IIR

Internet
Infrastructure
Review

Nov.2024
Vol. 64

Periodic Observation Report

Broadband Traffic Report Looking Back on the Past Five Years

Focused Research(1)

Evolution of Virtualization Technology and IIJ's Initiatives

IIJ
Internet Initiative Japan

Internet Infrastructure Review

November 2024 Vol.64

IIR Nov.2023 Vol. 64

Executive Summary

Executive Summary

You will no doubt be aware of the extensive global disruption sparked on July 19 (US time) due to an error in a channel file distributed by CrowdStrike. The error affected Windows devices (PCs, servers, etc.) that use CrowdStrike's Falcon sensor, and according to Microsoft (in a post titled "Helping our customers through the CrowdStrike outage" on the Official Microsoft Blog), the incident affected 8.5 million Windows devices, or less than one percent of all Windows machines.

That there was a huge impact on financial, aviation, medical, and other systems that underpin our society despite only one percent of all Windows devices being affected may seem surprising, but this is evidence of how widely CrowdStrike's security products were deployed on devices used in critical operations.

According to the Root Cause Analysis published by CrowdStrike (Channel-File-291-Incident-Root-Cause-Analysis-08.06.2024.pdf (crowdstrike.com)), there were, broadly speaking, two problems. First, insufficient checks were performed. As it is not realistically possible to reduce human error to zero, it goes without saying that it is crucial to perform multi-stage checks to prevent errors being missed, thereby minimizing their impact. Second, there is a need for staged deployment, or canary releases. Our observations also indicate that the malfunctions occurred in sequence starting with the devices that received the new channel file first. So if the problem had been detected early via a staged deployment—by limiting the number of devices the file was initially deployed to, for example—then it may have been possible to take countermeasures of some kind before it had such a large impact.

As an additional mitigation, CrowdStrike noted that it has engaged independent third-party reviewers. Multi-stage checks, canary releases, and third-party reviews are something that we always strive for when developing and operating our own systems as well.

Today's computer systems continue to grow in complexity. This incident has reminded us that it is becoming increasingly important for us to be mindful that errors can and will happen, and to think about how to prevent them from being missed and how to reduce the scope of their impact when designing systems and operations.

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.

The periodic observation report in Chapter 1 is our broadband traffic report for the year, providing our analysis of IIJ's fixed broadband and mobile traffic. Our observations indicate that overall traffic volume on both broadband and mobile services continues to grow steadily, and that the proportion of traffic accounted for by TCP port 443 (HTTPS) and UDP port 443 (QUIC) is rising, consistent with the trends we observed over the past few years. In this edition, we also take another look back at the past five years. While the changes observed from year to year are not all that notable, the data reaffirm that the cumulative changes over that past five years have been large enough to have a decent impact on the infrastructure.

Chapter 2 presents a focused research report on the evolution of virtualization technology and IIJ initiatives in this area over the years. Virtualization technology has a long history, but the technology has made significant advances alongside the spread of Intel Architecture servers and cloud computing since 2000, resulting in explosive growth in its use. The report traces the history of virtualization technology, which underpins today's cloud computing systems, back to the 1960s, and also looks at what sort of virtualization technologies are used on IIJ's cloud services and the features that IIJ has implemented.

Through activities such as these, IIJ continues striving 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 that our customers can take full advantage of as infrastructure for their corporate activities.

Junichi Shimagami

Mr. Shimagami is a Director and Senior Managing 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's MVNO Council, stepping down from that post in May 2023. In June 2021, he also became a vice-chairman of the association.

Broadband Traffic Report
Looking Back on the Past Five Years

1.1 Overview

In this report, we analyze traffic over the broadband access services operated by IIJ and present the results each year[*1*2*3*4*5]. Here, we again report on changes in traffic trends over the past year, based on daily user traffic and usage by port. Then in the latter part of this report, we take a look back at changes in traffic over the past five years, encompassing the COVID-19 pandemic.

Overall, traffic continued to grow steadily this year, similar to last year. We see no notable changes in the trends at this point.

Figure 1 plots 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 OUT observations for January 2020, just before the pandemic, for both services to 1.

Over the past year, broadband IN traffic increased 14% and broadband OUT traffic increased 12%. The corresponding year-earlier figures were 11% and 18%.

The broadband figures include IPv6 IPoE traffic. IPv6 traffic on IIJ's broadband services comprises both IPoE and PPPoE traffic. As of June 2024, IPoE accounted for a bit under 50% of all traffic, at 43% of IN and 48% of OUT broadband traffic overall, year-on-year increases of 1 percentage point for IN and 4 points for OUT. As is evident from the graph, PPPoE traffic has been range-bound since 2020, and IPoE is driving the increase in traffic.

Mobile services traffic was largely range-bound in the first year or so of COVID as people went out less, but it has subsequently been in an uptrend. Over the past year, mobile IN traffic increased 29% and mobile OUT traffic increased 20%. The year-earlier figures were 27% and 31%. This year, OUT reached 2.1x the January 2020 level, and as Figure 1 shows, post-COVID growth has caught up with that for broadband.

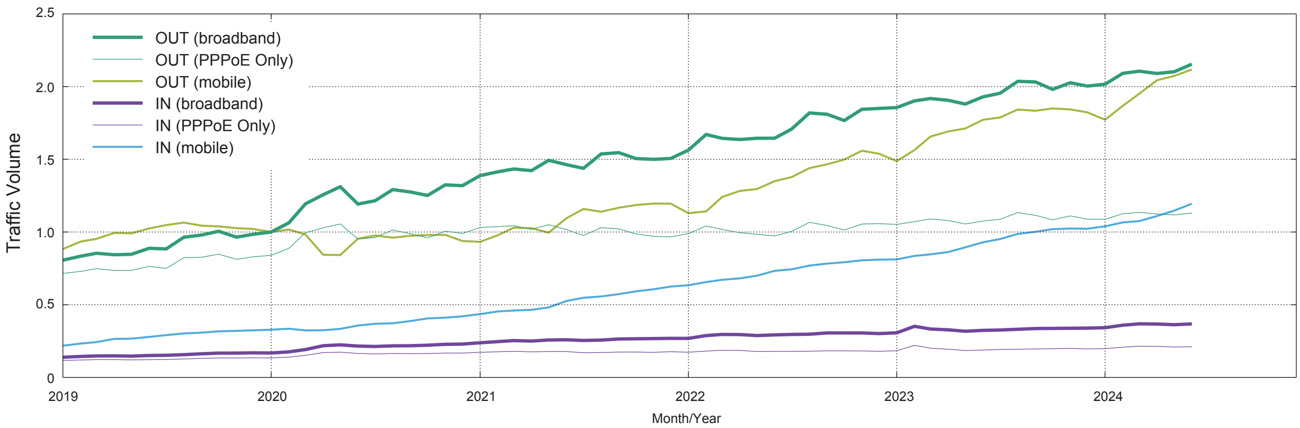

Figure 1: Monthly Broadband and Mobile Traffic

*1 Kenjiro Cho. Broadband Traffic Report: Broadband Traffic Report: Traffic in a Stable Uptrend Post-COVID. Vol. 60. pp4-9. September 2022.
*2 Kenjiro Cho. Broadband Traffic Report: Broadband Traffic Report: COVID's 3rd Year Brings Lull in Traffic. Vol. 56. pp4-11. September 2022.
*3 Kenjiro Cho. Broadband Traffic Report: Broadband Traffic Report: COVID-19's Impact in its 2nd Year. Vol. 52. pp4-11. September 2021.
*4 Kenjiro Cho. Broadband Traffic Report: The Impact of COVID-19. Vol. 48. pp4-9. September 2020.
*5 Kenjiro Cho. Broadband Traffic Report: Moderate Growth in Traffic Volume Ongoing. Vol. 44. pp4-9. September 2019.

Mobile services IN traffic accounts for a high proportion of total because of the high volume of uploads on services for enterprise customers. Looking solely at personal services, IN accounts for around a tenth of the total, similar to the situation for broadband.

We now look at broadband traffic by time of day on weekdays over the past year. Figure 2 plots hourly average traffic volume for Monday–Friday for four one-week blocks selected at intervals of roughly four months since May 2023. Weekday daytime traffic volumes have increased during school holiday periods in recent years, so we selected school weeks. Traffic volume here is the sum of PPPoE and IPoE. The dotted lines in the lower part of the plot represent uploads for each week, but focusing again on download volumes in this edition, we see that traffic volumes were up across all times of the day.

