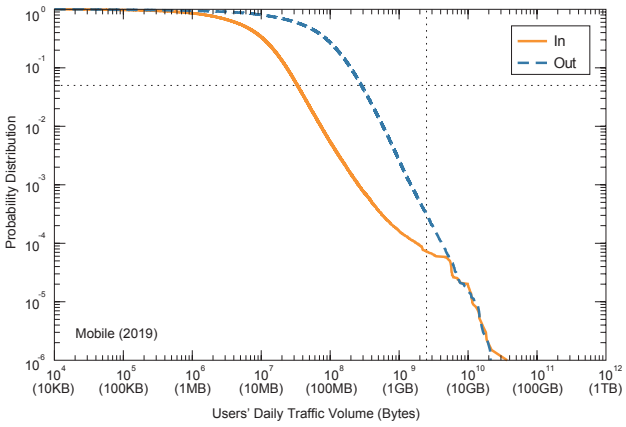



Figure 7: Complementary Cumulative Distribution of Mobile Users' Daily Traffic Volume

account for 52% of total OUT and 82% of total IN traffic, while the top 1% of users account for 17% of OUT and 58% of IN traffic. As the proportion of heavy users has declined over the past few years, the skew has also decreased, albeit only slightly. As for mobile, the top 10% of users account for 43% of OUT and 47% of IN traffic, while the top 1% account for 12% of OUT and 18% of IN traffic. The skew is less pronounced than indicated in our reports up to last year because we are now looking only at data on individuals.

## 1.4 Usage by Port

Next, we look at a breakdown of traffic and examine usage levels by port. Recently, it has become difficult to identify applications by port number. Many P2P applications use dynamic ports on both ends, and a large number of client/server applications use port 80, which is assigned to HTTP, to avoid firewalls. Hence, generally speaking, when both parties are using a dynamic port numbered 1024 or higher, the traffic is likely to be from a P2P application, and when one of the parties is using a well-known port lower than 1024, the traffic is likely to be from a client/server application. In light of this, we take the lower of the source and destination port numbers when breaking down TCP and UDP usage volumes by port.

Table 3 shows the percentage breakdown of broadband users' usage by port over the past five years. In 2019, 81% of all traffic was over TCP connections. The proportion of traffic over port 443 (HTTPS), which was down a little last time, rose sharply from 41% to 52%. The proportion of traffic over port 80 (HTTP) fell from 27% to 20% here, and the figure for UDP port 443, which is used by Google's QUIC protocol, fell from 10% to 8% after having risen up until the previous edition of this report. These figures demonstrate that the shift from HTTP to HTTPS is ongoing, while QUIC has tapped the brakes on growth a little.

TCP dynamic port traffic, which has been on the decline, fell to 8% in 2019. The proportion accounted for by individual dynamic port numbers is tiny, with the most commonly used port 8080 only accounting for 0.5%. Port 1935, which is used by Flash Player and has also been in decline, fell to around 0.3%. Almost all other traffic here is VPN related.

Table 4 shows the percentage breakdown by port for mobile users. The figures are close to those for broadband on the whole, suggesting that mobile users use applications in a manner similar to broadband users.

Table 3: Broadband Users' Usage by Port

| year | 2015 | 2016 | 2017 | 2018 | 2019 |
|---|---|---|---|---|---|
| protocol port | (%) | (%) | (%) | (%) | (%) |
| **TCP** | **80.8** | **82.8** | **83.9** | **78.5** | **81.2** |
| (< 1024) | 63.3 | 69.1 | 72.9 | 68.5 | 73.3 |
| 443 (https) | 23.3 | 30.5 | 43.3 | 40.7 | 51.9 |
| 80 (http) | 37.9 | 37.1 | 28.4 | 26.5 | 20.4 |
| 993 (imaps) | 0.1 | 0.1 | 0.2 | 0.2 | 0.3 |
| 22 (ssh) | 0.2 | 0.2 | 0.1 | 0.1 | 0.2 |
| 182 | 0.4 | 0.3 | 0.3 | 0.3 | 0.2 |
| (>= 1024) | 17.5 | 13.7 | 11.0 | 10.0 | 7.89 |
| 8080 | 0.3 | 0.2 | 0.3 | 0.3 | 0.5 |
| 1935 (rtmp) | 1.8 | 1.5 | 1.1 | 0.7 | 0.3 |
| **UDP** | **11.4** | **11.1** | **10.5** | **16.4** | **14.1** |
| 443 (https) | 0.9 | 2.4 | 3.8 | 10.0 | 7.8 |
| 4500 (nat-t) | 0.2 | 0.2 | 0.2 | 0.2 | 0.3 |
| **ESP** | **7.4** | **5.8** | **5.1** | **4.8** | **4.4** |
| **IP-ENCAP** | **0.2** | **0.2** | **0.3** | **0.2** | **0.2** |
| **GRE** | **0.2** | **0.1** | **0.1** | **0.1** | **0.1** |
| **ICMP** | **0.0** | **0.0** | **0.0** | **0.0** | **0.0** |

