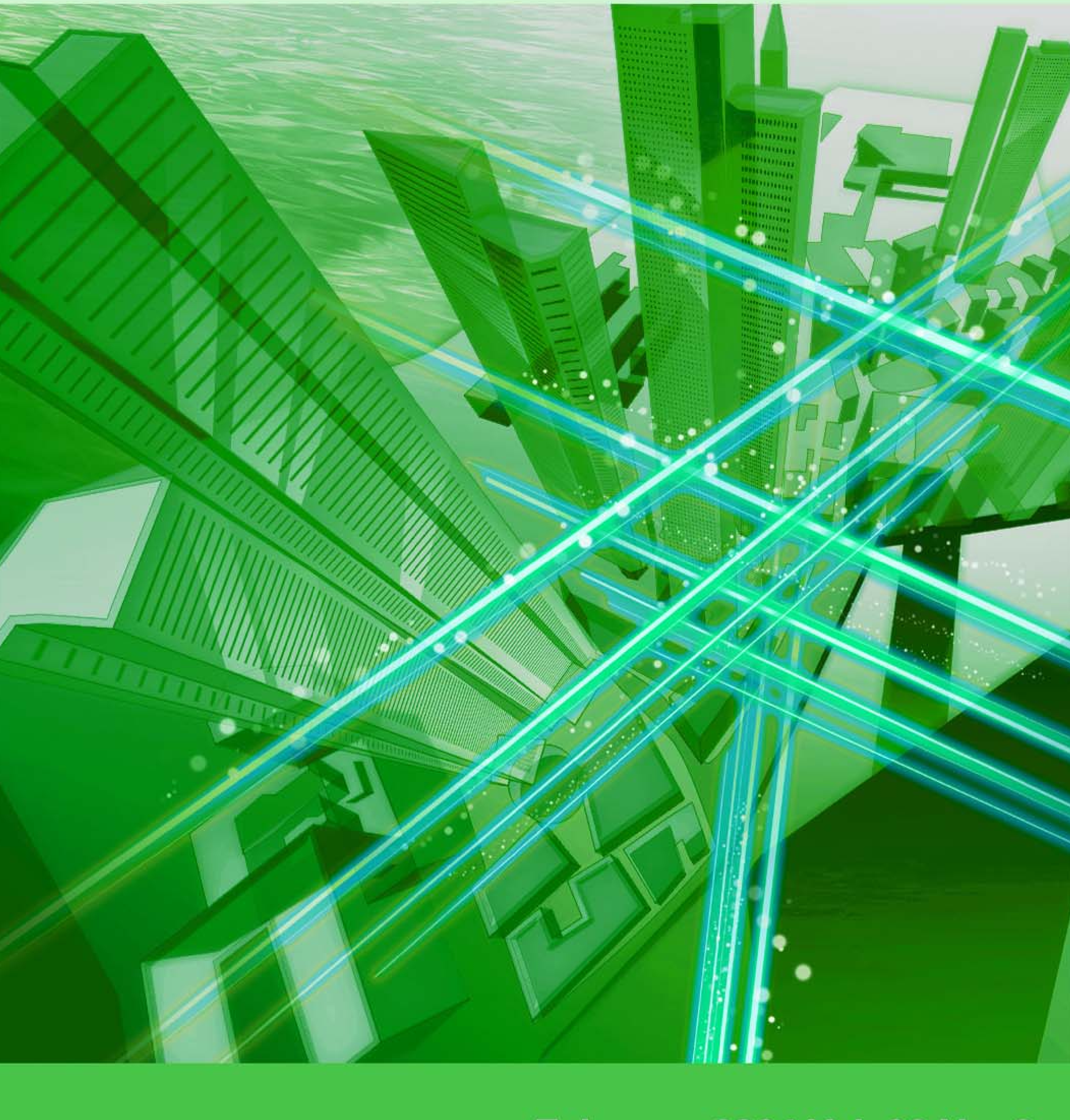
NTT Technical Review

2
2024



February 2024 Vol. 22 No. 2

NTT Technical Review

February 2024 Vol. 22 No. 2

Front-line Researchers

Rising Researchers

Feature Articles: Recent Developments in the IOWN Global Forum

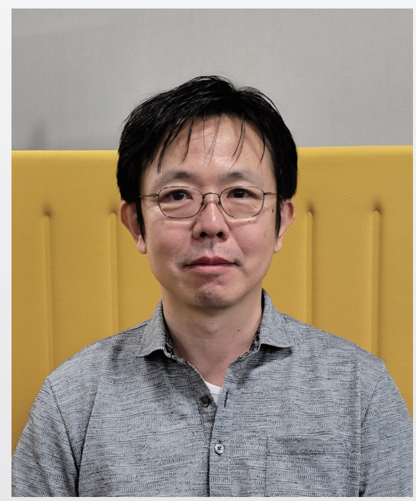
Global Standardization Activities

Practical Field Information about Telecommunication Technologies

External Awards/Papers Published in Technical Journals and Conference Proceedings

First Experimental Demonstration of a Phenomenon That Had Not Been Verified for More Than 20 Years Since Its Theory Was Proposed

Koji Muraki
Senior Distinguished Researcher,
NTT Basic Research Laboratories



Abstract
In the field of condensed-matter physics, it is often the case that a theory about a phenomenon is experimentally verified 20 or 30 years after it was proposed. Koji Muraki, a senior distinguished researcher at NTT Basic Research Laboratories, and his co-researchers were the first in the world to experimentally observe a phenomenon similar to Andreev reflection, which occurs at the interface between a superconductor and normal metal, in a material other than a superconductor. This observation was achieved through research on a many-body effect, which is a phenomenon by which electrons acquire the properties that an individual electron does not possess by interacting with each other. We asked him about this achievement, research on the fractional quantum Hall effect (one of many-body effects) that led to this achievement, and his mindset as a researcher engaged in basic research.

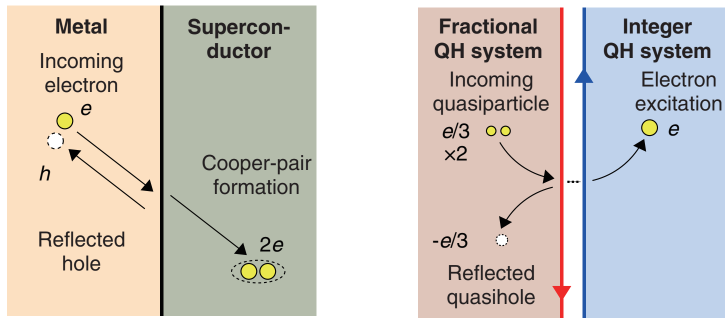
Keywords: many-body effect, quasiparticle, Andreev reflection

Exploring quantum devices with new functions that cannot be obtained with individual electrons by applying the fractional quantum Hall effect

—Would you tell us about the research you are currently conducting?

I'm researching many-body effects and correlation effects caused by interactions between electrons as well as quantum-mechanical properties of electrons such as wave nature, superposition state, and spin. By engineering and controlling these properties of electrons using the heterostructure and nanostructure of semiconductors and atomic-layer materials, I'm exploring quantum devices with new functions that cannot be obtained with individual electrons.

In my previous interview (March 2021 issue), I talked about our group experimentally demonstrating that (i) a topological insulator, which is a material that conducts electricity on its surface but is an insulator inside, can be formed using a semiconductor heterostructure and (ii) by applying a voltage to the gate electrode of the topological insulator, it can be electrically switched between a normal insulator and



(a) Andreev reflection at a conventional superconductor junction

(b) Andreev reflection of quasiparticles at a fractional-integer QH junction

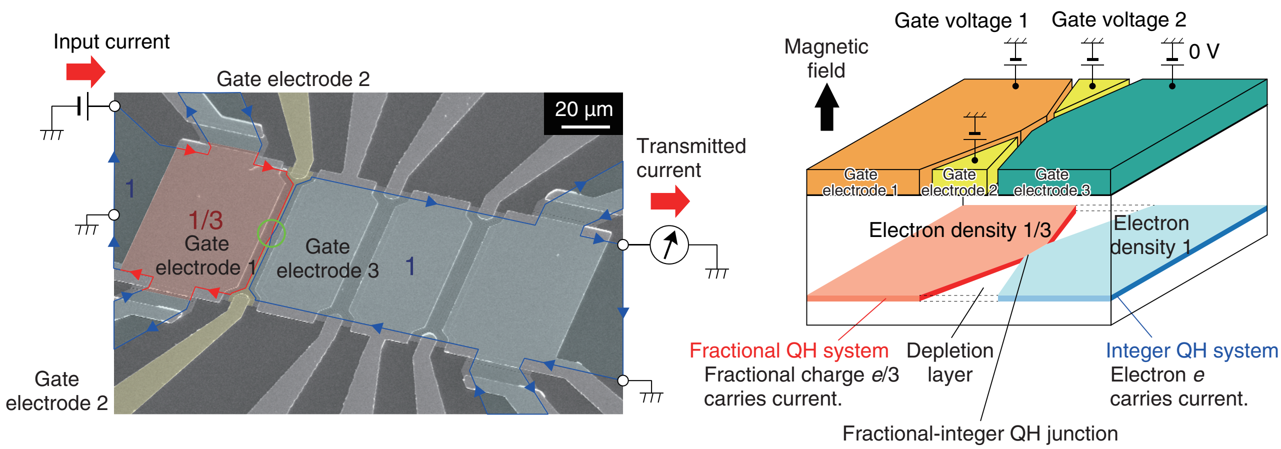
Fig. 1. Conceptual diagram of Andreev reflection.

topological insulator. In this interview, I'd like to discuss many-body effects.

A many-body effect is a phenomenon by which electrons interact with each other and acquire the properties that an individual electron does not possess. Typical examples are superconductivity and ferromagnetism. Among many-body effects, our research group focuses on a phenomenon called the fractional quantum Hall (QH) effect, in which electrons with electron charge e (< 0) try to avoid each other in a strong magnetic field; as a result, a certain kind of order is created—a uniform state like liquid without the density fluctuation. Interestingly, deviation from this uniform state of electrons occurs in units of one-third the charge on the electrons ($e/3$). The convexity of the electron density is called a "quasiparticle" and the concavity is called a "quasihole," and they behave like particles with charges $\pm e/3$. The fractional QH effect has a long history; it was discovered in 1982 and its underlying theory was published the following year, which were awarded the Nobel Prize in Physics in 1998. In the early 2000s, it was theoretically proposed that quasiparticles and quasiholes in certain states could be used for quantum computing, since then, research to verify that proposal has been ongoing.

It was experimentally clarified that quasiparticles in the fractional QH state have charge $e/3$ from measuring and analyzing the weak noise in the current. As a fundamental technology for quantum computing using quasiparticles, measuring the quasiparticle interference with a minute device fabricated in the form of a ring is necessary. In 2020, a research group in the USA reported—as a world's first—an experiment in which they successfully conducted such measurement, and that report has gained much attention. Such experiments investigate the properties of quasiparticles with charge $e/3$. The electrodes and lead wires connected to the measurement instruments are normal metals, and the electrons flowing through them are electrons with charge e. Therefore, what happens when a quasiparticle with charge $e/3$ flows into an electrode composed of a normal metal? Two theoretical papers published in 1998 considered such a question. According to the theory, when two quasiparticles with charge $e/3$ are injected into a metal, an electron with charge e flows into the metal, and a quasihole with charge $-e/3$ (> 0) is reflected back into the injection side to conserve charge. This phenomenon is similar to the Andreev reflection (**Fig. 1**) that occurs at the interface between a superconductor and normal metal, and we succeeded in observing it with a material other than superconductors in an experiment for the first time [1]. In addition to the significance of elucidating charge conservation at the level of quasiparticles, which are smaller than electrons, and the elementary processes of scattering and reflection of quasiparticles, this achievement also provides a new perspective to consider in regard to quantum circuits using quasiparticles. That is, not only quasiparticles but also quasiholes carrying charge are flowing in a quantum circuit using quasiparticles, and the effect of the quasiholes must also be considered when discussing the presence or absence of quasiparticle interference.

I was very happy to be able to demonstrate experimentally for the first time a phenomenon that had not been verified for over 20 years since its theoretical prediction. The researcher who wrote one of the above-mentioned theoretical papers in 1998 read our



(a) False-color scanning electron micrograph of the sample

(b) Fractional-integer QH junction

Fig. 2. Conceptual diagram of the experiment.

paper and sent us an email saying, "Thank you for verifying my theory." Subsequent research has also helped us to understand the noise and heat generated during these scattering and reflection processes of quasiparticles [2]. It may sound easy when I only write about the result, but it was a long journey from the start of the research until we got the satisfying result.

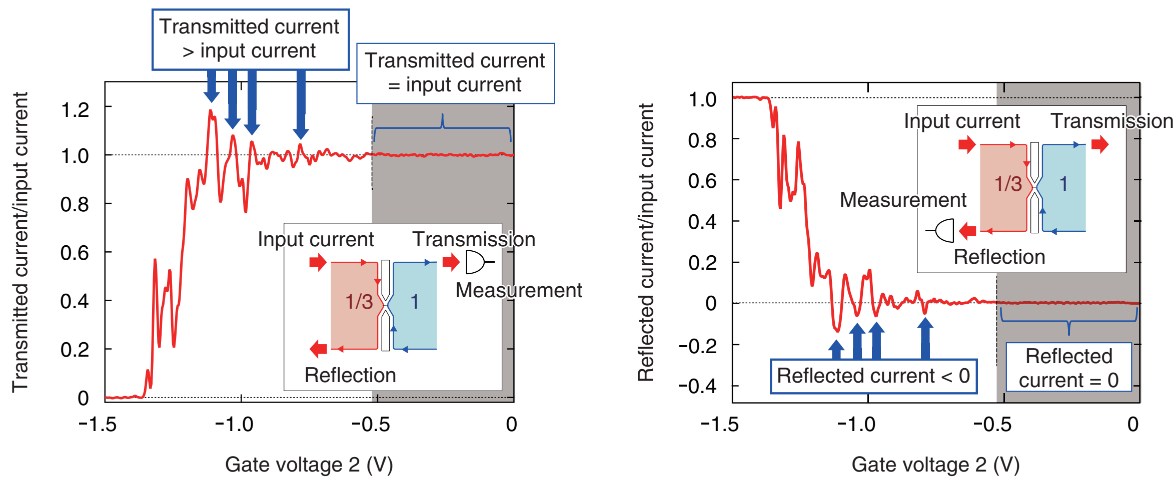
Investigating the properties of quasiparticles, which have a smaller charge than electrons, requires extremely advanced experimental techniques, and only a handful of research groups in the world can conduct such experiments. It was also a significant change to have people from these research groups pay attention to our work because of this achievement. Although my role in these studies was supervising, and my main contribution was to keep asking questions to the member leading the research, being involved in research that gives a young researcher an opportunity to spread their wings and be recognized around the world gave me a different sense of accomplishment than I ever had before.

—Could you tell us about the experiment that produced impressive results in more detail?

To investigate the properties of quasiparticles, more-precise measurement techniques were needed. Thus, I asked a researcher to join our group who had similar interests as us and the necessary skills and knowledge, and he came up with an idea of this experiment.

The overview of the experiment is illustrated in **Fig. 2**. First, a magnetic field is applied perpendicular to a two-dimensional electron system in a GaAs (gallium arsenide) semiconductor heterostructure at low temperature. The magnetic-field strength is then adjusted so that the entire electron system is the integer QH state with a Landau-level filling factor* of 1. A false-color scanning electron micrograph of the sample is shown in Fig. 2(a). The red region is a fractional QH state with a Landau-level filling factor of 1/3, which is formed by applying gate voltage 1 to gate electrode 1. The blue region is an integer QH state with a Landau-level filling factor of 1, where the electrons with charge *e* play the role of a normal metal carrying current (gate electrode 3 is fixed at 0 V). The yellow regions are gate electrode 2 applied to narrow the interface between the fractional QH and integer QH states. The dark-gray regions are areas where the semiconductor is etched away, and the light-gray regions are gate electrodes not used in this experiment. A fine fractional-integer QH junction is formed in the center between gate electrodes 1 and 3 (marked with the green circle). In a QH region, current flows in one direction along the channels at the edge of the sample (red and blue arrows), so transmitted (blue) and reflected (red) currents can be measured.

* Landau-level filling factor: The ratio of electron density to a magnetic field (magnetic-flux density). When this value approaches an integer, an integer QH effect occurs, and when it approaches a certain fractional value (such as 1/3), a fractional QH effect occurs.



(a) Transmitted current

(b) Reflected current

Fig. 3. Measurement results when the fractional-integer QH junction is narrowed by applying gate voltage 2.

Figure 2(b) shows a conceptual diagram of the fractional-integer QH junction, namely, the region marked with the green circle in Fig. 2(a). When gate voltage 2 is applied to gate electrode 2 and the two-dimensional electron system directly below the gate electrode is depleted, a junction (width less than 1 μm) is formed at which the fractional and integer regions contact each other. The width of the junction narrows as the negative voltage of gate voltage 2 increases and the depletion region widens.

In the experiment, a current was applied to the junction from the fractional QH side, and the transmitted current flowing out of the junction to the integer QH side was measured. We observed that the transmitted current exceeds the input current at certain values of gate voltage 2, which corresponds to changing the junction width (**Fig. 3(a)**). When we measured the reflected current returning from the junction to the fractional QH side, the current of the opposite sign to that of the input current was observed (**Fig. 3(b)**). The increase in transmitted current and a negative reflected current indicate that Andreev reflection of quasiparticles occurred. This result indicates that Andreev reflection is not a phenomenon unique to superconductors but a universal phenomenon, demonstrating an important achievement in condensed-matter physics.

—What kind of research will you be focusing on in the future?

Topological insulators and the fractional QH effect both exhibit interesting physical properties due to the nature of the sample (material) interior; however, it is at the edges and surfaces of the sample (i.e., at the interface with the vacuum or with ordinary material) that they appear as phenomena observed in experiments. Therefore, to observe theoretically predicted, interesting phenomena experimentally, it is important to create not only pure crystals but also high-quality interfaces. I haven't much time to talk about it in this interview, according to theory, the interface between a topological insulator and superconductor is also a platform where new physical properties appear. Although researchers worldwide are working to verify the predictions of the theory, a unified understanding has not yet been reached because either the results are not as the theory predicts or they are as the theory predicts but do not exclude the possibility of other interpretations. We are currently focusing on creating various interfaces, including superconducting junctions, by using methods that nobody has attempted before. The young researchers in our group are actually conducting the research, and I support them. I hope that our research will result in an ideal interface and bridge the gap between theory and experiment. This research is unprecedented, so it is a continuous process of trial and error, but it is both rewarding and fun.

Basic research is like creating a map of science. By taking on the right challenges, I want to make a positive impact on the research community and society as a whole

—What do you keep in mind as a researcher?

As I mentioned in the previous interview, I think that the basic research I'm involved in is like creating a map of science. In the field of physics, theory alone does not determine whether a proposition is correct. At this phase, a rough outline of the map is created. To update the map to be more accurate, experiments must be conducted to clarify and verify the theory. While certain Nobel Prize-winning studies significantly rewrote the map, it is also important to accumulate small updates to the map. I want to contribute to the creation of maps through my research on controlling the interface between materials and explore new properties of electrons. The social significance of that research, i.e., the new properties of electrons may be used for fault-tolerant quantum computing, motivates me.

I used to think of research as a means of self-realization, so whether it was my own idea, whether I carried it out, and how I was evaluated by others as a researcher were all important indicators for me. However, as I get older, I find myself becoming more appreciative of the importance of all the people and things that surround me. This appreciation includes not only "standing on the shoulders of giants" but also extending to all ordinary people and things. Regardless of whether it is my idea, and regardless of how I am evaluated, I want to use my energy to do something that will have a positive impact on the research community and society as a whole. In other words, I want to do things that I'm not sure that I can do and things that seem difficult but are worth doing, rather than things that I think I can do faster than others. Even if I take on a challenge that no one else is doing and the results are not what I expect, if I am doing it for the right reasons, the knowledge I gain will be useful to me and the world outside. I still want to do research that will make people exclaim, "That was a breakthrough!" but to achieve such breakthroughs, I want to continue to take on the right challenges.

—What is your message to younger researchers?

Undoubtedly, young people are more likely to accomplish great things in the future, so I want them to believe in their own potential and strive forward. In the past few years, there have been several major breakthroughs in my field of research, and all have been achieved by young researchers. I have thought about some of the things I was close to doing, but when I actually observe their breakthrough research, I realize that it was not so much that I was close but rather that I had not pursued my ideas far enough. Even if you are progressing toward an achievement, if you don't have the last piece in place, you won't know if you are close to making an achievement, and if you stop there, the possibilities that may have existed will be closed off. It is important to keep your eyes open and your ears to the ground to be sensitive to information from the outside; however, if you really have something you are aiming for, rather than convincing yourself that you have already worked hard for years or that you have done so much, you should keep asking yourself whether you have really done everything you need to do and whether there is a better way to do it. This message is not only for young people but also for myself as a reminder.

References

[1] M. Hashisaka, T. Jonckheere, T. Akiho, S. Sasaki, J. Rech, T. Martin, and K. Muraki, "Andreev Reflection of Fractional Quantum Hall Quasiparticles," Nat. Commun., Vol. 12, 2794, 2021. https://doi.org/10.1038/s41467-021-23160-6

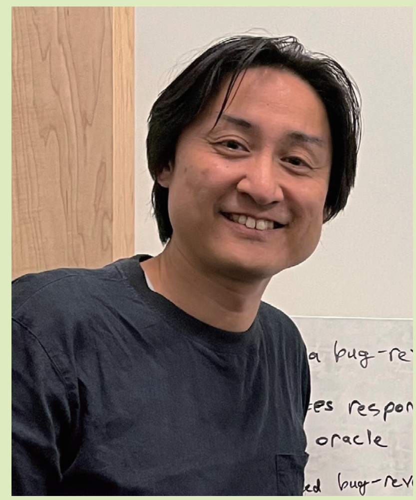
[2] M. Hashisaka, T. Ito, T. Akiho, S. Sasaki, N. Kumada, N. Shibata, and K. Muraki, "Coherent-Incoherent Crossover of Charge and Neutral Mode Transport as Evidence for the Disorder-Dominated Fractional Edge Phase," Phys. Rev. X, Vol. 13, 031024, 2023. https://doi.org/10.1103/PhysRevX.13.031024

■ Interviewee profile

Koji Muraki received a B.E., M.E., and Ph.D. in applied physics from the University of Tokyo in 1989, 1991, and 1994. He joined NTT Basic Research Laboratories in 1994. From 2001 to 2002, he was a visiting researcher at the Max Planck Institute for Solid State Research, Stuttgart, Germany. His research interests are focused on many-body effects in low-dimensional semiconductor structures. He is a member of the Physical Society of Japan and the Japan Society of Applied Physics.

Software Digital Twins for Reducing Software Waste and Achieving a Sustainable IT Society

Shinobu Saito
Distinguished Researcher,
NTT Computer and Data Science Laboratories



Abstract

The global information and technology services market for 2022 was reported to be 1.3 trillion U.S. dollars (about 150 trillion yen), and it's been pointed out that the massive amount of software produced by this market demand, while bringing great value to our daily lives, has also led to a variety of problems. There are expectations that an approach that not only seeks to increase the amount of software (production) but to also reduce that amount (reduction) could help solve those problems. In this interview, we asked NTT Distinguished Researcher Shinobu Saito about his research on "Software Digital Twins."

