Evolution of the IIJ Cloud—Commemorating 30 Years

3.1 Introduction

Over IIJ's history, we have provided Internet connectivity alongside a range of related services with a focus on communications. We have continued to enhance and expand the service hosts used to provide these services over the past 30 years, which now encompass several thousand servers. We also provide computing resources to customers through the IIJ GIO cloud service, the infrastructure for which has also grown to comprise tens of thousands of servers in the decade since launch.

In this article, we take a look back at IIJ's 30-year journey. The first half describes how IIJ's service hosts have changed with the times and the innovations we have made in the process. The second half looks at how the IIJ GIO cloud platform has evolved through successive generations to become the large-scale service infrastructure it is today.

3.2 1990s: Where it all Started

■ The era of dedicated systems for each service

The number of service hosts we deployed to provide email and web services increased rapidly from the time IIJ was founded in 1992. By the end of the 1990s, we had over 200 racks housing several thousand servers distributed across multiple data centers in Tokyo to provide IIJ's services. We maintained separate racks for each service and procured and built the equipment on that basis, which meant that work had to be performed onsite whenever the configuration changed. With this in mind, we sought to optimize server operations by using conveniently located data centers in and around Tokyo that offered easy physical access.

In the early days, we used UNIX workstations as our servers, but we later transitioned to PC/AT-compatible machines and PC/AT-based industrial PCs for cost performance reasons.

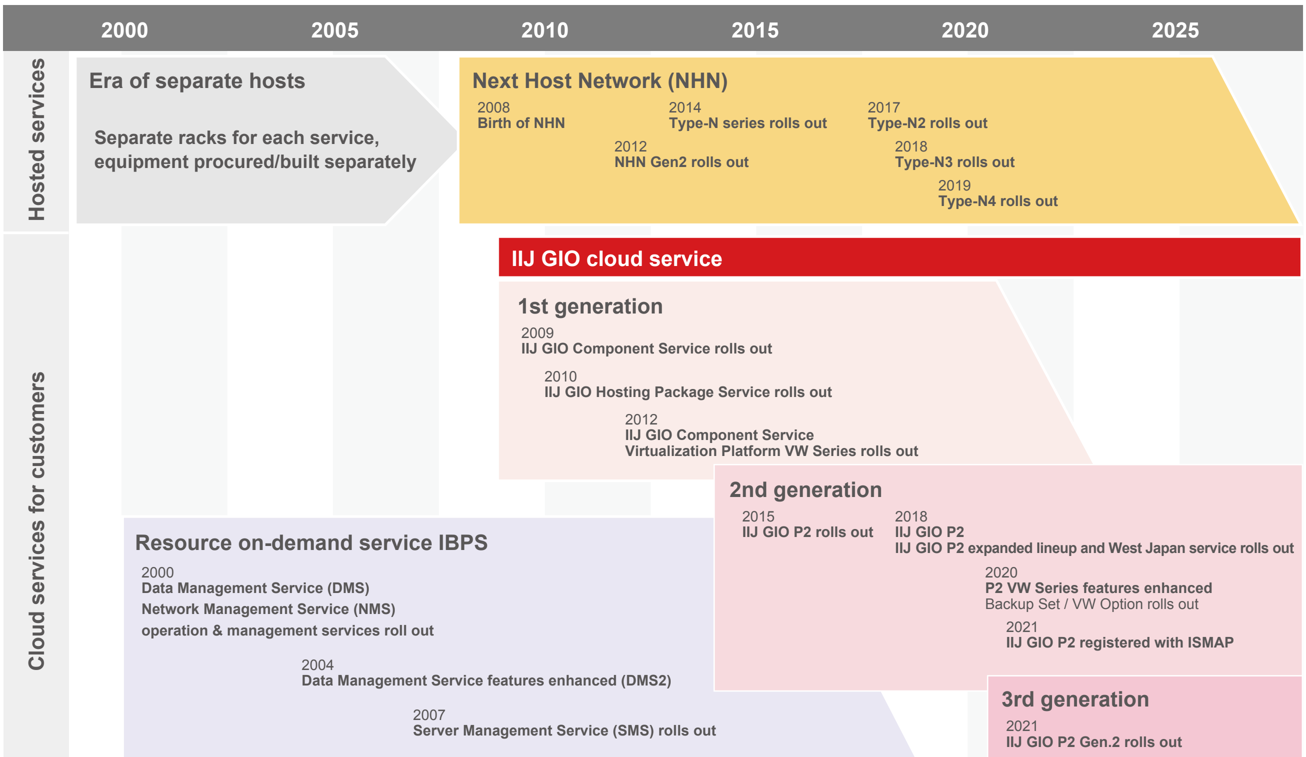

Figure 1: Evolution of IIJ's Cloud Services

As rack-mounted PC servers were not yet commercially available, IIJ staff often assembled the servers themselves using off-the-shelf parts. A variety of work was performed at the Otemachi data center: we would replace server motherboards and CPUs, IIJ staff would assemble RAID arrays themselves, and we had teams and systems in place to ensure that machines could be replaced immediately if they went down. While we were able to set up our equipment on a low budget, we also encountered more and more physical problems like power supply and cooling fan failures, which can in turn lead to hardware failures.

The OSes installed on commercial UNIX workstations at the time were not all that featureful given the price, and it was difficult to make improvements on our own. And Linux had only just appeared on the scene and was not yet production ready. So we decided to obtain a license for PC-BSD, a derivative of BSD (a collection of software created by a group of developers at the University of California, Berkeley), and reconfigure it for use as a server ourselves. This greatly increased the flexibility of our servers. Both our software and hardware were handcrafted. IIJ was probably the only company in Japan at the time that was quite so particular about the cost and quality of its servers.

This was at a time when the TeleHodai service, which offered fixed-rate phone charges after 11pm, was in widespread use in Japan, and 11pm was thus the peak for personal website traffic. So while everything worked smoothly during the day, once 11pm rolled around, the great influx of web server traffic ramped up server loads, and this would sometimes result in circuit breakers being tripped due to excessive power consumption and entire racks thus losing power.

The mainstream storage options at the time were on-board server storage and, where high capacity was required, DAS, whereby external storage was connected to the servers via SCSI cables. In the 1990s and 2000s, we had to plan and budget around questions like how much physical capacity we needed and how much capacity per server, how much it would cost if we installed some number of servers in a location and configured them in a certain way, which hardware vendor's servers we should use to make sure we stayed within budget, and where we should locate our data centers. Experiencing major storage failures taught me firsthand about the absolute need for backups and adequate system sizing.