Table 4: Mobile Users' Usage by Port

| year | 2015 | 2016 | 2017 | 2018 | 2019 |
|---|---|---|---|---|---|
| protocol port | (%) | (%) | (%) | (%) | (%) |
| **TCP** | **93.8** | **94.4** | **84.4** | **76.6** | **76.9** |
| 443 (https) | 37.4 | 43.7 | 53.0 | 52.8 | 55.6 |
| 80 (http) | 52.5 | 46.8 | 27.0 | 16.7 | 10.3 |
| 31000 | 0.0 | 0.2 | 1.8 | 2.9 | 6.4 |
| 993 (imaps) | 0.5 | 0.5 | 0.4 | 0.3 | 0.3 |
| 1935 (rtmp) | 0.5 | 0.3 | 0.2 | 0.1 | 0.1 |
| **UDP** | **5.2** | **5.0** | **11.4** | **19.4** | **17.3** |
| 443 (https) | 1.0 | 1.5 | 7.5 | 10.6 | 8.3 |
| 12222 | 0.0 | 0.1 | 0.1 | 2.3 | 3.4 |
| 4500 (nat-t) | 0.3 | 0.2 | 0.2 | 4.5 | 3.0 |
| 53 (dns) | 0.1 | 0.2 | 0.1 | 0.1 | 0.1 |
| **ESP** | **0.7** | **0.4** | **0.4** | **3.9** | **5.8** |
| **GRE** | **0.3** | **0.1** | **0.1** | **0.1** | **0.0** |
| **ICMP** | **0.0** | **0.0** | **0.0** | **0.0** | **0.0** |

Figure 8 compares overall broadband traffic for key port categories across the course of the week from which observations were drawn in 2018 and 2019. We break the data into four port buckets: TCP ports 80 and 443, dynamic ports (1024 and up), and UDP port 443. The data are normalized so that peak overall traffic volume on the plot is 1. By comparison with 2018, the proportion of traffic over TCP port 443 has risen even further whereas TCP port 80 has seen a decrease. The overall peak is between 19:00 and 23:00 hours. Traffic increases during the daytime on Saturday and Sunday, reflecting household Internet usage times.

Figure 9 shows the trend for TCP ports 80 and 443 and UDP port 443, which account for the bulk of mobile traffic. In the mobile space as well, TCP port 443 has seen an increase while the proportion of traffic over TCP port 80 has declined. When compared with broadband, we note that mobile traffic levels remain high throughout the day, from morning through night. The plot shows that usage times differ from those for broadband, with three separate mobile traffic peaks occurring on weekdays: morning commute, lunch break, and evening from 17:00 to 22:00 hours.

## 1.5 Conclusion

Traffic volume has been growing moderately over the past few years. Although I say “moderately”, it’s only moderate in relation to past growth. At its annual rate of 20%, traffic is set to more than double over a four-year period. Both broadband and mobile traffic have been increasing, albeit with some ups and downs. The fact that both tend to rise and fall around the same time suggests that common factors are at play, but we have not been able to pinpoint what the specific factors are.

Both broadband and mobile usage volume by user have not changed much in the past few years. No new services that would drive traffic upward have appeared over that time, and it is clear that users’ Internet usage has not changed much as a result. Video resolutions are definitely on the rise, but it looks like the accompanying rise in codec compression rates is keeping total traffic growth in check.