Keywords: software waste, Software Digital Twins, Software 3Rs

Achieving an ideal environment for users and developers through Software 3Rs (Reduce, Reuse, and Recycle)

—Dr. Saito, what kind of technology is "Software Digital Twins"?

"Software Digital Twins" that I am now researching is a technology that reproduces the internal conditions of software in a digital world on the computer to discover and predict locations (programs) where software should be revised or deleted. It's been pointed out that the increase in software that is not used at all called "software waste" is a major problem that has come to encircle software in recent years. As it turns out, software, which was originally produced with the aim of making the world a better place, is also giving rise to a variety of problems. These problems are called "big code" or "software bloat." The former means that software that has become too big is difficult to maintain by a limited number of people (developers). In this regard, it is predicted that the amount of software will eventually exceed the range in which programmers are capable of managing it. In overseas surveys, more than half of all programmers said "The amount of program code that one must deal with has increased by more than 100 times compared with that 10 years earlier." The latter, meanwhile, means that software that has become overly abundant (bloated) is consuming system resources such as central processing units and memory on a large scale. This results in excessive power consumption and carbon dioxide emissions, which has a negative impact on the global environment.

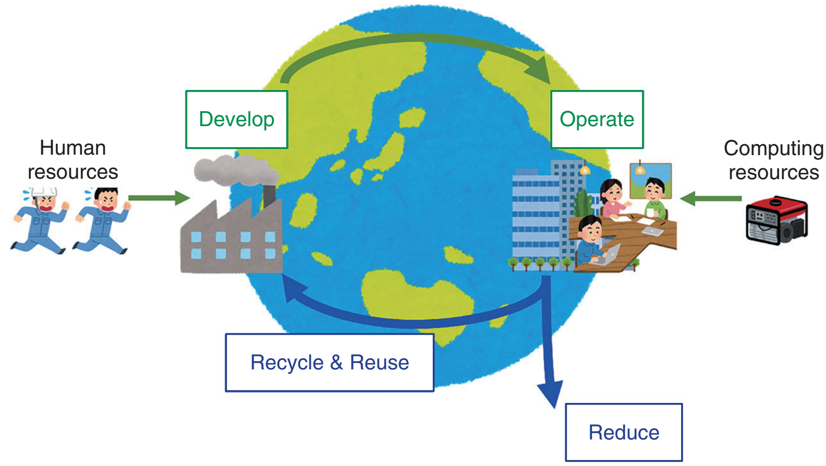


Fig. 1. Vision of Software 3Rs.

After joining NTT DATA Corporation, I myself thought for many years "What can be done to make it easier for developers to create good software for users?" Several years ago, however, I came across two articles in a business magazine titled "Software not used by companies and ignored by users is waste" and "One-third of an average company's information technology (IT) budget is consumed as software waste," both of which shocked me. Indeed, it has also been reported that more than half of the functions in a developed system are rarely used. I therefore began to feel "There is the possibility that my research in the end will actually have a negative impact on people and the Earth," and I began to think of an approach that could break through this situation.

Then, one day while thinking about this, when I was helping my child in elementary school with summer-vacation homework, I noticed that the concept of the 3Rs in relation to environmental problems was introduced in the textbook. The 3Rs refer to Reduce, Reuse, and Recycle. Seeing this, I thought "Shouldn't this concept also be applicable to IT society that is constantly producing software?" I considered that, from here on, "developing" software that can enrich everyone's life is necessary, but it is also important for people and the environment to "eliminate" software that has been overly developed and become excessive. The concept born out of these thoughts is "Software 3Rs" (**Fig. 1**), which promotes not only the development of software but also the reduction of unnecessary software and the reusing of existing software. In this way, I believe that we can achieve a sustainable IT society for both users and developers of software. Last year, I presented this concept at an international conference on software engineering and received comments like "very interesting" from participating researchers.

—In particular, what kind of approach are you researching?

I am currently researching Software Digital Twins as an approach to implementing the Software 3Rs concept. This idea came to me from research on digital twins that reproduce data collected from real-world objects such as cities and human bodies in a digital world on the computer. For example, since it is difficult to directly observe an actually functioning brain or heart, research is progressing on using the digital twin of a human body to express such organs in digital space with the aim of investigating the causes of disease, predicting the occurrences of disease, and even preventing disease. Software, as well, consists of an aggregation of programs, and directly observing the internal conditions of working software is said to be a difficult objective. As a result, correctly grasping "where the frequently used programs are" and conversely "where unused programs are" can be extremely difficult. Consequently, my aim with Software Digital Twins is to discover and predict locations (programs) that are unnecessary and non-urgent or in need of improvement by reproducing with high accuracy the internal conditions of "black-box" software and making that content "fully visible" (**Fig. 2**).

This research is closely related to Digital Twin

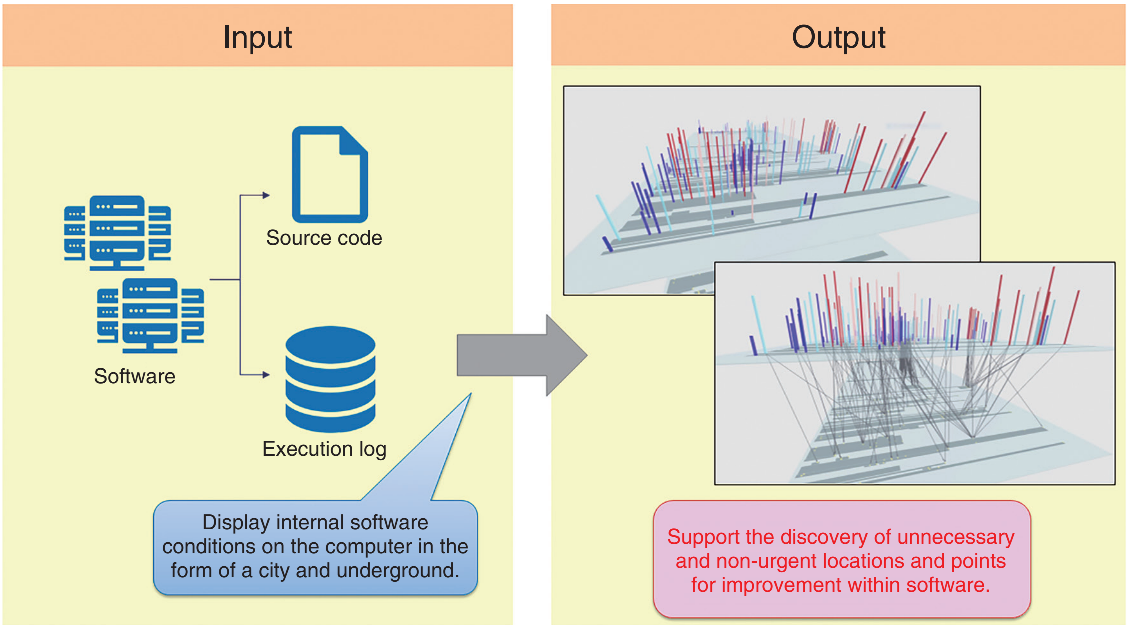


Fig. 2. Overview of Software Digital Twins.

Computing (DTC), a major pillar of the Innovative Optical and Wireless Network (IOWN) concept proposed by NTT. At present, studies of DTC are proceeding in areas targeting the real world such as urban management and medical treatment. In contrast to this approach that attempts to reproduce the real world in the cyber world, my research attempts to reproduce a cyber world in another cyber world, that is, to reproduce the interior of software running on a computer on another computer. Here, reproducing a cyber entity as another cyber entity may appear at first glance to be somewhat strange, but I believe that this approach can enable humans such as software architects to accurately understand the internal execution conditions of software for the first time and facilitate predictions and simulations.

Main data used for accurately reproducing the execution conditions of software are of two types: "source code" and "execution log" that records program operation. In the latter, the collection of the execution log presents a major hurdle to making advances in research and achieving practical implementation. The ideal method would be to record the operation of "all locations (programs)" inside software during "all periods" in which the software is executed. This, however, would require that software engineers create a new program to record that operation at all locations. In addition, attempting to collect the execution log in all periods of execution would consume an excessive amount of computing resources, which in itself would have a negative impact on people and the environment, thereby defeating our original purpose.

Given the above, the technique that I am studying in my current research will estimate and predict execution conditions throughout the software from execution log data obtained in limited periods at limited locations within the software. My aim here is to establish a technique that applies fields such as statistical analysis and machine learning through joint research with universities.

Continue with "solitary" and "interaction" activities to ensure that research does not end with basic research

—Please tell us about your research vision going forward.

One research achievement in Software Digital Twins to come out so far is ProcessCity, which is technology for visualizing the usage frequency of each function in a business system and the operation of the entire system in a three-dimensional manner. We have already provided this technology to several NTT operating companies, and at NTT R&D Forum

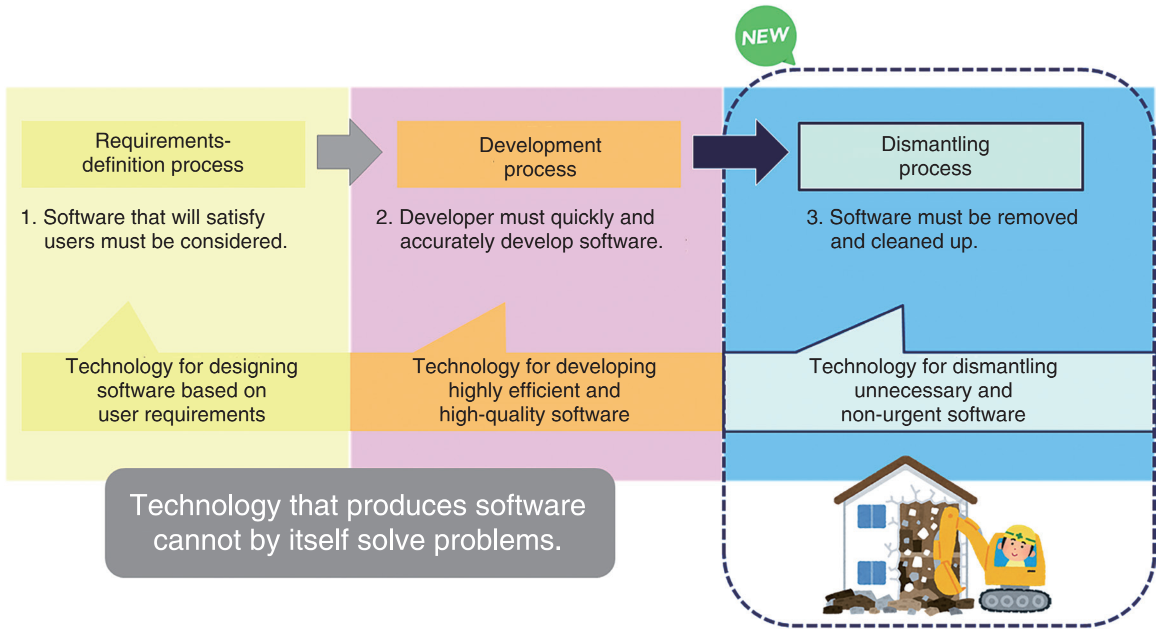


Fig. 3. Dismantling process in software development.

2022, we held a joint exhibit and a press release was issued with one of those companies. I also published a paper with members of another operating company to which this technology had been provided, and we were honored to receive a best paper award at the largest software-engineering international conference in the Asia-Pacific region held in 2022.

At present, it is said that the cost of maintaining aging and complex old systems called "legacy systems" has risen to 60–70% of a company's IT budget. In addition, there have been cases in which surveys conducted for discovering system functions that could be scaled down or discontinued took months to complete. Consequently, the first thing that could be achieved by using the results of this research is a shortening of the survey period and improvement in the accuracy of the survey itself thereby reducing the labor required in the survey step by 20–30% and eventually to as much as 10% of the current amount.

It was reported that the global IT services market for 2022 was 1.3 trillion U.S. dollars (about 150 trillion yen), and according to the business magazine that I mentioned earlier, it's been calculated that about one-third of those IT services, or 50 trillion yen, is being wasted. Of course, it would be difficult for NTT by itself to reduce such software waste throughout the world, but I would nevertheless like to create technology that could help reduce such a huge amount of waste on the scale of the national budget and to connect this effort to providing great value to NTT's corporate customers.

Additionally, on looking at the IT services industry going forward, I anticipate that an industry for the demolition of software in a safe and secure manner will probably come to be formed. In this sense, if producing software is a "construction industry," then removing (deleting) software is a "demolition industry." By making further progress in my research, my ultimate goal is to form a new industry and create new business opportunities and to make the software dismantling business one of NTT's core business areas (**Fig. 3**).

—What is your impression of NTT laboratories?

As the name implies, NTT Computer and Data Science Laboratories that I belong to conducts research related to computers and data, and in this capacity, it is involved in a wide variety of research across all sorts of businesses and industries. NTT Group companies include both companies that develop software and companies that use software, which is why I believe that NTT laboratories are especially strong in both user and development fields. The scope of stakeholders is broad, and I am given opportunities for exchanging opinions with group companies from the standpoint of using software, so I think I am blessed to be in such an environment that helps me make real progress in my research.

Moreover, on meeting with people from companies outside the NTT Group, there are many occasions in which I can feel the name value of NTT. For example, when attending an international conference for the first time 10 years ago, I knew nobody there, but on walking by myself to the venue after receiving my name tag at the reception desk, someone walking next to me noticed my name tag and said in a very friendly manner: "Ah, NTT! Do you know Mr. XX?" This person is a very famous professor in the software engineering community, and thanks to this chance happening, I still talk with him and discuss research whenever we meet up at an international conference. Of course, the name value of NTT does not guarantee that research will always go well, but it certainly gives me a reassuring feeling whenever I go to an unfamiliar environment or place.

—Dr. Saito, please leave us with a message for researchers, students, and business partners.

As I conduct my research at NTT, I have made the words "there is no research that ends with basic research" into a guiding principle. I first learned about this concept on joining NTT DATA from researchers that had come to my department from NTT laboratories. What this meant was that "Even basic research that seems unlikely to have an immediate application must be meaningful in the sense of being useful in some way. If it has no meaning (if no meaning can be discussed), it is not basic research, and in fact, it is not research in the first place." I keep these words in mind all the time, and since research is funded, I make an effort to present "the meaning of research" in terms of how those research results will someday be useful.

At around the same time, I remember well the words of a senior colleague, who said, "Research, in the end, is a solitary endeavor." Research requires originality, so in some respects, I think "solitary" is correct here. Looking back at my own research when I was at the stage of thinking about a certain basic concept, there were many times when I could not verbalize it well and just kept worrying about it by myself. However, at the stage of actually implementing that concept, I found that it was necessary to gain the cooperation of many people. NTT provides an environment and the opportunities to interact with other people and other companies, so there are many uplifting (exciting) times that occur when acquiring new knowledge from outside one's field and seeing a new world that one could not reach by oneself. Going forward, I plan to pursue my research with great feeling and conviction and to make up for what I am lacking by interacting with many of my fellow researchers with the hope that we can conduct research together.

■ Interviewee profile

Shinobu Saito received his M.S. degree from Keio University in 2001 and joined NTT DATA Corporation in the same year. He received his Ph.D. degree (engineering) from Keio University in 2007. He is engaged in the research of new software engineering for the digital era. He is a recipient of the 2022 IPSJ-CS Outstanding Achievement and Contribution Award and the APSEC2022 Best SEIP (Software Engineering in Practice) Paper Award among other awards. He was a visiting researcher at the Institute for Software Research (ISR), University of California at Irvine, from 2016 to 2018 and was a distinguished researcher at NTT Software Innovation Center from 2018 to 2021. He has been affiliated with NTT Computer and Data Science Laboratories since 2021.

Latest Activities in the IOWN Global Forum

Yoshitake Tajima and Yosuke Aragane

Abstract

With a view to creating a smart world, the concept of the Innovative Optical and Wireless Network (IOWN) aims to create a next-generation communication and computing infrastructure that features ultra-high capacity, ultra-low latency, and ultra-low power consumption. With input from global member companies, the IOWN Global Forum has been developing technical architectures and use cases and conducting proof of concept (PoC) activities and implementation verification. This article introduces the latest activities of the IOWN Global Forum, as well as the PoC Reference documents and reference-implementation-model documents that reflect the results of these efforts.

Keywords: IOWN, Global Forum, optical network

1. IOWN concept and the IOWN Global Forum

In May 2019, with a view to creating a sustainable and prosperous smart world, NTT proposed the concept of the Innovative Optical and Wireless Network (IOWN), which will break through current limitations and build a communication and computing infrastructure that features ultra-high capacity, ultra-low latency, and ultra-low power consumption. NTT is conducting research and development with the goals of increasing power efficiency by a factor of 100, transmission capacity by a factor of 125, and end-to-end low latency performance by a factor of 200 by the year 2030. Achieving these goals will call for a major transformation in information processing, communication, and network infrastructure, which in turn will require the development and combination of many innovative technologies. Such endeavors cannot be achieved by the NTT Group alone. Thus, in January 2020, Intel Corporation, Sony Corporation, and NTT Corporation established the IOWN Global Forum, an international non-profit organization dedicated to the future of communications. During the four years since then, the forum has been developing the innovative technical frameworks, technical specifications, and reference architectures needed to achieve the above goals and pressed forward to actualize a new communication and computing infrastructure, IOWN. **Figure 1** shows the overall technical structure of IOWN [1]. What is immediately apparent is that its network, the Open All-Photonic Network (Open APN), enables communication with deterministic transmission rates and latency. Placed on top of that is an architecture called a data-centric infrastructure (DCI), which achieves efficient data transfer and processing using accelerators. Via the APN, two remotely located DCIs can transfer data at high speed and with low latency and combine multiple distributed datacenters to form one large virtual datacenter. The IOWN Global Forum is developing the IOWN Data Hub (IDH) to enable scalable and highly available database storage that is based on the distributed database architecture within a cluster by using the above capabilities of the DCI. The forum will implement various use cases on top of the IDH and enable Digital Twin Computing by linking digital twins that digitally represent various objects in the real world. It will feed the results of Digital Twin Computing back to the real world to create a smart world.

The activities of the IOWN Global Forum are characterized by the fact that it studies not only technology but also use cases so that the desired smart world can be envisioned in concrete terms from the

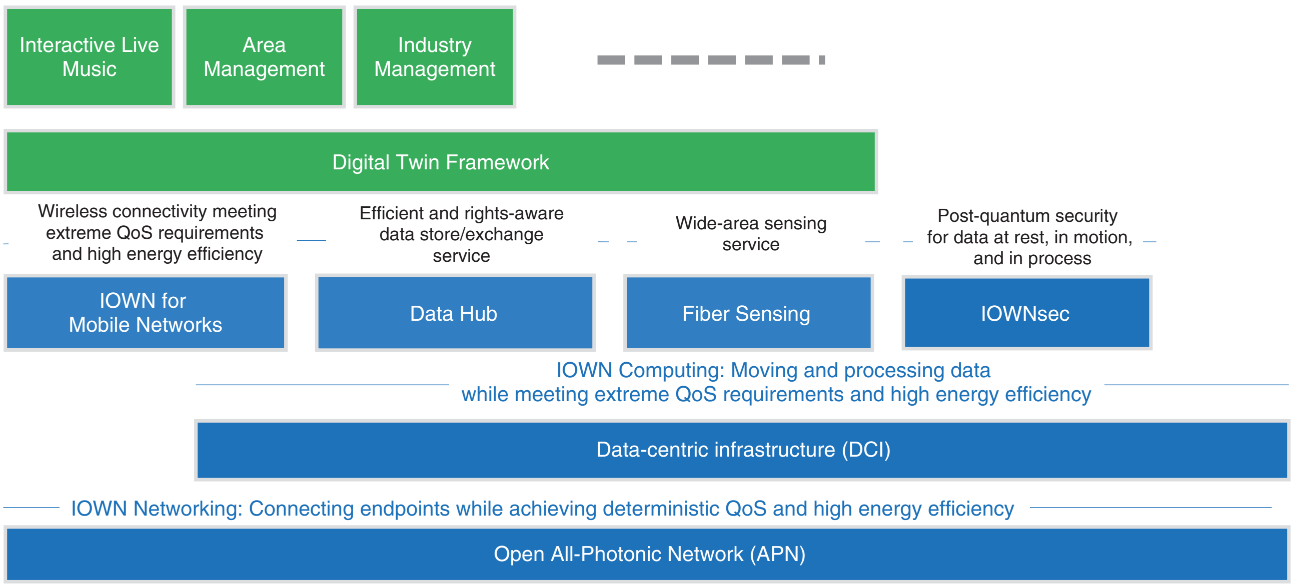


Fig. 1. Structure of IOWN technologies.

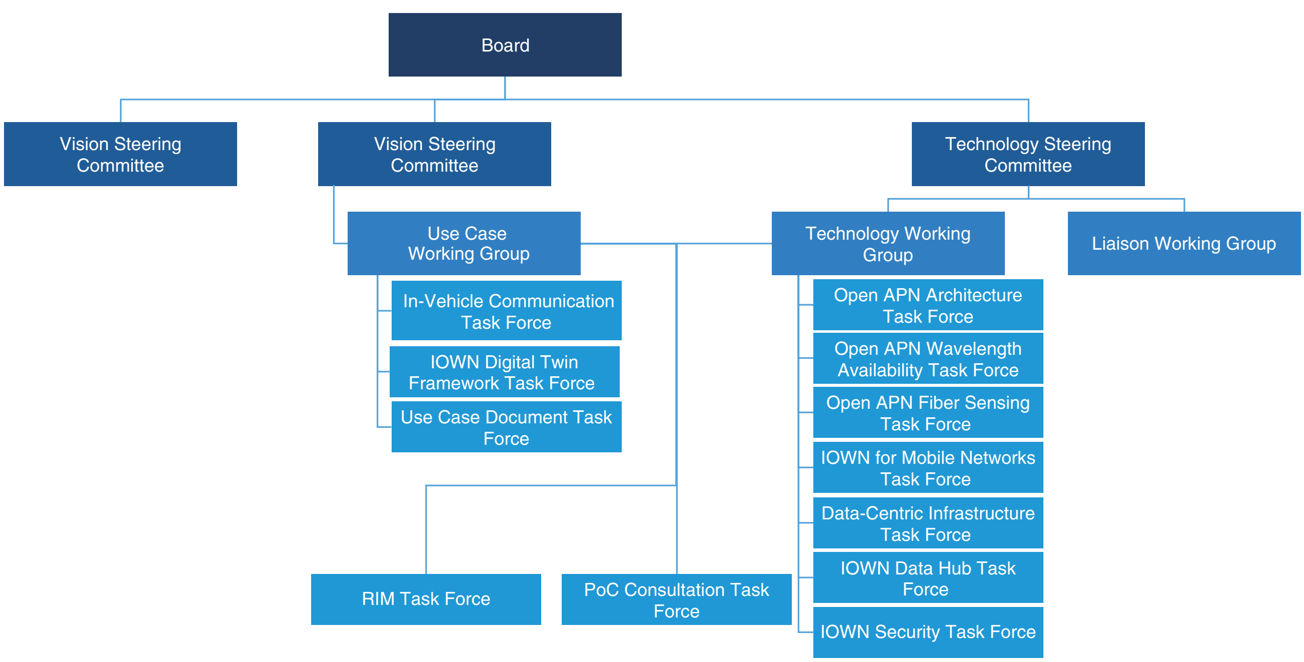


Fig. 2. Organization chart of IOWN Global Forum.

beginning. Two working groups (WGs) have been established for this purpose: the Use Case WG and Technology WG (**Fig. 2**). The Use Case WG develops the details of applications in line with the vision of a smart world, estimates IOWN's potential impact on business, and discusses technical requirements, while the Technology WG discusses technical solutions, such as reference architectures, protocols, interfaces, and other specifications. The two WGs work in close collaboration with each other (**Fig. 3**).