3.3 2000s: Shedding the Dedicated-service-hosts Approach

■ Advent of the Next Host Network

The Next Host Network (NHN) is a platform for IIJ services that we developed in 2008 to streamline service host operations, and it represented a redesign of our server configurations and server operation network. Up to this point, we had procured equipment and built systems on a service-by-service basis, as discussed above, and this meant that overall infrastructure costs were continuously on the rise. Any configuration change triggered the need for onsite work, and the operation of any one system was defined in isolation from the others, resulting in a convoluted setup. And while we had installed sufficient capacity to meet anticipated growth in demand for each individual service, it was difficult to repurpose equipment when our demand forecasts failed to pan out, and maintenance costs were also rising as we generally entered into maintenance contracts on the equipment we selected. The cost of maintaining our physical facilities also continued to soar because we were using conveniently located data centers in urban centers that we could get to quickly whenever system problems occurred.

We needed to fundamentally reconsider the nature of our infrastructure if we were to break away from this chronically high cost structure. This led to the implementation of NHN, designed to solve the myriad problems we faced by consolidating our disparate systems and making it possible to provide an equivalent level of service even while running our operations remotely. The basic concept we adopted was to provide flexible service hosts, and we embarked on a four-year plan to progressively replace thousands of servers distributed across our data centers in Japan, consolidating data that had been stored on DAS systems into iSCSI storage and making this available over an IP network, along with newly installed diskless servers, in a flexible manner according to service demand.

The basic design principle, based on the assumption that servers would inevitably fail, was to redesign storage to be highly reliable and the network to be simple and uncomplicated. We wanted to reduce onsite work and switch to a model that would allow operations to basically be made remote. Server equipment has a wide variety of failure points, and because failures are unavoidable and there are limits to the extent to which failure rates can be reduced, we designed the system so that the inevitable failures would only have a small impact. We moved away from having assigned racks for each service and adopted a server pool system to improve storage efficiency, making it possible to consolidate onsite work and outsource it to contractors. We standardized server configurations to reduce the configuration management workload, and added virtualization technology into the mix to enhance flexibility, enabling post-installation work to be accomplished remotely. Individual HDD failures are inevitable with any storage device, but we needed to adopt a more reliable configuration to avoid failures that would lead to service outages. We used PXE Boot and iSCSI storage with controller and path duplexing to create an inexpensive, highly reliable diskless server environment, making it possible to recover by simply switching server devices in the event of failures. In our experience, edge switch failures had not been all that common, so we adopted a simple configuration that only used NIC redundancy settings and the like where necessary. We introduced a mechanism for automatically rewriting switch VLAN settings based on config data to reconfigure the network without changing the physical wiring, making it possible to reduce configuration errors and operational costs. NHN primarily used x86 servers from hardware vendors with strong cost performance. The only OSes available were Linux distributions likeCentOS, but Windows Server and VMware platforms also later appeared. We selected low-end but highly maintainable storage options to simplify storage maintenance tasks for which we had previously relied on experts, making it easier for IIJ staff to perform maintenance themselves and thus helping to reduce annual maintenance costs. The key characteristics of each NHN generation are outlined below.

NHN Gen1: The first generation of NHN infrastructure, launched in 2008. Gen1 adopted iSCSI storage in order to eliminate DAS, and provided two storage areas: basic (60GB) and expanded (160GB) storage. NHN Gen1 units were configured as a set of four server racks and four storage racks comprising a single unit, but once inventory resources were exhausted, it was impossible to expand in the same L2 area, so it was difficult to adjust inventory on a unit-by-unit basis. Server instances also could not be moved between units. Uplink speed was 1Gbps, and so traffic growth also started to put pressure on upstream bandwidth.

NHN Gen2: The second generation of infrastructure, launched in 2012 and designed to solve the problems of NHN Gen1. The units had the same configuration as Gen1, but rack storage efficiency was doubled (to accommodate 40 units per rack). Characteristically, Gen2 adopted Juniper's Virtual Chassis (VC), had a 10Gbps uplink speed, had racks that would accommodate non-NHN standard service-specific equipment, redundant power supplies, and NIC bonding. However, NHN Gen2 did have a problem in that failures in switches used for top-of-rack (ToR) switching would cause the entire network to be disconnected. The failure would spread to the entire VC unit, resulting in a situation that took quite a bit of time to recover from.

NHN Gen2 Container: A version of NHN optimized for the IZmo/S containerized data-center modules deployed at Matsue Data Center Park. Congruent with the number of racks installed in IZmo/S modules, each unit comprised eight racks and 288 servers (later, 16 of the servers were removed and four storage units added). Although NHN Gen2 Container's slim container dimensions make it space efficient, the narrowness of the aisles inside the containers made maintenance particularly difficult, which caused problems on the operations front.

3.4 2010s – Present: Type-N, a Next-generation Service Platform

■ Shifting to Type-N service hosts

In October 2014, we launched the new Type-N series, the third generation of NHN infrastructure. Until this point, we had provided the servers in bare-metal form, but starting with the Type-N series, we also started providing them as virtual machines (VMs). This meant we could also support reduced spec applications that did not require a full bare-metal server. The characteristics of each Type-N generation are outlined below.

Type-N: NHN's third generation of infrastructure, launched in 2014. Until this point, L2 switches were housed in the router. With Type-N, the L2 switches were housed in the L2 core switch, making it possible to provide the same L2 surface between units. Connectivity with users' racks improved, and inventory resources no longer needed to be adjusted on a unit basis. In addition to ordinary servers, Type-N also made available high memory servers and storage servers equipped with a large number of HDDs. There were problems, however: performance of the switches used for the servers' 1GbE NICs was poor, and the maximum number of connectable devices was low.

Type-N100: A broadband infrastructure option based on Type-N and developed for content streaming services. The name refers to the fact that this option provided 100Gbps infrastructure. Key features were that it used dedicated load balancer units and MPLS and allowed connections to IIJ backbone routers, which handle direct IX connections. Systems had to do diskless PXE booting and run the entire OS in memory (no iSCSI storage provided), and a 1TB SSD was provided for temporary logging and caching purposes.

Type-N2: The fourth generation of NHN infrastructure, launched in 2017. Notably, we did away with the 1G network and went with full 10G all the way to the edge. High-memory servers were set up to support the addition of units for clustered applications based on VMware, and hardware performance enhancements led to improved system consolidation rates. A problem with Type-N2, however, was that storage performance did not meet the requirements of some applications.