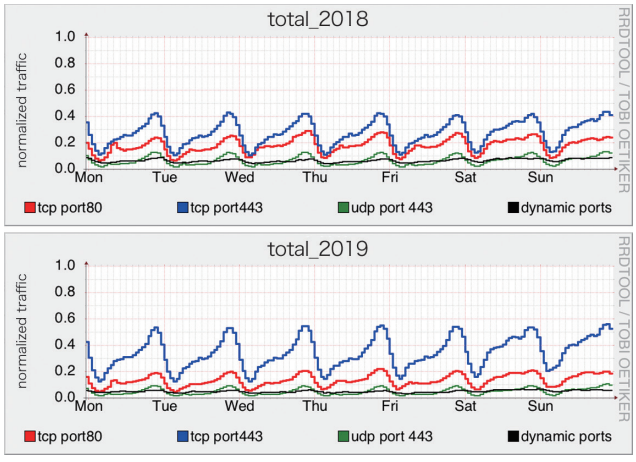



Figure 8: Broadband Users’ Port Usage Over a Week
2018 (top) and 2019 (bottom)

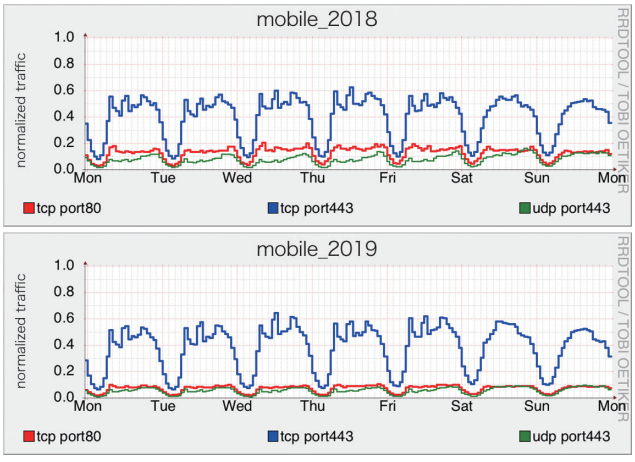



Figure 9: Mobile Users’ Port Usage Over a Week
2018 (top) and 2019 (bottom)

**Kenjiro Cho**
Research Director, Research Laboratory, IIJ Innovation Institute Inc.

# About the IIJ Public DNS Service

## 2.1 Introduction

IIJ released the beta version of its IIJ Public DNS Service in May. It is a caching DNS service that only accepts DNS over TLS (DoT) and DNS over HTTPS (DoH), meaning that it does not support name resolution via the usual UDP/TCP setup, and it is available to anyone, not just IIJ users.

This report explains how DoT/DoH differs from the usual DNS setup and describes key considerations and future challenges for IIJ in providing this service.

## 2.2 What is DoT/DoH

### 2.2.1 DNS and Privacy

Information registered to the DNS is assumed to be publicly and widely available. For a long time, therefore, the focus of DNS security has been to ensure that the information is not tampered with (i.e., that its integrity is maintained), whereas ensuring that the information is not intercepted (i.e., that it remains confidential) has not been a priority.

In 2013, however, the Snowden affair revealed the existence of PRISM, an extensive communications monitoring and information collection program carried out by the US National Security Agency (NSA). This prompted the IETF to declare that "Pervasive Monitoring Is an Attack" (RFC 7258) and call for protocols to be designed to withstand pervasive monitoring going forward.

As it became clear that DNS was also being monitoring under PRISM, the IETF began developing mechanisms to ensure DNS privacy, until then a perfunctory affair, through its new DPRIVE (DNS PRIVate Exchange) Working Group. DPRIVE has published various protocol extensions/revisions for DNS, including Qname Minimisation (RFC 7816) and EDNS(0) Padding Option (RFC 7830, RFC 8467), with transport encryption being of relatively high importance among these efforts.

### 2.2.2 DNS Transport Encryption

Traditional DNS mainly uses UDP for the lower-level protocol (transport), supplementing this with TCP. However, plain UDP/TCP, and DNS over UDP/TCP, do not have a mechanism to provide confidentiality, and communications are easily eavesdropped since they take place in the clear. A decision was therefore made to insert an encryption layer between DNS and the lower layers to provide protection.

Various encryption layers have been proposed, with the following having been standardized so far: DNS over TLS (RFC 7858), DNS over DTLS (RFC 8094), and DNS over HTTPS (RFC 8484). A draft of DNS over QUIC has also been submitted to the IETF and is now under discussion. And if HTTP/3, which is also currently under discussion, is standardized, DoH will automatically support HTTP/3 (Figure 1).