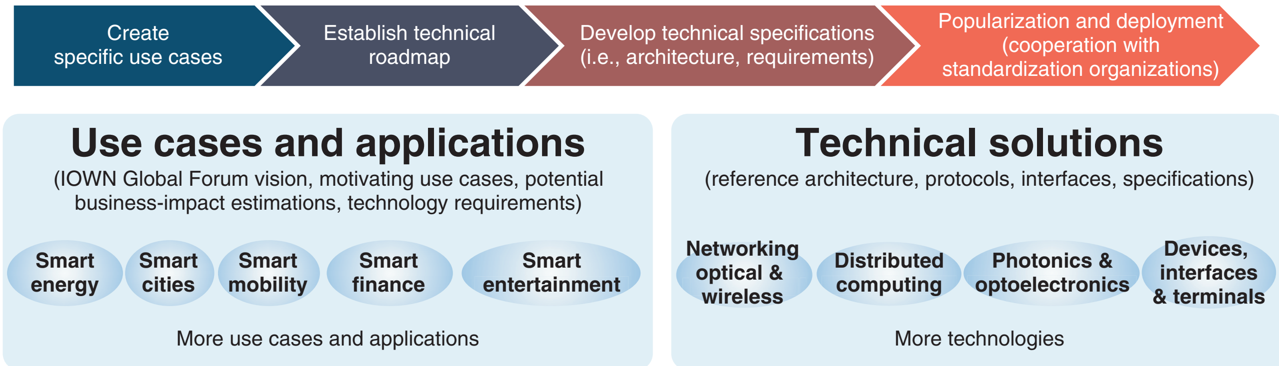


Fig. 3. Forum activities.

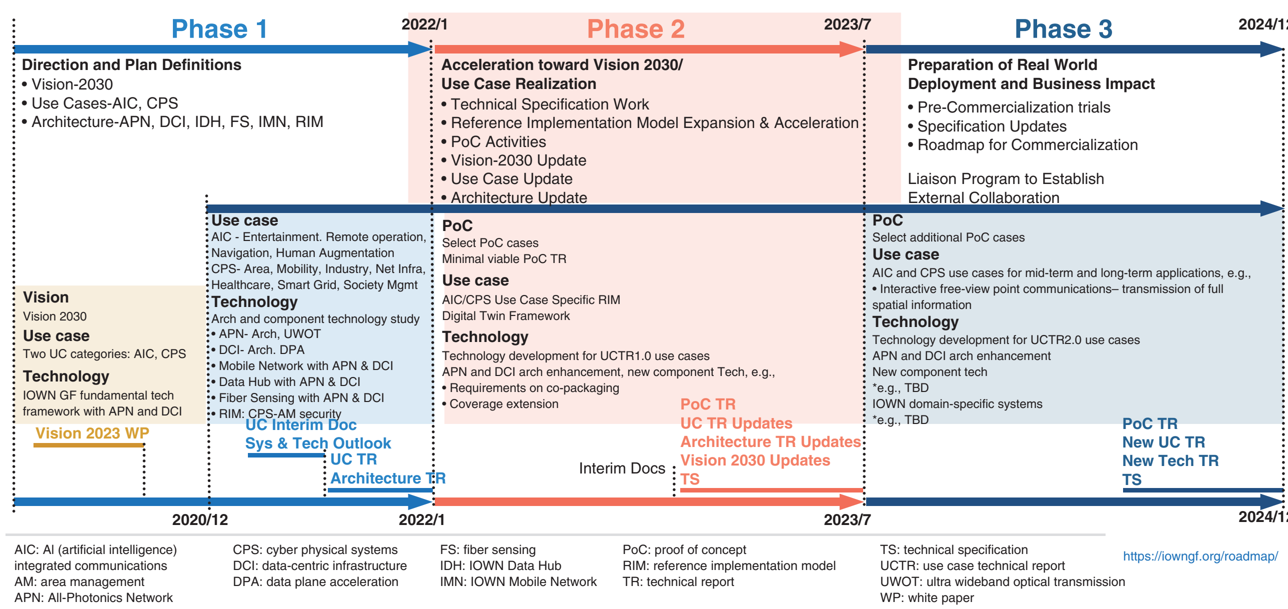


Fig. 4. IOWN Global Forum Vision 2030 Roadmap.

The IOWN Global Forum has developed and followed a roadmap consisting of phases, each of which spanning about eighteen months. In Phase 1 (up to January 2022), entitled "Direction and Plan Definition," the forum formulated a vision, developed major use cases, established a basic technical framework, and identified the issues that need to be considered. In Phase 2 (up to July 2023), entitled "Acceleration toward Vision 2030/Use Case Realization," the forum updated the vision, use cases, and architecture and developed technical specifications and reference implementation models (RIMs), and conducted proof of concept (PoC) activities. Currently, in Phase 3, entitled "Preparation of Real-World Deployment and Business Impact," the IOWN Global Forum is accelerating its efforts toward the actualization of IOWN (**Fig. 4**).

As a way of presenting the main initiatives of the IOWN Global Forum, this article introduces the results of Phase 2, the forum's organization, the status of its activities, and its planned activities for the future.

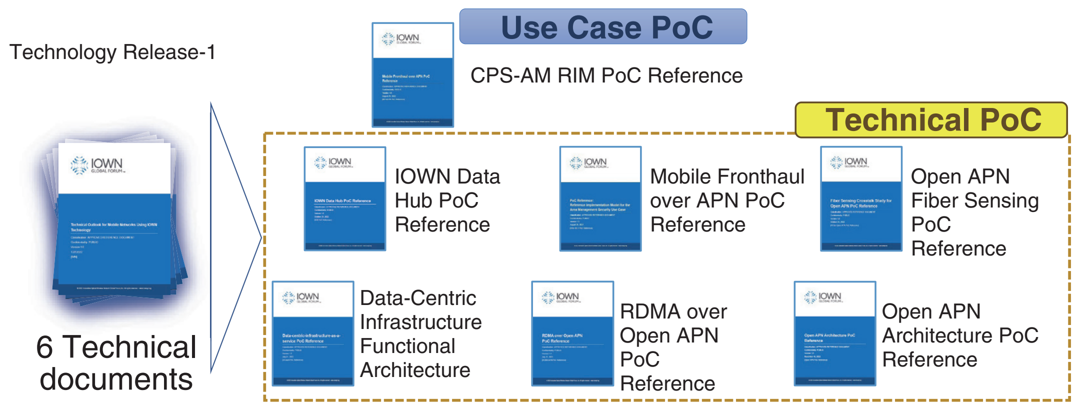


Fig. 5. Seven released PoC Reference documents.

2. Phase 2 activities and key results

2.1 PoC References

In Phase 1, the IOWN Global Forum studied the architectures of the Open APN, DCI, IDH, and other technical elements of IOWN, and released six technical documents in January 2022.

The main activities in Phase 2 were to study implementation details and conduct PoC activities on the basis of the architectures developed in Phase 1. To advance PoC activities and promote market deployment and collaboration among the parties concerned beyond simple technical demonstration, the IOWN Global Forum compiled PoC Reference documents, which summarize reference use cases, the features that key technologies must deliver, and evaluation metrics for each technical theme. By November 2022, the forum had released the following seven documents (**Fig. 5**):

(1) Open APN Architecture PoC Reference
(2) RDMA over Open APN PoC Reference
(3) Data-centric-infrastructure-as-a-service PoC Reference
(4) Mobile Fronthaul over APN PoC Reference
(5) IOWN Data Hub PoC Reference
(6) Fiber Sensing Crosstalk Study for Open APN PoC Reference
(7) PoC Reference: Reference Implementation Model for the Area Management Security Use Case

The main PoC Reference documents that will be discussed in the Feature Articles in this issue are introduced below.

The Open APN Architecture is an open architecture for building the multi-vendor APN that enables communications with deterministic transmission rates and latency. PoC Reference (1) describes dynamic optical path design, open interfaces, and support for multi-vendor environments. PoC Reference (4) describes applying the Open APN to mobile fronthaul and presents interface requirements, dynamic path switching, etc. to enable mobile network operators to use the Open APN.

The IDH is an architecture for database storage that is hierarchically distributed to meet stringent requirements. PoC Reference (5) presents smart factory, smart grid, and metaverse scenarios in which the IDH is applied and describes the configuration and evaluation of the implementation.

A DCI is an architecture for streamlining data transfer by incorporating network interfaces into data-processing modules, such as accelerators. This enables multiple remotely located DCIs to share information technology (IT) resources by transferring data at high speed and with low latency via the APN. PoC Reference (3) shows use cases, such as area management, virtualized radio access networks, and disaggregated IDH, and describes the implementation and evaluation of the PoC.

2.2 RIMs

The IOWN Global Forum has developed RIMs for promising use cases. It is important to develop full-stack engineering of these models by taking the characteristics of each use case into consideration. RIMs provide best practices for overall implementation. The following RIM documents have been released:

(1) RIM for the Interactive Live Music Entertainment Use Case
(2) RIM for the Area Management Security Use

Fig. 6. Increase in IOWN Global Forum members.

Case

(3) RIM for the Remote-Controlled Robotic Inspection Use Case

Interactive Live Music Entertainment is a use case selected as an AI-Integrated Communication. This use case delivers the musical performances of artists with volumetric video and audio to the audience in real time and reflects the audience's reactions in the virtual space. RIM document (1) provides the technical specifications of the devices used in the use case and functional requirements of the system and presents guidelines for the implementation of the overall configuration.

Area Management Security is a use case selected as a cyber-physical system. In this use case, a large number of surveillance cameras are installed in urban areas and facilities to prevent crime and ensure security. The Open APN, DCI, and IDH are applied to this use case. RIM document (2) shows an analysis of the data flows and workload of camera-captured videos and other data in an implementation configuration and describes the configuration and effects of each technology when it is applied.

Remote-Controlled Robotic Inspection is a use case for the remote control of a factory, as used by petrochemical plants in their production processes. RIM document (3) presents the requirements to be met by the data to be acquired, remote-control requirements, an analysis of data flows, and the implementation configuration and its effects.

As described above, the IOWN Global Forum has been analyzing the requirements, studying the effects of the PoC-implementation models for specific use cases on the basis of the PoC Reference documents for each IOWN component technology, releasing the results of these studies, and conducting actual PoC verification projects on the basis of these results. The reports on the results of these PoC verification projects will be released in the coming years.

3. Organization and activities of the IOWN Global Forum

3.1 Expansion of membership

The IOWN Global Forum was established in January 2020 by three companies (Intel, Sony, and NTT). The number of members has since been rapidly increasing, reaching 39 organizations by January 2021, 88 by January 2022, and exceeding 130 by September 2023. New members are joining almost every month (**Fig. 6**).

Reflecting the characteristics of the forum's activities

Fig. 7. IOWN Global Forum Member Meeting in Munich.

wherein use-case studies and technical studies proceed hand in hand, not only those who study and develop IOWN Global Forum technologies but also many of those who use these technologies participate in the forum. Participation in the IOWN Global Forum is not confined to enterprises. Research institutes, universities, and academic institutions are also actively involved. Municipalities are also taking part, and major cities around the world have expressed their interest [2].

3.2 Recent activities of the IOWN Global Forum

Since its inception, due to the COVID-19 pandemic, the IOWN Global Forum has conducted its activities online. However, as the pandemic has now subsided globally and there has been a revival of social activities from the latter half of 2022, the forum has gradually shifted to holding a hybrid of online and offline meetings. In October 2022, the 5th Member Meeting—the first locally held member meeting—took place in New York City in a hybrid format, with 136 local participants and more than 300 online participants from 44 member organizations located in various countries.

In April 2023, after the peak of the 8th coronavirus wave in Japan, the 3rd Annual Member Meeting—the first locally held annual member meeting—was held in Osaka in a hybrid format. More than 400 participants from 78 member organizations from different countries gathered in Osaka while many others participated online. There was lively discussion around updating the forum's vision and use cases, and the related technologies were intensively considered.

In September 2023, the 6th Member Meeting was held in Munich [3]. This was the first member meeting held in Europe and was attended by more than 400 people (190 locally) from 89 member organizations from various countries (**Fig. 7**). All meetings were operated in a hybrid format, which allowed many people to participate in meetings online while enabling local participants to hold in-depth discussions. Prior to the meeting, a call for proposals regarding activities in Phase 3 was issued, and 34 proposals were received. In the Munich meeting, breakout workshops, each dedicated to a specific subject category, were organized and active technical discussions on these proposals took place. The meeting also officially released a reference document that identified key values and presented a roadmap for technological evolution [4]. The values identified included the importance of a holistic approach, an implementation-driven approach, and system-level solutions. A roadmap for evolution in networking and computing was presented. “An Implementation of Heterogeneous and Disaggregated Computing for DCI as a Service” was also approved as the first recognized PoC report. More PoC reports will be published in the coming years to demonstrate the value and performance of IOWN technologies.

3.3 Other activities

In parallel with the Use Case and Technology WGs, the Liaison WG of the IOWN Global Forum has been active, promoting collaboration with other technical and standardization organizations [5]. The Liaison WG exchanges various types of information with collaborating organizations, jointly studies specifications, and seeks to have the IOWN Global Forum’s technologies adopted as standards. The Liaison WG is currently collaborating with the International Telecommunication Union (ITU) Radiocommunication Sector, the ITU Telecommunication Standardization Sector, the Linux Foundation, OpenROADM, and other organizations, and plans to expand the scope of collaborating organizations.

The IOWN Global Forum is also actively engaged in external activities as part of its marketing initiative. In addition to giving lectures and demonstrations at events held by member companies, the forum participates in events hosted by the information and communication industry and academic conferences, giving speeches and organizing exhibitions.

In 2023, the IOWN Global Forum jointly set up an exhibit with other companies at the Mobile World Congress (MWC) held in Barcelona in February, gave a speech at the Optical Fiber Communication Conference and Exposition (OFC) held in San Diego in March and at the Institute of Electrical and Electronics Engineers (IEEE) International Conference on Communications held in Rome in June. At the NTT booth at Interop Tokyo 2023 held in Chiba, Japan in June, NTT exhibited the remote control of a robot and a virtual-reality table-tennis application using remote graphics processing units to demonstrate the ultra-low latency achieved using the Open APN.

4. Activities and prospects for Phase 3

In Phase 3, entitled “Preparation of Real-World Deployment and Business Impact,” the IOWN Global Forum aims to further develop technology, expand the organization and activities, and create business cases toward the deployment of IOWN. The forum will also conduct PoC and technical evaluation of specific use cases and release PoC-result reports. It will develop technical-specification documents, deepen collaboration with other organizations, and work toward standardization of the forum’s specifications. It will also seek to deploy IOWN by expanding support for the forum members, such as holding local events and seminars, to accelerate studies tailored to each country or region so that many user companies can use IOWN in their respective businesses.

References

[1] IOWN Global Forum technology documents, https://iowngf.org/technology/

[2] IOWN Global Forum members, https://iowngf.org/members/

[3] IOWN Global Forum 6th Member Meeting September 2023 – Highlights, https://iowngf.org/event/iown-global-forum-6th-member-meeting/

[4] IOWN Global Forum, “IOWN Global Forum: Key Values and Technology Evolution Roadmap,” https://iowngf.org/wp-content/uploads/formidable/21/IOWN-GF-RD-KVaTE_Roadmap-1.0-1.pdf

[5] IOWN Global Forum’s Liaison Program, https://iowngf.org/liaisons/

Yoshitake Tajima
Director, IOWN Development Office, Research and Development Planning Department, Research and Development Market Strategy Division, NTT Corporation.

He received a B.E in electronic engineering and M.E. in information and communication engineering from the University of Tokyo in 1997 and 1999. He joined NTT Information Sharing Platform Laboratories in 1999. He has been mainly engaged in the research and development (R&D) of network security and computing architecture. He also received an MSIA from School of Business, Purdue University, USA in 2014. He is currently engaged in the promotion of IOWN R&D at NTT and administrative operation of the IOWN Global Forum. He is a member of IEEE, the Institute of Electronics, Information and Communication Engineers (IEICE), and Information Processing Society of Japan (IPSJ).

Yosuke Aragane
Vice President, Head of IOWN Development Office, Research and Development Planning Department, Research and Development Market Strategy Division, NTT Corporation.

He received an M.S. and Ph.D. in engineering from the Tokyo Institute of Technology in 1997 and 2005. Since joining NTT in 1997, his major research area was cyber security, especially in CSIRT (computer security incident response team) operations. Since 2016, he has been the head of Cyber Security R&D Promotion in NTT Research and Development Planning Dept. In 2019, he became a vice president of IOWN Development Office. He is an alternate director, leader of Recruitment Taskforce, and co-leader of Officer team in the IOWN Global Forum. His current responsibilities are managing IOWN R&D in NTT labs and leading the IOWN Global Forum with diverse partners. He has the CISSP certification and is a member of IEEE, IEICE, and IPSJ.

Activities for Detailing the Architecture of the Open APN and Promoting Its Practical Application

Yoshihiko Kainuma, Tomonori Takeda, Jun-ichi Kani, and Hideki Nishizawa

Abstract

The IOWN Global Forum has proposed the Open All-Photonic Network (Open APN) as a network infrastructure that achieves low latency and low power consumption and issued an architecture document entitled "Open All-Photonic Network Functional Architecture," in early 2022. This article introduces the activities carried out since the release of the above architecture document to detail the Open APN architecture and promote its practical application.

Keywords: IOWN, Open APN, PoC

1. Introduction

The Open All-Photonic Network (Open APN) is an open architecture for photonics networking proposed by the IOWN Global Forum (IOWN GF), which is studying ways that would enable service providers to integrate photonics networking capabilities with their overall computing and networking infrastructure at a finer level of granularity. Since the architecture document, Open All-Photonic Network Functional Architecture release 1 [1], was issued in early 2022, the forum conducted a proof of concept (PoC) based on this architecture document and issued the Open All-Photonic Network Functional Architecture release 2 in October 2023 [2], which expands and details the architecture. This article introduces the forum's activities since the publication of the architecture document release 1.

2. Open APN

Let us first briefly explain what the Open APN is. It is a network that enables different locations to be directly connected using optical wavelength paths and consists of Open APN Transceivers (APN-Ts), Open APN Gateways (APN-Gs), and Open APN Interchanges (APN-Is). An APN-T is an endpoint for an optical wavelength path and transmits and receives optical signals. An APN-G is a gateway for an optical wavelength path and sets up control channels to communicate with the connected APN-Ts, executes user plane admission control, multiplexes/demultiplexes, turns back, and adds/drops optical wavelength paths. An APN-I is an interchange for an optical wavelength path, executes wavelength cross-connect, and provides adaptation between interfaces (**Fig. 1**).

For more details on the architecture, please refer to the article, "Study on Open All-Photonic Network in IOWN Global Forum" in the May 2022 issue of this journal [3].

3. Open APN PoC

IOWN GF planned the Open APN PoC [4] activities to show the global market that the Open APN Functional Architecture is not a technology of the future but can be built and operated using technologies already available. These activities have prompted many organizations around the world to aim at building and introducing the Open APN, create use

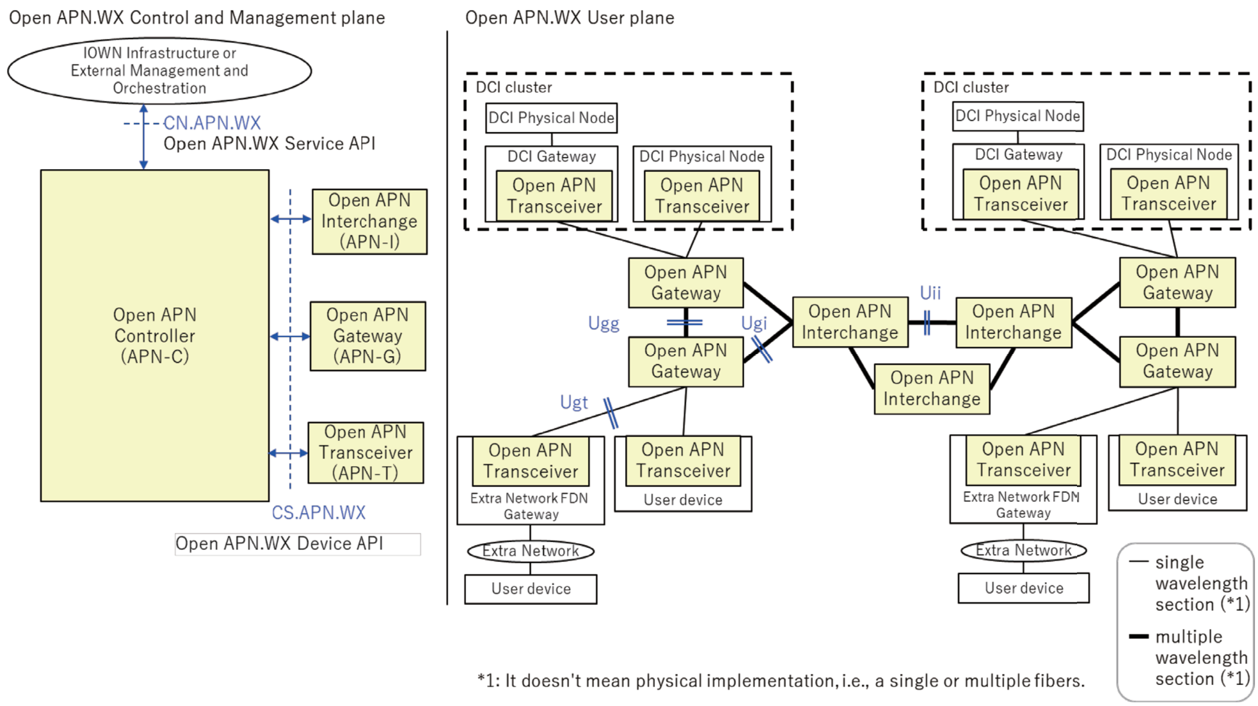


Fig. 1. Open APN reference architecture for optical wavelength path services.

cases, and deploy them in the global market. IOWN GF has published a PoC Reference document, which provides guidelines for conducting PoC, enabling anyone to undertake PoC under a common set of guidelines. The PoC Reference document for the Open APN provides guidelines for the following items to enable those interested to conduct a comprehensive demonstration of the architecture defined not only by the two task forces (TFs) studying the Open APN architecture, namely the Open APN Architecture TF and Open APN Wavelength Availability TF, but also by the Open APN Fiber Sensing TF, which is studying technology for sensing using optical fibers. The functions subject to evaluation as defined in the PoC Reference document are introduced below.