Type-N3: The fifth generation of NHN infrastructure, launched in 2018. We adopted what were at the time the latest Intel CPUs and installed flash storage (NVMe SSDs) for a huge jump in I/O performance, and also aimed for significant improvements in network performance. With Type-N3, we did away with the separate basic and expanded storage areas and provided the entire storage area as a single PV (physical volume).

Type-N4: The sixth generation of NHN infrastructure, launched in 2019. We again used the latest Intel CPUs at the time and also adopted new OCP-compliant servers (described below), reducing procurement costs and power consumption. We increased network speeds between the edge and core from 10–20Gbps previously to 40Gbps, and we used chassis switches for the core switches to future-proof against future capacity expansions.

From 2019, we adopted mid-range-class storage systems for our main storage. In 2021, we adopted the successor model of storage (from the same manufacturer) that we had used in NHN's first iteration, naturally for its performance and reliability, and also because, on the operations front, we knew it would fit in nicely with our existing procedures and frameworks. Storage products that support self-maintenance operations, including OS and firmware self-updates, also offer strong cost benefits for NHN, where our basic philosophy is to build and run the systems ourselves.

■ Adoption of OCP servers

The OCP (Open Compute Project), formed in 2011, is an organization that proposes new server standards from hyperscalers like Facebook. Compared with traditional servers primarily developed and produced by manufacturers, the OCP uses a different scheme whereby the cloud service providers that use the servers are heavily involved in the selection, procurement, and manufacturing of the components.

While we were interested in OCP servers, the purchase prices were too high to justify the expense when compared with products from ordinary server manufacturers, and this remained the case up until about 2017. But as 2018 rolled around, with the Japanese yen appreciating and the semiconductor industry experiencing an inventory glut, the price of DIMMs, SSDs, and other such components slipped into decline. And in light of these market conditions, we started to give OCP servers serious consideration. We saw a multitude of issues and concerns: we would be directly exposed to exchange rates; on the hardware front, we would need new OCP-compatible racks and central power supplies, and BIOS and BMC tests had to be performed by the buyer; and on the operations front, maintenance was limited to parts replacements (self-maintenance), and we would have to deal with the ODM vendors in English.

And yet we determined that the benefits would outweigh all these concerns, leading us to the decision to deploy OCP servers. When we actually deployed OCP servers, we saved up to 35% on procurement costs and cut power consumption by up to 30%. For procurement costs in particular, these savings were far beyond the 10% we had anticipated, and this spurred us to expand our deployment of OCP servers. At present, however, there are some enterprise applications and areas to which they are clearly not suited. In the summer of 2019, we made multi-vendor arrangements for OCP servers like we had for existing equipment to diversify the technological and procurement risks we face.

Next, let's look at the evolution of IIJ GIO, our cloud platform for customers.

3.5 2000s: Pioneering IaaS (Infrastructure as a Service)

Launch of the IBPS resource on-demand service

In the late 1990s, the advent of server-side Java and ASP made Internet payment infrastructure a possibility, and thus a lot of e-commerce sites based on dynamic website technology began to appear. The system architecture of e-commerce sites was fairly similar, regardless of how the customer's site was configured, so we realized that it would be more efficient to set up the equipment and resources in advance to have them ready to deploy in the form of system components so that we could provide the necessary combination of those components whenever a customer made a order. With this in mind, we created the resource on-demand service IBPS (Integration & Business Platform Service)[*1], the predecessor to our present-day cloud service, IIJ GIO.

IBPS was launched in March 2000 by then group company IIJ Technology Inc. (absorbed into IIJ in April 2010). The service offered everything an Internet business would need, from server equipment through to software, payment/logistics components, and monitoring, operations, and management. The necessary components were combined in accord with the customer's needs to form a complete system, which was provided as an outsourced service. This made it possible to more than halve (relative to IIJ's conventional offerings) development times and initial investments and meet the demands of companies looking to rapidly deploy Internet businesses.

IBPS's four main services

- Data Management Service (DMS): IIJ's storage resources provided on-demand. DMS optimized storage costs by providing storage that allowed the connection type (NAS/SAN), capacity, and performance to be selected according to the customer's service level and usage patterns. DMS also provided high-speed backups on separate disks, freeing customers from the need to perform the old labor- and time-intensive tape backups.
- Network Management Service (NMS): IIJ's network resources provided on-demand. NMS pooled load balancers, firewalls, etc. and provided them as functions. Typical configurations of these were also provided as low-cost, packaged options.
- Server Management Service (SMS): IIJ's server resources provided on-demand. SMS employed a provisioning tool to manage customer-specific server configurations, thus semi-automating the server building task and making systems tailored to customer-specific needs available with minimal lead times.
- Operations and Management Service: This service provided operation and monitoring of customer systems built on IBPS to help reduce total cost of ownership.

Customers were able to use the resources they needed when they needed them, and cancel resources they no longer needed. Users were freed from the risk of ownership. IBPS was what we would now call an IaaS offering. We originally had UNIX servers (SPARC Solaris) on the frontend, and although we later migrated to x86 servers, UNIX servers remained on the middle/backend for database purposes. We had been using rack-mounted servers, but in August 2007, we introduced blade servers—a first for large-scale server infrastructure in Japan—for SMS, through which we offered the first resource on-demand servers in Japan. We liked the ability to save space with blade servers, and the fact that you simply had to insert the CPU blade into a pre-configured chassis to get them working.

*1 The service was called iBPS at launch but rebranded to IBPS in 2002.

In October 2003, we expanded the Data Management Service we had been providing since launching IBPS and added a large-scale storage service with a total capacity of 40TB. For midrange storage, we adopted storage virtualization software to enable flexible resource management and reduce costs. In 2007, we upgraded the Data Management Service, halving the per-GB storage costs*2. The Data Management Service originally provided SAN and NAS options, but as services offering high-quality Fiber Channel (FC) at reasonable prices became available, the use of FC-SAN storage grew. On the NAS front, we initially used a combination of commodity servers and cluster software or dedicated NAS storage appliances. Few of the NAS products around in the early 2000s supported volumes over 1TB, so if you wanted large-capacity NAS, you would build a NAS system by combining commodity servers, FC-SAN storage, and cluster software. In the late 2000s, large-volume support started coming to NAS products, prompting a shift toward NAS-specific storage appliances.

2008 onward: L2 ring protocol

From 2008, we used an L2 ring protocol configuration for our service platform network technology. The multi-stage ring physical configuration was suited to the scale of data centers in Japan. This had the advantage of making it possible to build an efficient system while also making it easy to expand the system, whether at the edge or the core. With node interfaces of 10GbE+, however, there were fewer models of equipment to choose from than with 1GbE or lower, so a disadvantage was that this limited the number of accommodated nodes. Many L2 ring configurations are manufacturer-specific, and they are not inter-compatible. The problem with this was that it made vendor lock-in likely and thus limited the range of available equipment options.