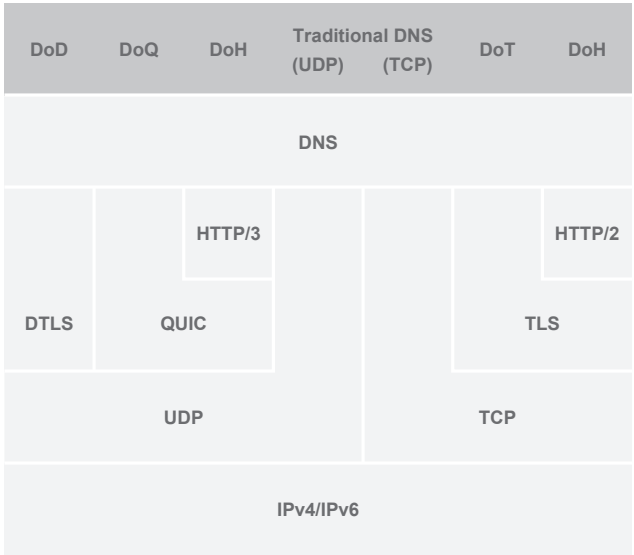



Figure 1: DNS Transport Protocols

These various encryption layers are not integrated into any one protocol; at present, you simply select whichever suits your users' circumstances. But flooding the space with a bunch of different protocols can have its drawbacks, so it is quite conceivable that a subset will be selected, and the rest deprecated, at some point in the future. (At present, DNS over DTLS is a specification only, with no existing implementations, and it seems unlikely that it will become available for use.)

### 2.2.3 Transport Encryption and DNSSEC

DNS already has a mechanism for verifying DNS information via digital signatures called DNSSEC. So why do we need a new method of transport encryption when we have DNSSEC? And will transport encryption eventually make DNSSEC unnecessary?

Before answering this question, let's look at the scope of transport encryption. With DNS, clients do not directly query the server on which the master DNS information is registered (authoritative server); instead, they query caching servers provided by ISPs and other parties. In general, the caching servers are responsible for querying the authoritative server.

Transport encryption is currently only performed between the user and the caching server. Communications between the caching server and the authoritative server use traditional DNS and are not encrypted.

In general, encryption guarantees both integrity and confidentiality, but when it comes to DNS transport encryption, encryption only happens between the user and the caching server. Since communication between the caching server and the authoritative server use traditional DNS (not encrypted), the integrity of the information obtained by the caching server cannot be guaranteed, and its integrity cannot be guaranteed even if the information is protected by encryption. That is, unlike encryption in general, the DNS transport encryption protocol only guarantees confidentiality between the user and the caching server (Figure 2).

DNSSEC, meanwhile, was introduced to protect against data forging and manipulation. It uses digital signatures to guarantee integrity, but communication itself takes place in the clear and is not confidential.

So both transport encryption and DNSSEC are mechanisms for improving DNS security, with one focused on guaranteeing confidentiality without integrity and the other focused on guaranteeing integrity without confidentiality. In that sense, they are each other's complement, so one cannot replace the other. Each protects something different, which means that transport encryption does not obviate the need for DNSSEC, and vice versa.

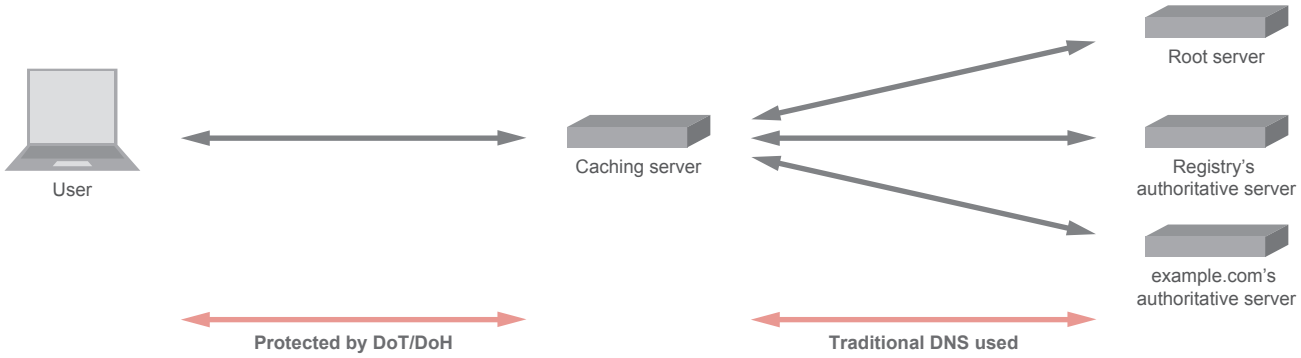



Figure 2: Scope of Transport Encryption

## 2.3 IIJ Public DNS Service and DoT/DoH

DoT and DoH differ from traditional DNS only in terms of transport-layer protocol. They both use the same DNS protocol as traditional DNS. Yet we still had a number of barriers to overcome to provide them as a service. Let's look at each in turn.

