Rolling Out OCP Servers: Our Path to Putting ORv3 into Production

2.1 Introduction

This article takes a look back at how we at Internet Initiative Japan Inc. (IIJ) came to view servers based on the Open Compute Project (OCP), and how, based on the lessons we drew from Open Rack v2 (ORv2), we came to put Open Rack v3 (ORv3)-generation servers into commercial use.

OCP is an effort to publish specifications for servers, racks, power supplies, and the like so that multiple manufacturers can build products to the same spec. Open Rack refers to the part of those specs that defines racks and power systems.

IIJ weighed ORv2 against a number of considerations before rolling it out in production: Was it suited to the power and rack specs in our own data centers? Would the hardware be available domestically on an ongoing basis? Could our operations teams handle it without trouble? Was the cost reasonable?

With ORv3, too, we waited until the full range of products was available and the path to deploying the hardware in Japan was clear, and we then started a proof-of-concept study (PoC) and moved to commercial deployment in 2025.

Below we walk through changes in our thinking on server procurement, what we learned from ORv2, and how evaluating ORv3 led to its commercial deployment in 2025.

2.2 How IIJ's Thinking on Server Procurement Has Evolved

During the mid-1990s, IIJ ran servers that combined PICMG CPU boards and I/O boards. Rather than using prebuilt servers, we procured custom servers made up of the components we needed.

Once the 2000s arrived, the options for PICMG-related products shrank while the number of servers we were procuring grew, so our old approach became hard to sustain in terms of both lead times and cost.

We then shifted to Sun's SPARC servers and enterprise x86 servers like NEC's Express5800 series and COMPAQ's ProLiant DL series. From around this time, we began to put more priority on ease of procurement, maintenance arrangements, and product completeness rather than the idea of optimizing individual components.

While it had become harder to sustain the approach of assembling servers to suit our own needs the way we had with PICMG, the need to think about the optimal server configurations for IIJ had by no means gone away. The reality was that, considering our procurement volumes and maintenance needs, we had little choice but to work within the specs that enterprise server vendors offered. Building servers to the specs we needed in-house was an option, but given the size of our operations, we could see that such an approach would not scale.

We reached a major turning point in 2009 with the launch of the IIJ GIO service. The initial rollout alone meant procuring roughly a thousand servers, and we had to design a system that held together as a whole, covering not just individual server specs but also power, racks, thermal design, and maintenance.

Subsequently, to handle the installation of over 5,000 units a year, we began bringing racks into the factory to mount the servers and do everything including cabling before then delivering complete racks. This made it possible to handle large rollouts while keeping on-site work time down.

We later learned about OCP servers, but at IIJ's scale of procurement, we were already able to keep costs low enough with enterprise servers, so we had no strong reason to opt for OCP right away. So at this point, we worked within our existing procurement model and put priority on finishing builds as fast as possible.

2.3 Why We Were Able to Take On ORv2

Around spring 2018, we received a proposal on ORv2 from Itochu Techno-Solutions Corporation (CTC), and this illuminated the potential for procuring the hardware more cheaply than conventional enterprise servers. So in September 2018 we visited an ODM in Taiwan with CTC to inspect its product lineup, development roadmap, and maintenance setup.

We also learned how the hyperscalers handle procurement. The motherboard and chassis do not account for all that much of a finished server's price, so for major components like CPU, memory, and SSDs, it often ends up being cheaper negotiating terms directly with the parts makers.

We had been introduced to a chipmaker's distributor right around that time, so during a trip to South Korea the following month, we also visited the company's headquarters there and worked out terms for memory and SSDs on the basis that the ODM would procure them on IIJ's terms. With another chipmaker, we ended up going through a distributor instead, but we were able to work out terms in the same way.

What makes this procurement model distinct is that you negotiate terms down to the component level on the premise of buying in volume and have the ODM source a configuration that fits your requirements. We believed this approach would work well at the scale of IIJ's operations and in the context of IIJ's data center operations, and that it would result in a reasonable cost structure.

2.4 The ORv2 System We Deployed and What We Learned Running It

The ORv2 system we deployed at our Shiroi Data Center Campus in 2019 accommodates 42 nodes per rack, and the centralized power supply draws single-phase 230V via four feeds. We based this configuration on an earlier deployment at a data center in Japan that we had toured around that time.

