IP over DWDM

3.1 Introduction

Since its founding, IIJ has designed, built, and operated its own Internet backbone. This is key to ensuring service quality as an ISP and enables IIJ to make decisions regarding traffic dynamics and technology selection entirely in-house.

In recent years, both domestic and international Internet traffic has been on a constant rise, growing faster than ever before, with an increasing number of traffic spikes due to sudden demand. Because the backbone is the core of IIJ's service infrastructure, allowing it to become congested is not an option. It is therefore crucial that we make timely capacity expansions in response to demand and plan our architecture with the future in mind.

That said, the previous architecture presented some challenges. Procuring additional carrier circuits to connect the backbone sites required long lead times, and costs scaled linearly. Moreover, the use of transponder-based WDM (Wavelength Division Multiplexing) systems came with operational complexity, and capacity upgrades had to be made in large increments, limiting flexibility.

It was against this backdrop that IIJ decided to evaluate IP over DWDM as a next-generation technology for expanding its backbone, and it began deploying this in its commercial systems in 2025. This article covers the pre-deployment validation process, the commercial network architecture and its benefits, and the future outlook.

3.2 WDM and the IIJ Backbone

IIJ's backbone connects domestic and international sites using high-capacity circuits. The technology used for these inter-site connections is known as WDM. WDM transmits multiple optical signals at different wavelengths simultaneously over a single fiber, and this greatly increases both capacity and transmission distance.

DWDM primarily uses wavelengths in the C-Band and L-Band, which exhibit low propagation loss in optical fiber and are easily amplified by EDFAs (Erbium-Doped Fiber Amplifiers), making them particularly well-suited for high-capacity long-haul transmission. Modern systems also employ coherent transmission, a more advanced technique than IMDD, and this uses phase and polarization information to achieve per-wavelength bit rates of 100G, 200G, 400G, and 800G over distances ranging from several hundred to several thousand kilometers. Carriers that own their own fiber, such as NTT, KDDI, and SoftBank, use WDM-based transport equipment to provide dedicated line services, and IIJ builds its backbone by using these services to connect its sites.

Since around 2006, IIJ has been designing, building, and operating its own WDM-based transport equipment for intra-Tokyo inter-site connections where particularly high capacity was anticipated.

At the time, 10 Gigabit Ethernet was the mainstream, and with IIJ using an architecture known as Backbone Fabric (BF)[*1], it

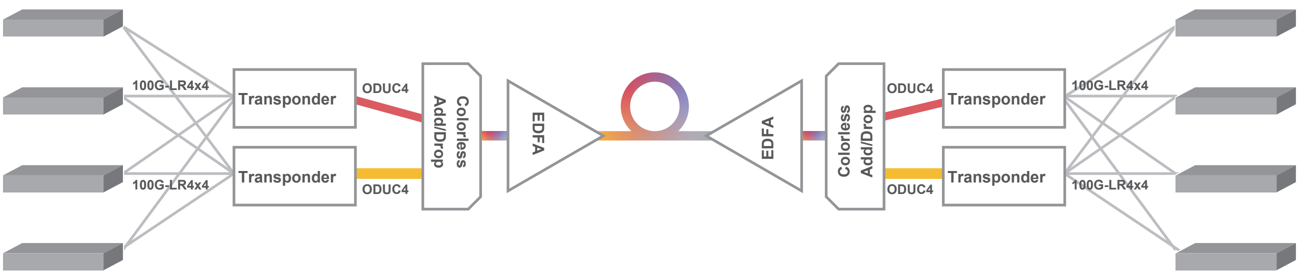


Figure 1: Conventional DWDM-based Inter-site Connectivity on the IIJ Backbone

*1 For more information on Backbone Fabric (BF), see "2. Focused Research (1): VX—IIJ's New Backbone Network" in IIR Vol. 57 (https://www.iij.ad.jp/en/dev/iir/057.html).

required a large number of 10G links between sites, and this is why we deployed 10G DWDM equipment. From the 2010s onward, the introduction of 100 Gigabit Ethernet and the migration to an MPLS L2VPN architecture called WARP enabled consolidation of both circuit utilization and circuit counts, reducing the importance of 10G DWDM within IIJ's backbone.