3.1 Basic functions of an APN-G

(1) Setting up a control channel for an APN-T
(2) Executing user plane admission control of optical signals
(3) Multiplexing/demultiplexing optical wavelength paths
(4) Turning back optical wavelength paths
(5) Adding/dropping optical wavelength paths
(a) Flexible Bridging Service: A service that aggregates multiple data flows with extremely stringent quality-of-service requirements and forwards them into a single optical wavelength path
(b) Open interface: An open interface used to connect devices from different vendors to the Open APN Controller (APN-C) or used for control by the APN-C
(c) Multi-vendor support: Interoperability between an APN-T and APN-G supplied by different vendors
(d) Automatic provisioning: Automatic provisioning of optical wavelength paths, including physical design, on the basis of link parameters such as fiber type and power profile
(e) Mechanism for transceiver connection at user sites: Automatic registration of transceivers at user sites that terminate optical wavelength paths
(f) Monitoring: Automatic and periodic collection of monitoring data from an APN-T and APN-G by the APN-C
(g) Fiber sensing: Confirmation of fiber sensing at an APN-G under the following conditions:
- The interrogator is an optical time-domain reflectometer for general loss measurements. It uses a wavelength of 1550 nm.

- The circulators and couplers are general commercial products.
- The optical switch is a commercial switch that supports 1550 nm.
- The sensing fiber is a commercial single-mode optical fiber with a length of more than 10 km.

Many IOWN GF member companies have conducted PoC on the basis of this PoC Reference document. IOWN GF will publish the results of PoC implementation with a view to promoting further study and expanding the number of participants.

4. Updating of architecture document

IOWN GF issued the Open APN Functional Architecture release 2 document in October 2023 [2]. This document is an update of the release 1 architecture document issued in early 2022 [1]. Its main focus is expanding and detailing the basic architecture, control plane architecture, and user plane architecture.

4.1 Basic architecture

The basic architecture of the Open APN has evolved into a two-layer structure with the definition of a new fiber layer in addition to the wavelength layer defined in the release 1 document. The wavelength-layer architecture is called Open APN Wavelength Exchange (Open APN.WX) while the fiber layer architecture is called Open APN Fiber Exchange (Open APN.FX). Open APN.FX newly defines Open APN Fiber Cross-connect (APN-FX), which is a gateway for setting up a fiber path using a physical fiber to connect endpoints in the Open APN, and various interfaces on the fiber path. It also clearly describes how these layers are combined to set up fiber paths and wavelength paths (**Fig. 2**).

While the architecture in the release 1 document had a point-to-point (PtP) configuration, the release 2 architecture additionally incorporates a point-to-multipoint (PtMP) configuration, in which a single signal is branched into multiple points. This new document describes both PtMP in Open APN.WX and PtMP in Open APN.FX.

4.2 Control and management plane architecture

The new document further details control and management plane architecture for optical wavelength paths in a PtP configuration. Assuming a reference model in which an APN-T is deployed at the user's location, the document systematically presents the functions required for the APN-C to set up an end-to-end optical wavelength path including user devices and describes the application programming interfaces (APIs) used for interfacing with the orchestrator and each APN node. It also clarifies how the APN-C cooperates with each entity to control and manage optical wavelength paths (**Fig. 3**).

The use of existing open interfaces, such as OpenConfig, OpenROADM (reconfigurable optical add-drop multiplexer), and Telephony Application Program Interface (TAPI), in implementing the southbound interface is also described (**Fig. 4**).

4.3 User plane architecture

The release 1 document defined the concept of Group of Optically Interoperable Ports (GOIP), which is a subnetwork within the Open APN that can connect to other GOIP subnetworks through optical paths. The release 2 document newly describes the interworking between GOIPs, which was treated as an item for future study in the release 1 document. The release 2 document defines a method for setting up an optical connection when a wavelength can be assigned and the required transmission characteristics are satisfied and a method for setting up a connection that involves electrical conversion using optical-electrical-optical (O/E/O) repeaters or Ethernet/Optical Transport Network (OTN) connections. These methods make it possible to construct the Open APN that scales beyond GOIPs (**Figs. 5(a)**, **(b)**).

In addition to these updates, annexes have been added to the release 2 document. The annexes describe methods for implementing each Open APN node, use cases, and PtMP implementation examples, thus providing more detailed guidelines for implementing and using the Open APN.

5. Activities at NTT

NTT's contribution to these IOWN GF activities has not been confined to the submission of documents but also includes other activities. Two of these are introduced below.

5.1 Demonstration test of Open APN function group

Prior to the publication of the release 2 document, NTT laboratories conducted, within a laboratory, a demonstration test of "a group of functions that enable high-capacity, low-latency communications between required locations when needed," which is a mandatory condition for further expanding use cases in the Open APN architecture [5]. This function group enables Open APN endpoints to be installed at

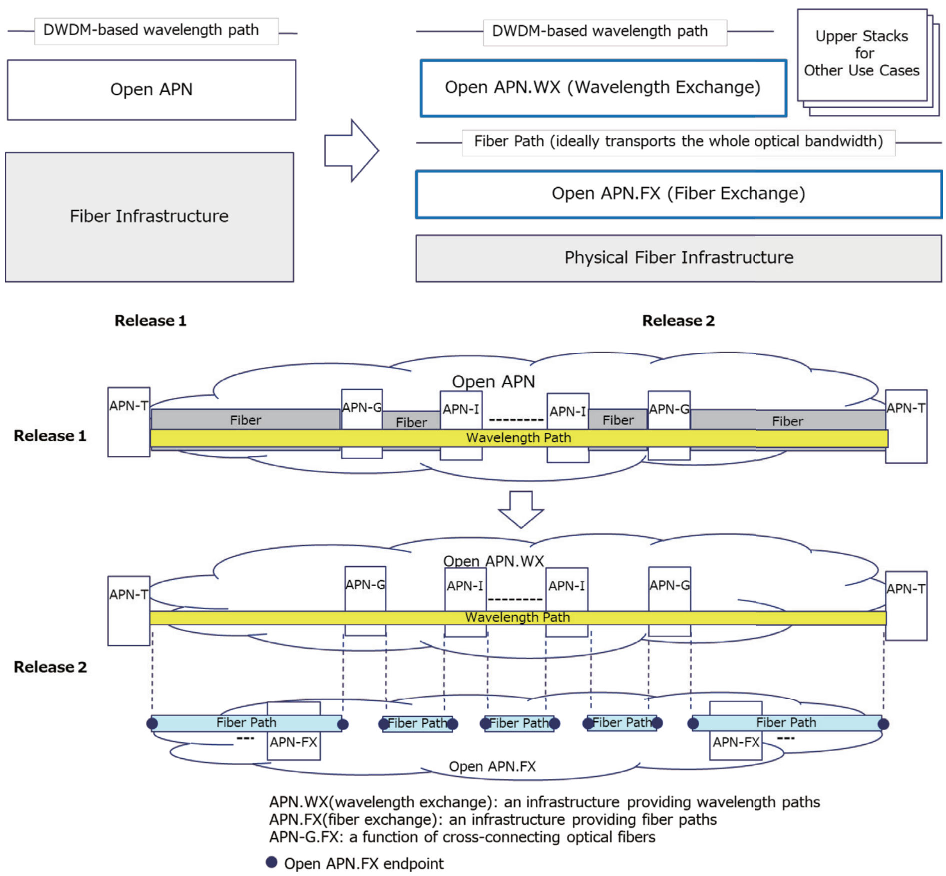


Fig. 2. Basic architecture of the Open APN.

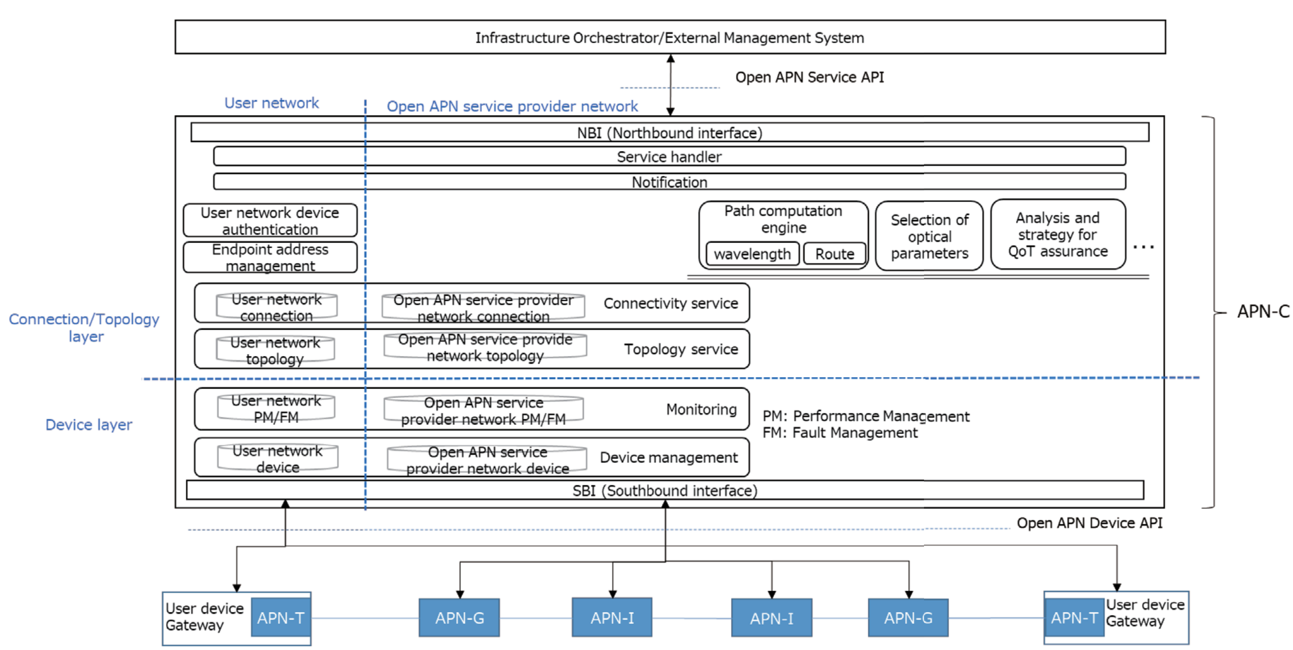


Fig. 3. Example reference of APN-C functional model.

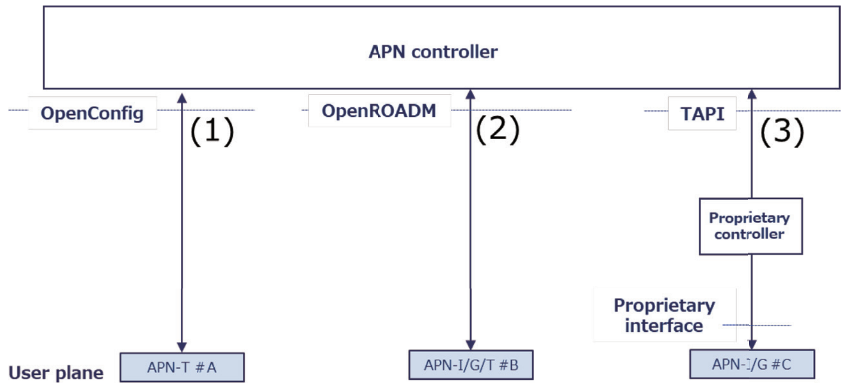
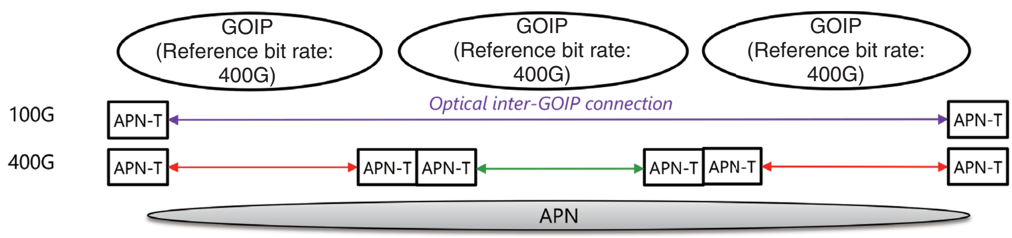
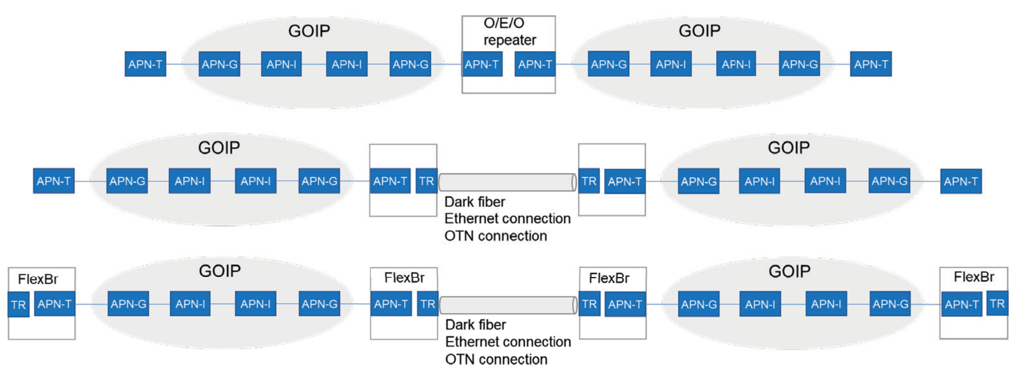


Fig. 4. Recommendation set of southbound interfaces for the APN-C.



(a) Method for setting up an end-to-end optical connection



(b) Method for setting up a connection that involves electrical conversion

Fig. 5. Inter-GOIP connections.

different sites, including user sites. This function group consists of the following three functions (**Fig. 6**):

(1) A function for automatic design and provisioning of optical wavelength paths that satisfy service conditions, implemented through coordination and collaboration between user site terminals and telecommunications operators' devices, and function for adjustment, such as selection of an optimal mode

(2) A function for setting and managing optical wavelength endpoints at a user site terminal and function for passing/stopping optical signals at telecommunications operators' devices on the basis of the authentication status of the terminal

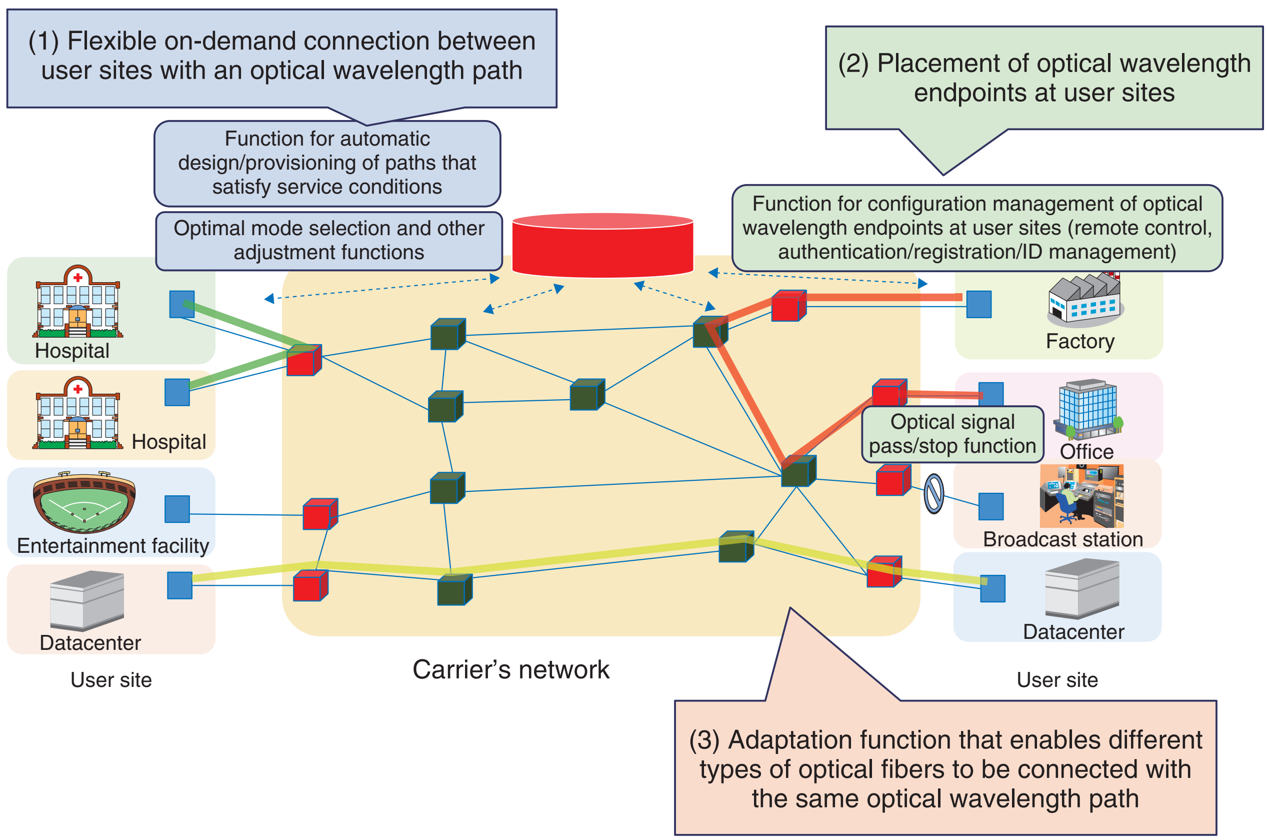


Fig. 6. Demonstration test of an Open APN function group for on-demand wavelength connections between user sites.

(3) An adaptation function that enables different types of optical fibers to be connected with each other while keeping the optical wavelength path unchanged

Using the results of the OpenROADM Multi-Source Agreement and Telecom Infra Project Open Optical & Packet Transport, which are organizations promoting flexible and highly interoperable open optical transmission systems, NTT is conducting a field test in the Tokyo metropolitan area to demonstrate that the system can operate under a variety of conditions, including fiber lengths and loss levels that occur in real-world environments. On the basis of the results of this field test, NTT submitted documents for the compilation of the release 2 document.

5.2 Use of the Open APN in a real environment through a dynamic exhibit at a trade show

At Interop Tokyo 2023 held in June 2023, a dynamic exhibit of the Open APN was staged [6]. This dynamic exhibit featured the following three points:

(1) A network connecting multiple remote locations was constructed.

(2) Various services, such as GPU (graphics processing unit) as a service, public cloud access, and internet access, were provided.

(3) Multi-vendor connections of APN-G/Is and connections of multi-vendor APN-Ts were presented.

Through this dynamic exhibit, we were able to show that the Open APN is a low-latency, low-jitter network that can be built using multi-vendor devices. By providing services that involve connections to sites outside the exhibition site at a large trade show, such as Interop, we were able to show that the Open APN is a cost-optimal solution in such cases because it does not require separate fibers for each service. Previously, the Open APN had been mainly constructed in restricted environments for demonstration purposes. The exhibit at the Interop was the first version of the Open APN that had been constructed as a network able to provide a variety of services in a situation that corresponded closely to a practical implementation. A report on the exhibit was presented at the IOWN GF September 2023 Member Meeting held in Munich as the event demonstrated that a new step forward had been taken in the transition of the Open APN from the PoC stage to practical application.

6. Future development

This article introduced the Open APN-related activities undertaken after the publication of the release 1 architecture document. The Open APN is expected to be used for various use cases in combination with the data-centric infrastructure, which is also introduced in another article in this issue. As a network infrastructure for IOWN Data Hub and IOWN Mobile Network, which were also introduced in another article in this issue, the Open APN plays a significant role in meeting extremely stringent requirements. Together with many IOWN GF partners, NTT will continue to engage in a series of discussions and demonstrations to contribute to the further development of technologies and social implementation.

References

[1] IOWN GF, "Open All-Photonic Network Functional Architecture," Ver. 1.0, Jan. 2022. https://iowngf.org/wp-content/uploads/formidable/21/IOWN-GF-RD-Open-APN-Functional-Architecture-1.0-1.pdf

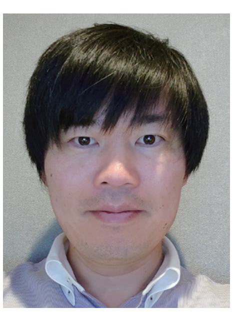
[2] IOWN GF, "Open All-Photonic Network Functional Architecture," Ver. 2.0, Oct. 2023. https://iowngf.org/wp-content/uploads/formidable/21/IOWN-GF-RD-Open_APN_Functional_Architecture-2.0.pdf

[3] H. Nishizawa, J. Kani, T. Hamano, K. Takasugi, T. Yoshida, and S. Yasukawa, "Study on Open All-Photonic Network in IOWN Global Forum," NTT Technical Review, Vol. 20, No. 5, pp. 18–23, May 2022. https://doi.org/10.53829/ntr202205fa2

[4] IOWN GF, "Open APN Architecture PoC Reference," Ver. 1.0, Nov. 2022. https://iowngf.org/wp-content/uploads/formidable/21/IOWN-GF-RD-Open_APN_Architecture_PoC_Reference-1.0.pdf

[5] NTT press release, "NTT conducts an Open APN PoC jointly with global optical product vendors, conducts a field test of on-demand wavelength connection technologies, and enhances its capability to develop photonics-electronics devices," Nov. 14, 2022. https://group.ntt/en/newsrelease/2022/11/14/221114a.html

[6] NTT press release issued on June 12, 2023 (in Japanese). https://group.ntt/jp/newsrelease/2023/06/12/230612b.html



Yoshihiko Kainuma

Senior Manager, NTT IOWN Development Office, Research and Development Planning Department, Research and Development Market Strategy Division, NTT Corporation.

He received an M.E. in information and computer science from Keio University, Tokyo, in 2007. He joined NTT WEST in 2007 and NTT in 2021. He has been engaged in the research and development management of a next-generation communications infrastructure and joint work with partners to commercialize the infrastructure.

Tomonori Takeda

Senior Research Engineer, Supervisor, Network Architecture Project, NTT Network Service Systems Laboratories.

He received an M.E. in electronics, information, and communication engineering from Waseda University, Tokyo, in 2001. He joined NTT the same year and has been engaged in research and development on the next-generation transport network architecture and next-generation mobile network architecture.

Jun-ichi Kani

Senior Distinguished Researcher, Group Leader, Optical Access Systems Project, NTT Access Network Service Systems Laboratories.

He received a B.E., M.E., and Ph.D. in applied physics from Waseda University, Tokyo, in 1994, 1996, and 2005. He joined NTT Optical Network Systems Laboratories in 1996, where he researched optical multiplexing and transmission technologies. He has been with NTT Access Network Service Systems Laboratories since 2003, where he is engaged in the research and development of optical communications systems for metropolitan and access network applications and currently heads the Access Systems Technology Group. He has been participating in the International Telecommunication Union - Telecommunication Standardization Sector (ITU-T) and the Full Service Access Network (FSAN) initiative since 2003.

Hideki Nishizawa

Senior Research Engineer, Supervisor, Frontier Communication Laboratory, NTT Network Innovation Laboratories.

He received an M.S. in physics from Chiba University in 1996 and joined NTT the same year, where he has been engaged in research on ubiquitous networking, photonic switching systems, and photonic transmission systems. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE).