Data center floor space matters a lot in Japan, so IIJ focused on increasing rack utilization. It is not uncommon with typical enterprise server racks to limit the number of installed units to allow for peak power draw, but real workloads rarely use up that much power. As our racks are mostly for our own use, and because we had the ability to work with the user side to move hardware to other racks when needed, we avoided having too much design headroom for ORv2 and instead designed our system to fill the racks from top to bottom. We used a centralized power supply from Murata Manufacturing, and we had Nitto Kogyo Corporation build our racks to ORv2 spec at 700mm wide with front and rear doors. This fit the needs of IIJ's machine rooms while also leaving space for cabling inside the rack.

Running ORv2 in production also delivered clear results. We found that the configuration deployed at IIJ cut power consumption by roughly 30% versus off-the-shelf servers with similar specs. While there is no single factor behind this, within the scope of what we evaluated, the servers the ODM builds made it easy to choose the necessary and sufficient configuration for the task at hand, making it easy to keep power consumption down. OCP servers also use centralized power supplies, which raise power conversion efficiency, and we believe this made further power savings easier still. On the maintenance side, meanwhile, we found that a bit more work was involved in on-site fault isolation and replacement compared with pre-built servers, and that staff needed the skills to gauge the condition of individual components as they worked.

Some tasks, like CPU replacement, did mean more on-site work, but for management functionality, a serial console over IPMI was sufficient for us, so we encountered no major problems even with the simple management functionality provided by the ODM.

2.5 Why We Didn't Deploy ORv3 Right Away

The original plan was to keep using the same racks after deploying ORv2 and simply refresh the servers in them. In reality, though, socket form factor changes accompanying Intel's CPU generation updates, along with a market-wide shift in development resources toward AMD EPYC, meant that the range of ORv2-compatible products was gradually narrowing. It thus became harder to keep the ORv2 racks and simply refresh the servers.

Meanwhile, although the ORv3 spec had been published in 2020, the key supporting products, such as racks, centralized power supplies, and compatible servers, were not yet fully available at the time, so we were not in a position to make a decision on adopting it for our production environment. What mattered to IIJ was not that a standard existed, but whether we could procure the hardware within Japan on an ongoing basis, source maintenance parts in the event of failures, and handle everything comfortably within our data center operations. So when procuring new servers during this period, we chose enterprise rather than OCP servers. With ORv3, we decided to wait until a full range of products was available and the path to deploying the hardware in Japan was clear.

2.6 OCP Global Summit as a Turning Point

The turning point came at the OCP Global Summit held in San Jose in 2024. IIJ was looking for the ability to procure hardware within Japan on an ongoing basis, source maintenance parts, and comfortably deploy the equipment in its own data centers.

We had heard that Nitto Kogyo would be exhibiting ORv3 racks and Murata Manufacturing ORv3 centralized power supplies, so at that point we could already see the key building blocks coming together. Seeing the actual hardware on display at the summit helped us envision the prospect of actually assembling and evaluating the components in Japan, and we moved straight into a PoC upon returning home.

We borrowed a rack and centralized power supply of the same spec as those we had seen at the show, and we evaluated ORv3 servers from several manufacturers. We compared one-node-per-1OU and two-nodes-per-2OU configurations, ultimately adopting the latter for its cooling efficiency.

What we were looking at here was not just whether the servers held up on their own but whether we could handle them comfortably within our existing data center operations.

2.7 IIJ's Key Considerations When Evaluating ORv3

What we focused on most in our PoC was how the change in power delivery would affect real-world operations, whether the rack width and cabling would be viable, and whether we could put the servers into a production configuration as-is.

With ORv3, the voltage supplied to the servers changed from 12V to 48V. This was a natural change to make as power consumption climbed with CPUs, memory, and accelerators growing more powerful. The centralized power supply also no longer needs to sit in the middle of the rack, allowing a greater degree of design freedom.

For the PoC, we had Nitto Kogyo prepare a 600mm-wide rack with front and rear doors, but we found that plugging and unplugging cables and securing cable routes was a struggle. So we had them redesign the rack to 700mm wide, in line with ORv2.

The service outlets on Murata's centralized power supply disappeared when it moved to the ORv3 version. So for IIJ's setup, we needed a separate PDU to provide the power we previously drew from the service outlets. These decisions carried straight over into the production configuration described in the next section.

2.8 The ORv3 Production Environment We Deployed in 2025

With the PoC done, we began building out ORv3 at IIJ in July 2025. Our design accommodates 34 nodes per rack at 13kVA, with power supplied via two three-phase feeds. With ORv3, we again worked on the premise of racks being filled from top to bottom, so we kept rack utilization high while supporting higher power density than in the ORv2 era.