As traffic subsequently grew, some segments were fielding over 100G of traffic, rendering statistical multiplexing ineffective on L2VPNs and links between MPLS routers. So we deployed 100 Gigabit Ethernet-capable DWDM systems and proceeded with migration.

This is how IIJ expanded its backbone using WDM, but the striking growth in traffic in recent years has created problems for the conventional approach in terms of speed and cost of expansion. IIJ thus turned its attention to IP over DWDM as a means of enabling more flexible backbone build-out, and began validating the technology and considering its deployment on the backbone.

3.3 Pre-Deployment Validation of IP over DWDM

3.3.1 IP over DWDM

IP over DWDM enables routers and switches to use DWDM without requiring separate transport equipment. In the past, you needed to have a mechanism between the routers and the DWDM optical transport network called an OLS (Optical Line System), comprising dedicated transponders, wavelength filters, and amplifiers. But with IP over DWDM, the transponder is replaced by pluggable Digital Coherent Optics (DCO) modules inserted directly into the router, facilitating tighter integration between the router and the physical layer. As a result, the network architecture is simplified, and shorter lead times can be expected when building out and expanding the capacity of the backbone. We discuss these beneficial aspects of IP over DWDM below. IIJ conducted a range of tests in preparation for deploying this technology on its commercial systems.

3.3.2 Rigorous DCO/OLS Validation

In recent years, DCOs available in form factors such as QSFP-DD and OSFP have rapidly gained traction, enabling long-haul transmission directly from router ports. Standards such as 400ZR and OpenZR+ have been established, attracting attention as alternatives to the costly dedicated transponders of the past.

Implementations of 400ZR are developed under implementation agreements and common technical specifications defined by the OIF (Optical Internetworking Forum)[*2], while implementations of OpenZR+ (400G-ZR+) are developed under those defined by the OpenZR+ MSA[*3]. But just because they are standardized does not mean they are plug-and-play at all. In practice, the following challenges exist.

- Compatibility issues arising from specification differences between router/switch vendors and DCO vendors
- Interoperability among DCOs from different vendors
- Performance variations across DCO vendors

IIJ began validating these modules incrementally from around 2021, and in 2024 it conducted hands-on testing of OLS and 400ZR modules from multiple vendors.

3.3.3 Cross-Vendor Interoperability Validation

Validation tests combining 400ZR/ZR+ modules from Juniper and Cisco brought up the following issues.

1. Certain combinations failed to link up even when based on the same standard
2. Issues arising from differences in vendor implementations
3. Cases in which tunable settings were not applied and link-up failed after a wavelength change

Issue 1 is not unique to DCOs; transceivers can also exhibit compatibility issues. Specifically, different transceiver vendors have varying tolerances for signal waveform quality, and certain combinations can result in unstable links. As part of pre-deployment validation, we performed integration testing for each combination of router, OS version, and DCO vendor that we anticipated using. In the course of this testing, one particular vendor combination showed degradation in optical signal quality, including OSNR (Optical Signal-to-Noise Ratio) and PRS (Polarization Rotation Speed), resulting in link instability.

*2 OIF, OIF-400ZR-03.0 (https://www.oiforum.com/wp-content/uploads/OIF-400ZR-03.0.1.pdf).
*3 OpenZR+, OpenZR+ Specifications, version 3.0, 12 September 2023 (https://openzrplus.org/resources/openzr-specifications-v-3-0/).

Regarding issue 2, the main problem involved the Application Select Code (AppSel) values. AppSel is a mechanism by which DCO modules advertise to the router the data rates, modulation schemes, and FEC (Forward Error Correction) methods they support. Because these are not explicitly defined in the relevant implementation agreements or specifications, they depend on the implementations used by the router and transceiver vendors, and this resulted in cases in which the expected modulation scheme was not set or could not be changed.

As for issue 3, because DCO support in router OSes was not yet fully mature, we encountered timing-related issues in which tuning settings were not applied or link-up could not be achieved.

Because such behavior poses risks in commercial deployment, IIJ reproduced multiple patterns, isolated the issues and documented the conditions under which they occurred, and worked with vendors to request fixes and share findings.