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 June 3 – 9, 2024, and compare those data with data for the week of May 29 – June 4, 2023, which we analyzed in the previous edition of this report.

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 addresses assigned to users with the IP addresses observed. Note that IPoE traffic is not included in the analysis of traffic by port, as detailed data are not available because we use Internet Multifeed Co.'s transix service for IPoE.

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.

Since our 2019 report, we have used daily usage data only on services provided to individuals. The distribution is heavily distorted if we include enterprise services, where usage patterns are highly varied. So to form a picture of overall usage trends, we determined that using only the personal user data would yield more generally applicable, easily interpretable conclusions. Note that the analysis of usage by port in the next section does include enterprise data because of the difficulty of distinguishing between individual and enterprise usage. Note also that we have included IPoE user data in the broadband figures since 2021, so the broadband data comprise both PPPoE and IPoE[*6].

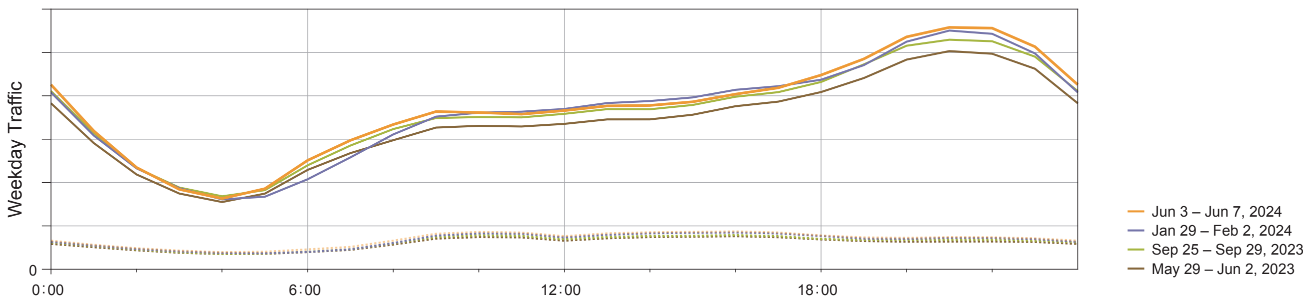

Figure 2: Hourly Average Broadband Traffic on Weekdays in the Past Year

*6 The PPPoE and IPoE usage figures of users who use both protocols are treated as coming from separate users.

Figures 3 and 4 show the average daily usage distributions (probability density functions) for broadband and mobile users. Each compares data for 2023 and 2024 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 (10^4) and 100GB (10^{11}) using a logarithmic scale. Most users fall within the 100GB (10^{11}) range, with a few exceptions.

The IN and OUT traffic distributions in the figures 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. 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.

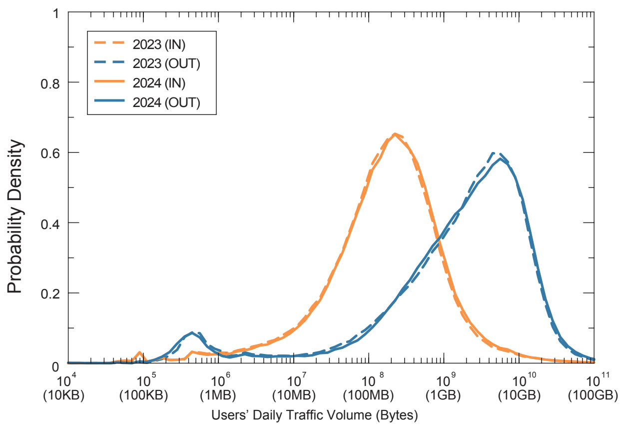

Figure 3: Daily Broadband User Traffic Volume Distribution Comparison of 2023 and 2024

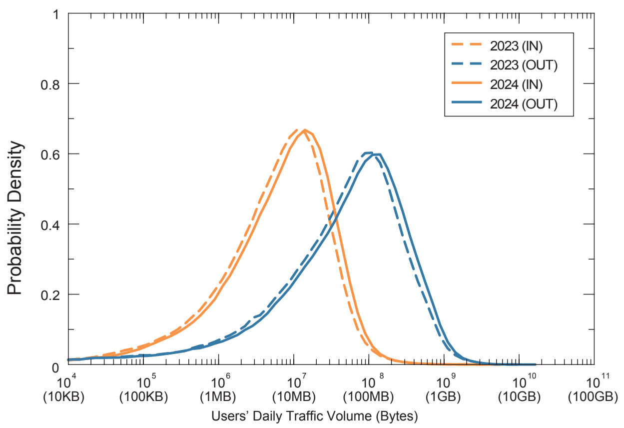

Figure 4: Daily Mobile User Traffic Volume Distribution Comparison of 2023 and 2024

First, we look at the broadband distributions in Figure 3. Comparing 2023 and 2024, both the IN and OUT distributions have moved ever so slightly to the right, but you can see that overall traffic volume is largely unchanged.

The peaks of the mobile distributions in Figure 4 have moved a little to the right since last year, indicating that overall traffic has increased. Mobile 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. 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.

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). When the peak is slightly off the center of the distribution, the mode is adjusted to bring it toward the center. Comparing 2023 and 2024, the IN mode remained

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
2020	609	122	158	3810	1638	3162
2021	714	143	200	4432	2004	3981
2022	727	142	178	4610	2010	3981
2023	804	166	224	5456	2369	5012
2024	834	178	224	5743	2372	5620

at 224MB while the OUT mode rose from 5,012MB to 5,620MB, translating into growth factors of 1.00 for IN and 1.12 for OUT. Meanwhile, because the means are influenced by heavy users (on the right-hand side of the distribution), they are significantly higher than the corresponding modes, with the IN mean at 834MB and the OUT mean at 5,743MB in 2024. The 2023 means were 804MB and 5,456MB, respectively. As mentioned, up to 2020 the data covered only PPPoE users, and since 2021 the data have covered both PPPoE and IPoE users.

Table 2 shows the mobile traffic metrics. In 2024, the IN mode was 14MB and the OUT mode was 112MB, while the means were IN 16MB and OUT 150MB. The 2023 modes were IN 11MB and OUT 100MB, and the means were IN 14MB and OUT 129MB.

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
2020	10.4	4.5	7.1	79.4	35.1	63.1
2021	9.9	4.7	7.9	85.9	37.9	70.8
2022	12.8	6.0	10.0	113.7	49.2	89.1
2023	14.1	6.8	11.2	129.2	56.0	100.0
2024	16.3	8.2	14.1	150.4	66.7	112.2

Figures 5 and 6 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. Variability between users in terms of usage levels and IN/OUT ratios is wide, indicating that there is a diverse range of usage styles. For mobile traffic, the pattern of OUT being an order of magnitude larger also applies, but usage volumes are much lower than for broadband. For both broadband and mobile, there appears to be almost no difference between these plots and those for 2023.

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 account for 50% of total OUT and 75% of total IN traffic, while the top 1% of users account for 15% of OUT and 46% of IN traffic. On mobile, the top 10% of users account for 48% of total OUT and 45% of total IN traffic, while the top 1% of users account for 12% of OUT and 13% of IN traffic.

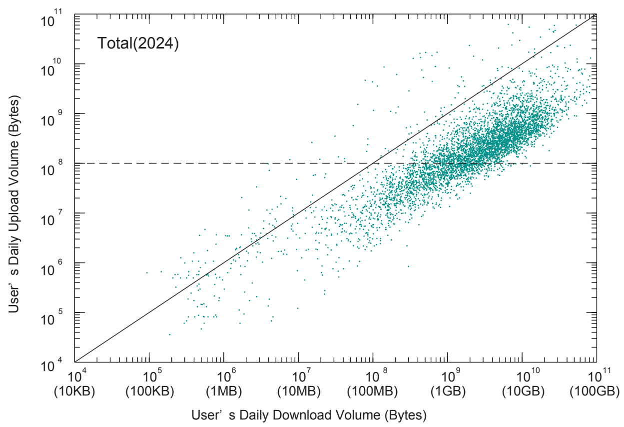

Figure 5: IN/OUT Usage for Each Broadband User

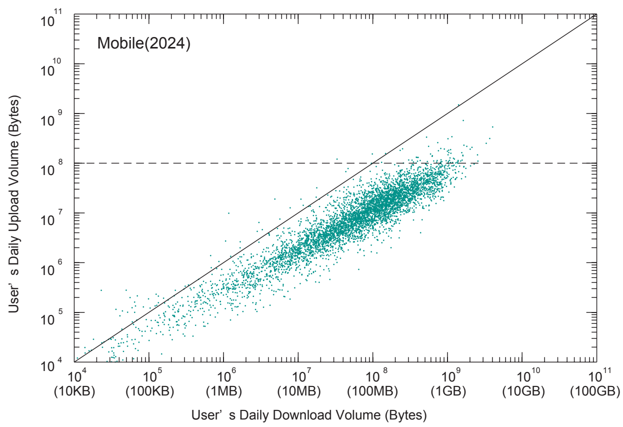

Figure 6: IN/OUT Usage for Each Mobile User

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 HTTP ports like port 80 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 2024, 68% of all traffic was over TCP connections, down 3 points from 2023. The proportion of traffic over port 443 (HTTPS) was 54%, a 3-point drop from last year. The proportion of traffic over port 80 (HTTP) was 7%, having declined ever so slightly. The figure for UDP port 443, which is used by the QUIC protocol, was up 3 points to 21%.