Storage array systems for a variety of platforms

In terms of service platform storage technology, we have used storage array systems since launching IBPS in 2000.

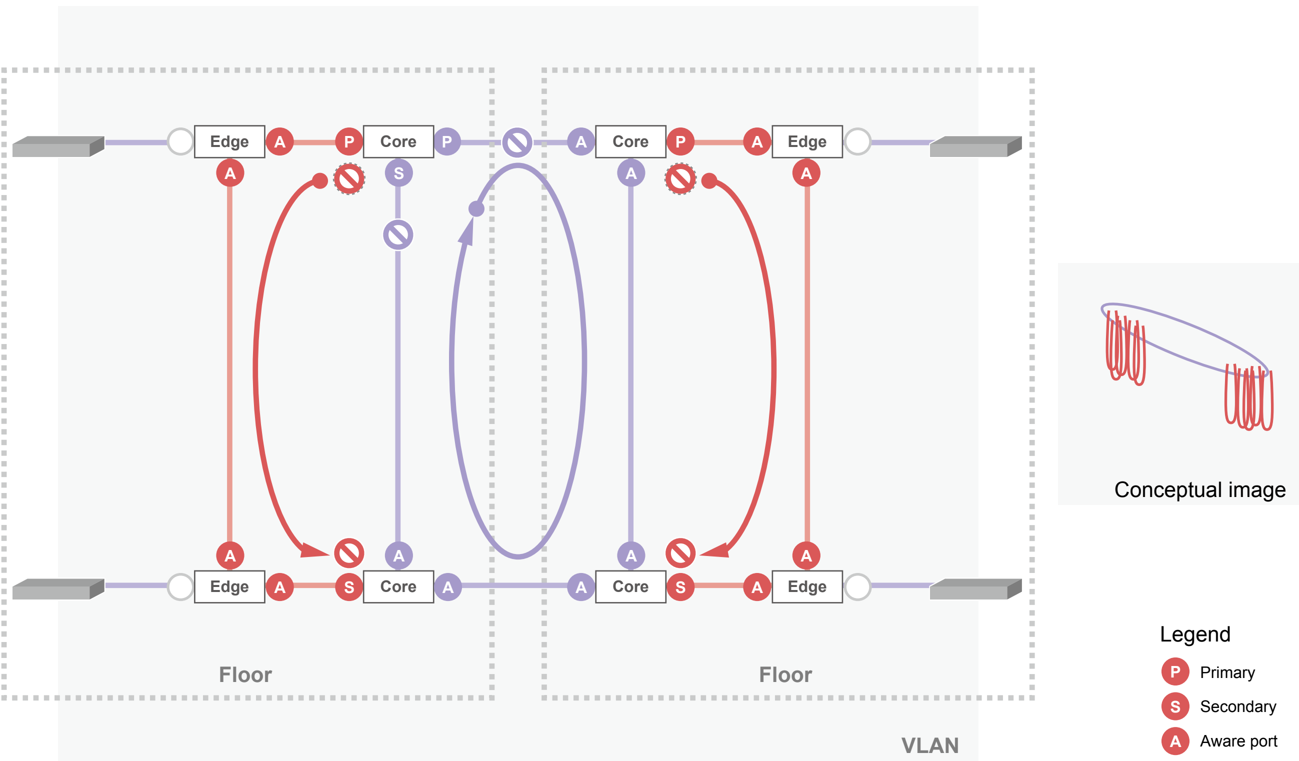

Figure 2: L2 Ring Protocol Configuration (2008–)

*2 This was also at a time when Moore's Law worked well.

We used several types of mainly high-end – mid-range storage systems to suit performance requirements, and also to diversify procurement and technology risks. Today, we also use storage array systems on the IIJ GIO and NHN platforms, with storage capacity continuing to increase.

When first introducing a storage option, we would contract all of the configuration changes and monitoring out to the vendor, partly because of our initial lack of knowledge. But this style of operations resulted in a fair amount of cash outlay and lead time just to create, for example, a single LUN, so IIJ staff learned how to operate the storage products so that we could make storage configuration changes in-house, which reduced operating costs. Initially, we used IIJ's standard monitoring service to monitor storage, but as the service continued to grow, the inevitable faults started to overwhelm this standard service, and it ended up taking longer for the storage operations team to recognize faults whenever they occurred. We therefore built a dedicated monitoring system for our storage equipment.

3.6 2010s: Cloud Computing Takes Off

■ Birth of the IIJ GIO cloud service

We have overcome a myriad of challenges through our experience in the IaaS business and the evolution of our services platforms spanning more than a decade, including the deployment of virtualization and provisioning systems in IBPS services, the development of in-house monitoring systems, and large-scale upgrades to server, network, and storage equipment. This process ultimately culminated in the development of the IIJ GIO cloud service, which embodies everything we learned. IIJ GIO harnesses the advantages of cloud computing to make high-service-level system environments with high availability and strong security available to users who need business-ready infrastructure at a lower cost, without the need for them to own IT assets themselves.

Our first step was the November 2009 launch of the IIJ GIO Component Service, a private cloud service geared to the fine-grained needs of business customers. This was followed by the June 2010 launch of the IIJ GIO Hosting Package Service, an inexpensive public cloud service with packaged options.

■ IIJ GIO Hosting Package Service

The IIJ GIO Hosting Package Service makes it easy for users to build information systems infrastructure for e-commerce sites and high-performance online businesses simply by selecting pre-packaged plans online according to the needs of their application. The service architecture makes full use of virtualization technology to provide the flexibility to select server resources according to system requirements, and control functionality developed by IIJ makes it possible to automatically allocate multiple server resources on a Layer 2 network. We also adopted open source software (OSS) and developed in-house provisioning tools for controlling resource allocations and management tools for automating complex operations and monitoring, leading to substantial efficiency gains on both the operations and cost fronts. By consolidating our infrastructure, we were also able to reduce hardware procurement costs and set affordable service prices, starting from 8,000 yen per month for a virtual server. In September 2013, we released an IaaS API to meet the needs of users wanting to create their own programs to automate routine tasks, such as the deployment of multiple virtual servers at once.