3.3.4 Heat Dissipation and Power Consumption Issues

400ZR modules typically generate significant heat, around 20 W. This heat is dissipated by drawing air in from the front of the device, passing it over the transceiver heatsink, and exhausting it via fans at the rear. QSFP-DD 400ZR/ZR+ modules commonly use a thicker-than-standard heatsink known as Type 2A. We found that these could impede front-panel intake airflow and affect cooling performance.

Figure 2 shows photographs of a 400ZR module installed in a router, illustrating how the heatsink interferes with the front air intake. Advisories regarding the airflow characteristics of such equipment are sometimes issued by router/switch vendors.

Care must also be taken in commercial environments as transceivers may be installed not only above and below but also to the left and right, so heat from adjacent ports can propagate to the DCO, causing increased heat generation.

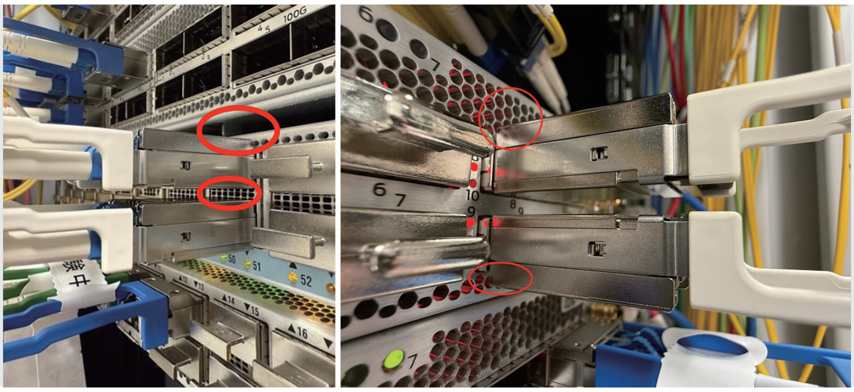


Figure 2: Examples of 400ZR Cooling Mechanism

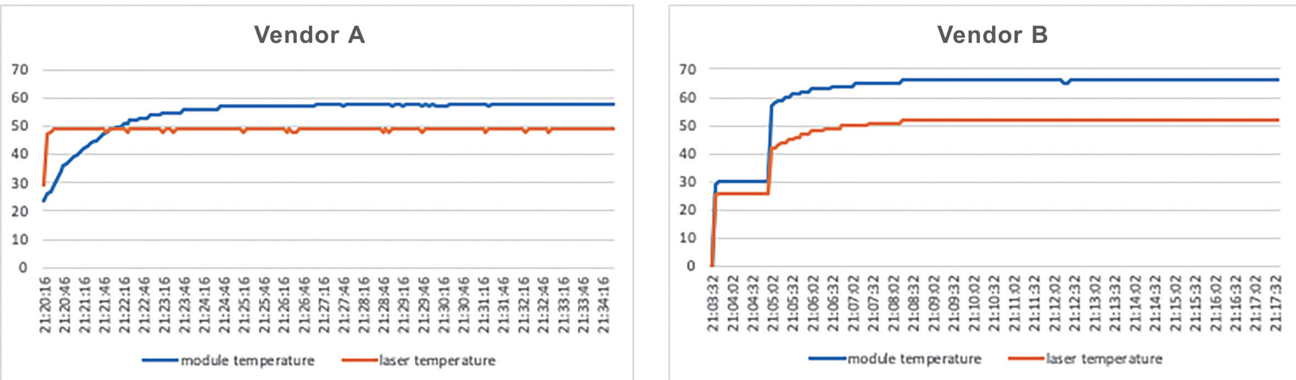


Figure 3: Thermal Performance Comparison of Two Vendors' DCO Modules

Performance differences between DCO vendors were also quite noticeable. Figure 3 compares the thermal performance of DCOs from two vendors under identical conditions. We observed a difference of around 8°C between the two. Caution is needed because if the temperature exceeds the threshold of either the host device or the DCO, a forced shutdown is triggered.

3.3.5 Challenges and Countermeasures for OLS Integration

cAs described in Section 3.2, IIJ had an existing 100G DWDM network, so we conducted validation by connecting third-party optical signals (alien wavelengths) to the existing OLS (Figure 4).