TCP dynamic port traffic rose ever so slightly to 6%. Individual dynamic port numbers account for only a tiny portion, with the most commonly used port 31000 only making up 1.2%.

Table 4 shows the percentage breakdown by port for mobile users. The figures are close to those for broadband on the whole. This is possibly because apps similar to those for PC platforms are now also used on smartphones, and because the proportion of broadband usage on smartphones is rising.

The broadband port data only include PPPoE, not IPoE, and so do not necessarily reflect the trend in fixed broadband overall. Comparing IPv4 and IPv6 on mobile, port 443 accounts for a higher proportion of both TCP and UDP usage on IPv6, and there is probably a similar trend in the case of IPoE.

Table 3: Broadband Users' Usage by Port

year	2020	2021	2022	2023	2024
protocol port	(%)	(%)	(%)	(%)	(%)
TCP	**77.2**	**71.9**	**71.6**	**70.5**	**67.5**
(< 1024)	70.5	65.8	65.4	64.8	61.1
443 (https)	52.4	53.5	55.7	56.9	53.8
80 (http)	17.2	11.6	8.9	7.2	6.5
183	0.0	0.1	0.2	0.2	0.2
993 (imaps)	0.2	0.1	0.1	0.1	0.1
22 (ssh)	0.2	0.2	0.1	0.1	0.1
(>= 1024)	6.7	6.1	6.2	5.7	6.4
31000	0.4	0.6	0.9	1.1	1.2
1935 (rtmp)	0.4	0.2	0.2	0.2	0.3
8080	0.4	0.4	0.3	0.4	0.3
UDP	**19.4**	**24.5**	**24.3**	**25.4**	**28.2**
443 (https)	10.5	15.9	16.3	18.2	21.0
4500 (nat-t)	0.6	0.8	0.8	1.0	0.9
8801	1.1	0.9	0.6	0.4	0.4
ESP	**3.2**	**3.3**	**3.8**	**3.8**	**4.0**
GRE	**0.1**	**0.2**	**0.2**	**0.1**	**0.2**
IP-ENCAP	**0.1**	**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	2020	2021	2022	2023	2024
protocol port	(%)	(%)	(%)	(%)	(%)
TCP	**75.5**	**70.3**	**71.6**	**71.0**	**71.0**
443 (https)	50.7	44.4	42.3	42.1	42.2
80 (http)	7.4	5.0	4.1	3.5	1.8
993 (imaps)	0.2	0.2	0.1	0.1	0.1
1935 (rtmp)	0.1	0.1	0.1	0.2	0.1
UDP	**18.0**	**23.8**	**24.4**	**26.5**	**27.5**
443 (https)	9.3	16.3	17.9	20.9	22.5
4500 (nat-t)	1.8	3.7	2.7	2.5	1.8
51820	0.0	0.0	0.1	0.2	0.3
53 (dns)	0.1	0.2	0.2	0.2	0.2
8801	1.4	0.7	0.3	0.2	0.1
ESP	**6.4**	**5.8**	**3.9**	**2.4**	**1.4**
GRE	**0.1**	**0.1**	**0.0**	**0.0**	**0.0**
ICMP	**0.0**	**0.0**	**0.0**	**0.1**	**0.0**

Figure 7 compares overall broadband traffic for key port categories across the course of the week from which observations were drawn in 2023 and 2024. We break the data into four port buckets: TCP ports 80 and 443, dynamic TCP ports (1024 and up), and UDP port 443. The data are normalized so that peak overall traffic volume on the plot is 1. The overall peak is around 19:00–23:00. There are no major changes overall relative to 2023, but traffic on UDP port 443 increased a little.

Figure 8 shows the trend for TCP ports 80 and 443 and UDP port 443, which account for the bulk of mobile traffic. As was the case with broadband, mobile traffic on UDP port 443 was up slightly compared with 2023. Comparing the plots with those for broadband, usage times evidently differ, with mobile having three separate traffic peaks on weekdays: morning commute, lunch break, and evening.

1.5 Looking Back on the Past Five Years

Let's take a look back at trends in broadband traffic volumes over the past five years, which encompassed the COVID-19 pandemic.

First, we'll look at average monthly traffic volume on IIJ's fixed broadband services overall, as shown in Figure 1. Over the five years from June 2019 to June 2024, IN traffic increased 1.72x and OUT traffic increased 2.01x, which work out to annual growth rates of 1.11x and 1.15x.

Next, Figure 9 displays five years of data for the weekday broadband traffic volumes shown in Figure 2. Since we only have hourly data remaining for weeks we have analyzed in the past, Figure 9 only compares those specific weeks, starting with February 2020 just before the COVID-19 pandemic, and then weeks covering periods in late May to mid-June in each year thereafter. It is evident from the graph that traffic has increased fairly evenly across all time periods.

Figure 10 is an enlarged view of the upload data from Figure 9. The proportion of traffic occuring during the daytime on weekdays has clearly increased vs. before COVID. While download traffic peaks from evening through to nighttime, uploads peak in the early afternoon. While it is difficult to tell from the graph because the values for each hour are

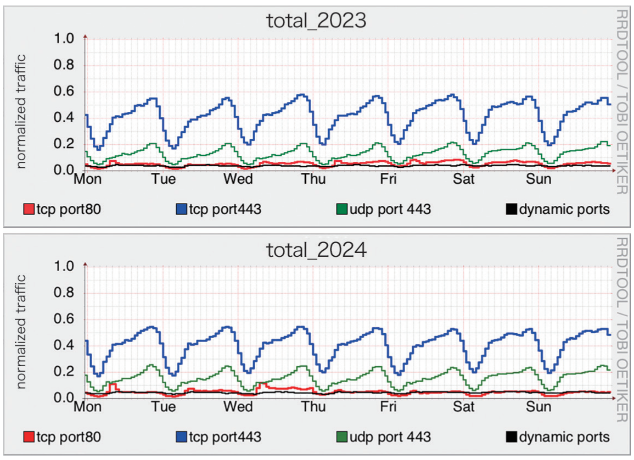

Figure 7: Broadband Users' Port Usage Over a Week
2023 (top) and 2024 (bottom)

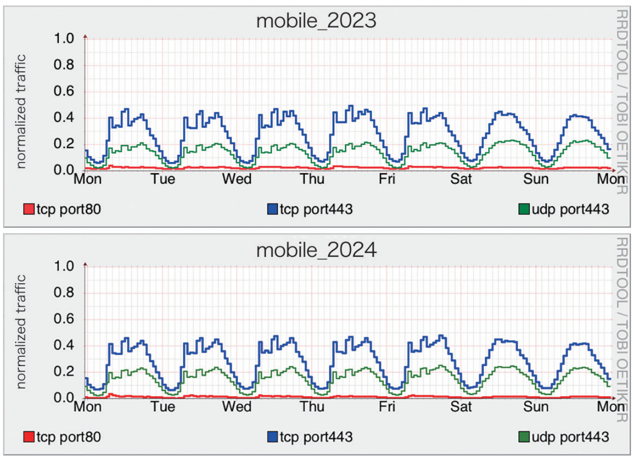

Figure 8: Mobile Users' Port Usage Over a Week
2023 (top) and 2024 (bottom)

displayed using a discrete line plot, traffic declines from 12:00 to 13:00 on weekdays. This drop during lunch-time probably reflects a drop in remote work-related usage, especially video conferencing. The annual increase in upload traffic has also been fairly constant.