The ToR switch we adopted for this configuration is made by Juniper and is not ORv3 compliant. It thus needs a separate single-phase 230V feed, so we drew single-phase power off the two three-phase feeds coming into the rack. We had Harting make custom cables for this, and we connected a single-phase PDU downstream to supply power to the ToR switches. Our design balances the phases across the entire facility rather than within each individual rack.

2.9 How We Made OCP Server Maintenance Work

With OCP servers, it's not uncommon for recovery efforts to involve isolating faults at the component level, as opposed to the equipment being maintained as an indivisible finished product. Working within the ODM's warranty, IIJ handles initial on-site fault isolation and parts replacement in-house, while CTC takes care of returning failed parts and coordinating with vendors. Failed parts are sent to a repair center in Taiwan as needed.

This differs from how maintenance works with enterprise server vendors, but with this clear delineation of roles, we have made OCP server maintenance work in an actual production environment at IIJ.

2.10 Why We Limit OCP Server Deployment

At IIJ, we limit OCP server deployments to data centers we operate ourselves, such as the Matsue Data Center Park and the Shiroi Data Center Campus. This is because deploying OCP servers and Open Rack requires detailed advance coordination with the data center side on power systems, rack specs, and installation conditions. Also, while total rack power consumption is not pinned to its ceiling under real workloads, densely packed racks do produce localized hot spots. So to take advantage of OCP servers' high rack utilization, we had to work with the data center side, including on cooling and placement conditions, to figure out just how far we could take the build-out. In some cases this means locally lowering cooling setpoints or adjusting airflow direction.

This equipment is not something you can install anywhere you want under the same conditions like you would with typical 19-inch rack servers. So it is more efficient to deploy at a decent scale than in small, one-off installations. This ties back to something IIJ has been aware of since around 2009: procuring in bulk increases the degree of design freedom.

2.11 Power Density Challenges in the SPR Generation and Planning the Next Configuration

While the 2025 commercial rollout was underway, IIJ was also working on its next configuration. The 4th Gen Intel Xeon Scalable Processor (Sapphire Rapids-SP, or SPR) generation, which IIJ had been evaluating since autumn 2024, came with performance gains but also increased power consumption, so per-rack power density and power delivery design mattered more than the performance of any single server.

So the question was not how the performance of individual CPUs compares but how to make the rack function properly as a whole as power density climbs.

For the CPUs we will adopt in 2026, we have chosen Intel Xeon 6 (Granite Rapids, or GNR) and 5th Gen AMD EPYC (Turin). Turin leads on power efficiency, but given CPU generation compatibility, we also needed GNR as a live-migration target from our current Intel Xeon servers.

For the chassis, we selected the 2U2Node format, which delivered good power efficiency in the SPR generation, for the GNR generation as well. Our decision was based not just on CPU performance in isolation but also on whether the configuration would work properly given our per-rack node count, power requirements, and the characteristics of our services.

2.12 Conclusion

In this article, we have taken a look back at how IIJ took on ORv2 and tested ODM-based procurement and whole-rack configurations, including centralized power supplies, in real operations. We then walked through how we built on the lessons from ORv2 and moved on to the ORv3 PoC and commercial deployment once the key building blocks were in place and the prospect of being able to handle the hardware comfortably under our operating conditions had emerged.

Along the way, we had to change how we work, not just the way we chose servers but also our approach to procurement and system building. On the procurement side, rather than simply comparing products from enterprise server vendors, we increasingly had to think about configurations and pricing with ODMs in mind. On the system building side, we also had to make the whole system work—not just the individual servers but the racks, power delivery, cabling, and maintenance operations as well.

The end result is that adopting OCP servers has broadened IIJ's options for procurement and system building. It has become clear, however, that one needs to be selective about where OCP servers are deployed and the conditions under which they are operated. Looking ahead, and with those constraints in mind, we hope to continue taking advantage of OCP servers as one of the options available to us at IIJ.

Masahiro Takahata
Manager, System Infrastructure Engineering Section, Infrastructure Engineering Department, Infrastructure Engineering Division, Network Services Business Unit, IIJ
Since joining IIJ in 2001, Mr. Takahata has been involved in the design and construction of server infrastructure. In 2008, he served as a core team member in the launch of NHN, the infrastructure underlying IIJ's services, and in 2009 he was responsible for building the infrastructure for IIJ GIO. He subsequently focused on operating the NHN infrastructure, and in 2011 he carried out the deployment of server infrastructure facilities in the containerized data center at Matsue Data Center Park. In 2019, upon completion of IIJ's Shiroi Data Center Campus, he worked on the rollout of ORv2 OCP servers, and since 2025 he has been working on the rollout of ORv3 OCP servers.