The advantage of this approach was that it required no additional investment and allowed spare wavelengths to be used effectively. Our integration testing revealed several issues, however.

First, while transponders typically have an optical transmit power of around +1 dBm to +3 dBm, DCOs are generally designed for output levels of around -10 dBm. When there are significant power differences across wavelengths, this makes amplifier tuning considerably more difficult.

Additionally, the existing OLS used Colorless Add/Drop. Colorless Add/Drop has high insertion loss, and because the degree-side signals are not filtered, the multiplexed optical signal is delivered directly to the DCO. The DCO can process its own wavelength, so link-up and operation are possible, but many DCOs have a maximum receive level of around 0 dBm, so even if the system is tuned to provide the proper level for the DCO's own channel, the total received optical power will often exceed that limit.

The combination of low DCO transmit power and the characteristics of Colorless Add/Drop also resulted in higher

Figure 4: Third-Party Signal Interconnection Using Existing OLS (During Deployment Testing)

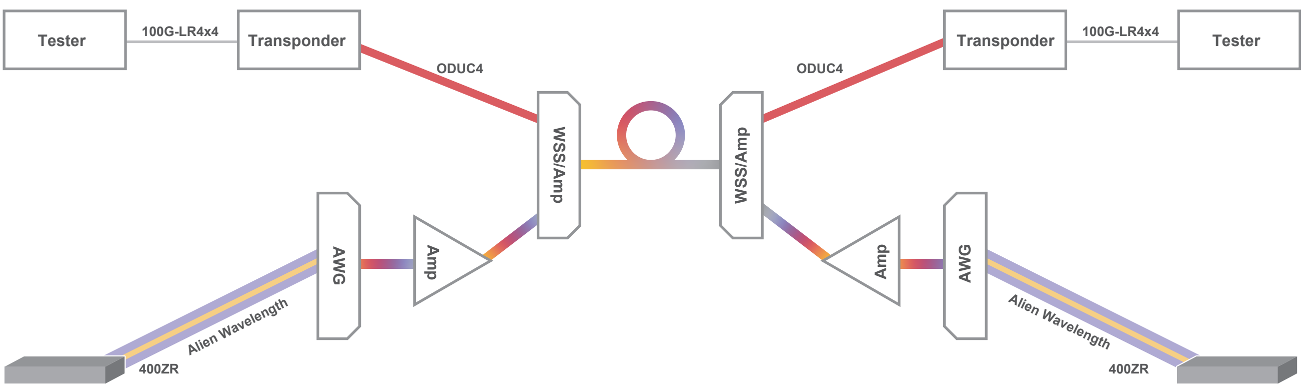


Figure 5: OLS Configuration Using IP over DWDM (During Deployment Testing)

performance demands on the booster amplifier (on the transmitter side). Around the same time, we were also evaluating High Tx output power (±0 dBm) variants, which would have mitigated both the amplifier tuning difficulty and the maximum receive power issue.

Our validation work confirmed that these challenges posed significant barriers to deploying 400ZR using the existing OLS. Accordingly, we selected a new OLS optimized for IP over DWDM. Ultimately, the validation environment used the OLS architecture shown in Figure 5. This enabled the coexistence of transponders and DCOs. The use of AWGs also meant that signals on channels other than the one in use were filtered out, which also solved the problem of the DCO's maximum receive level being exceeded.

During final testing with this configuration, however, we encountered a new issue. As part of failure testing, when light was looped back to the AWG at one end, the link on an adjacent channel that was not under test began flapping.