■ IIJ GIO Component Service

The IIJ GIO Component Service provides servers, storage, and networks in the form of system "components" to create a highly flexible IaaS-based enterprise cloud service that allows users to combine the best components for their needs from a variety of options. The service can also be used as a private cloud by connecting it directly to the user's on-premises environment via a wide area network. The main components of IIJ the GIO Component Service are the base servers and, launched in August 2012, the Virtualization Platform VW Series (the VW Series). Two types of base servers were made available: V Series virtual servers that allow users to share physical servers with other customers, and X Series physical servers that give users exclusive use of server hardware units. With the VW series, we provided customers with a private environment with VMware vSphere ESXi (US-based VMWare's virtualization software) on the physical servers along with a VMware vCenter Server management server, all with administrator privileges. This provided a level of flexibility in systems configuration comparable to that with on-premises environments, ensuring it could be used with confidence not only by those looking to integrate servers or build a cloud from scratch, but also by users who already maintained and operated virtual infrastructure using VMware.

We also provided functions other than server resources as IaaS offerings. The first is the Network Add-on feature, which enhances the network functions provided as standard on base servers and the VW Series. Switching from a shared Internet connection line to a dedicated line (private connection) lets users safely connect to their on-premises environment via an Internet VPN or closed network (wide area network). They can also use multi-carrier configurations that divide WAN lines among carriers. The second is the Storage Add-on feature, which broadly offers two options: Standard, which provides high-end storage of the sort that financial institutions use, and Basic, a mid-range offering suitable for data management on ordinary web systems and the like. These storage options are provided as NAS, FC-SAN, or iSCSI-SAN storage over the network. The third is the Database Add-on feature, which provides Oracle Database and MySQL as DBaaS (DataBase as a Service) offerings. When we launched this feature in July 2012, we became the first operator in Japan to offer Oracle Database—which commands a dominant share of the Japanese database market—for a monthly fee. We made it possible for users to lower their initial investment in database licenses, reduce maintenance costs, and avoided investment risks. With this service, IIJ draws on its extensive experience installing and operating relational databases to design and operate database instances and make them available on IIJ GIO virtual servers. We also made it possible to connect to existing on-premises environments via a closed network or Internet VPN. In May 2014, we revised our pricing to reduce monthly fees by up to 56%, and we continued to actively develop the services, adding Microsoft SQL Server to our lineup in October of the same year. The fourth is the License Add-on feature, which provides software licenses for use on the cloud for a monthly fee. We offer Microsoft SPLA licensing as well as licenses for in-demand products and services from vendors like Red Hat, VMware, Arcserve, and Trend Micro.

In January 2014, we bolstered the IIJ GIO Component Service lineup to add Base Server V Series G2 (V Series G2), which built on and enhanced the existing Base Server V Series. V Series G2 was compatible with Windows Server 2012 R2, then the latest version, and compared with the previous series' Windows lineup, offered increased CPU, memory, and disk capacity, and had up to triple the CPU performance at 24 ICU[*2] and up to six times the memory at a maximum of 48GB. We also made the server equipment available at sites in both East and West Japan, so it was viable for disaster recovery applications as well.

Release of IIJ GIO Infrastructure P2, our 2nd generation cloud service

In October 2015, we overhauled IIJ GIO's IaaS offerings and launched a new lineup under the IIJ GIO Infrastructure P2 (IIJ GIO P2) banner. Up to this point in our IaaS lineup, we had provided the public cloud IIJ GIO Hosting Package Service, which users could easily deploy online, and the bespoke IIJ GIO Component Service, which allowed users to combine a variety of IT resources to configure a complete system. With IIJ GIO P2, however, our aim was to provide a service covering the full gamut of user needs by combining improved public cloud reliability and processing performance with private cloud offerings that can be requested and deployed immediately online. IIJ GIO P2 comprises public resources

*3 ICU is a measure of CPU performance. 24 ICU is equivalent to six cores x 2.

offering shared resources primarily through virtual servers, private resources offering dedicated VMware virtual environments and physical servers, and storage resources available on all servers in both the public and private resources offerings. Users can select the optimal combination of resources to build their systems. IIJ GIO P2 also offers excellent external connectivity, with multi-carrier support and private segment extensibility. It makes it possible for users to seamlessly integrate their own on-premises systems and third-party cloud service environments.

Public resources: These are shared resources primarily available on virtual servers, constituting a public cloud adaptable to a wide variety of use cases—from development environments and simple web services to platforms for online games and e-commerce sites requiring high I/O performance. Users can choose the best server resources for their needs from three characteristic types.

- Performance guarantee type: For users that require stable processing performance, this resource type provides virtual servers to which CPU resources will always be reliably allocated. It offers peace of mind for a fixed monthly fee.
- Best effort type: This resource type uses CPU cycle distribution to provide low-cost virtual servers. Pricing is based on usage in one-hour increments so that users can optimize cost outlays.
- Dedicated type: This type provides dedicated virtual servers to users that require high I/O performance capable of withstanding heavy workloads. These virtual servers run in secure server environments physically separated from other users' servers and are equipped with SSDs or SanDisk high-speed flash storage.

Users can combine the three server types and switch among them at will via an online interface. They can also capture OS images from running virtual servers and keep them in a dedicated storage area, enabling the rapid deployment of new virtual servers based on those images. This reduces operational workloads, and enables rapid scaling out during times of heavy workloads and rapid patch deployment. Pricing for the custom OS image storage area is based on usage, so users avoid unnecessary costs.

Private resources: A highly reliable private cloud, suitable even for mission-critical systems, to which business customers can easily migrate systems built in on-premises environments. We offer a lineup of user-dedicated resources with a focus on VMware virtual environments and physical servers. Users can also request the resources using an online interface, something that was not possible with the IIJ GIO Component Service. Via the control panel, users can access services instantly (on the standard model) and self-manage their server resources, adjusting the amount they need in daily increments. IIJ GIO P2 offers enhanced performance specs compared with the IIJ GIO Component Service. Users can select CPUs supporting up to 24 cores, 192GB of memory, and network bandwidth of 10Gbps. Disk and other server specs can also be customized online. By increasing installed memory capacity to facilitate greater server consolidation, we made it possible for users to design and build servers to their system requirements themselves.

■ Expanded IIJ GIO P2 lineup and West Japan launch

In June 2018, we enhanced the performance of the Virtualization Platform VW Series (P2 VW Series) servers provided as part of IIJ GIO P2, adding VW48-1024-FC-10G—which doubled the number of available CPU cores to 48—and VW96-1024-FC-10G—equipped with 96 cores. We released the former on June 1, 2018, and the latter in October 2018. The added options were geared toward anticipated demand for data-processing infrastructure with high core counts and high memory capacity for applications such as AI information processing and SAP S/4 HANA.