### 2.3.1 The TCP barrier

The major difference between DoT/DoH and traditional DNS is that all communications take place via TCP before being encrypted using TLS. While traditional DNS can use either TCP or UDP as the transport protocol, in most cases UDP is used, with TCP only being used in limited cases.

With TCP, a session must be established before communications on higher-layer protocols can start. TCP also uses various techniques to ensure the reliability of transmissions, which include checking that sent packets were in fact received and resending them if necessary.

DNS involves very little data exchange. In most cases, both the query and response rarely exceed a few hundred bytes. When TCP is used here, the process of setting up and later terminating the session accounts for vastly more of the communications exchanged than the actual DNS message, resulting in extremely poor efficiency. Possible solutions to load problems include the extravagant approach of simply adding more servers, but nothing can really be done about increased latency stemming from an increase in packet roundtrips.

UDP has no such mechanisms, making it fast and simple, and it is widely used in protocols that involve the exchange of small packets, such as DNS and NTP. Its reliability, on the other hand, is low, and many of the attacks on DNS that have so far been discovered, such as cache poisoning and DNS amp, really stem from the use of UDP in the lower layer rather than from any problems with the DNS protocol itself.

Although it is known that using TCP would preclude or greatly reduce the threat of such attack methods, the massive overhead that would result from the use of TCP with DNS—because DNS, by its nature, involves the exchange of a large number of small packets—has dissuaded the community from shifting to TCP as the main transport protocol. Even when the Kaminsky Attack[*1], which allows vastly more efficient cache poisoning that with previously known methods, was revealed in 2008, we still didn't switch to TCP and opted instead to make do with UDP and treat the symptoms with source port randomization[*2].

So there was considerable aversion to the TCP overhead with DNS, but requiring TLS naturally also means requiring TCP. The DNS over DTLS protocol, which uses UDP, does exist, but it entails a large overhead just like TCP, and moreover, no one can use it because it is only a specification; no implementations exist.

Nonetheless, we must abandon the conventional wisdom if we are to use it as a foundation for providing a secure caching DNS service. A major factor in enabling us to provide the IIJ Public DNS Service without restriction to anyone in the world is that we do not use DNS over the low-reliability UDP protocol, which dispels concern of the service being used as a launchpad for a DNS amp attack or such like. We should focus squarely on the benefits that using TCP actually provides.

### 2.3.2 The TLS Barriers

DoT and DoH eschew UDP and use TCP, resulting in a large overhead, and they employ a TLS encryption layer above the TCP layer. TLS is widely used in HTTPS and elsewhere, but it is certainly not lightweight, and it causes significant performance degradation with protocols that necessitate the high-speed exchange of small amounts of data, such as DNS.

*1 For example, JPRS Topics & Columns, "Aratanaru DNS cache poisoning no kyoui: Kaminsky Attack no shutsugen" [New DNS cache poisoning threat: Emergence of the Kaminsky Attack] (https://jprs.jp/related-info/guide/009.pdf, in Japanese only)

*2 Randomizing the source used by clients when sending queries increases the number of factors that an attacker needs to guess when seeking to forge packets, which reduces the probability of a successful attack.

Although logic says we will take a significant performance hit compared with traditional UDP-based DNS, we need tools if we are to measure how much performance actually degrades. TCP-based DNS has been in use for some time, albeit only in limited cases, and tools do exist. But DNS over TLS is completely new. And there are no satisfactory tools for measuring performance. We needed to measure how much performance degrades and what the processing load would be so that we could estimate just how much equipment we would need to provide the service, and to do this we had to start by developing performance measurement tools.

Since TLS involves a very high processing load, options are available to reduce the overhead by, for example, reusing information from a previous connection to resume a session (TLS session resumption) and, in TLS 1.3 (the latest version), adding the application data to the initial ClientHello/ServerHello exchange that takes place during the handshake (0-RTT).