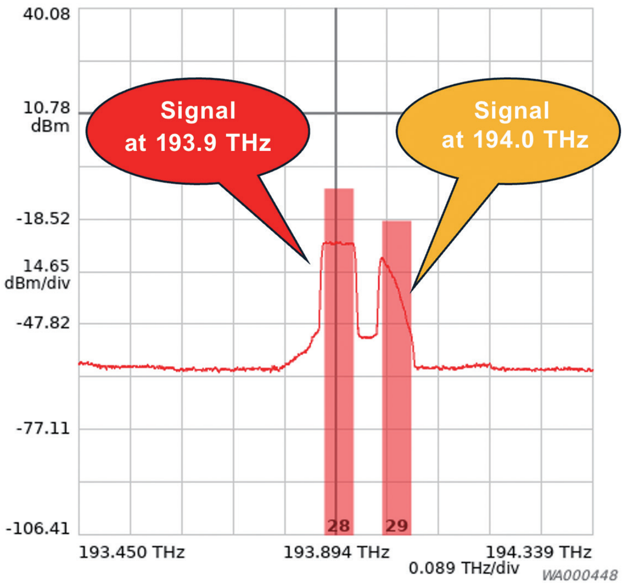


Figure 7: Signal Leakage into an Adjacent Channel

When we investigated this, we found that signals that were supposed to be filtered out by the AWG were leaking into adjacent channels more than expected. This leakage acted as noise on the main signal, causing quality degradation and link flapping. This issue can also arise if a user accidentally configures a wavelength offset by one channel from the intended one, or during wavelength scanning by a DCO with an automatic wavelength-setting function. So for the commercial deployment of this model, we decided to establish a rule prohibiting the assignment of adjacent channels and to recommend the use of multiplexers and demultiplexers with sufficient adjacent-channel isolation.

3.4 Deployment on Commercial Networks

3.4.1 The Current IIJ Backbone

As described in "Focused Research (3): The IIJ Backbone—30 Years of Transformations" in IIR Vol. 58 (https://www.iij.ad.jp/en/dev/iir/058.html), IIJ's backbone currently interconnects multiple sites in major cities in Japan and overseas, primarily using 100 Gigabit Ethernet. To date, the backbone has been expanded by adding 100G circuits to the required links in accord with traffic volumes. Yet there have been a number of inherent challenges in this approach.

First, the lead times for procuring carrier circuits are extremely long. Procurement requires sharing expansion plans, negotiating fiber routes and pricing, and coordinating construction schedules between companies. This can take several months at minimum and close to a year in some cases, making it difficult to respond swiftly enough to sudden traffic increases. Second, circuit costs rise in proportion to traffic increases, so we needed a more efficient approach.

Figure 6: Issues with IP over DWDM

As noted at the beginning, IIJ had long used its own 10G/100G DWDM transport equipment as a complementary measure to address these challenges, particularly on links where high traffic volumes were anticipated, thus maintaining backbone flexibility through a combination of carrier circuits and in-house transport. That said, expanding this in-house DWDM transport was by no means easy. Once the capacity of an existing deployment was exhausted, several months were needed to procure and build new equipment, and the minimum expansion increment was on the order of several hundred gigabits. This meant the company needed to make careful investment decisions that accounted not only for near-term demand but also future traffic projections. So with this approach of relying solely on in-house DWDM transport, it was gradually becoming more difficult to maintain the backbone.

It was against this backdrop that IIJ began evaluating IP over DWDM as a new approach to inter-site backbone connectivity.

3.4.2 Our Expectations for IP over DWDM

IP over DWDM has several features that set it apart from conventional DWDM systems, and we saw it as quite a promising way to address the challenges IIJ was facing.

The biggest advantage was shorter lead times for capacity expansion. Once the OLS has been put in place, expanding capacity simply means installing 400ZR transceivers (which can be procured more quickly than transponders) in router ports and applying the necessary configuration. There is no need to add new modules for DWDM transport equipment, which had traditionally involved long lead times. This can be expected to significantly reduce both procurement time and deployment time.

This shift in architecture also has a positive effect on backbone operating costs. Compared with conventional DWDM equipment, it promises lower hardware costs as well as reduced power consumption and rack space requirements. There are operational benefits as well. Conventional optical transport equipment requires dedicated OSes and proprietary management frameworks, making it necessary to have operations staff with specialized DWDM expertise. With IP over DWDM, on the other hand, optical quality checks and transmission monitoring can be handled entirely on the router side. That makes it a good fit for existing operational processes, and we also expected it to reduce operational overhead.

For these reasons, we felt IP over DWDM was a promising way to address the challenges IIJ faced in building its backbone, and we moved forward with validation to determine how best to apply it to the commercial network.