Technical Study on IOWN for Mobile Networks

Manabu Sugihara and Yoshitsugu Shimazu

Abstract

The IOWN for Mobile Network Task Force has been investigating technical topics related to mobile and transport networks in the Innovative Optical and Wireless Network (IOWN), including the study of the key requirements for transport networks in the transition from the current 5th-generation mobile communications system (5G) to 6G, analysis of the gaps in technology hindering actualization of the uses cases aimed at by the IOWN Global Forum and proposal of solutions to them, and exploration of a future transport network architecture that leverages the All-Photonics Network and data-centric infrastructure. This article introduces the activities of this task force.

Keywords: Technical Outlook for Mobile Networks Using IOWN Technology, PoC Reference, IMN

1. Technical Outlook for Mobile Networks Using IOWN Technology

The IOWN for Mobile Network Task Force (IMN-TF) has been studying, proposing and discussing, together with global vendors and carriers, use cases and technologies required for mobile and transport networks that use IOWN Global Forum (IOWN GF) infrastructures. The IMN-TF released "Technical Outlook for Mobile Networks Using IOWN Technology" [1] in January 2022 and "Technical Outlook for Mobile Networks Using IOWN Technology - Advanced Transport Network Technologies for Mobile Network" [2] in April 2023. This article focuses on the topics that have been discussed in the IMN-TF activities to date.

Many advanced use cases for the 5th-generation mobile communications system (5G), including video-centric applications and latency-sensitive cases, such as area management, telesurgery, augmented reality/virtual reality, and industrial automation, demand huge end-to-end bandwidth capacity, high reliability and availability, and extremely low latency on mobile networks. Among the various technologies that can be used to provide transport networks for mobile cell sites, fiber is the best solution for fixed network sites because of its high capacity and widespread deployment. Therefore, the IMN-TF is proposing a method for implementation using fiber-based optical transport networks for mobile networks that use the data-centric infrastructure (DCI) and Open All-Photonic Network (Open APN) architectures.

5G can achieve high data rates using a wide bandwidth available in millimeter wave frequency bands but requires a high-density installation of cell sites to cover a given area. Mobile-network infrastructures are evolving toward cloud-native network function virtualization, which opens up a variety of transport network options and configurations. We are studying such options and configurations from both technical and economic perspectives, including whether they are in line with technical trends, will improve service deployment and provisioning agility, and will reduce operating costs.

New services and applications enabled by 5G will use enhanced mobile broadband, ultra-reliable low latency communication, massive machine-type communication, and fixed wireless access technologies. Such new services and applications generate larger amounts of data and require latency to be extremely low. As the volume of data in mobile networks increases, the entire transport network, from the backhaul to the datacenter, is required to have high

Table 1. Target fronthaul bandwidth and delay requirements of various RAN split options.

KPI	5G (2020)	6G (2030) (Projected)
E2E		
Peak data rate	< 10 Gbps	< 100 Gbps–1 Tbps
User plane latency (ms)	1	0.1
Transport: Lower Layer Split (Option 7)		
Bandwidth	< 25–50 Gbps	< 250 Gbps–5 Tbps
Frame delay (one-way)	0–160 µs Fiber delay: 0–150 µs (0–30 km) Packet delay variation (PDV): 0–10 µs (0–2 switches)	0–larger than 160 µs Fiber delay: 0–larger than 150 µs (0–larger than 30 km) PDV: 0–less than 10 µs
Transport: Higher Layer Split (Option 2)		
Bandwidth	< 10 Gbps	< 100 Gbps–1 Tbps
Frame delay (one-way)	Up to ms order (up to 100 km order)	Up to ms order (up to 100 km order)

E2E: end to end
KPI: key performance indicator

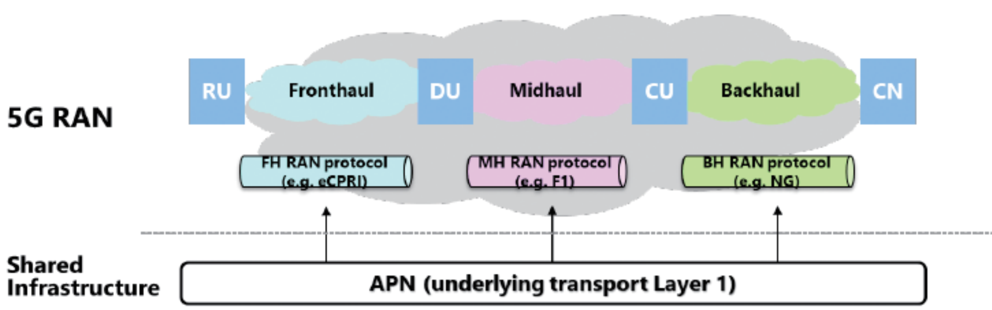


Fig. 1. RAN and the APN.

capacity. People are increasingly using time-sensitive applications, which demand ultra-low latency on the order of milliseconds. Among the technical reports released by the 3rd Generation Partnership Project (3GPP), 3GPP TR 38.801 [3], 3GPP TR 38.806 [4], and 3GPP TR 38.816 [5] specify eight possible radio access network (RAN) functional split options. The most stringent requirements considered by mobile carriers are listed in **Table 1**. The IMN-TF will conduct a proof of concept (PoC) on the feasibility of building a transport-network solution that can meet these stringent requirements (**Fig. 1**).

Considering that IOWN mobile networks use the All-Photonics Network (APN) and DCI technologies, the task force is discussing scenarios for RAN deployment on the APN/DCI, high availability and fault tolerance enabled by the DCI and APN, and use cases involving elastic load balancing, which reduces power consumption by flexibly operating base stations in response to daytime and nighttime variations in mobile traffic.

1.1 RAN deployment scenarios on APN/DCI

In studying scenarios for RAN deployment on the APN/DCI, it is necessary to consider the connectivity and performance of the transport layer. The definition of a transport-connection service typically includes topology, a user network interface, traffic/bandwidth profiles/characteristics, protocols, and a set of their service attributes. When setting up connections to RAN functions, it is important to support the characteristics of the RAN domain, such as connections using the standardized enhanced Common Public Radio Interface of the Fronthaul (FH) RAN protocol.

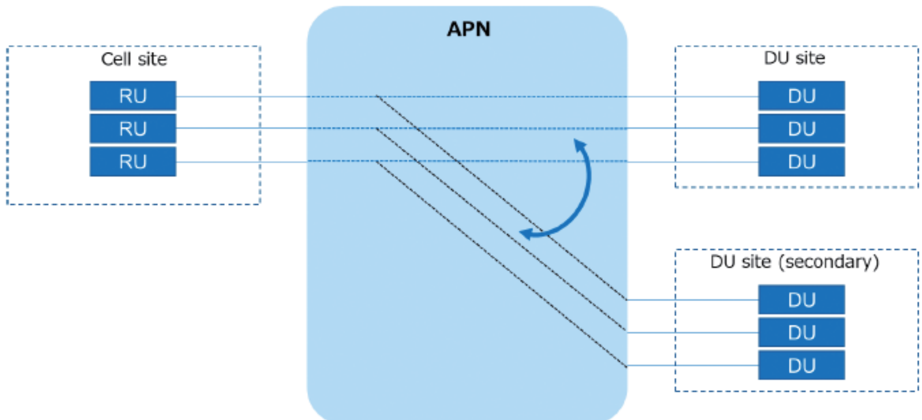


Fig. 2. Dynamic path switching to achieve high availability and elastic load balancing.

There are two methods for connecting each RAN function to the APN: one that directly terminates a wavelength path with an optical transceiver and another that uses a bridging function such as a flexible bridge service.

Passive optical network (PON) technology can be used to directly connect the radio unit (RU) side to the APN. For example, time division multiplexing (TDM)-PON can be used. However, when using TDM-PON, it is necessary to limit the distance between an RU and distributed unit (DU) to 10 km or less to compensate for any packet delay that exceeds the packet delay variation (PDV) requirement (10 microseconds). An advantage of using a flexible bridge service is that multiple types of services can be defined to support multiple quality of service levels (delay, jitter, etc.) depending on the given use case. The task force is discussing various APN-mediated solutions, including a flexible bridge service, and developing a PoC Reference [6] for a demonstration model for Mobile FH over APN that satisfies both stringent mobile fronthaul requirements and a wide variety of base-station installation configurations.

1.2 High availability and elastic load balancing

By combining the APN, DCI and a subsystem, namely, a RAN intelligent controller, it becomes possible for the FH network to flexibly switch DUs within a DCI resource pool. A virtualized DU (vDU) is a virtual node that runs on a network functions virtualization infrastructure (NFVI) and can be freely reconfigured on multiple NFVIs installed at different geographic locations. We believe that we can leverage this feature to reduce power consumption. The associations between physical servers, DUs, and RUs in an FH network may change as a result of external triggers such as the failure state of the components involved, load balancing requirements, and energy saving requirements. High availability of RAN systems can be achieved by changing the physical associations between RUs and DUs using the APN-wavelength path-switching function when triggers such as the following occur: traffic fluctuation or detection of a problem such as a failure or service degradation detected by the network. For example, all user devices (user equipment: UE) connected to an RU reconnect to a new DU that has been reconfigured on a physically different server, or some of the RUs connected to a DU reconnect to a physically different DU taking the balancing of resources into account. We are studying an elastic load-balancing function to improve flexibility and resource-utilization efficiency in these FH networks. This function is expected to improve the availability and power efficiency of RAN systems. For example, by shutting down unused DCI resources (vDUs), thus reducing their power consumption, we can improve the power efficiency of RAN systems (**Fig. 2**).

1.3 Enhanced CTI

With a view to providing end-to-end low-latency and low-jitter services using IOWN technology, the IMN-TF is also studying an extended cooperative transport interface (eCTI) between a mobile network and the APN.

When using an optical access system as a transport network for a mobile system, it is necessary to reduce the latency of the optical access system because the

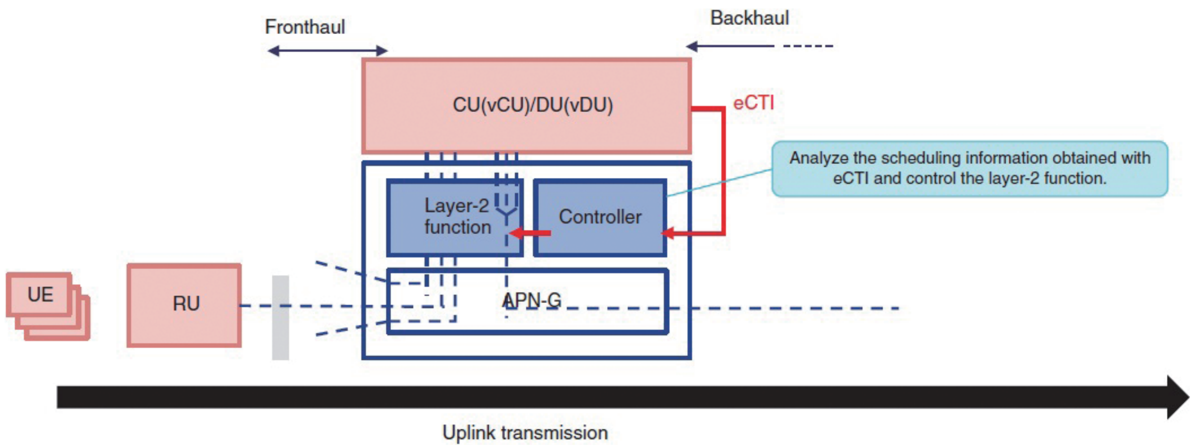


Fig. 3. eCTI overview.

allowable latency for a mobile system is very short. Therefore, cooperative dynamic bandwidth allocation (CO-DBA), a function to reduce latency by determining the communication timing of the optical access system in cooperation with the mobile system, was specified. The Open Radio Access Network (O-RAN) Working Group (WG) 4 has specified a CTI, which is the interface between the mobile system and optical access system for exchanging the information required for operating CO-DBA. Since the current CTI specification cannot be applied to short transmission time interval and mobile midhaul (MH)/backhaul (BH) that provide services demanding low latency in a mobile network, it is necessary to extend the CTI specification as an eCTI in the IOWN GF architecture. We are studying a latency-control method with which an orchestrator coordinating the mobile network and APN operations collects mobile information, such as the aggregate traffic load, and transfers the data to the APN controller. In implementing high-precision remote control and automatic operation of robots and other devices, stable control is difficult if the fluctuations (jitter) of packet arrival time are large. Therefore, it is necessary to keep jitter, which has not been specified as a performance requirement in conventional networks, within a certain time frame.

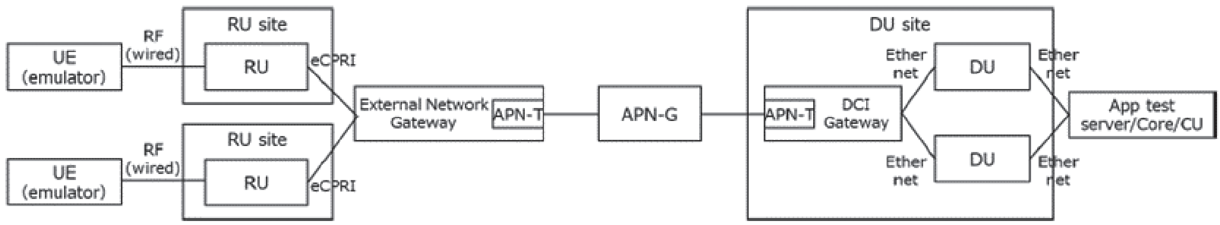
To suppress delay fluctuations, we are discussing a method for controlling the devices that make up the mobile system and transport network on the basis of information about transmission timing and the volume of transmitted data in the wireless section (**Fig. 3**).

2. PoC Reference

The IMN-TF recognizes that, to support the use cases envisioned by IOWN GF, it is necessary to satisfy the very stringent O-RAN FH requirements and that the Open APN is a promising solution for meeting these requirements. It is therefore necessary to demonstrate the feasibility of implementing mobile FH to which the APN is applied. While various technologies and architectures have been considered for FH solutions, the task force considers that one of its missions is to demonstrate the value of applying the APN to FH, thereby having the APN recognized widely as a major FH solution.

The purpose of this PoC is to demonstrate, to mobile network operators and other operators who wish to offer FH as a service, the benefits of applying the Open APN as an FH solution and the feasibility of O-RAN key performance indicators (KPIs) and present the demonstration test results. We expect that the recognition of Mobile FH on APN as a viable and promising solution will encourage mobile network operators worldwide to adopt this solution, which in turn will lead to further advancement of IOWN GF-related technologies.

There are two aspects that we want to demonstrate in this PoC. The first is a benchmark test to demonstrate that the energy efficiency of the APN is higher than that of other alternative solutions. An example of its benefits is that the operating cost can be reduced because the APN eliminates the conversion from electricity to light and vice versa. The second is to demonstrate the feasibility of implementing network functions that can support high-availability services



APN-G: APN gateway
APN-GT: APN transceiver
RF: radio frequency

Fig. 4. Network configuration for PoC Step 1.

and elastic load balancing. Elastic load balancing is a term used in the definition of a use case in which connections of RUs to a set of vDUs can be actively and dynamically switched on the basis of actual load on the DUs. Although directly connecting RUs to vDUs using dedicated dark fibers without multiplexing is more efficient than using the APN, dedicated dark fibers do not give mobile network operators the flexibility to dynamically allocate computing and other necessary resources to DUs in response to traffic fluctuations. The APN, however, allows dynamic switching of wavelength paths, which enables mobile network operators to flexibly use computing resources at the destination in response to the current traffic volume and other factors. We believe that the APN will enable mobile network operators to optimize their DU computing resources, thus reducing both operating costs, including power costs, and base-station costs.

In this PoC, we plan to prove the feasibility of these two possibilities separately in two steps.

In Step 1, we will evaluate the feasibility of Mobile FH over APN. It will demonstrate that the very stringent O-RAN mobile FH requirements can be met. These requirements are specified in an O-RAN specification (O-RAN Fronthaul Interoperability Test Specification (IOT)) and in the technical requirements in IMN documents.

In Step 2, we will evaluate the energy efficiency of elastic load balancing. We already compared the energy efficiency of a model in which packets are multiplexed using layer 2/layer 3 (L2/L3) switches over dark fiber in normal operation mode with that of a model in which elastic load balancing is applied, and demonstrated the extent to which the latter reduced power consumption compared with the former. We will evaluate the impact of using the APN on services and operations. Specifically, we will verify the extent of end-to-end throughput degradation, the total time needed to switch vDU hosts, and changes in end-to-end throughput during switching of DU hosts to determine whether there are any service-continuity issues or other technical issues.

In Step 1, we plan to show that the O-RAN KPI is satisfied by collecting the latency and throughput data of the synchronization plane, control plane, and user plane using virtualized RAN (vRAN) devices and APN devices in the following network configuration.

Mobile FH between RUs and DUs is located at a certain distance apart and vDUs are concentrated at the same site. The minimum configuration is defined as follows. There are multiple RU sites in one cell site. Since signals from these RUs are to be multiplexed into one fiber on the APN, at least two RU sites are needed* (**Fig. 4**).

In Step 2, we plan to confirm the feasibility of the concept of elastic load balancing in the following network configuration. The environment in Step 1 will be extended so that, in addition to collecting data as in Step 1, the effectiveness and usefulness of the concept can be evaluated. In the extended environment, multiple vDUs will be deployed to allow switching from distributed processing on multiple hosts to centralized processing on a single host (**Fig. 5**).

The evaluation of two dynamic path-switching methods for wavelength switching on the APN is defined as follows.

Method 1 is based on optical switching. When the traffic aggregated for each vDU varies, this method switches the optical path between the extra network gateway and DCI gateway so that the APN gateway

* Bandwidth per RU: 10G, 25G, or 50G; distance between RU and DU: 2 patterns, 10–30 km.

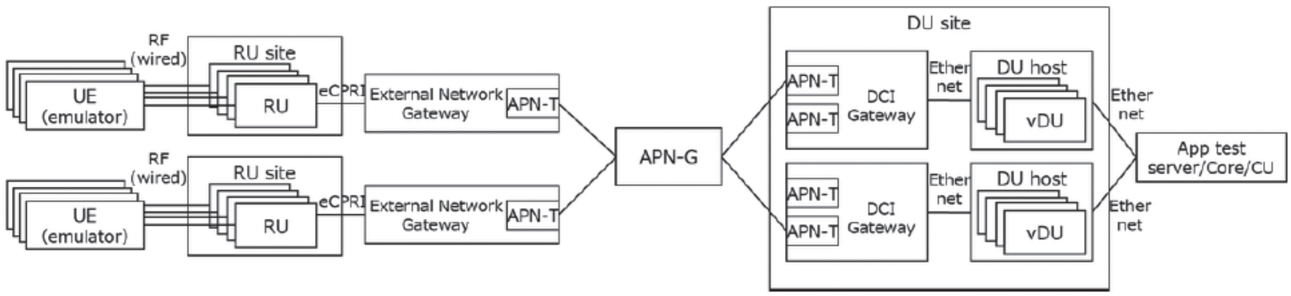


Fig. 5. Network configuration for Step 2.

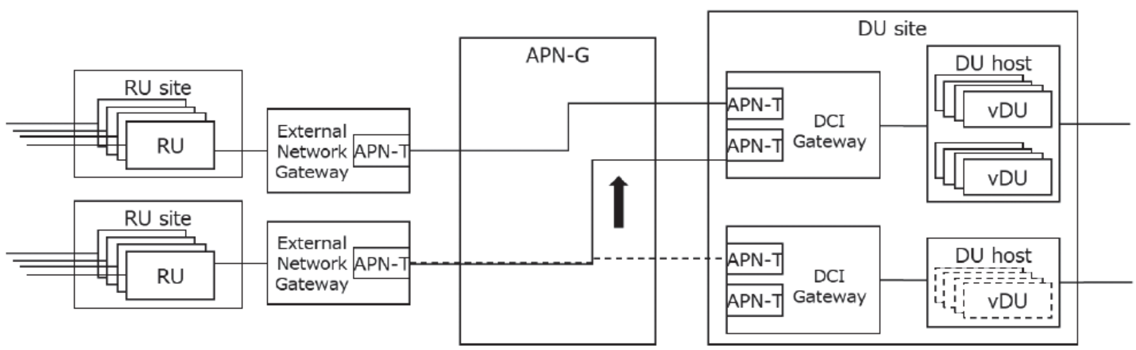


Fig. 6. Method based on optical switching.

(APN-G) will be reconfigured to optimize the number of vDUs used. More specifically, suppose that both RU sites and DU hosts are respectively connected. When the workload on the DU side decreases, this method switches the optical path that has been connected to the lower DU host to the upper DU host so that the upper DU host will handle the entire workload (**Fig. 6**).

Method 2 is based on packet switching. When the traffic varies, the DCI gateway switches the L2/L3 path between the DCI gateway and a DU host so that the vDUs will be reconfigured. More specifically, suppose that both RU sites and DU hosts are respectively connected. When the workload on the DU side decreases, this method switches the L2/L3 path to the upper DU host so that the upper DU host will handle the entire workload (**Fig. 7**).

In Step 2, the measurement of operational data and service level data is defined, including energy consumption, switchover time, and service-interruption time of the Open APN and external networks between the RU sites and DU hosts in addition to latency and throughput included in the measurements in Step 1.

3. Status of PoC implementation

IMN-TF member companies are holding discussions in preparation for both Steps 1 and 2, which are scheduled to be conducted up to FY2023. Member companies are stepping up efforts for Step 1, which includes five PoC projects that are either in process or in a preparatory stage. IMP-TF has received and discussed various proposals for detailed network configurations of Mobile FH over APN from different companies and finally agreed on two models based on the APN technology in accordance with the PoC Reference. The task force members continue to discuss whether the proposed solutions are promising and will clarify possible use cases and benefits. Subsequent PoC projects and evaluation will then be carried out.

For Step 2, the task force aims to discuss detailed use cases, operating-condition specifications, KPIs, and data to be collected and revise the PoC Reference toward the second half of FY2023.

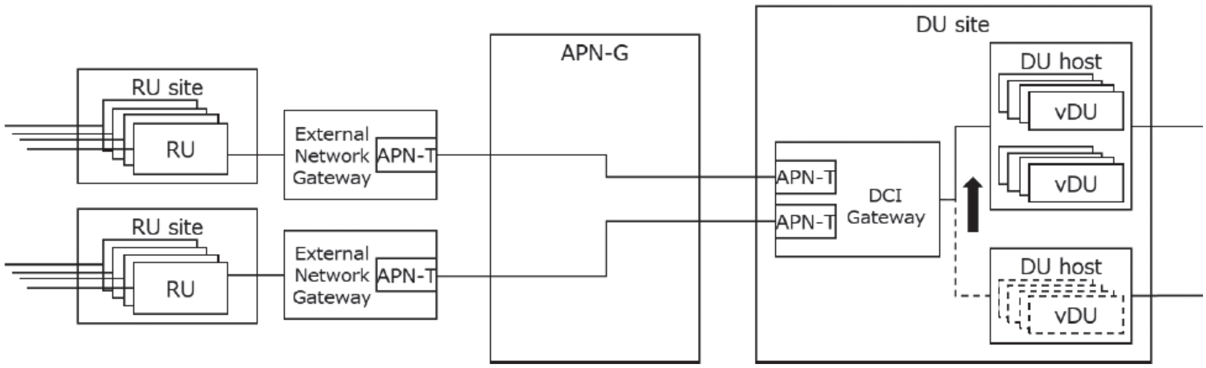


Fig. 7. Method based on packet switching.