Also in June 2018, we launched IIJ GIO P2 public resources in the West Japan region, adding private resources and storage resources in October that year. The East and West Japan regions are connected by a private backbone service provided by IIJ, and the inter-regional broadband network is made available free of charge. This expansion also broadened the range of viable use cases. For example, users could now configure systems based on their business continuity plans (BCP) by using services across sufficiently geographically separate regions.

In July 2020, we enhanced the P2 VW Series to add the Backup Set / VW Option, facilitating easy, cost-effective VM backups. The Backup Set / VW Option uses RCDM (Rubrik Cloud Data Management) to provide the components needed to back up VMs in a single package. With this option, users can simply select the required plan from the backup settings menu to enable easy backup and restore operations. The acquired data is encrypted and stored on backup servers in IIJ's cloud, eliminating the need for users to build and operate backup systems, thus reducing costs and workloads.

■ From L2 MLAG to L2 over L3

From 2014, we used L2 MLAG as our service platform's network technology. An advantage of this is that decent scale can be achieved even with 10GbE+ node interfaces. MLAG is implemented by the data center switch manufacturer, so both the hardware and software are optimized for data centers, but it must be noted that the designs are not suited to all types of facilities. They are basically designed for fairly large floor areas, so depending on the area of the floor on which you intend to install your equipment, as well as power supply and cooling capacity, it might not be possible to install enough equipment to use up all switch ports. So unless you can design and configure the floor to use chassis (assuming the number of accommodated nodes is determined), it's easy to see how you can get into a situation in which you can't use the equipment efficiently.

From 2017, we adopted an L2 over L3 configuration. While this offers the same benefits as MLAG, its strength lies in being able to design a flexible physical topology using L3 technology, which has a long and proven track record on the Internet. The VxLAN implementation at that time had design constraints, however. The standard implementation was not mature enough to enable dynamic control of, for instance, which VLAN each VTEP is configured for, so our design positioned the VTEPs in locations that would allow them to realistically be managed with static settings.

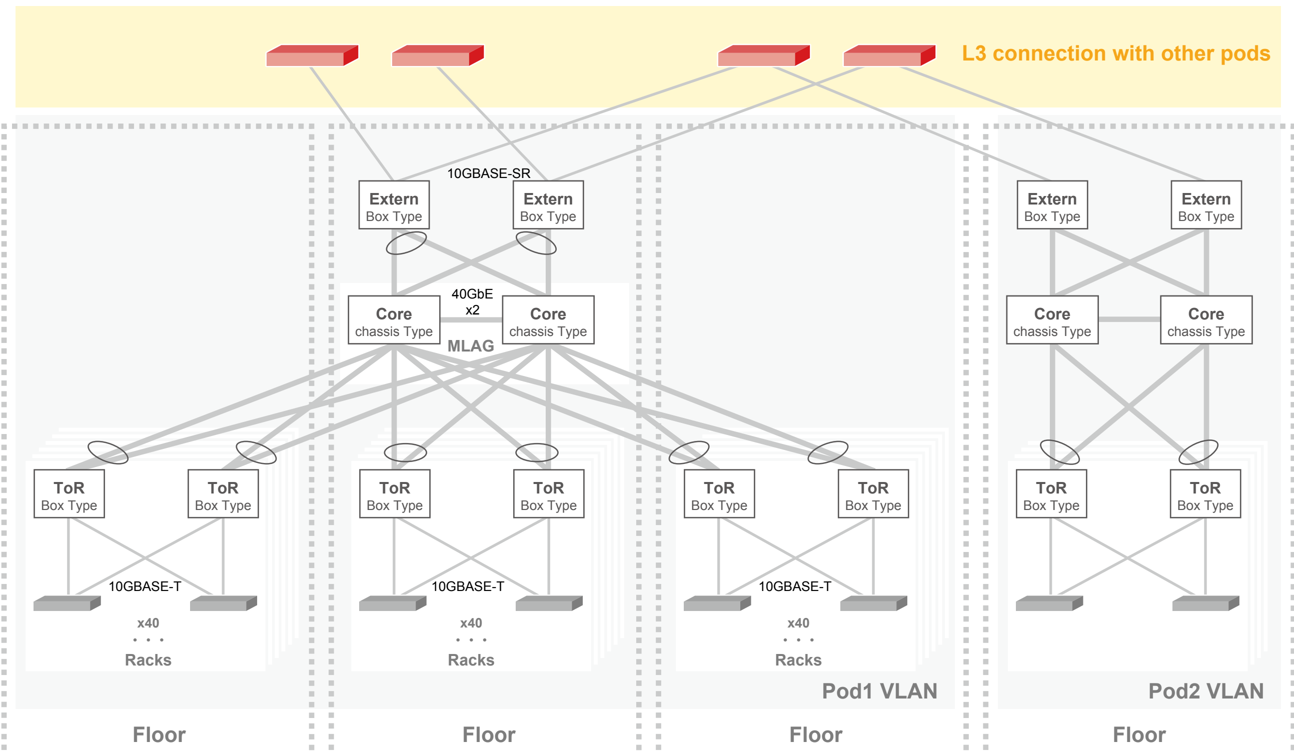

Figure 3: L2 MLAG Configuration (2014–)

Automating storage configuration

Until around 2010, whenever a user requested storage on IIJ GIO, the engineer responsible for running the storage had to manually change the storage and FC switch settings. Up to this point, that only happened infrequently, at most once a week, so even with the engineers performing these changes manually, we were quite capable of fulfilling the lead times set in the service specifications. After 2010, however, that once a week turned into several times a week, and then several times a day, and we could tell it was going to be difficult to sustain things manually, so we designed and built a system to completely automate storage configuration changes.

Storage settings can easily be automated with commercially available applications, but many such applications offer more than just storage configuration automation features (and are thus expensive), so we created the storage control system ourselves. Storage control is performed using scripting languages like Python and Ruby. When using these sorts of languages to manage storage configurations, it is generally preferable to do it via the Storage Management Initiative Specification (SMI-S, a storage management standard) or a proprietary API. The storage and FC switches we were using at the time did not support SMI-S or an API out of the box and instead required separate (expensive) applications, so we used the command line applications that came as standard with the storage products. The input and output capabilities of such command line applications were not really created with humans in mind, and the output in particular was often in a format that is extremely difficult to handle with a scripting language, so the programs you create have to do a lot of cumbersome string processing. Plus, with some of the devices we used, the scripts would return a code of 0 (completed successfully) even when the command or parameters contained an error, so we had to add our own error handling.