Figure 11 shows five years of data for the distribution of daily traffic volume, which we looked at in Figure 3. These are the values for end-May through early June each year that we analyze in this report. The largest increase was from 2019 to 2020, with a particularly large increase in uploads. Since 2020, however, we have seen a relatively stable increase. Over the five years from 2019 to 2024, the mode (the most frequent value, which represents the peak of the distribution) has increased 2.8x for OUT (download), from 2.0GB to 5.6GB, and 2.5x for IN (upload), from 89MB to 224MB.

In past reports from the COVID days[*3][*4], we reported that traffic volumes fluctuated heavily in response to changes in rates of people staying at home amid the spread of infections. During the first State of Emergency Japan declared during COVID, restrictions on leaving the home meant that people were only able to engage in activities online, and the rapid uptake of video conferencing and video streaming resulted in a large increase in traffic volumes. At first, there were concerns this would put a strain on the Internet infrastructure. Looking at the long-term trend, however, traffic volume has increased almost uniformly each year, and the rate of increase is by no means large. Yet it is evident that some changes in our daily lives have also taken hold—for instance, traffic that may reasonably be attributed to remote working on week-days and video streaming during long school holidays is increasing year after year.

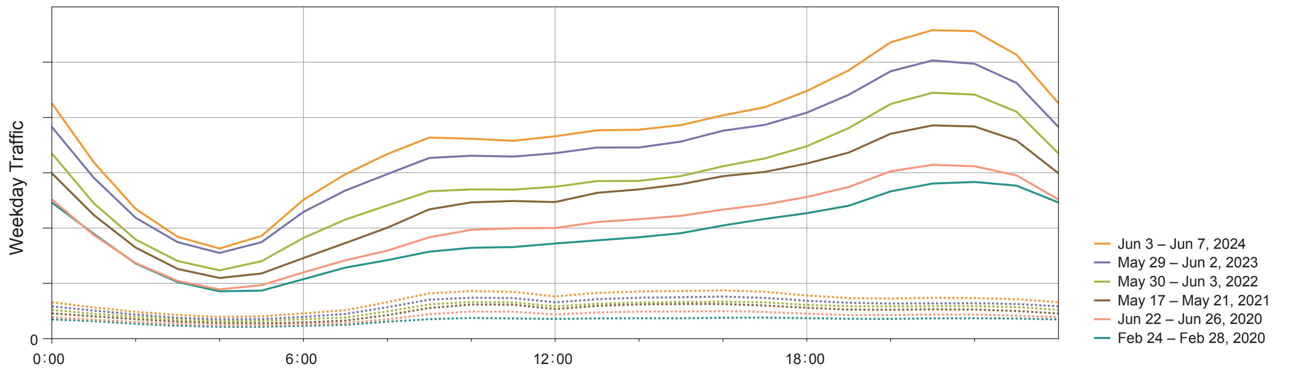

Figure 9: Hourly Average Broadband Traffic on Weekdays in the Five Past Years

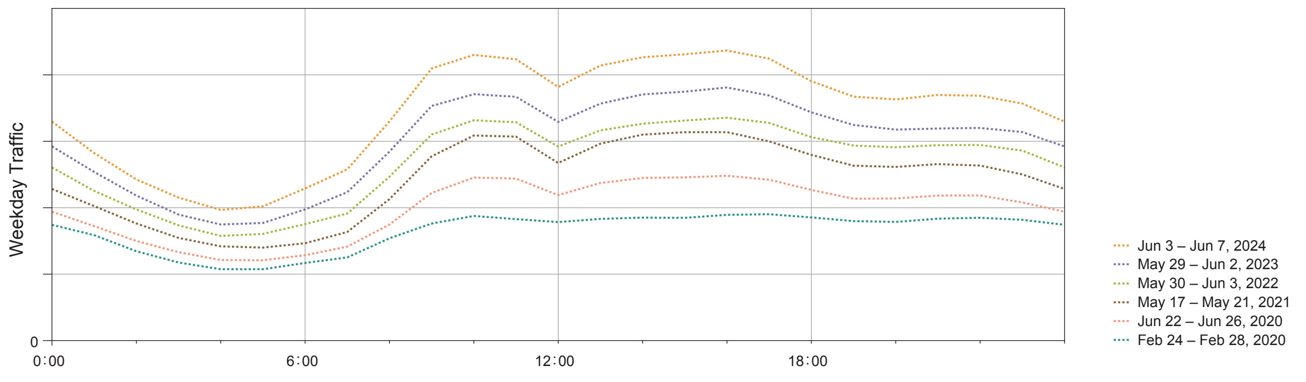

Figure 10: Hourly Average Broadband Traffic (Uploads) on Weekdays in the Five Past Years

1.6 Conclusion

Broadband traffic has been increasing relatively steadily over the past few years, with very little change in the overall trend. But while the changes from year to year may be small, over the span of five years, the cumulative changes do have an impact.

I am writing this report in mid-July, just before the Paris Olympics. When writing my installment of this report five years ago, I never dreamed that something like COVID would befall us. I just had a vague idea that online streaming would probably increase the following year with the Tokyo Olympics set to take place. Today we take things like online sports broadcasts, remote working, and the idea of doing all sorts of clerical tasks online for granted, but that wasn't the case five years ago. We are now able to do more over the Internet than we could have imagined back then. Much still needs to be improved, of course, but as I write this I am again reminded that the Internet continues to change the way we live.

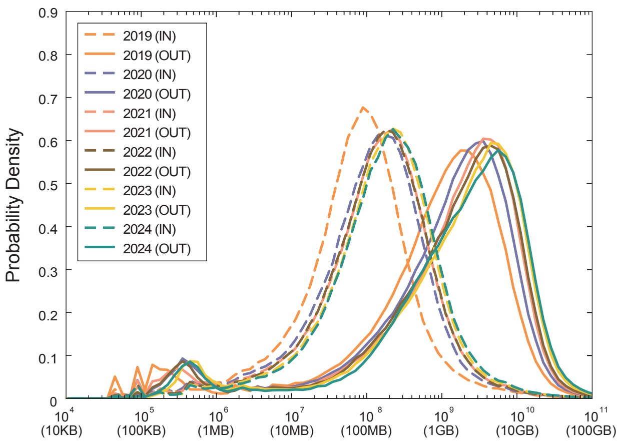

Figure 11: Daily Broadband User Traffic Volume Distributions in the Past Five Years

Kenjiro Cho
Research Director, Research Laboratory, IIJ

Evolution of Virtualization Technology and IIJ's Initiatives

2.1 Introduction

The concept of the "cloud" has become such an ordinary part of our lives that we barely give it a second thought anymore. Today, companies all over the world offer cloud services, and the infrastructure underpinning those services continues to expand.

We covered the evolution of IIJ's cloud services platforms in a previous article[*1]. In this edition, we look at the ever evolving virtualization technology that has made the concept of cloud computing possible.

2.2 History of Virtualization Technology

2.2.1 Early Days of Virtualization

Virtualization is the abstraction of computer resources, and encompasses various technologies for providing an abstraction layer between software and physical hardware and managing those resources. Its history goes all the way back to the 1960s.

The word "virtualization" can be traced back to earlier stages of computer development. In the 1940s and 50s, different computer models commonly had different architectures, and so to mitigate the risks associated with new designs, computer designers began to use past models as a reference to create "compatible machines" that used instruction sets compatible with those of previous models and that shared the same logical design.

The late 1950s saw the advent of computer models that emulated the instruction sets of previous-generation models in microcode so as to provide backward compatibility, and the 1960s brought a rising trend toward standardizing computer architectures and maintaining compatibility. It was around this time that what were called "virtual machines" (the term was actually used earlier than this) came to be used to provide compatibility by emulating the instruction sets of different models of computer, and thus the word "virtualization" entered the vernacular.

It was then in 1964 that the first "hypervisor," capable of running multiple virtual operating systems on a single computer, appeared. As you may know, however, we would then have to wait until the 2000s before virtualization became further widespread.

Virtualization at the time was intended to allow multiple users to use a single large and expensive computer. The use cases revolved around corporate departments that used it to run thousands of routine tasks, such as batch payroll processing, at high speed.

Over the next few decades, a number of other approaches were taken to solve the problem of allowing multiple users to use a single device. One such approach was time sharing, which offered the solution of separating users within the one operating system (OS), rather than using hardware-based virtualization to logically separate users into different virtual machines. This meant the virtualization of the time was relegated to certain niche areas, and hardware-based virtualization thus faded from the forefront for a time. Time sharing is what created the impetus for the creation of UNIX and Linux, which we use in our systems today.