3.4.3 Commercial Deployment at the New Osaka Core Site

For commercial deployment of IP over DWDM, IIJ selected the backbone link between Osaka-Kita, a new core site built in 2025, and the existing Osaka-Chuo site. This link met the requirements identified during field validation: a distance of 30 km or less and span loss of 25 dB or less; and it had been confirmed as a suitable environment for 400ZR deployment. And as a newly established site, it allowed relatively high flexibility in equipment selection and placement, making it an ideal link for the initial deployment of the new technology.

For the deployment, one of the two fiber paths was provisioned using conventional 100G DWDM equipment, while the other used IP over DWDM. This configuration was chosen with redundancy in mind, taking into account the possibility of unforeseen issues during the early stages of deployment and ensuring that the backbone could be maintained even if problems occurred with IP over DWDM.

When selecting DCOs, we considered both 400ZR and 400ZR+ (OpenZR+) as candidates. Because both satisfied the distance and OSNR requirements, we selected the lower-cost 400ZR. At the same time, we also performed a comparison of the High Tx output power version with an optical output level of around 0 dBm and the Normal Power version at around -10 dBm. We selected the Normal Power version as there were concerns that troubleshooting would be more difficult with the High Tx output power version, for which no router-vendor-supplied modules were available.

For the OLS design, we adopted the commonly used C-band and used 100 GHz grid passive filters for wavelength multiplexing. To avoid the risk of adjacent-channel interference identified during validation, channels were spaced at 200 GHz intervals in the production environment. As this was a simple two-site connection, we selected an OLS that did not require a proprietary controller, thus avoiding vendor lock-in, and that provided automatic amplifier gain adjustment to reduce operational overhead. We worked closely with engineers at equipment manufacturers to determine which equipment would be optimal. IIJ places importance on this kind of direct dialogue with manufacturers and vendors as well as on establishing a structure that enables technical questions to be resolved quickly even after deployment.

Integrating IP over DWDM into our operations required the establishment of new operational workflows, including new optical quality monitoring metrics, 400ZR-specific quality thresholds, and redesigned fault-isolation procedures. Previously, the teams responsible for backbone routers and optical transport equipment were clearly separated, but with the introduction of IP over DWDM, routers also came to handle optical processing. This lowered the barrier by allowing operators to check status using the routers they were already familiar with, while at the same time creating a need for router operators to understand the basics of optical transport. Standardizing monitoring settings and developing troubleshooting procedures was therefore key.

Following this process of evaluation and preparation, IIJ deployed IP over DWDM in its backbone, and as of January 2026, it remains in stable operation. We attribute this success to the thorough validation and careful design work that preceded it.

3.4.4 Benefits of Deployment

This commercial deployment confirmed that IP over DWDM delivers substantial benefits for the IIJ backbone. The IP over DWDM configuration using two 400ZR wavelengths delivered a roughly 52% reduction in costs versus building the same 800G of capacity using conventional 100G DWDM transport equipment. This is attributable not only to the low cost of the 400ZR transceivers themselves but also to the reduction in transport equipment components.

From a procurement perspective, 400ZR transceivers can be procured with shorter lead times than conventional DWDM equipment, significantly shortening the lead time for backbone expansion. In addition, the OLS deployed in this case was commercially designed to support up to 22 multiplexed wavelengths, meaning that for the foreseeable future, capacity can be increased simply by adding transceivers, without the need to build additional OLS infrastructure on the same link. Although 400ZR was selected in this case because high traffic volume was expected on the link from the outset, DCOs that enable capacity expansion in 100G increments have recently become available. Going forward, we can therefore expect this approach to be applicable to links using routers that do not support 400G and those for which 100G capacity is sufficient.