The most time-consuming task is that of updating the firmware on storage array systems and FC switches. With some products, such updates can change command line output, so we needed to perform tests in a development environment before upgrading in production. More recent storage products and FC switches increasingly offer support for configuration management software like Ansible, so there has been progress on device API implementations,

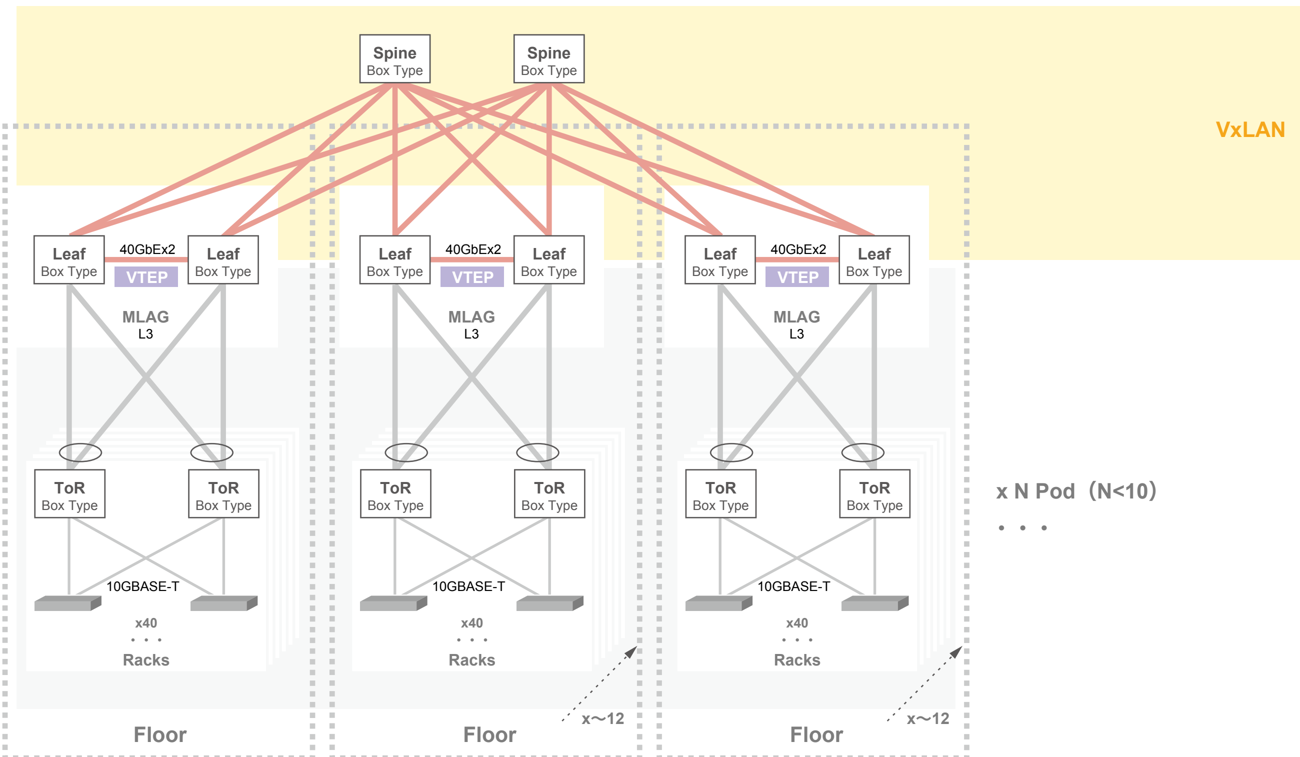

Figure 4: L2 over L3 Configuration (2017–)

and vendors now offer modules that can be used to manage storage using Python and the like, which has made the programming task much easier than it used to be. From 2010 onward, we have continued to follow the procurement policies we used for previous generations, but as performance requirements increase and the range of storage array products meeting our selection criteria grows, the number of different storage products we use is also increasing.

3.7 2020s – Present: Ongoing Evolution

■ Launch of IIJ GIO Infrastructure P2 Gen.2

On October 1, 2021, we launched IIJ GIO Infrastructure P2 Gen.2 (GIO P2 Gen.2), our next-generation IaaS model, which fully integrates the public IaaS and private IaaS offerings developed and provided under the IIJ GIO brand and makes it easy to migrate systems from on-premises environments.

A key characteristic of GIO P2 Gen.2 is that it uses VMware as its virtual infrastructure and allows the design concepts and operational systems of on-premises environments to be migrated as is, and thus, like GIO P2, continues to target demand for the migration of systems from on-premises VMware environments to the cloud. As a successor to GIO P2, it is also naturally designed to serve as a potential new home for existing users. With GIO P2 Gen.2, users are free to create VMs with a minimum of 1vCPU / 4GB memory from a virtual resource pool, instead of on a per-server basis like in ordinary private clouds. In other words, users can migrate the machine specs that they run in their current environment to GIO P2 Gen.2 as is. This setup means that users can migrate physical machines and VMs from their existing environment by using images or performing P2V or V2V migration. In addition to providing a range of managed services with components that are essential in corporate IT environments, such as file servers, Active Directory, and databases, GIO P2 Gen.2 also abstracts out the hypervisors and hardware such as servers, storage, and networks, so users need not worry about differences between devices. This greatly reduces the user workloads that came with the GIO Virtualization Platform VW Series, where users had to perform software updates to deal with vulnerabilities arising from them managing their hypervisors, and migration tasks in the case of hardware upgrades. Linking GIO P2 Gen.2 with IIJ's other various services, such as network services that provide closed connections to third-party public clouds, also makes it possible to use the platform in anticipation of