2.2.2 x86 Virtualization Sparks a Resurgence in Virtualization Technology

In the 1990s, many companies adopted vertically integrated systems (mainframes) comprising physical servers and software from a single vendor. Existing applications could not be run on hardware provided by other vendors.

Meanwhile, Intel Corporation, which rose to prominence with its x86 architecture being used in IBM's PCs in the 1980s, unveiled the Pentium CPU in 1993. It was also

*1 Internet Infrastructure Review (IRR) Vol. 60, Focused Research(2) "Evolution of the IIJ Cloud—Commemorating 30 Years" (https://www.iij.ad.jp/en/dev/iir/060.html).

around this time that the market for personal computers underwent a rapid expansion amid the spread of commercial Internet services and the release of Microsoft Windows. One after another, manufacturers who had previously developed computers based on their own proprietary architectures began developing compatible machines based on Intel CPUs.

Intel followed up in 2001 with the release of Xeon CPUs for servers. This prompted moves to replace corporate systems, which had until then been dominated by mainframes and UNIX, with low-cost, general-purpose IA (Intel Architecture) servers.

Intel's CPU architecture, however, was not designed for use with multiple operating systems and thus rendered existing forms of virtualization inapplicable. It was in this context that VMware Virtual Platform, released by VMware Inc. on February 8, 1999, made it possible to achieve virtualization in software.

The earliest form of software virtualization for the x86 architecture (x86 virtualization) worked by running a virtual layer on the host OS (Windows or Linux). The need to go through the host OS resulted in a large overhead, such that performance was not suitable for commercial services. This prompted the release of hypervisor products with their own custom kernels capable of providing a virtual layer without the need for a general-purpose host OS.

When operating without hardware virtualization support, x86 virtualization of this era always used a dynamic instruction conversion mechanism to capture the execution of specific instructions and dynamically replace them. This technique inherently came with some overhead in terms of performance compared with virtual machines on virtualization-friendly architectures.

Paravirtualization then arose as a means of solving this performance problem. Instead of emulating the hardware, paravirtualization involves modifying the guest OS to provide a special API. This was used in early forms of server virtualization. A key example of this approach was the open-source Xen (up to 2.x).

In contrast to this, full virtualization (native virtualization) worked by converting privileged CPU instructions in the virtualization layer to allow the guest OS on the virtual machine to run as is without any special modifications. This made it possible to run OSes like Windows that did not support paravirtualization at the time. Key products in this space included Windows Virtual Server, VMware Server, and VMware ESX.

2.2.3 Limitations of Software Virtualization and Emergence of Hardware-Assisted Virtualization

As discussed, the problem with past forms of virtualization was that operations performed by the CPU in hardware had to be executed in software, which created a non-trivial virtualization overhead. The I/O delays caused by the exchange of data between devices, OSes, and applications, in particular, were unsatisfactory. This is why the earliest forms of hypervisor-based virtualization were used somewhat sparingly for the purpose of running guest OSes migrated to run applications coming up against their limits due to equipment being too old, or for testing during application development. It did not have the power to replace the high-spec, high-speed mainframes and UNIX servers of the time.

IA servers were gaining market share in the late 2000s, and this is when Intel, struggling to increase the performance of its processors, switched to 64-bit multi-core CPU architecture. It was around the same time that Advanced Micro Devices (AMD) implemented hardware virtualization support, sparking rapid progress in the field of virtualization. This support made it possible to run the virtual machine monitor (VMM), part of the functionality that had previously been run in software, on the CPU, and this greatly reduced the overhead.

To describe its role in simple terms, the VMM does for the guest OS what the OS does for user applications. The OS manages physical resources such as devices and memory, and isolates processes (user applications) from each other in memory. Similarly, the VMM emulates physical resources for each guest OS, coordinates requests from the guest OSes, and isolates the guest OSes from each other in memory.

Speeding up these VMM tasks and thus improving the performance of the guest OSes helped hypervisor-driver server virtualization gain traction.

2.2.4 Maturation of Server Virtualization and Advent of the Cloud

Hypervisors that greatly reduced the virtualization overhead thanks to the implementation of virtualization support functionality in CPUs thus made significant inroads into commercial use, and products that previously provided guest OSes via paravirtualization mechanisms (such as Xen 3.0 and later) also started providing full virtualization of guest OSes, and thus the development of virtualization products suddenly became increasingly more competitive.

It was around this time that Eric Emerson Schmidt, CEO of Google Inc., first used the term "cloud computing" in a speech at the Search Engine Strategies Conference[*2], and thus the term "cloud" came to be used in the sense we are familiar with today.

Following this, Microsoft implemented Hyper-V, which provides support for CPU virtualization support functionality, as a Windows Server feature. Similarly, in the open-source world, version 1.0 of the Kernel-based Virtual Machine (KVM) was released, integrated into the Linux kernel. And alongside the earlier products in this space, these offerings helped accelerate the development of the virtualization market.

In 2008, Google unveiled Google App Engine (GAE), a PaaS offering designed for cloud computing. And alongside Amazon Web Services EC2 (AWS EC2), a service Amazon had launched earlier for the purpose of renting out surplus resources present after replacing IA servers within its own systems, this really opened up the cloud market. IIJ also launched a commercial cloud service under the IIJ GIO brand in 2010, and we continue to develop this and make it available today.

Once the hypervisor performance issue was resolved, attention turned to the difficulty of operating and managing virtualization products. Similar to hardware servers, when it came to configuring system elements such as network and storage, different manufacturers and devices all adopted different architectures (unlike the case with IA servers, where this had gradually grown more consistent over time), and so specialized knowledge was required to operate the products. Virtual machines also had to be moved between physical resources to improve availability, efficiency, and stability. Then came products called orchestrators, which abstract and allow overarching control of system elements that had previously been furnished separately in the form of physical resources. Key products here include VMware vRealize Orchestrator, OpenStack, and Apache CloudStack.

2.2.5 Microservices Architecture and OS-Level Virtualization

Thus, various technical solutions were deployed to overcome the challenges associated with hypervisors, and the business environment, which benefited from these technologies, was changing at an ever-increasing pace.

*2 Google Press Center, "Search Engine Strategies Conference Conversation with Eric Schmidt hosted by Danny Sullivan," August 9, 2006 (https://www.google.com/press/podium/ses2006.html).

While mechanisms for rapidly deploying cloud resources had been developed, the great number of infrastructure elements that software developers had to deal with, beyond the required libraries, meant that prodigious effort was needed to keep up with the demands of business.

To solve this problem, developers devised a software development methodology called microservices, which involves building applications out of collections of small, independent pieces of software. As only a small amount of resources were required—the environment to run the application—container technology was adopted to virtualize some, not all, of the OS functionality, and this led to the creation of container management systems, a new type of product incorporating virtualization capabilities. Key products here include Kubernetes and a derivative of it called Red Hat OpenShift Virtualization.

So as the technology has advanced through the years, user perceptions of the issues that needed to be addressed have changed, and virtualization technology has accordingly evolved and taken on new roles.

2.3 Technology Selection and Service Development

Technologies have solved various issues, and products incorporating those technologies have expanded the market, allowing us to diversify and differentiate our services, but as you know, not all products work to the desired level.

So just because a new product becomes available does not mean it will immediately be suited to the user environment. As discussed in the previous section, however, virtualization technology has been advancing rapidly in recent years, and given the lack of sufficient time to fully evaluate new technologies, we now face the need to start investigating such technologies at an earlier stage, rather than waiting for service requirements to be finalized, to identify elements that will determine the direction of development efforts. And even when it comes to products that have not currently been adopted for use in any services, there is always potential for the market to expand rapidly due to technological advances of one kind or another, so we continue to keep tabs on the markets around those products as well.

The actual product evaluation process consists of two steps: check the basic functionality, and conduct checks for each individual service. The basic functionality means the functions or features that the vendor claims have been implemented in the product, as well as their expected performance. It depends on the product being evaluated, but for the infrastructure layer, we look at the operation of physical equipment and the behavior of the catalog values or limit values, and for the virtualization layer, we look at the combination of equipment that makes up the product, the operation and performance of the hypervisor, and so on. The way we evaluate products is also not set in stone; instead, we incorporate knowledge and insight gained from operating services (mainly availability and completeness as it relates to quality) and gradually increase the number of common evaluation criteria. When it comes to conducting checks for each individual service, our basic approach is to evaluate the functional and non-functional aspects for fulfilling SLAs based on service requirements. This involves looking at the parts that, when assembled into a system, will be provided to users and the parts that we will be operating and managing (not just maintenance but also the mechanisms for distributing resources to users etc.). Our main focus is integration with relatively higher-level software and applications.