4. NTT Group's activities in IMN-TF

NTT Group's activities in the IMN-TF are twofold. First, NTT laboratories are proposing to use the results of their research on coordinated optical wireless control technology [7]. Mobile systems, which constitute wireless sections, use a wired network for the transport network that establishes connections between base-station devices and connections with the core network. However, mobile systems and the transport network have conventionally been designed, constructed, and operated independently of each other. Therefore, they cannot operate as a single unit capable of dynamically changing the operation of the transport network in accordance with the state of the mobile systems. To provide services that require low latency in a mobile network, it is necessary to control and reduce latency through optical paths and DCI gateways on the basis of mobile information provided by network devices, such as DUs. The function for reducing latency is specified as CTI by O-RAN WG 4. NTT laboratories have extended the CTI specification to define an eCTI for the IOWN GF architecture and studied a method with which the APN controller reduces both end-to-end latency and jitter by obtaining mobile information from CUs/DUs through the eCTI. On the basis of this study, NTT laboratories are leading discussions for producing a proposal on a latency/jitter-reduction method and for demonstrating its feasibility.

Second, NTT DOCOMO has been operating and supporting the IMN-TF as a co-coordinator since April 2022. On the strength of its experience in the commercial deployment of Open RAN and in the promotion of vRAN through OREX, and through its communication with other task force members, NTT DOCOMO is leading efforts to ensure smooth progress of task force activities. Examples are formulating activity plans for the IMN-TF, setting goals and targets for each topic, proposing agendas and summarizing discussions held in member meetings, and inputting the trends of international standardization bodies such as 3GPP and O-RAN Alliance. The IMN-TF will continue to work on solving problems in the mobile domain, including the actualization of elastic load balancing.

References

[1] IOWN Global Forum, "Technical Outlook for Mobile Networks Using IOWN Technology," Ver. 1.0, 2022. https://iowngf.org/wp-content/uploads/formidable/21/IOWN-GF-RD-Technical-Outlook-for-Mobile-Networks-1.0-1.pdf

[2] IOWN Global Forum, "Technical Outlook for Mobile Networks Using IOWN Technology - Advanced Transport Network Technologies for Mobile Network," Ver. 1.0, 2023. https://iowngf.org/wp-content/uploads/formidable/21/IOWN-GF-RD-IMN-PHSE2-1.0.pdf

[3] 3GPP TR 38.801: Study on new radio access technology: Radio access architecture and interfaces, https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3056

[4] 3GPP TR 38.806: Study of separation of NR Control Plane (CP) and User Plane (UP) for split option 2, https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3307

[5] 3GPP TR 38.816: Study on Central Unit (CU) – Distributed Unit (DU) lower layer split for NR, https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3364

[6] IOWN Global Forum, "Mobile Fronthaul over APN PoC Reference," Ver. 1.0, 2022. https://iowngf.org/wp-content/uploads/formidable/21/IOWN-GF-RD-MFH_over_APN_PoC_Reference_1.0.pdf

[7] T. Hatano, Y. Sakai, K. Asaka, and T. Shimada, "Low-latency Function-dedicated Network Providing Sustained and Stable Services," NTT Technical Review, Vol. 21, No. 12, pp. 43–49, 2023. https://doi.org/10.53829/ntr202312ra2

Manabu Sugihara

Senior Manager, IOWN Development Office, Research and Development Planning Department, NTT Corporation.

He received an M.E. in informatics and computer engineering from the University of Electro-Communications, Tokyo in 1998 and joined NTT DOCOMO the same year. He was engaged in the development of wireless networks, voice/packet switching equipment, subscriber database systems, and NFVIs. He has been with IOWN Development Office in NTT since 2021 and is engaged in activities of the IOWN Global Forum IMN-TF.

Yoshitsugu Shimazu

Vice President of Innovative Technology Office, Technology Planning, NTT Corporation.

He received a B.E and M.E from Kyoto University in 1997 and 1999. He joined NTT DOCOMO in 1999 and has over 21 years of experience in development of RAN equipment for commercial network. He was awarded the Maejima Hisoka Award in 2017, and Commendation for Science and Technology by the Minister of Education, Culture, Sports, Science and Technology in 2020. He has been director of Innovative Technology Office, Technology Planning, NTT since 2020 and vice president since 2023. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) of Japan.

IOWN Data Hub to Become a Reality

Kunihiko Minamihata, Tomohiro Inoue, and Kazuyuki Takaya

Abstract

The IOWN Global Forum is working on the social implementation of the Innovative Optical and Wireless Network (IOWN) technology. This article introduces the progress of the storage service study in the IOWN Global Forum, focusing on the newly published IOWN Data Hub Functional Architecture version 2.0 and the Proof of Concept Reference and NTT laboratories' activities related to the IOWN Data Hub.

Keywords: storage, database, virtual data lake

1. Background of the need for IOWN Data Hub

Integrating the real and virtual worlds and feeding the results of simulations in the virtual world back to the real world in real time can provide many benefits, such as the ability to detect and quickly respond to hazards and maintain a comfortable environment with minimal changes. However, to achieve this integration, it is necessary to synchronize in real time the data and behavior of a vast number of objects sensed in the real world with those in cyberspace. Due to their high latency, current networks cannot synchronize vast amounts of data over a wide area in real time, making it difficult to construct a high-definition digital twin that covers a vast space. To solve this problem, we need the IOWN Data Hub (IDH), a new data-sharing platform built on the Open All-Photonic Network (APN) and data-centric infrastructure (DCI), which are both infrastructures of the Innovative Optical and Wireless Network (IOWN). This article focuses on the update of the IDH architecture document and proof of concept (PoC) activities undertaken to enable the IDH.

2. IDH activities at the IOWN Global Forum

In January 2022, IOWN Global Forum members, including NTT, Oracle, and NEC, collaborated to develop and release IOWN Data Hub Functional Architecture version 1.0. For the basic mechanism and concept of the IDH, refer to "Study of Storage Services at IOWN Global Forum" [1] that can be found in the NTT Technical Review. After the release of the above architecture document, the IOWN Global Forum held a series of discussions toward social implementation, and in July 2023, updated the document to release the IOWN Data Hub Functional Architecture version 2.0 [2]. The main updates are as follows:

(1) Specific description of use-case requirements: Identification of data types, bandwidth requirements, etc. for each use case presented in the Cyber-Physical System (CPS) and AI (artificial intelligence) Integrated Communications (AIC) use-case documents [3, 4].

(2) Identification of inherent gaps: Gap analysis based on more specific implementation models.

Some of the requirements for the use cases described in update (1) are listed in **Table 1**. These future use cases will need to meet diverse and demanding data processing and sharing requirements. For example, to meet real-time requirements, it is necessary to process data as-is without any conversion, close to the original data. To handle a large number of data flows, it is necessary to significantly improve distributed data management engines so that they can operate over ultra-high-speed, high-quality networks. Robust data security and data-usage control mechanisms will be required to enable fast and

Table 1. IOWN Global Forum use cases and high-level data requirements.

Use case	Expected requirements			
	Data type	Data producers and consumers	Data collection bandwidth	Producer-consumer latency
Area management security	• Video cameras • LiDAR sensors	Data Producers • Video cameras: 1,000 per monitored area (e.g., building) x 100–10,000 monitored areas Continuously sending data • LiDAR sensors: 1,000 per monitored area (e.g., building) x 100–10,000 monitored areas Continuously sending data Data Consumers • Security officers: Up to 1,000 per monitored area Receiving alerts when detected • Security monitors: Up to 50 per monitored area Displaying the selected cameras	• 45–60 Mbit/s per camera (assuming Full HD motion JPEG) • 600–800 cameras i.e., 27–48 Gbit/s per building • Thousands or millions of buildings (depending upon the required control)	From data collection through analysis to alerting: 1 sec, ideally 100 ms
Industry management Remote factory control	• Robot/drone's high-definition (8K) cameras • Robot/drone's sound/odor sensors	Data Producers • Machines and/or robots Up to 10,000 per factory Continuously sending data when they are in operation Data Consumers • Machines and/or robots Up to 10,000 per factory Receiving control signals when required • Factory Engineers Up to thousands Receiving alert signals and other information when required	• 2.35 Gbit/s per robot/drone	From data collection through analysis to feedback controls: 10 ms, sub-second, or a second (depending upon the use case scenario assumed)
Smart-grid management Renewable-energy optimization	• Sensor data from prosumer • External data sources (weather data, human flow data, etc.)	Data Producers • Prosumer devices (e.g., smart meters, electric vehicles, etc.) and power grid system equipment (e.g., switchgear, transformer, etc.) More than 100,000 Continuously sending data Data Consumers • Prosumer devices and power-grid-system equipment More than 100,000 Receiving control signals when required	• 24 Tbit/s (received from data producer) • 128 Pbit/s (data transmission from electric vehicles in 2030) • 3.2 Gbit/s (other data sources)	From data collection to display: 50 ms From data collection to analysis and feedback control: 200 ms

HD: high definition
JPEG: Image compression formats created by the Joint Photographic Experts Group.
LiDAR: light detection and ranging

reliable data exchange between multiple parties.

Update (2) identified gaps in the computing environment as well as problems in datacenters (DCs) and models for data sharing and processing models that make up the computing environment.

2.1 Fundamental problems in DCs

The current networks and DC technology have several fundamental problems in the following areas. These must be solved if we are to build IDH services that meet the use-case requirements (**Fig. 1**).

(1) Inter-DC network quality

The network quality between DCs used in today's cloud is focused on current use cases. Since packet reordering and loss occur in networks, the Transmission Control Protocol (TCP) is used more frequently than other protocols, such as the User Datagram Protocol (UDP), to ensure reliability. These networks do not provide sufficient performance for the workloads of geographically distributed databases and storage systems that are required by the use cases in the IOWN era.

(2) Intra-DC network quality

The quality of the networks used by hyperscale

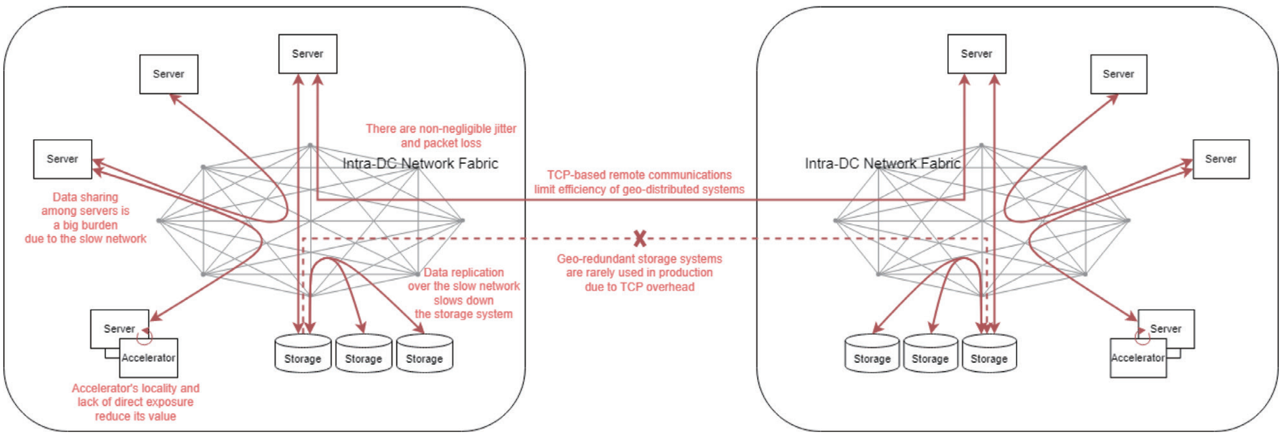


Fig. 1. Origins of gaps in today's technology landscape.

DCs that make up a cloud are not high enough. For example, one-way latency can exceed 1 ms, a latency sufficiently high to cause packet loss to exceed 1% during severe network congestion. The bandwidth is similarly unstable. This is essentially due to oversubscription, a scheme that allows reservation of bandwidth beyond the capacity of the network devices. Therefore, the network that connects computing resources in a DC cannot guarantee all the bandwidth allocated to individual resources.

(3) Storage performance

To ensure data persistence, data are typically synchronized across multiple servers. However, synchronization requires a large workload, rendering performance very unstable. Thus, storage systems that are connected via today's cloud-based storage area networks tend to be slower than directly connected storage devices. For example, the response time of storage systems can exceed dozens of milliseconds, making such systems unsuitable for typical database workloads or redundant storage systems.

(4) Data-sharing performance

In cases where data are processed cooperatively across multiple servers, a certain amount of data exchange is required between servers, for example, to redistribute data. This exchange significantly reduces overall performance. To speed up such data servers, some implementations connect servers using remote direct memory access (RDMA) fabric connections. However, since the use of RDMA in production environments imposes strict requirements, such as no packet loss and no packet reordering, RDMA is used only for very limited purposes in today's cloud environments.

(5) Accelerator usage

Accelerators, such as graphics processing units (GPUs) and field programmable gate arrays (FPGAs), have the potential to achieve high performance, cost-effectiveness, and energy efficiency at least 10 times greater than software-based processing on general-purpose central processing units (CPUs). However, such an accelerator must be deployed within each server and cannot directly access external clients for the purpose of streamlining data loading. Therefore, with current network and DC technologies, it is not possible for multiple general-purpose computing resources to share accelerator resources at high speed and with low overhead.

2.2 Today's implementation model and inherent gaps

Considering the aforementioned issues, we must accept the following constraints if we are to use today's technology to provide data services such as the IDH.

(1) Completion of data processing within one place

If processing efficiency could drop sharply due to poor quality of the network that connects DCs, it would be difficult to implement geographically distributed data services. Therefore, current data-processing services are designed to be implemented within a single DC or within nearby availability zones. In fact, today's clouds are an aggregation of such services. The size of cloud DCs has become so enormous that the number of available cloud regions is very limited, 30 or so at most worldwide. This means that, regardless of where data were generated, all data must be transferred to one of the cloud DCs and all data processing must be completed there, making the system's energy use highly inefficient.

(2) Asynchronous data replication as a foundation to build a scalable transactional distributed relational database

In building a scalable transactional relational database (RDB) system, it is a common practice to place replicated read-only data near each RDB server. Such data replication tends to be asynchronous because the network is slow. To reduce the load imposed on slow networks by data replication, often only the transaction or change logs, rather than the data, are propagated. This improves performance but requires reconstruction of table data from the log data on each server, which inevitably increases the system cost and power consumption.

(3) Sharded data processing to build a scalable key-value store, message broker, and/or analytical distributed RDB

Sharding is a technology for reducing response time and improving scalability by grouping related data together and distributing it across multiple databases. It is used to build scalable systems, such as a key-value store (KVS). However, it does not eliminate data transfer needed for data relocation and server-to-server communication. Rather, the volume of such data transfer and communication may increase in advanced digital transformation (DX) services that need to handle data from many different angles. Since sharded systems require a certain amount of data transfer between scale-out and scale-in operations, data services in the cloud today experience constraints on dynamically changing scalability.

(4) Usage of a cache layer

The constraints imposed by low-speed networks and storage systems often make it necessary to manage data on a process-by-process basis and rely heavily on asynchronous distributed processing. To streamline such implementations, various cached data management services are provided in today's cloud. However, such services increase end-to-end latency and cause the same data to be copied unnecessarily multiple times in the cloud, which is undesirable in terms of cost and energy consumption.

(5) Accelerator in use only in a laboratory

One possible approach to solving the aforementioned problems is to use accelerators, such as GPUs and FPGAs, to speed up queries and other data processing in distributed RDBs. Software has been developed and validated for such purposes, but in actual practice, the use of such software rarely improves and often diminishes performance and cost-effectiveness. This is mainly due to the need to preload data at the startup of the accelerator.

The IDH must be designed to address the fundamental and inherent challenges of today's cloud system and must be further developed to adapt to use cases of the wide-area and real-time CPSs that will be increasingly demanded in the future. To this end, it is necessary to verify the IDH iteratively in actual use cases and reflect the results to refine the required functions and design methodology of the IDH.

3. PoC activities on the IDH

In October 2022, the IOWN Global Forum formulated the IOWN Data Hub PoC Reference [5] to help implement the IDH. The use of the Open APN will result in a quantum leap in the speed of hybrid and multi-cloud connections, which in turn will cause a shift from concentration in hyperscalers to the use of enterprise DCs and edge clouds. Thus, it will be necessary to innovate data sharing architecture so that it can accommodate this shift. The IDH disaggregates the functions and processing that have previously been executed by hyperscalers in the central cloud and relocates them to edge clouds. To validate and deploy the IDH architecture, in which connections are set up by the Open APN and DCI, in many use cases, the IOWN Global Forum has released this PoC Reference and is seeking participation from a wide range of organizations, not limited to IOWN Global Forum members.

3.1 The IDH PoC Reference

This PoC Reference has been formulated to validate the IDH architecture that meets the stringent requirements of various DX services. Smart factory, smart grid, and metaverse have been selected as use-case examples and five PoC scenarios have been defined.

To build a metaverse service, for example, it is typically necessary to collect information about the participants, represent them as avatars in the virtual space, and allow the avatars to interact with each other. Thus, provision of a high-quality metaverse service requires the following.

- Collect motion data from up to millions of participants and reflect them in the movements of their avatars in cycles of dozens of milliseconds.
- Move avatars in the virtual space, which contains many virtual buildings and other structural elements, and enable them to communicate with each other.

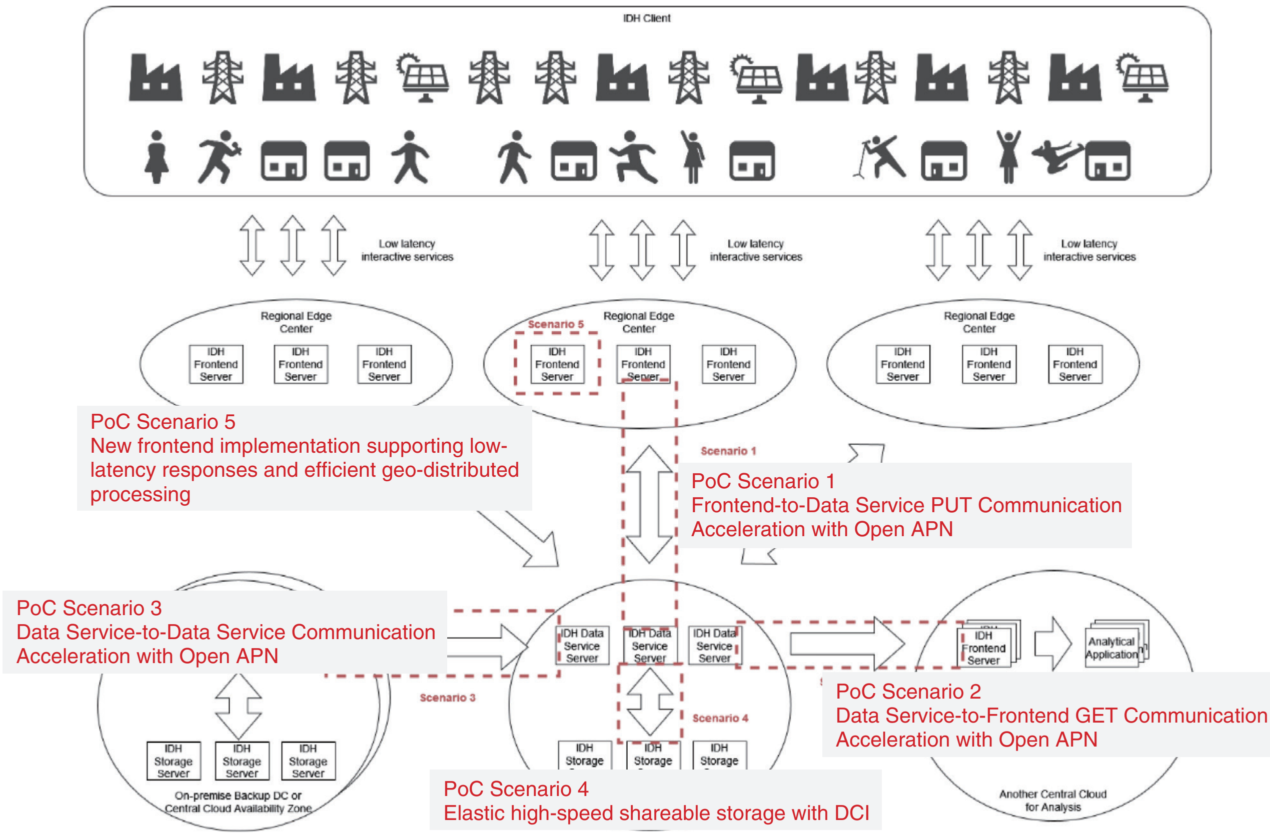


Fig. 2. Geo-distributed IDH adoption model and IDH PoC scenario mapping.

- Render virtual space scenery from each avatar's perspective with motion-to-photon latency of less than dozens of milliseconds.

To demonstrate the IDH architecture for supporting such use cases, the PoC Reference defines a test environment that reflects the geographically distributed deployment of assumed IDH services and identifies the items that PoC projects should validate (**Fig. 2**).

In this model, a subgroup of front-end servers of the IDH architecture is deployed at a regional edge center to provide low-latency services to connected Internet of Things (IoT) devices. The data service servers and storage servers, which are part of the IDH architecture, are deployed at remote DCs or the central cloud to support data persistence and usage.

On the basis of this model, the PoC Reference defines the following five PoC scenarios:

- Scenario 1: Frontend-to-Data Service PUT Communication Acceleration with Open APN
- Scenario 2: Data Service-to-Frontend GET Communication Acceleration with Open APN
- Scenario 3: Data Service-to-Data Service Communication Acceleration with Open APN
- Scenario 4: Elastic High-speed Shareable Storage with DCI
- Scenario 5: New frontend implementation supporting low-latency responses and efficient geo-distributed processing

For each scenario, the PoC Reference describes an overview, required functions, variable conditions, and expected benchmarks to enable the readers to proceed with their PoC projects (**Table 2**).

IOWN Global Forum member companies are currently conducting PoC projects for their respective scenarios on the basis of this PoC Reference. Thus, the IDH is making steady progress toward its social implementation.

4. Activities in the NTT Group

To contribute to the implementation of the IDH, the NTT Software Innovation Center is developing virtual data lake, which is defined as a set of functions in the IDH reference architecture.

Table 2. Benchmark requirements for each scenario.