There are also benefits from a power-consumption perspective. Compared with building 800G of capacity using conventional DWDM transport equipment, the IP over DWDM configuration reduces power consumption by around 10%. Because power consumption on the OLS side does not depend on the amount of capacity in use,

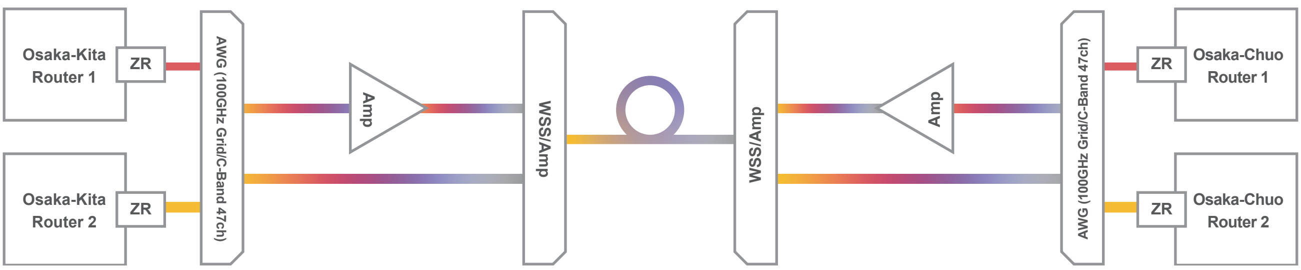


Figure 8: IP over DWDM Configuration at the New Osaka Core Site

even greater energy savings can be expected as the number of wavelengths in use increases in the future.

Space efficiency, however, remains a challenge. The configuration selected for this deployment required somewhat more rack space than conventional equipment, but this reflected the decision to prioritize reliability, functionality, and a proven validation track record when selecting equipment. Looking ahead, installation space may be limited when this technology is rolled out at existing sites, so we believe it will be important to select more space-efficient equipment.

Overall, we achieved significant benefits in terms of cost reductions, lead time shortening, operational efficiency, and scalability, thus confirming that IP over DWDM is a viable technology for inter-site backbone connectivity at IIJ.

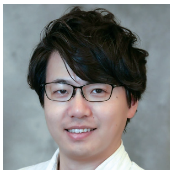
3.5 Future Outlook

IIJ is currently migrating to 400ZR-capable routers, and we expect the number of links to which IP over DWDM is applicable to continue growing. Traffic between major sites within Tokyo and Osaka in particular is in an uptrend, and we expect the use of this technology to be essential for efficiently handling these growing volumes of high-capacity traffic.

Although this commercial deployment targeted a short-distance link that did not require inline amplifiers, going forward we will also be looking at introducing inline amplifiers to expand applicability to medium- and long-distance links exceeding 30 km. This should unlock the benefits of IP over DWDM across a broader range of links.

As IIJ is not a carrier, it does not have an operations team specializing in optical technology, but we intend to draw on the knowledge gained through this deployment to expand our validation environment and promote internal knowledge sharing, thus building a foundation for IIJ to become an ISP with strong optical technology capabilities as well. The optical expertise developed through initiatives like this also has broader implications for the future of networking. As demand for AI grows, data centers are becoming geographically distributed, increasing the need for low-latency connections between urban centers and suburban/regional areas. If infrastructure capable of provisioning flexible optical paths, such as the All Photonics Network (APN) under discussion in recent years, becomes widespread, it could enable ultra-low-latency services that, for instance, connect customer equipment to service infrastructure entirely via optical signals, without conversion to electrical signals.

Going forward, IIJ will continue to proactively adopt the latest technologies while maintaining a backbone that provides both stability and quality, delivering even better network infrastructure to its customers as part of the infrastructure that underpins society.



Jun Sugahara

Manager, Planning & Development Section, Network Engineering Department, Infrastructure Engineering Division, Network Services Business Unit, IIJ

Mr. Sugahara joined IIJ in 2014. He has since been involved in the design, construction, and operation of IIJ's Internet backbone as well as IX (JPNAP). He currently works on initiatives in backbone design and operational efficiency improvements within the Planning & Development Section, with a particular interest in optical transmission technology.

Tomoya Takezaki

Network Engineering Section 1, Network Engineering Department, Infrastructure Engineering Division, Network Services Business Unit, IIJ

Mr. Takezaki joined IIJ as a new graduate in April 2020. Since joining, he has been involved in backbone network operations, serving as lead on physical and logical design and construction projects at domestic and international locations. Since around 2023, he has served as a peering coordinator, negotiating interconnection agreements with other operators, while also working on improvements to the backbone's physical design, evaluating new technologies, and testing equipment such as optical transceivers. His hobbies include travel as well as tracking down and visiting structures related to telecommunications infrastructure, such as communications facilities and manholes.