But having certain functions available in the TLS protocol is meaningless unless applications actually use them. Running services in a large-scale production environment will be difficult unless you make full use of these options.

The IIJ Public DNS Service uses Unbound, a DNS implementation developed by NLnet Labs in the Netherlands. It is quite old and supported TLS before DoT became an IETF draft. When we investigated Unbound's TLS support, however, we learned that it lacked those mechanisms for reducing the TLS overhead; specifically, it lacked TLS session resumption capabilities. And performance measurements also indicated that performance was inadequate. Further, performance differs greatly depending on what encryption algorithm is used, but this was hardcoded into the source. So IIJ decided to implement the necessary features. The results were passed back to NLnet Labs, and the latest available version now includes our code.

Aside from performance issues, TLS poses one more barrier, namely that communications are encrypted.

To ensure a stable DNS service, you need a mechanism that allows you to collect statistics to make sure that a large number of abnormal queries are not being sent, and that abnormal responses are not increasing despite the queries being normal, so that you can investigate and take action if abnormalities do arise. With traditional DNS, in most cases this sort of statistics gathering and troubleshooting was done not on the DNS server itself but by capturing DNS packets. Since this process can be performed independently of the DNS server, the same method can be used regardless of what DNS server implementation is used.

But with TLS, the captured packets are encrypted. Perfect Forward Secrecy (PFS) is now commonplace, so packets cannot be decrypted even if you have the server's private key. This means that the tools so far used to collect information are no longer viable. The ability to collect statistics is needed for private testing of course, and it is indispensable if you plan to make something widely available as a service, so much so that you may as well shelve the service without it. So we had to rebuild the statistics collection functionality from the ground up to enable us to gather the same sort of information from the DNS server without relying on packet capture, and with this in place we were finally ready to launch the service.

### 2.3.3 The HTTP barrier

The barrier posed by HTTP was actually not all that high.

With DoH, once TLS has been applied, the DNS message must then be encapsulated in an HTTP message. Getting the DNS server itself to speak HTTP would be quite a

challenge, but a two-stage setup where an ordinary HTTP server receives the queries, converts the message format, and passes them to a DNS server behind it would be pretty much just like that used by any other Web application out there, aside from the fact that the backend is a DNS server.

Clearly because we are using TCP instead of the traditional UDP, as well TLS, the latency and other performance issues are unavoidable. But because we take care of the hard parts in the HTTP layer, which has a proven track record, and not in the DNS layer, which would require us to fumble around in unfamiliar territory, this should not be seen as too much of a setback.

#### 2.3.4 Have we overcome the barriers?

As of this writing, the service has been live for about six months.

We did create a bit of a stir when we published the press release since this is the first DoT/DoH service in Japan, and the Android DoH client Intra[*3] subsequently added an IIJ Public DNS list option for users to select (we did not even have to ask for the option to be added).

I wish we could give a glowing account of how everything went off without a hitch post launch, but reality does disappoint.

As explained, Unbound did not have TLS session resumption, so we implemented it ourselves, but TLS session validity appears to be extremely short on Android 9, and our session resumption implementation has proved ineffective in many cases. When Android smartphone users visit a Web page and then follow a link from that page to another one, often the TLS session established to resolve the previous domain name has already timed out and the whole process has to start again with a handshake. This handshake frequently fails if the network is congested, which means that users are constantly unable to use the Internet because a name cannot be resolved.

Moreover, it looks like performance is even worse on Android 10 (currently under development as of this writing) than it is on Android 9.

Other than this, we have had not major problems. Latency should theoretically be worse than with traditional DNS, but it does not appear to be a problem in practice and we have not received any complaints.

DoT and DoH are still new technologies and the basic specifications have only just been finalized. A lot of the peripheral specifications are yet to be sorted out (for instance, DoT/DoH servers presently can only be configured by hand; network administrators are unable to distribute configurations for automatic deployment).

Going forward, we will continue to investigate issues with the aim of making improvements and implement new specifications that are on the road to becoming standards in the hopes that people will be able to use the latest technologies with peace of mind. And we will continue to give back by communicating the insights we glean from operating our service to the community.

## 2.4 Public DNS and DoT/DoH

In closing, we look at developments beyond the IIJ Public DNS Service.