	Selected features	Variable conditions	Expected benchmark
PoC Scenario 1 Frontend-to-Data Service PUT Communication Acceleration with Open APN	• PUT Communications over Open APN • Data preprocessing before data transfer	• Effect of the network type, latency, jitter, bandwidth, packet loss, and packet order change • Network-protocol parameters • Data-object size and presence/absence of data compression	• Throughput per unit system resource • Energy consumption per unit system resource
PoC Scenario 2 Data Service-to-Frontend GET Communication Acceleration with Open APN	• Federation function in cloud DC • Get Communications over Open APN • Data filtering and preprocessing in the on-premise DC	• Effect of the network type, latency, jitter, and bandwidth • Network-protocol parameters • Data-object size • Ratio of relevant data	• Time for required data transfer
PoC Scenario 3 Data Service-to-Data Service Communication Acceleration with Open APN	• Inter-server communications over Open APN • Continuous data ingestion and data • Pull-access of data • Push-delivery of data • Synchronous data replication across multiple data-service servers of message broker, distributed RDB, and KVS • Scalability	• Effect of the network type, latency, jitter, and bandwidth • Effect of the transaction size • Effect of the data-record size	• Data-hub response time and jitter • Throughput per unit system resource • Energy consumption per unit system resource • Scaling factor
PoC Scenario 4 Elastic High-speed Shareable Storage with DCI	• Establishment of communication paths between data-service servers and storage servers, which may go through the gateway server, made by DCI • Usage of standard protocols such as iSCSI • Shared storage • Scalability	• Effect of the network type, latency, jitter, and bandwidth • Effect of the data-block size • Random/sequential access • Ratio of read/write operations • Storage system topology	• Storage-system response time • Throughput per unit system resource • Energy consumption per unit system resource • Scaling factor • Data-hub performance
PoC Scenario 5 New frontend implementation supporting low-latency responses and efficient geo-distributed processing	• Data Ingestion through Open APN • Convergence of message broker and other data hub services • In-memory-based data processing • Memory-space sharing across message-broker (buffering) role and other data-hub roles • Continuous data ingestion and subsequent data processing • Resource controls (both test environments) • Scalability	• Effect of the network type, latency, jitter, and bandwidth • Effect of the transaction size • Effect of the data-record size	• End-to-end response time and jitter • Throughput per unit system resource • Energy consumption per unit system resource

There are various stakeholders in both the real world and virtual world, which is built by Digital Twin Computing. To enable the feedback between the two worlds to loop in real time, data must not only be transmitted at high speed but also be securely exchanged between stakeholders who have different standpoints. Therefore, it is essential to enable stakeholders to handle the latest data as if they were at hand, while ensuring permanent governance of the shared data.

The virtual data lake being developed by NTT Software Innovation Center virtually aggregates and centralizes omnipresent data that are managed by different organizations and geographically dispersed across multiple locations, enabling data users to acquire and use only the needed data efficiently and on-demand. For this purpose, the virtual data lake provides functions that enable data users to search for and discover the data they need from among a vast amount of data based on metadata (semantic and format information of data) and functions that maintain governance by enabling data users to display and use only those items of data they are permitted to use based on the policies set by data providers. These functions enable organizations to mutually use a large volume of widely varied data quickly, easily, securely, and reliably across organizational boundaries.

To make the IDH platform fit for practical use, we will continue to conduct demonstration tests of the platform by referring to the IDH PoC Reference and other reference materials.

5. Future development

As a "database is a network" platform that supports Digital Twin Computing that runs on the Open APN and DCI, the IDH is expected to be used for mission-critical use cases such as wide-area automated driving. Together with many partners, NTT will verify various IDH implementation models that satisfy the requirements imposed by use cases in the IOWN era and promote its social implementation.

References

[1] T. Inoue, "Study of Storage Services at IOWN Global Forum," NTT Technical Review, Vol. 20, No. 5, pp. 32–36, 2022. https://doi.org/10.53829/ntr202205fa4

[2] IOWN Global Forum, "Data-Centric Infrastructure Functional Architecture," Ver. 2.0, 2023. https://iowngf.org/wp-content/uploads/formidable/21/IOWN-GF-RD-Data_Hub_Functional_Architecture-2.0.pdf

[3] IOWN Global Forum, "Cyber-Physical System Use Case Release-1," Ver. 1.0, 2021. https://iowngf.org/wp-content/uploads/formidable/21/IOWN-GF-RD-CPS_Use_Case_1.0.pdf

[4] IOWN Global Forum, "AI-Integrated Communications Use Case Release-1," Ver. 1.0, 2021. https://iowngf.org/wp-content/uploads/formidable/21/IOWN-GF-RD-AIC_Use_Case_1.0.pdf

[5] IOWN Global Forum, "IOWN Data Hub PoC Reference," Ver. 1.0, 2022. https://iowngf.org/wp-content/uploads/formidable/21/IOWN-GF-RD-IDH-PoC-Reference_1.0.pdf

Kunihiko Minamihata

Senior Manager, IOWN Development Office, R&D Planning Department, NTT Corporation.

He received a B.E. and M.E. in information and knowledge engineering from Shizuoka University in 1998 and 2000. He then joined NTT WEST Corporation in 2000 as a system and security engineer for corporate business. He is engaged in spreading IOWN technologies and supporting IOWN Global Forum operations in his current position.

Tomohiro Inoue

Director, Data Sharing Infrastructure Project, NTT Software Innovation Center.

He received a B.E. and M.E. from the University of Tokyo in 1999 and 2001. He joined NTT in 2001 and is currently engaged in the research and development of data hubs, trusted data-sharing infrastructure, and global data-space interconnection as the leader of Information Sharing Architecture Group.

Kazuyuki Takaya

Senior Research Engineer, Supervisor, Social Information Sharing Research Project, NTT Social Informatics Laboratories.

He received an M.E. from Waseda University, Tokyo, and joined NTT in 2000. He is engaged in the research and development of technologies and platforms for data sharing.

DCI Architecture Promoted by the IOWN Global Forum

Hiroshi Yoshida

Abstract

The IOWN Global Forum has proposed an information and communication technology (ICT) infrastructure, called data-centric infrastructure (DCI), that leverages the high-speed and low-latency features of the Innovative Optical and Wireless Network All-Photonics Network (IOWN APN). This article describes the DCI on the basis of the Data-Centric Infrastructure Functional Architecture version 2.0 released by the forum in March 2023 and explains the advantages of the DCI in tackling challenges in scalability, performance, and power consumption that confront current ICT infrastructures. It is hoped that this article will help readers deepen their understanding of the DCI.

Keywords: data-centric, IOWN Global Forum, DCI

1. Introduction

An article entitled "Data-centric Infrastructure for Supporting Data Processing in the IOWN Era and Its Proof of Concept" [1], in the January 2024 issue of this journal, presented an overview of the data-processing infrastructure and accelerator in the data-centric infrastructure (DCI), a new information and communication technology (ICT) infrastructure proposed by the forum, and described its proof of concept activities using video analysis as a use case.

As a follow up to the above issue, this article explains the overall picture of the DCI with a focus on the Data-Centric Infrastructure Functional Architecture version 2.0 [2] released by the IOWN Global Forum in March 2023. We hope that readers will peruse this article together with the Feature Articles on efforts toward the early deployment of IOWN in the January 2024 issue to deepen their understanding of the DCI.

2. Challenges confronting current ICT infrastructures

2.1 Challenges in scalability

ICT infrastructures need to support various types of data processing with different requirements. For example, online transaction processing needs to respond to a large number of inquiries while data batch processing must handle vast amounts of data. The data-processing requirements of ICT infrastructures can also change rapidly over time. For example, an event, such as a sale, generates a large number of access attempts to an Internet shopping site, requiring ICT infrastructures to respond to these changes within minutes or hours. Therefore, ICT infrastructures must have sufficient scalability to flexibly respond to such changes in demand.

2.2 Challenges in performance

Some types of data processing, such as communicating the movements of individual participants in a virtual space to the people around them or processing high-speed transactions in financial applications, must meet stringent response-time requirements. Current ICT infrastructures cannot adequately satisfy such exacting requirements, especially in data transfer.

2.3 Challenges in energy consumption

Today's ICT infrastructures have various bottlenecks. For example, if data transfer is a bottleneck, the central processing unit (CPU) consumes most of its processing power simply waiting for data transfer.

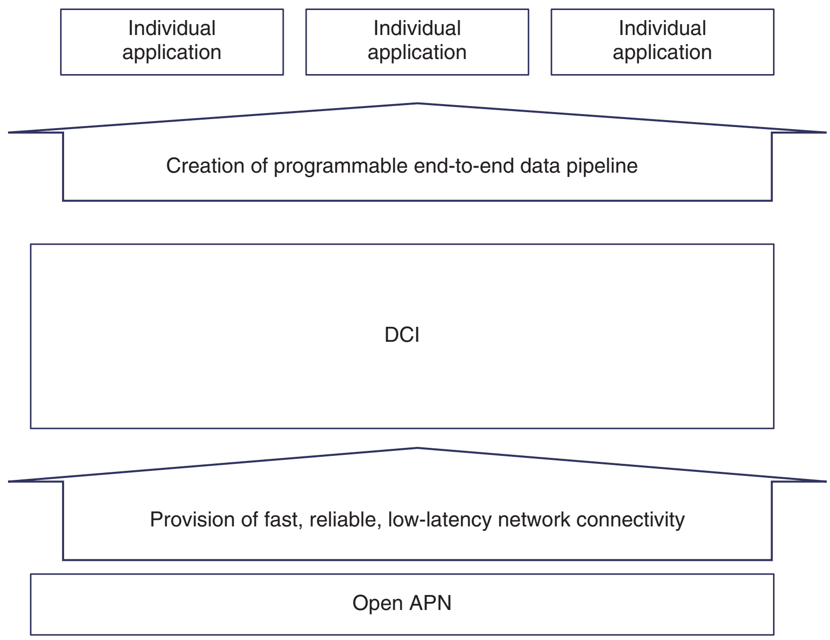


Fig. 1. Overview of ICT infrastructure centered on DCI.

It has become common to allocate processing to accelerators, which are adept in areas that are different from those of the CPU, such as graphics processing units (GPUs), which specialize in image processing to achieve high-speed processing, and data processing units (DPUs), which specialize in specific types of computation such as network processing and security processing. Efficient use and appropriate role sharing of computing resources can optimize the power efficiency of the entire ICT infrastructure.

3. The DCI

The DCI is a new ICT infrastructure proposed by the IOWN Global Forum to address these challenges. **Figure 1** gives an overview of the ICT infrastructure centered on the DCI. In the hierarchical structure of this ICT infrastructure, the Open All-Photonic Network (APN) is located at the bottom and applications at the top. Leveraging the high-capacity, low-latency network environment of the Open APN, the DCI provides programmable end-to-end data pipelines to help applications perform.

This article first explains the concepts of the Open APN and data pipeline then describes how the DCI helps applications perform.

3.1 What is the Open APN?

The Open APN is a high-speed network that consists entirely of optical communication. It is characterized by its ability to provide a high-capacity network with low latency, particularly with predictable latency. Therefore, it can transfer data between various devices connected to the network as if these devices were close to each other, regardless of their locations and the distances between them. More information on the Open APN and NTT's related activities are given in the article, "Activities for Detailing the Architecture of the Open APN and Promoting Its Practical Application," in this issue [3].

3.2 What is a data pipeline?

A data pipeline is a uniform stream that executes functions, such as acquisition, processing, conversion, and presentation of the data required by various applications. **Figure 2** shows an example of a data pipeline for a system that aggregates and analyzes surveillance-camera footage on the basis of metadata [4]. In this example, the footage from a surveillance camera is sent to an analysis server, which analyzes information contained in the metadata, such as information about what is captured in the video. The metadata are sent to the analysis server for the purpose of data aggregation, and raw video data are stored in storage units. Using a terminal, analysts can analyze

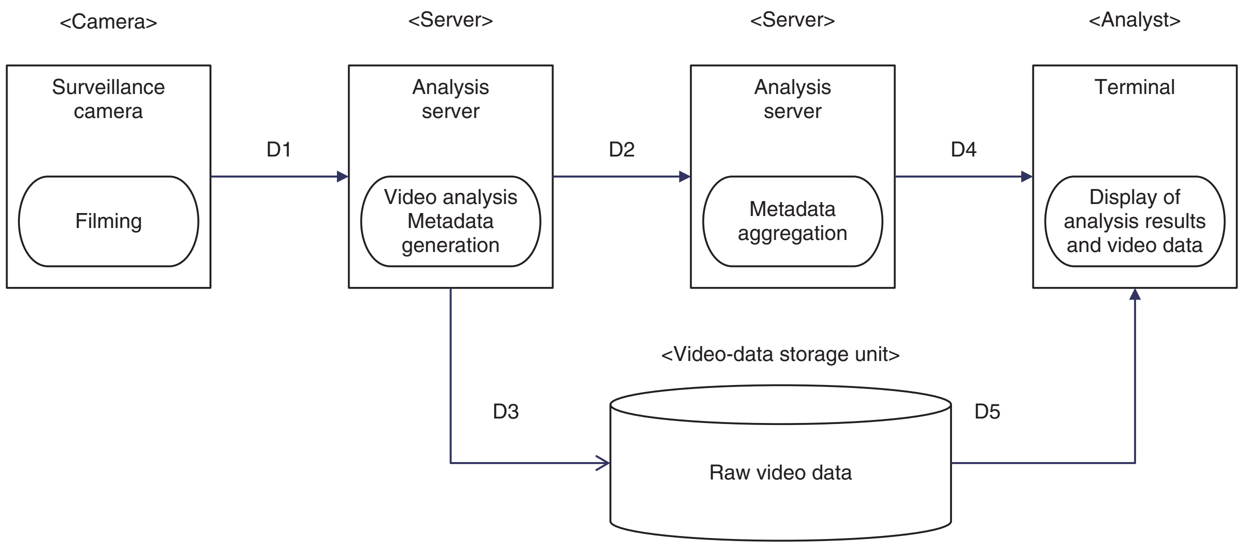


Fig. 2. Example of a data pipeline of a system that aggregates and analyzes surveillance-camera footage on the basis of the metadata.

data on the basis of the metadata and refer to the actual footage data in the storage units. In Fig. 2, each oval represents data processing, i.e., a set of autonomous actions, such as data collection and analysis. A process is executed by a functional node, represented with a rectangle, which exists outside the process. A database/storage, represented with a drum, is typically a file system or database. It is a function for storing data. Connecting these are data flows, which are functions to transfer data to data processing or storage. The IOWN Global Forum has defined this representation of a data pipeline as a data-pipeline diagram [4]. Therefore, a data pipeline is an abstraction of resources used to enable an application to function.

4. Various DCI resources

How does the DCI create a data pipeline? Before answering this, let me explain how the DCI manages various resources.

4.1 Functional cards, physical nodes, and DCI clusters

A schematic diagram of the various resources managed by the DCI is shown in **Fig. 3**.

As mentioned above, data processing refers to the function of collecting or analyzing data in the data pipeline and is typically implemented using CPUs. GPUs, which specialize in image processing to achieve high-speed processing, and DPUs, which specialize in specific types of computation, such as network processing and security processing, have been attracting attention. Storage is implemented using solid-state drives (SSDs) and hard disk drives.

In the DCI, data processing and storage are classified into two categories: host board and functional card. A host board typically consists of a CPU and memory unit. It corresponds to a motherboard in a typical server. Functional cards are other components, such as GPUs, DPUs, network interfaces, and storage units. A functional card corresponds to an expansion board in a general server, such as a network interface controller (NIC) or a graphics card, or to a storage unit such as an SSD. A DCI physical node consists of a host board and multiple functional cards. An intra-node interconnect connects the host board and functional cards within a DCI physical node. Peripheral Component Interconnect Express (PCIe), which is widely used in servers, is an example of an intra-node interconnect. However, more flexible standards, such as Compute Express Link (CXL), are being studied.

A DCI cluster is a computing infrastructure consisting of multiple physical nodes. It also includes the characteristic mechanisms of an inter-node interconnect and DCI gateway.

The inter-node interconnect is a network that connects DCI physical nodes. Although general clusters also have a network that connects servers that make

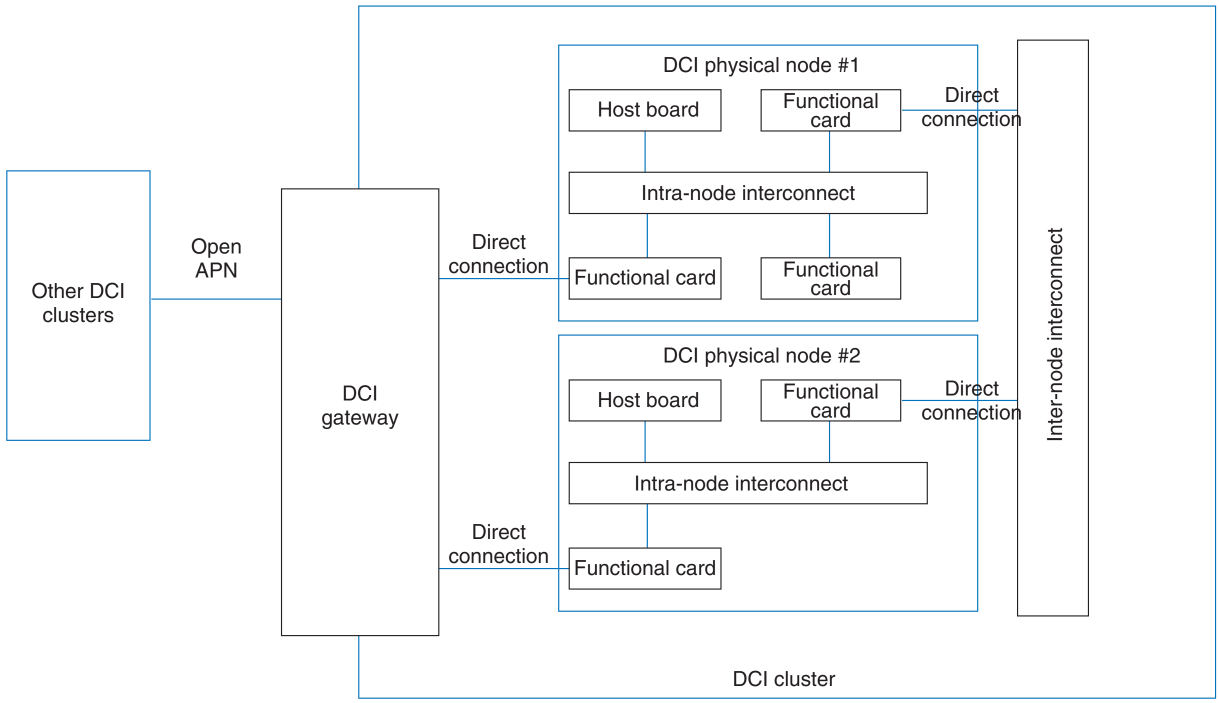


Fig. 3. Schematic of various resources managed by the DCI.

up the cluster, the inter-node interconnect differs from such a network in that some functional cards can directly access inter-node interconnects. Functional cards that can directly access inter-node interconnects are called "network-capable" functional cards. Such a card can directly transfer data to and from functional cards and host boards residing in other physical nodes without going through the intra-node interconnect of the physical node to which it belongs. An NIC is of course a network-capable functional card. It is assumed that some GPUs and storage units and some functional cards are also network-capable.

The DCI gateway interconnects DCI clusters. It is connected to other DCI clusters via the Open APN. Network-capable functional cards can also directly access the DCI gateway and transfer data to and from functional cards belonging to other DCI clusters via the Open APN at high speed.

4.2 Logical service node and data-pipeline implementation

How a DCI cluster can create a data pipeline for an application is explained. To create a data pipeline, a DCI cluster selects functional cards and host boards within itself and builds a logical service node (LSN) from them.

If an LSN consists of elements within a single physical node, all data flows in a data pipeline can be transferred via the intra-node interconnect in that physical node. However, if the LSN spans multiple physical nodes, an inter-node interconnect is required to transfer the data flow. A data pipeline can be created using multiple LSNs that belong to different DCI clusters. In this case, data flows across LSNs must go through a DCI gateway.

Figure 4 shows an example of a data-pipeline configuration. This DCI cluster has selected the functional cards shown in yellow, i.e., Functional cards A, B, and C, to configure the LSN. Not being network-capable, Functional Card A cannot directly connect to the inter-node interconnect. Consequently, it connects to the inter-node interconnect via the intra-node interconnect and Functional Card X.

Functional Card B of DCI physical node #2, however, is network-capable. Therefore, the inter-node interconnect can be directly connected to Functional Card B. Functional Card C is also network-capable and is connected to the DCI gateway. This means Functional Card C can connect to another LSN in another DCI cluster via the DCI gateway.

As described above, a feature of the DCI is the ability to configure flexible data pipelines across DCI

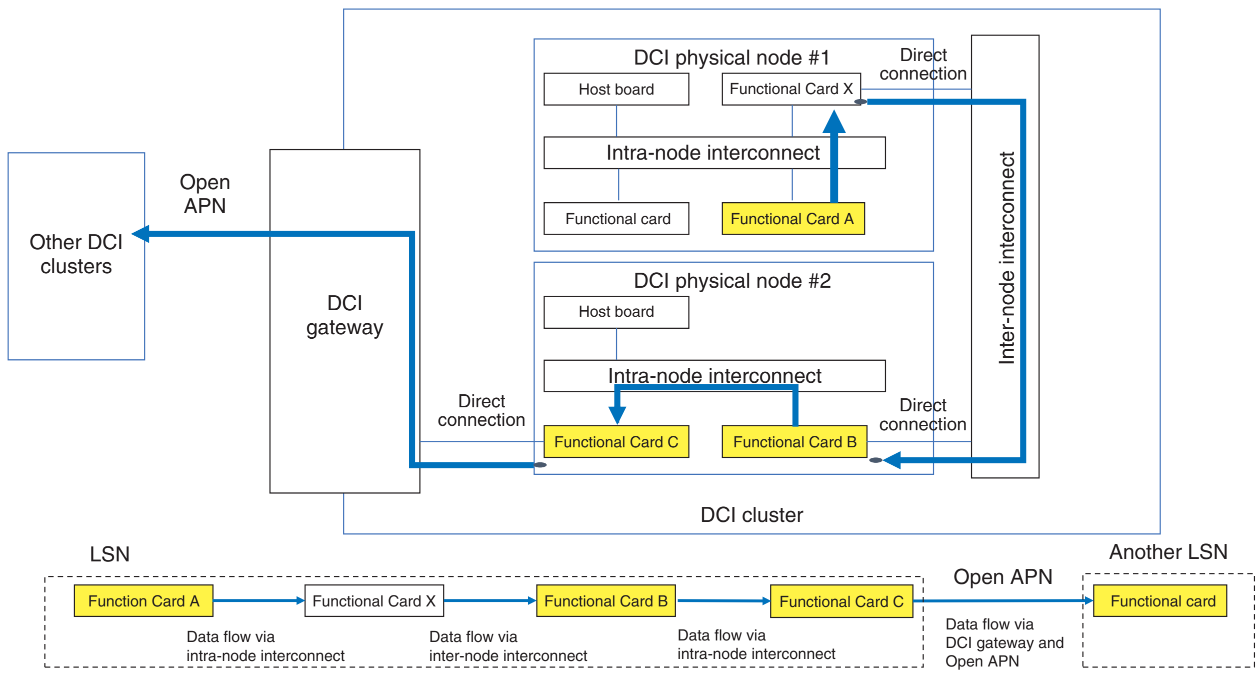


Fig. 4. Example of data-pipeline configuration.

physical nodes and DCI clusters.

5. Challenges with the DCI and solutions

How the DCI creates a data pipeline has been described. The following describes how the DCI solves the challenges in scalability, performance, and power consumption mentioned earlier.

5.1 Addressing challenges in scalability

In conventional ICT infrastructures, computing resources and wide-area networking resources for data transfer have evolved independently, making flexible scaling across physical nodes and clusters difficult. With the DCI, however, network-capable functional cards are directly interconnected via inter-node interconnects, thus eliminating data-transfer bottlenecks. LSNs in different DCI clusters can also be easily integrated via DCI gateways. The integration of resources across physical nodes and clusters in this way makes for high scalability.