There is another important consideration when it comes to operating a service and maintaining quality. The emergence of new products determines the lifespan of existing or older-generation products, so you must consider how you will handle changes to systems that use such products. All is well if successor products stick with established mechanisms and frameworks, but when they adopt a completely different architecture or a completely different license agreement, you have to determine whether you will be able to continue providing your service using the new product. To ensure you can deal swiftly with such situations, it is crucial to determine what level of compatibility exists between different products, in terms of the features a product shares in common with similar products and how they perform, rather than focusing on the unique characteristics of a given product. From these perspectives, identifying and understanding technologies early is crucial for the purpose of sorting out both the functional and non-functional requirements of competing products on the market.

Finally, when looking at products to provide as part of a service, we evaluate cost. Even if a product fulfills users expectations of functionality, we will not adopt it if the cost seems inappropriate for purpose. This applies both to third-party and IIJ products—if they are not suitable in this respect, we do not select them.

This is how we have evaluated products at IIJ over the years, adopting the technologies that had become established at the time and developing platforms suited to the application in question.

2.4 Virtualization in Action at IIJ

■ **IIJ GIO Hosting Package Services**

■ **IIJ GIO Component Service Base Server V Series Linux Type**

At IIJ around 2008, we reviewed the design of our internal service hosting infrastructure, consolidating uniformly configured resources into a pool so that they could be divided up and used according to demand. We adopted Xen as the virtual machine monitor for dividing up and using the resources of the physical servers. Drawing on our experience operating this system within IIJ, we also adopted Xen for the IIJ GIO Hosting Package Service and IIJ GIO Component Service V Series Linux Type released in 2010.

The Xen hypervisor fully isolates the privileged administrator domain, dom0, from each of the user domains, denoted domU, allocates resources to those domains, and provides memory protection. As the privileged domain, dom0 manages the hardware and operates the Xen hypervisor. The domU domains are virtual user machines—they run independently and are isolated from other domains by the Xen hypervisor. Each domain is thus fully isolated, ensuring security and stability.

While hardware-based performance support was still scarce at the time, we chose Xen because paravirtualization allowed for practically feasible performance when the guest OS was optimized for Xen, and the source code was open and it was thus being actively developed in a transparent manner.

To ensure service stability, we made modifications to the Xen hypervisor as needed for IIJ services, and to ensure continuity of operations, we froze the version of Xen, backporting patches from later versions, and we defined our network security specifications and used our accumulated knowhow to build in measures against security threats, including virtual servers engaging in packet spoofing or attacks on other virtual servers. For the IIJ GIO platform, which provides IT resources to our customers, we streamlined operations by developing our own orchestrator, based on our operational knowhow with IIJ service hosts, to automate the resource delivery process, from the securing and deployment of resources through to their return to the pool. Building an automated control mechanism to centrally manage configuration information and rewrite settings in a coordinated manner made it possible to deploy, without error, a huge amount of resources that it would have been impossible to deal with manually, and to calculate resource allocations in a fair and efficient manner (Figure 1).

Using simplified system configurations premised on the use of virtualization technology enabled, for instance, live migration, which let us perform maintenance without interrupting the equipment, and we were thus able to ensure sufficient quality even for services made up of thousands of servers. We also use the knowledge gained from this in running the successor service, IIJ GIO Infrastructure P2 Public Resources.

IIJ GIO Component Services Base Server V Series Windows Type

IIJ GIO Component Services are targeted at enterprise systems and were developed with the aim of providing the various components needed for system integration.

In many cases, users' internal IT environments use Windows Server components (Active Directory, file servers, WSUS, etc.), and so to migrate enterprise systems to the cloud, Windows Server needs to be made available in the cloud. We therefore built a separate virtualization platform using Hyper-V to enable us to operate and provide Windows Server with stability.

Microsoft released Windows Server 2008 R2 in 2009. Hyper-V 2.0, available on Windows Server 2008 R2, features Cluster Shared Volumes (CSV), which enables a single storage area (LUN) to hold multiple virtual machines and provide multiple virtualization hosts with simultaneous read/write access. It also enables live migration of virtual machines, among other features. It thus provides the minimum functionality required for a multi-tenant service platform.

The System Center products provided the functionality required for operating a service platform, such as virtual machine lifecycle management and component monitoring, but they are designed to be used through a GUI, so for automation and orchestration purposes, we had to develop tools using the Dynamic Datacenter Toolkit (DDTK) framework provided by Microsoft.

For the IIJ GIO Component Services Base Server V Series Windows Type platform, we used DDTK to implement the orchestration tool for IIJ's internal service operators and

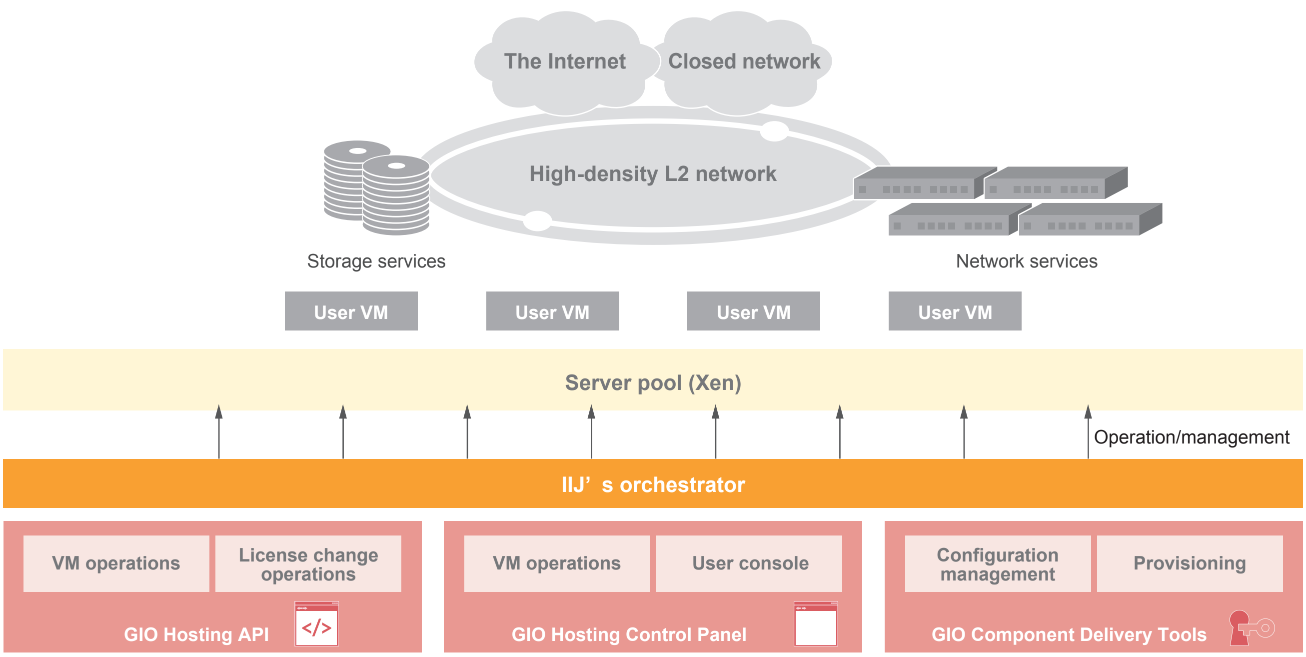

Figure 1: IIJ GIO Hosting Package Services / IIJ GIO Component Services Base Server V Series Linux Type

the control panel that lets users operate their virtual machines (Figure 2).

A lot of useful features for operating services have since been added with each new generation of Windows Server. One such feature is Storage Live Migration, which lets you relocate a virtual machine's storage even while the virtual machine is running.

As this was not available on the initial platform, we implemented our own migration tool (a proprietary tool that required us to stop the virtual machines but also supported rollback after migration work had begun). But Storage Live Migration became available on the service platform starting with Windows Server 2012, and this made it possible to reduce the impact on users when carrying out service maintenance and troubleshooting.