Originally, caching DNS servers were only available to users within a particular organization. They were not public in nature. But the realization that making them public would not really be all that harmful eventually led to most of them becoming available without restriction (open resolvers). Later, however, the DNS amp attack was discovered and attackers

*3 Intra (https://play.google.com/store/apps/details?id=app.intra)

began to use open resolvers to stage DDoS attacks. So since around 2010, the scope of access has generally been restricted to the minimum necessary.

Meanwhile, other services aimed at a wide range of users sought to foil attacks by implementing rate limitations and the like, rather than address-based restrictions. The pioneer here is OpenDNS[*4], but this was later followed by the Google Public DNS[*5] service, and ever since that gained traction, these sorts of explicitly open resolvers have generally been called "public DNS" services.

The DoT RFC became a standard in 2016. The public DNS service Quad9[*6], launched in November 2017, and Cloudflare[*7], launched in April 2018, supported DoT from the start, and Google also added support in January 2019.

DoH officially became an RFC in October 2018, but implementations based on the draft came out ahead of that. Cloudflare provided support upon launch in April 2018, Quad9 added support two weeks before RFC 8484 was released, and Google also later added support in June 2019.

On the client side of things, Android has supported DoT at the OS level since August 2018, and the DoH client app Intra was released in October 2018. Among Web browsers, Firefox added DoH support in August 2018. In Chrome's case, only the development version supported it as of this writing, but the stable version may also have it by the time this is published.

Mozilla has stated that it plans to make DoH the default for name resolution in Firefox[*8], meaning that the public DNS services selected by Firefox would automatically be used unless the user configures the browser to do otherwise. But public DNS services, because they are public, cannot resolve namespaces such as those found on intranets. And services like parental controls that use DNS will be ineffective if the browser selects its own DNS server instead of what the OS settings dictate. In light of these points, there is debate about the pros and cons of imposing DoH as default.

Google Public DNS supports DoT/DoH, and use of these protocols ensures confidentiality between the user and Google. Meanwhile, Google also supports EDNS0 Client Subnet (ECS; RFC 7871). With ECS, when a client queries a caching server, the caching server relays information about the network to which the client belongs to the authoritative server. This is intended to be useful in content delivery traffic management. But it is important to note that traditional DNS is always used between caching servers that use ECS and the authoritative server. If the user is using DoT/DoH, communications on the route between the user and Google will not be eavesdropped, but communications on the route between Google and the authoritative server can be eavesdropped because traditional DNS, which does not guarantee confidentiality, is used here, and this could result in a breach of user privacy since the ECS information traverses that route.

There is now a definite move toward encrypting DNS transport, and it is unlikely to be stopped at this point. But as providers of a new public DNS service, we have an important responsibility not to blindly accept every development that unfolds but to evaluate each one by one and determine if it really is the right way to proceed.

**Takanori Yamaguchi**
Application Service Section, Application Support Department, IIJ. Mr. Yamaguchi works on support for DNS services etc.

*4 OpenDNS (https://www.opendns.com/)
*5 Google Public DNS (https://developers.google.com/speed/public-dns/)
*6 Quad9 (https://www.quad9.net/)
*7 Cloudflare (https://developers.cloudflare.com/1.1.1.1/)
*8 Firefox Nightly News, "What's next in making Encrypted DNS-over-HTTPS the Default" (https://blog.mozilla.org/futurereleases/2019/09/06/whats-next-in-making-dns-over-https-the-default)

November 2019 Vol.44

Internet Initiative Japan

**About Internet Initiative Japan Inc. (IIJ)**

IIJ was established in 1992, mainly by a group of engineers who had been involved in research and development activities related to the Internet, under the concept of promoting the widespread use of the Internet in Japan.
IIJ currently operates one of the largest Internet backbones in Japan, manages Internet infrastructures, and provides comprehensive high-quality system environments (including Internet access, systems integration, and outsourcing services, etc.) to high-end business users including the government and other public offices and financial institutions.
In addition, IIJ actively shares knowledge accumulated through service development and Internet backbone operation, and is making efforts to expand the Internet used as a social infrastructure.


IIJ-MKTG020-0042

**Internet Initiative Japan Inc.**

Address: Iidabashi Grand Bloom, 2-10-2 Fujimi, Chiyoda-ku,
Tokyo 102-0071, Japan
Email: info@iij.ad.jp URL: https://www.iij.ad.jp/en/