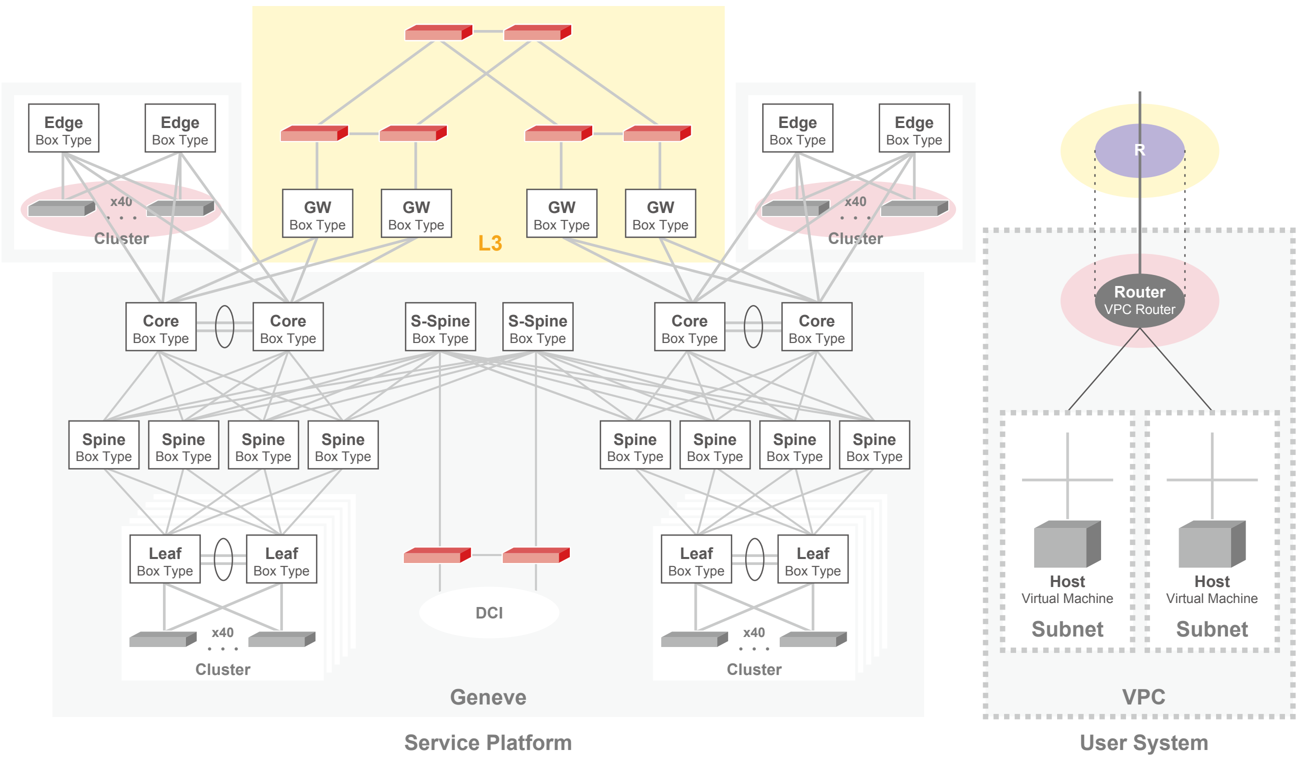

Figure 5: Complete Separate of the Control and Data Planes (2021–)

a future migration to a public cloud. In addition to providing Flexible Server Resources (FSR), which let users freely select resources and configure their system environment just like with a private cloud, we also continue to provide Dedicated Server Resources (DSR) equivalent to what was available on the P2 VW Series using VMware vCenter Server with full privileges.

For GIO P2 Gen.2's Flexible Server Resources, we use VMware Cloud Director (VCD), a product for service providers from VMWare, and hide the hypervisor (vSphere) layer from the user. This lets us provide users with flexible resource control privileges on par with vSphere while allowing IIJ as the service provider to manage the hypervisor and hardware lifecycles. This newly defined shared responsibility model makes it possible for IIJ to manage and operate the hypervisor network. We also provide migration functionality as a service feature, allowing customers to migrate to the cloud with minimal downtime and operational overhead. To operate the hypervisor layer network efficiently, we adopted VMware NSX-T Data Center (NSX-T), which allows comprehensive integration with vSphere, and significantly improved the IaaS network. GIO P2 Gen.2 uses NSX-T to configure an overlay network on top of a Layer 3 IP fabric underlay network, completely separating out the network for each tenant and providing it as a VPC (virtual private cloud). Combining this design with the knowledge we gained from operating a large-scale server pool with IIJ GIO has made it possible for us to allocate resources to users unfettered by the actual physical placement of computing resources (CPUs, memory, storage).

■ Overlay networks using SDN technology

The minimum design requirements for IaaS platforms that accommodate multiple users are physical resource sharing that ensures resources are deployed efficiently and inter-user security. The products on the market offering virtualization technologies for servers as a user computing resource were sufficiently mature, but given the need to ensure interoperability with existing protocols, the implementation of network virtualization is something that has only moved ahead with considerable caution. In recent years, with the improvement of hardware performance and the development of SDN technology, network virtualization technology has become a viable option in large-scale networks. We recognized the usefulness of overlay networks, a type of network virtualization technology, from an early stage, and we have used them through the GIO platform generations, selecting reliable technologies in each instance.

To ensure performance, we previously used network hardware functions to achieve overlay network termination, but with GIO P2 Gen.2, we switched to doing this using the virtual switches within the hosts. This, of course, ensured complete separation between users and also enabled the complete separation of user systems and the physical infrastructure systems. This reduced the impact of underlying system changes on user systems, making it easier to add new functionality.

The loose coupling with the physical layer is also advantageous when deploying IaaS across multiple sites. As with the existing GIO system, not only can equipment be deployed on a large scale across a small number of locations in East and West Japan and seamlessly connected, we also expect it to be easy to distribute equipment on smaller scales across various locations as a disaster preparedness measure.

On the other hand, whereas the settings for each device were previously managed separately due to hardware limitations, the shift to processing in software removed these limitations, resulting in a truly huge amount of configuration information. Since this exceeds what a human can possibly grasp, to ensure quality, we need to move away from the conventional practice of having humans manually make changes. And to ensure services are available to users in a timely manner, we also need to move away from static device configuration and enable dynamic configuration. To achieve these goals, we are using an orchestrator (based on a commercial product with missing or additional features developed in-house) to centrally manage configurations. Centralizing in this manner makes inter-system processing available to linked services via APIs as well, and makes it possible to start related services safely and quickly.

3.8 Conclusion

We've taken a look back at IIJ's 30-year history through a service infrastructure lens. The service host infrastructure underpinning IIJ's various services continues to evolve as we seek to balance stability and efficiency while keeping an eye on the relentless march of innovation and adopting the best technologies to keep up with the times. We mark IIJ GIO's 24th anniversary this year, and the platform's service lineup continues to expand to meet diverse business infrastructure needs. With demand for digital social infrastructure based on AI technology growing in recent years, we are working to provide ultra-high-density AI computing platforms to make this a reality. At IIJ, we will continue to develop and provide services and fundamental technologies to serve market needs.

Shinri Kimura

Head of Cloud Services Division 2, Cloud Division, and Director of Technology Development, Cloud Division, IIJ
Since February 2001, Mr. Kimura has been engaged in the development and operation of large-scale service and cloud infrastructure, and the provision of cloud services. He is also responsible for teams engaged in technology development, human resources development, etc.