We continued to use System Center as our management tool, and using Service Provider Foundation (SPF), which provides a RESTful API for service providers, made it possible to implement the orchestration functionality that IIJ needed while also keeping development man-hours down.

The Hyper-V base platform was originally developed as a service platform for Windows Server, but from a support and licensing standpoint, it came to be used in a variety of ways over time, such as to provide other OSes on the Hyper-V base platform.

We have continued to make fine-grained updates to ensure we can operate hundreds of Hyper-V virtualization hosts and provide increased service stability in step with the evolution of Windows Server and Hyper-V, but with the

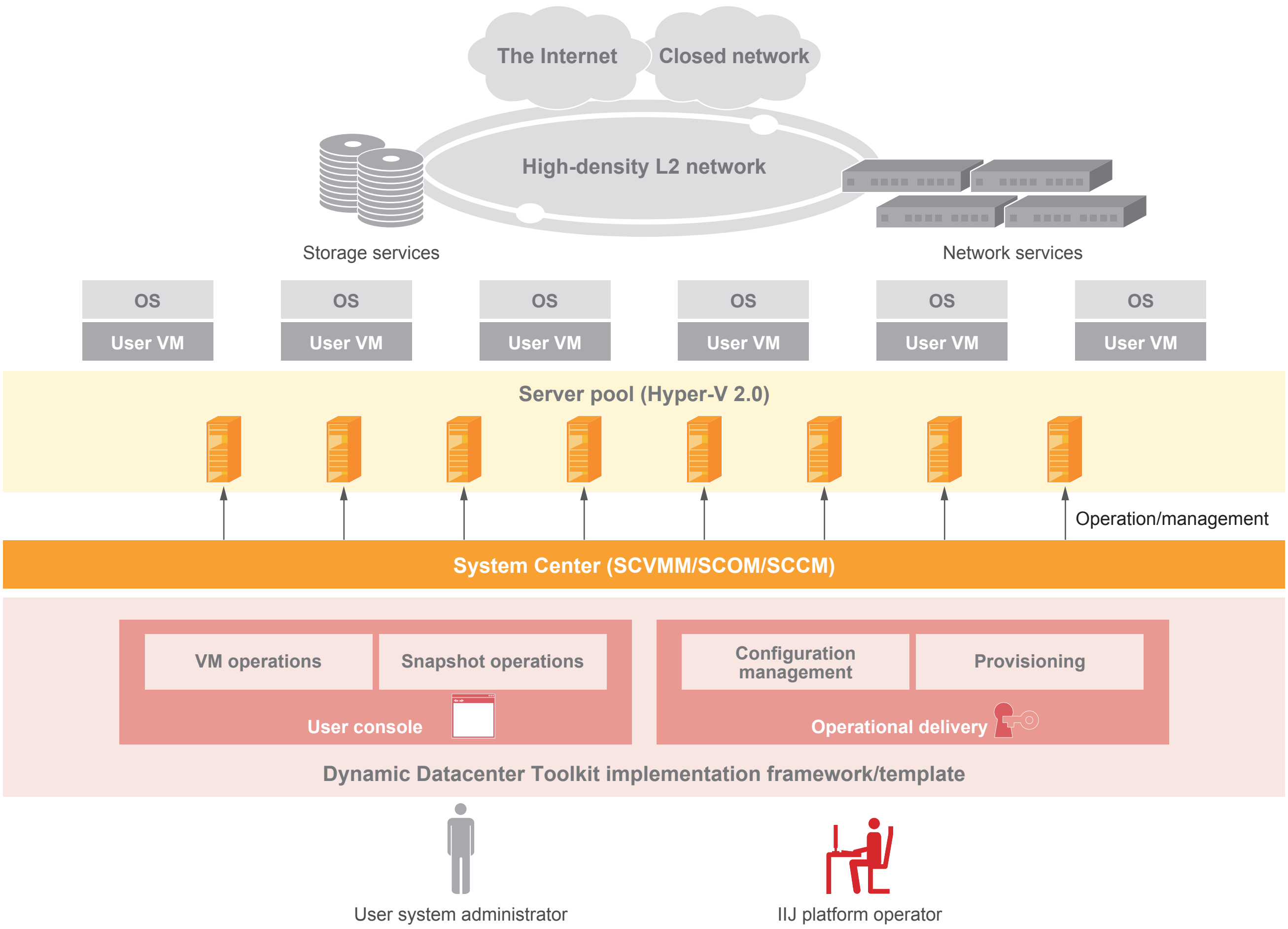

Figure 2: IIJ GIO Component Services Base Server V Series Windows Type

release of the successor services, IIJ GIO P2, we discontinued IIJ GIO Component Services in September 2023.

IIJ GIO Infrastructure P2 Public Resources

In 2015, IIJ released IIJ GIO Infrastructure P2 Public Resources, which can be scaled using software-defined networking (SDN) technology we developed in-house. We selected KVM as the virtual machine monitor and QEMU as the virtual machine emulator. The five years following the initial launch of IIJ GIO in 2010 were also a time during which the various technologies around server virtualization were refined and became increasingly widespread. In keeping with the times, we also switched to KVM+QEMU on the service host platform that provides virtual servers within IIJ.

KVM is tightly integrated with the Linux kernel, allowing direct use of Linux features and security updates. It can use hardware-assisted virtualization features such as Intel VT-x to enable the efficient use of I/O and CPU resources. The virtual machines run as independent QEMU processes, which ensures robust isolation from other virtual machines. Combining KVM and QEMU entails a high degree of difficulty in terms of configuration and management, necessitating a high level of expertise, particularly in large-scale environments where the intention is to provide resources to customers. Our decision to select KVM+QEMU rested on the fact that IIJ had built a track record and the knowhow to operate virtualization environments using KVM+QEMU, and the fact that advances in hardware virtualization support meant that KVM+QEMU provided practically feasible performance even with full virtualization, so it would be possible to provide SEIL/x86[*3] and Windows without the need for paravirtualization support in the guest OS.

When it came to providing services, given KVM's tight integration with the Linux Kernel, IIJ decided to avoid actively making any unique modifications. We added the functionality we needed, which included making our internally developed orchestrator compatible with our own SDN technology, utilizing development assets imbued with our pioneering services knowhow (Figure 3).

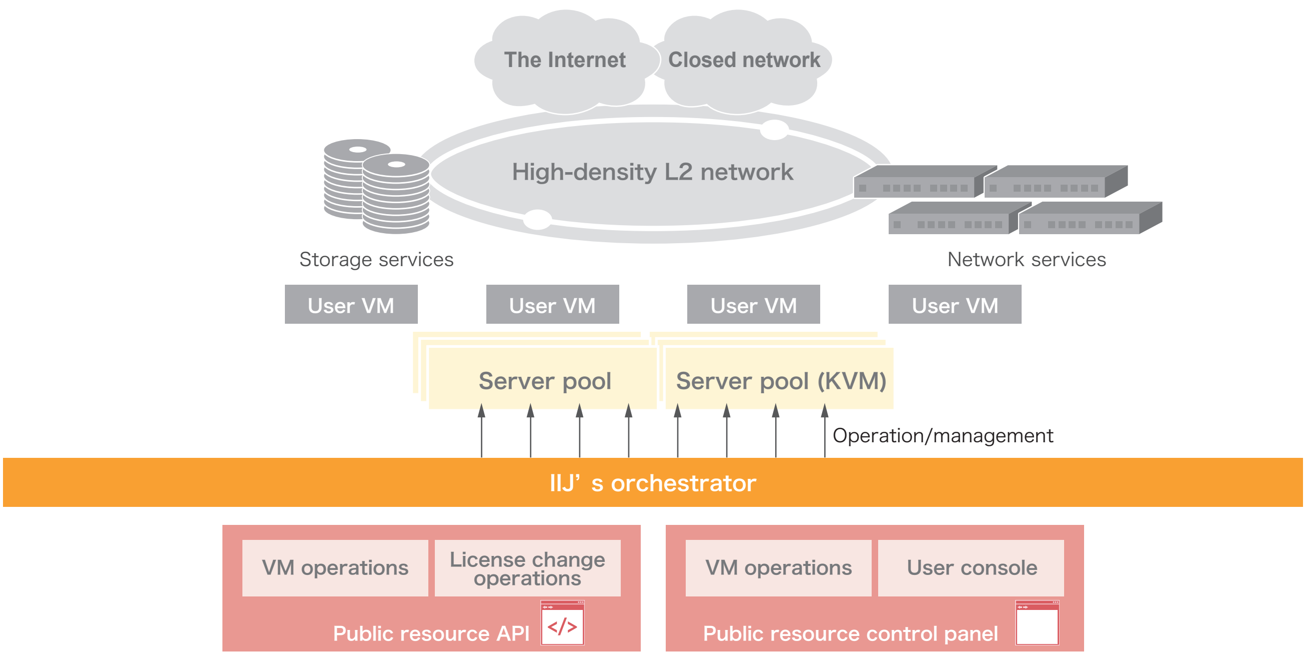

Figure 3: IIJ GIO Infrastructure P2 Public Resources

*3 SEIL/x86 is a high-performance software router developed by IIJ that runs on an x86 architecture-based platform.

While maintaining our established operational approach, making active use of live migration and so forth, we continue to operate our services so as to maintain consistently high quality even at larger scales.

■ **IIJ GIO Virtualization Platform VW Series**
■ **IIJ GIO Infrastructure P2 Private Resources**
□ **IIJ GIO Infrastructure P2 Gen.2 Dedicated Server Resources**

Since 2012, IIJ has provided hosted private cloud services to enterprise customers that use VMware vSphere as the hypervisor. When it comes to virtual machine specs on the sorts of cloud services generally available, user systems need to be configured based on the instance models (vCPU count, memory capacity, OS type, etc.) specified by the cloud service provider, and the difficulty of selecting the right menu options had been a problem for users unfamiliar with designing systems for peak resource usage. With this service from IIJ, the user has direct control over a dedicated hypervisor, allowing them to freely design their virtual machines and select which OS to use, making it possible to configure systems based on the resource allocations that the user has in mind (Figure 4).

Of course vSphere was in use by many users, but its applicability to a wide variety of hardware also made it an easy choice for us to provide as a service operator. On the technical side, as well, commonly implemented functionality was sufficient to provide isolation between users—by using VLAN for networking and functionality for carving out logical disk volumes for data stores, for instance—so development could be tailored to scale without the need for any specialized development efforts (development of special protocols etc.).

As a result, the service has now evolved through three generations and been in operation for over a decade, providing a platform that lets users who had built their own on-premises private clouds using vSphere migrate smoothly to the cloud without any major changes.

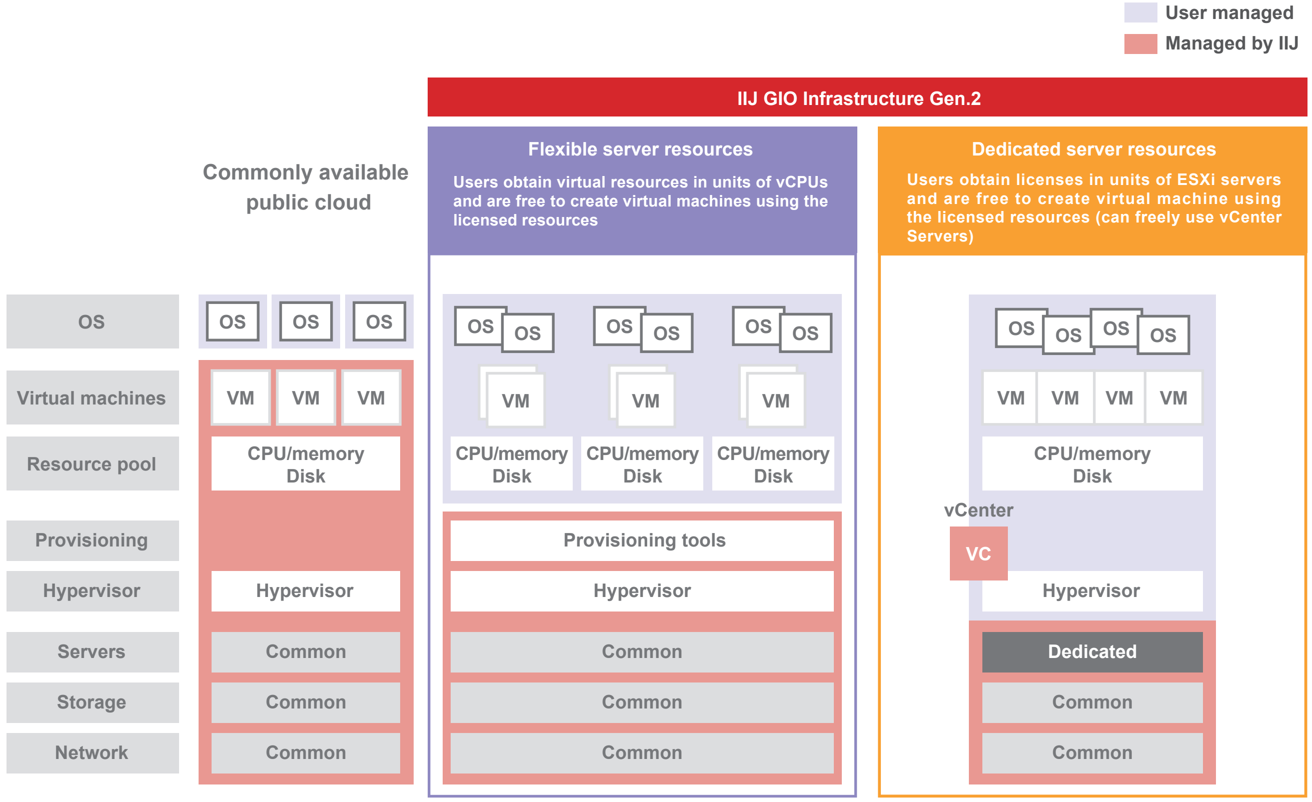

Figure 4: Service Demarcation of Responsibility

IIJ GIO Infrastructure P2 Gen.2 Flexible Server Resources

IIJ released IIJ GIO Infrastructure P2 Gen.2 Flexible Server Resources (FSR), its third-generation public resource service, in 2021. With the previously mentioned service on which users directly operated a dedicated hypervisor, they were able to operate vSphere directly, giving them considerable freedom in building their own systems, but on the flip side, users faced operational challenges, such as the need to upgrade and perform maintenance on the virtualization platform themselves.

For FSR, we selected VMware Cloud Director (VCD), making it possible to abstract individual resources, such as CPUs and memory, and present them to users as a resource pool, and provide a mechanism for dividing up resources and building systems in the same way as with vSphere. The hypervisor layer is hidden from users, but users have the authority to control resources just like with vSphere, while IIJ handles lifecycle management for the hypervisor and hardware, a setup that resolves the issues mentioned above.

This platform design not only benefits users but also completely separates the user environment (mainly software) from the service provider's operating environment (hardware), allowing IIJ to handle maintenance and other tasks that, while important for maintaining the platform, had in the past been a challenge due to user considerations, and it thus provides a more stable service platform from IIJ's perspective as well. The platform also lends itself to continuous development and modifications to meet the contemporary demands of business environments.

2.5 Conclusion

We have taken a trip through the history and evolution of virtualization technology, along with some examples of it in action at IIJ. In our current times, we are all being compelled to make choices about virtualization technology, and I hope this article provides some assistance to you in selecting and using the most suitable technology. IIJ has been proactive about adopting and operating virtualization technology, and this has improved business efficiency and reduced costs. Another consideration that we must pay attention to is that of potential threats to supply chains, with those in recent memory including the global chip shortage and open source vulnerabilities as highlighted in the media. As a service provider, it is incumbent upon us to prepare a whole range of options that we can turn to in order to continue to provide stable services to our users even when corporate acquisitions or other such events mean that products we use are no longer available in the same way as before, and we will thus continue to keep up with advances in virtualization technology and the products in this space.

Looking ahead, we will continue to monitor the latest trends in technology and adapt as virtualization technology evolves so that we are able to deliver even greater value to our users.

2.1 Introduction / 2.2 History of Virtualization Technology / 2.3 Technology Selection and Service Development
Kaoru Tanaka
Technical Manager, Technology Development, Cloud Division, IIJ

2.4 Virtualization in Action at IIJ
Takehiro Yamamoto
General Manager, UAG Development Office, Cloud Division, IIJ

Yasuhide Takahashi
Deputy Head of Cloud Services Division 2, Cloud Division, IIJ

Kazuki Takai
Technical Manager, Technology Development, Cloud Division, IIJ

2.5 Conclusion
Shinri Kimura
Deputy Head of Cloud Division, IIJ

November 2024 Vol.64

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-MKTG020-0062

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/