5.2 Addressing challenges in performance

With the DCI, network-capable functional cards are directly interconnected via inter-node interconnects. Thus, data can be transferred between them at extremely high speed without going through a CPU. Because data can also be transferred across DCI clusters at high speed via the Open APN, applications can have sufficient response performance with stringent response-time requirements.

5.3 Addressing challenges in energy consumption

With the DCI, because there are few bottlenecks in data transfer, CPUs are not kept waiting needlessly for data transfer. Accelerators, such as GPUs and DPUs, can also be selected flexibly depending on the computing requirements. This improves computing-resource utilization and makes it possible to optimize overall system energy consumption.

6. Key Values and Technology Evolution Roadmap and future development of the DCI

Finally, how DCI-related technologies will evolve on the roadmap is described by referring to the Key Values and Technology Evolution Roadmap [5] released by the forum in August 2023.

The current unit used for computing resource allocation management is a server. In Phase 2, the unit will be decreased in granularity to an element, such as a host board or functional card. In Phase 2 as well, network-capable functional cards will be able to directly transfer data between GPUs and DPUs.

Table 1. IOWN computing development roadmap.

Element	Phase 1 (now)	Phase 2 (3–5 years from now)	Phase 3 (by 2030)
Computing resource allocation management unit	Server level	Computing device level	Computing device level
Energy efficiency	Measurements made. Independent of workload, event, and traffic	Closely aligned with accurately measured and controlled workload, event, and traffic	We aim for significant reduction compared to Phase 2.
Data transfer	Supported by host board	Direct connection between XPUs (GPU, DPU)	Direct connection between XPUs (GPU, DPU)
Cache coherence memory space	COTS scale	Rack scale (via CXL)	Rack scale (via CXL)
RDMA	Application within datacenter	Application across datacenters (by RDMA over APN)	Application across datacenters (by RDMA over APN)
Logical server configuration	Server level	Module card level (PCIe card, CXL card, etc.)	Chip level
Security	Current cipher (e.g., RSA, DH)	Post-quantum cryptography	In addition to post-quantum cryptography, its management using zero-trust concept
Computing resource configuration time	Several hours	Several minutes	Several seconds
Applications that require very tight time constraints and deterministic processing times	Unfeasible	Processing time can be guaranteed deterministically at the millisecond level.	Processing time can be guaranteed deterministically at the millisecond level.
Storage/data transfer	Transfer at local level	Transfer at regional level	Transfer at national level

COTS: commercial off-the-shelf
DH: Diffie–Hellman key exchange
RDMA: remote direct memory access
RSA: Rivest–Shamir–Adleman

The time needed for computing resource configuration is currently in hours, but the use of LSNs will reduce the configuration time to minutes in Phase 2 and seconds in Phase 3 (**Table 1**).

7. Conclusions and summary

This article presented an overview of the DCI functional architecture.

For IOWN to enable advanced information technology to unleash its full potential in a data-driven society, a new ICT infrastructure that integrates the computing and networking infrastructures is required under the condition that a high-speed network with managed quality of service is available. IOWN's DCI serves as this new ICT infrastructure.

Due to limited space, this article does not describe some of the important concepts that support the DCI, such as infrastructure orchestrator, function-dedicated network, and resource pool. For these concepts, refer to the DCI Functional Architecture version 2.0 released by the IOWN Global Forum.

References

[1] R. Kurebayashi, T. Ishizaki, S. Priyankara, C. Schumacher, and S. Mizuno, "Data-centric Infrastructure for Supporting Data Processing in the IOWN Era and Its Proof of Concept," NTT Technical Review, Vol. 22, No. 1, pp. 64–72, Jan. 2024. https://doi.org/10.53829/ntr202401fa8

[2] IOWN Global Forum, "Data-Centric Infrastructure Functional Architecture," Ver. 2.0, Mar. 2023. https://iowngf.org/wp-content/uploads/2023/04/IOWN-GF-RD-DCI_Functional_Architecture-2.0.pdf

[3] Y. Kainuma, T. Takeda, J. Kani, and H. Nishizawa, "Activities for Detailing the Architecture of the Open APN and Promoting Its Practical Application," NTT Technical Review, Vol. 22, No. 2, pp. 19–25, Feb. 2024. https://ntt-review.jp/archive/ntttechnical.php?contents=ntr202402fa2.html

[4] IOWN Global Forum, "Reference Implementation Model (RIM) for the Area Management Security Use Case," Ver. 1, p. 64, Figure B.2-2: An Example of Data Pipeline Diagram (Big Mail System), Jan. 2022. https://iowngf.org/wp-content/uploads/formidable/21/IOWN-GF-RD-RIM-for-AM-Use-Case-1.0.pdf

[5] IOWN Global Forum, "IOWN Global Forum: Key Values and Technology Evolution Roadmap," https://iowngf.org/wp-content/uploads/formidable/21/IOWN-GF-RD-KVaTE_Roadmap-1.0-1.pdf

Hiroshi Yoshida
Senior Manager, IOWN Development Office, Research and Development Planning Department, Research and Development Market Strategy Division, NTT Corporation.
He received a Ph.D. from the Department of Human and Engineered Environmental Studies, the University of Tokyo, in 2011. Since joining NTT, he has been engaged in operational research, including network operation systems development and basic research on optimal staffing and safe work environments. He is currently in charge of arranging PoCs with other IOWN Global Forum member companies.

Standardization Trends in IEEE 802.11be/bn

Akira Kishida

Abstract

We introduce the latest trends in the Institute of Electrical and Electronics Engineers (IEEE) 802.11be (11be), which will be the main standard of the next generation of wireless local area networks (LANs), and the IEEE 802.11bn (11bn), the following primary standard of 11be. The IEEE 802.11 standard is the most used unlicensed wireless LAN system globally, recognized as Wi-Fi. Wi-Fi is implemented in laptop computers and smartphones, Internet of Things devices, and home appliances. Its usage models also extend from consumer markets to business solutions such as industrial automation and intelligent agriculture in the 11bn era.

Keywords: IEEE 802.11bn, IEEE 802.11be, Wi-Fi 7

1. Standardization of IEEE 802.11 WG

The Institute of Electrical and Electronics Engineers (IEEE) 802.11 Working Group (WG) [1] is a group that discusses and decides the technical specifications of the IEEE 802.11 wireless local area network (LAN) systems, recognized as Wi-Fi*. The WG is a subsidiary of the IEEE Standards Association (SA) that covers standardization activities of IEEE. The IEEE 802.11 WG consists of several subgroups, such as Task Groups (TGs), which create technical specifications, and Study Groups (SGs), which clarify the scope and target of TGs. Other subgroup types are Interest Topic Groups (TIGs), which discuss the possibility of standardization for specific topics, Standing Committees (SCs), which discuss general topics of wireless LAN systems, and Ad-hoc groups (AHGs) that address global issues. The technical specification defined by TGs is to be released as the IEEE 802.11xx. Usually, "xx" is a combination of alphabet. Hereafter, we assume that 11xx indicates the technical specification of the IEEE 802.11xx. **Table 1** lists each subgroup of the IEEE 802.11 WG and its topic as of December 2023.

2. Certification of wireless LAN devices by Wi-Fi Alliance

Wi-Fi Alliance [2] is an industrial organization that certifies the interoperability of devices and promotes marketing for wireless LAN systems and devices. Wi-Fi Alliance gives certification of a wireless LAN device on the basis of the IEEE 802.11 standard as Wi-Fi X, which passed the inspection of the test plan on the basis of the corresponding IEEE 802.11 standard. Here, X denotes the corresponding Wi-Fi generation number. For example, a wireless LAN device with Wi-Fi 6/6E certification means that the device has passed the Wi-Fi test plan of the 11ax. A test plan for Wi-Fi 7 is currently being discussed on the basis of the 11be draft standard.

3. IEEE 802.11be: Extremely High Throughput

TGbe is discussing and defining the 11be standard as the main technical specification next to 11ax. The project name of creating 11be is Extremely High Throughput (EHT), and the aims are enabling at least one mode of operation capable of supporting a maximum throughput of at least 30 Gbit/s at the media

* Wi-Fi is a registered trademark of Wi-Fi Alliance.

Table 1. Subgroups of the IEEE 802.11 WG.

Subgroup	Project topic	Completion date
TGbb	Light Communications	June 2023
TGbe	Extremely High Throughput	December 2024
TGbf	Wireless LAN Sensing	March 2025
TGbh	Randomized and Changing MAC Addresses	September 2024
TGbi	Enhanced Data Privacy	March 2026
TGbk	320 MHz Positioning	November 2024
TGme	802.11 Accumulated Maintenance Changes	September 2024
TGbn	Ultra High Reliability	May 2028
AMP SG	Ambient Power	-
AIML TIG	Artificial Intelligence / Machine Learning	-
ARC SC	Architecture	-
COEX SC	Coexistence	-
JTC1 SC	Support Liaison with Joint Technical Committee1 / SC6	-
WNG SC	Wireless LAN Next Generation	-
ITU AHG	International Telecommunication Union Liaison	-

access control (MAC) layer level and defining at least one mode of operation capable of improved worst-case latency and jitter. As well as 11ax, 11be will use the frequency band of 6 GHz in addition to those of 2.4 and 5 GHz. In Japan, a part of the 6 GHz band (5925–6425 MHz) was made available from September 2, 2022.

TGbe was established in May 2019 and will end their standardization activities in December 2024. Draft standard 5.0 was released in December 2023, and almost all specifications of various features are settled. Detailed functions are under review per comments from IEEE 802.11 WG members. The editorial review called Mandatory Draft Review (MDR), which is required to issue the formal IEEE standard, has begun. A ratification vote called Sponsor Ballot (SB) at the IEEE SA level will begin in January 2024. Thus, the status of the standardization process of 11be is in the final phase to be published.

3.1 Main technical features of IEEE 802.11be

New features for improving frequency-utilization efficiency, wider frequency bandwidth, and latency reduction will be defined in the 11be. Multi-link operation is a new feature that was never defined in the old standard before 11ax. Combining these features enables a 2.4 times improvement of the maximum data rate at the physical (PHY) layer from 9.6 Gbit/s (11ax) to 23 Gbit/s (11be). Applying the multi-link operation that enables simultaneous transmissions will enhance throughput at the MAC layer, and 11be aims to achieve a maximum throughput of at least 30 Gbit/s by using this feature. **Table 2** lists the effect of throughput enhancement of 11be compared with 11ax.

(1) Features of frequency-utilization efficiency and wider frequency bandwidth

The 11be standard defines 4096 quadrature amplitude modulation (QAM) as a modulation scheme of each subcarrier of the orthogonal frequency division multiplexing (OFDM) symbol, which achieves a 1.2 times higher modulation rate than 11ax (1024 QAM). In 11ax, a resource unit (RU) that indicates grouped subcarriers in a channel bandwidth of orthogonal frequency division multiple access (OFDMA), is assigned to each station (STA). Multiple RU (MRU) defined in 11be, however, enables the assignment of multiple RUs to an STA. This feature establishes several patterns of RU combinations, allowing for more flexible resource allocation.

(2) Multi-link operation

The 11be standard specified a new feature defined as multi-link operation that was never defined in the old standard before 11ax. In multi-link operation, a wireless device defined as a multi-link device (MLD) can support more than one affiliated STA or access point (AP) and operate using one or more affiliated STAs or APs. This feature enables an AP/STA to bundle multiple transmission paths (links), thus achieves a higher transmission rate in accordance with the number of links. **Figure 1** shows a conceptual image of multi-link operation.

Table 2. Effect of throughput enhancement of 11be.

	Functions	IEEE 802.11ax	IEEE 802.11be	Enhancement from 11ax
PHY layer	Maximum modulation level	1024 QAM (10 bit/subcarrier)	4096 QAM (12 bit/subcarrier)	1.2 times
	Maximum channel bandwidth	160 MHz	320 MHz	2 times
	Maximum number of spatial stream	8	8	Same
	Maximum data rate	9.6 Gbit/s	23 Gbit/s (without multi-link operation)	2.4 times
MAC layer	Multi-link operation	None	Maximum 15 links	In accordance with the number of links

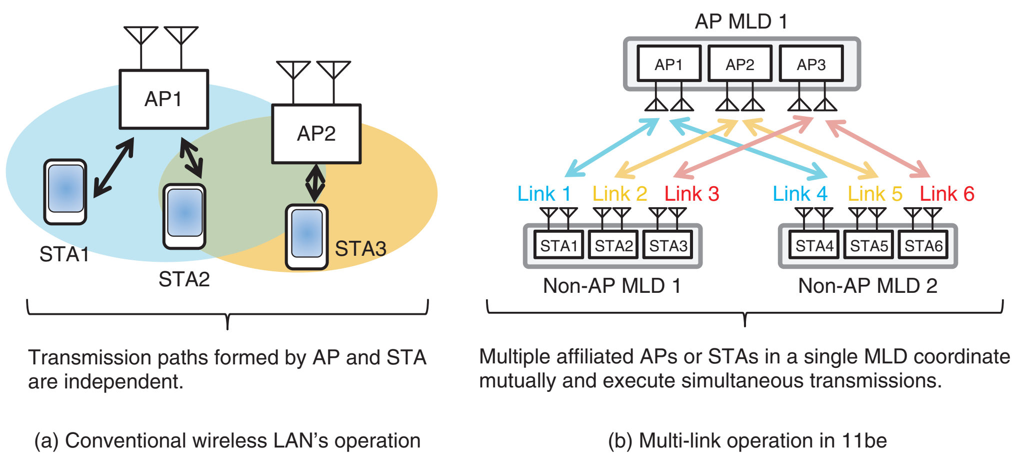


Fig. 1. Concept of multi-link operation.

(3) Feature of latency reduction

In 11be, the latency-reduction feature defined as restricted target wake time (TWT) will be supported. This feature ensures a specific period within the scheduled service period (SP) for latency-sensitive traffic. Therefore, contention between latency-sensitive traffic and non-latency sensitive traffic is reduced, and latency-sensitive traffic generated periodically can be transmitted more reliably.

(4) Other new features

A new peer-to-peer (P2P) feature without setting up any connections between STAs will be defined. In this new P2P feature, called triggered transmission opportunity (TXOP) sharing mode 2, an AP obtains time for P2P transmission representatively, and the designated STA can transmit data frames to the other STA directly. Emergency preparedness communications service (EPCS) will also be defined in 11be. This feature can prioritize the transmission of specific STAs authorized by networks in advance.

3.2 Standardization activities of NTT for 11be

The predecessors of TGbe were EHT TIG and EHT SG, which were launched in 2018 (TGbe was formed in 2019). The focus was initially only on improving maximum throughput, and reducing latency was not included in the scope of 11be. In the subgroup called Real Time Application (RTA) TIG, NTT and several companies discussed and considered latency and jitter reduction for industrial automation, cloud gaming, and so on around the same time as EHT TIG and SG. NTT collaborated with the RTA TIG's leading contributors to create a report document [3] that summarizes RTA TIG's activities. We input this document into EHT SG and emphasized the importance of latency-reduction features when EHT SG created the Project Authorization Request (PAR), which defines the scope and necessity of the standard, and Criteria for Standards Development (CSD), which describes the significance of the standard and market requirements. This document is referenced not only in 11be but also in the PAR and CSD of 11bn. After the establishment of TGbe, NTT continued to solicit that 11be should have some form of latency reduction and notification of latency related information features. Therefore, 11bn will have a latency-reduction feature, such as restricted TWT, and new signaling information, such as quality-of-service characteristic elements.

4. IEEE 802.11bn: Ultra High Reliability

In parallel with the standardization discussions for 11be, discussions were underway to formulate the next primary standard after 11be. In January 2022, the establishment of a new primary SG following 11be was proposed at the Wireless Next Generation (WNG) SC, which discusses the need for standardization regarding specific functions and topics. After several months of discussion, the Ultra High Reliability (UHR) SG was established and launched in September 2022. The main topic of UHR SG was to create the PAR and CSD for the next generation of wireless LANs following 11be. On the basis of the agreement of creating the PAR and CSD, TGbn was established in November 2023. The standardization of 11bn is scheduled to be completed in May 2028. Though the main scope of the existing primary standards, 11bn or before, is the enhancement of maximum throughput or improving channel efficiency, 11bn focuses on improving the reliability of wireless LANs. Specifically, in the same frequency band as 11be (2.4/5/6-GHz band), latency characteristics, packet loss, and throughput characteristics in received power according to distance will be improved by 25% compared with 11be in 11bn. Improved AP power saving and P2P will also be considered. As of December 2023, UHR SG mainly discusses functional proposals from various companies to achieve the scope described in the PAR and CSD.

4.1 Features under consideration for 11bn

In UHR SG and TGbn, many companies have proposed various features for 11bn as follows. Note that these features have not been decided for technical specifications but are under discussion.

(1) Multi-AP coordination

Multi-link operation defined in 11be enables coordination using multiple wireless interfaces equipped with a single MLD. In contrast, multi-AP coordination improves frequency efficiency and reliability through cooperation between multiple independent APs. This feature had been discussed in early TGbe; however, TGbe decided not to include this feature in 11be specifications due to the timeline compliance and implementation complexity. Therefore, discussions of multi-AP coordination have been taken over by TGbn. Several types of coordination forms are under consideration. Coordinated spatial reuse (Co-SR) controls transmission power to decrease interference to STAs. Coordinated beamforming (Co-BF) enables simultaneous transmission to each target STA of each AP without generating interference by coordinated beamforming and nulling. In joint transmission (JT), multiple APs transmit identical signals simultaneously with synchronization, and a receiver STA can combine multiple identical signals. **Figure 2** shows conceptual images of multi-AP coordination.

(2) Distributed RU

As described in the previous section, the MRU feature was defined in 11be. TGbn is discussing its enhancement to distributed-tone RUs (dRUs). This feature does not assign RUs by several patterns of combinations but freely, which offers more flexibility. For example, the same data block can be assigned to multiple RUs for redundancy, and the reliability of data transmission improves.

(3) Latency reduction

Restricted TWT defined in 11be can protect a specific period for latency-sensitive traffic generated periodically. In addition to the enhancement of restricted TWT, support for non-periodic (sporadic) latency-sensitive traffic is under consideration in 11bn. Preemption can interrupt latency-sensitive traffic into ongoing transmission by the frequency and/or time domain and transmit with priority. A function that ensures transmission of latency-sensitive traffic by using triggered TXOP sharing and enhancement of a buffer status report, which notifies an AP or an STA of the requirement of latency bound in real time, are also considered.

(4) Other features under consideration

Power-saving functions applying multi-link operation and/or multi-AP coordination, extension of area coverage of a wireless LAN network using relay architecture, and enhancement of security functions are under discussion in 11bn.

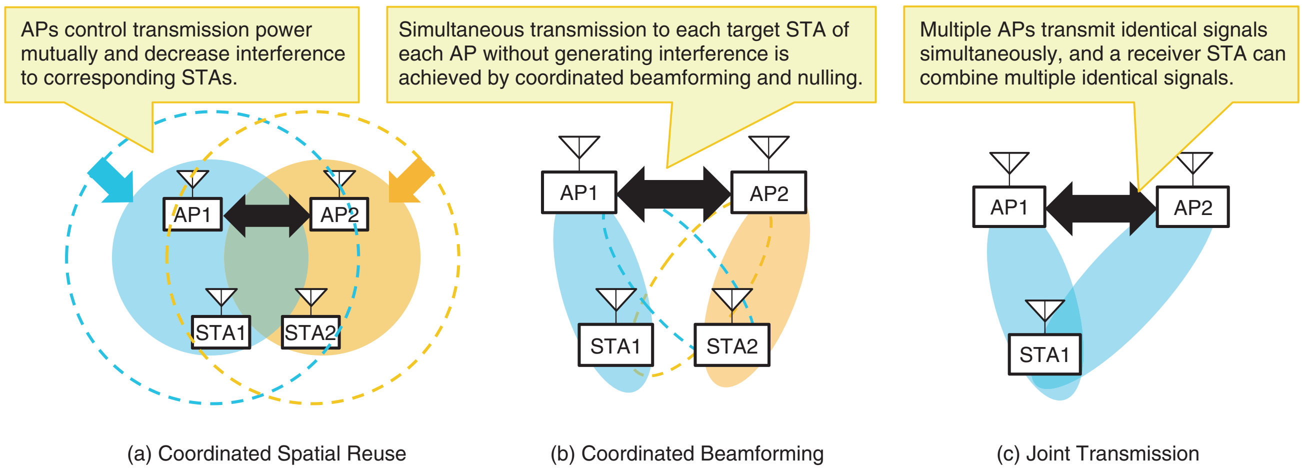


Fig. 2. Conceptual images of multi-AP coordination.

4.2 Discussion on utilizing millimeter wave

Whether to include the millimeter-wave frequency band (60 GHz) as the target frequency in addition to existing frequency bands such as 2.4/5/6 GHz was an important topic in UHR SG, which discusses the scope of 11bn. Although 11ad/ay are standards for wireless LAN systems that use millimeter waves, this discussion topic in UHR SG assumes using millimeter waves as additional links for multi-link operation features, which is unrelated to 11ad/11ay. As a result of the discussion, UHR SG decided that the millimeter-wave band should be discussed in a separate subgroup from TGbn, and the Integrated Millimeter Wave (IMMV) SG was launched in November 2023.

4.3 Standardization activities of NTT for 11bn

NTT has continued to promote the need to improve latency characteristics and reliability in wireless LAN since RTA TIG launched in 2018 and proposes to expand the Wi-Fi market not only for the consumer but also for business use cases requiring certain communication qualities such as low latency. This contribution changed the scope of mainstream wireless LAN standards from improving maximum throughput, such as 11be or before, to improving reliability in 11bn. NTT also proposed the PAR [4] and CSD [5] of 11bn with other companies and contributed to including important business use cases such as industrial automation, robotics, logistics, and smart agriculture in these prospective documents for establishing TGbn [6]. At the same time as the TGbn foundation, Yusuke Asai, a distinguished researcher of NTT, was elected 11bn leadership officer (Secretary).

5. Conclusions

This article explained the standardization trends in and NTT's activities for the next generation of extremely high throughput wireless LAN IEEE 802.11be and the next leading standard after 11be, the ultra-high reliability wireless LAN IEEE 802.11bn.

References

[1] IEEE 802.11 Working Group, https://www.ieee802.org/11/
[2] Wi-Fi Alliance, https://www.wi-fi.org/
[3] K. Meng, A. Jones, D. Cavalcanti, K. Iyer, C. Ji, K. Sakoda, A. Kishida, F. Hsu, J. Yee, L. Li, and E. Khorov, "IEEE 802.11 Real Time Applications TIG Report," IEEE 802.11-19/2009r6, Mar. 2019.
[4] L. Cariou, V. Erceg, L. Hervieu, C. Hu, B. Hart, A. Andersdotter, and A. Kishida, "802.11 UHR Proposed PAR," IEEE 802.11-23/0078r7, Mar. 2023.
[5] L. Cariou, L. Hervieu, C. Hu, B. Hart, and A. Kishida, "IEEE 802.11 UHR Proposed CSD," IEEE 802.11-23/0079r10, Mar. 2023.
[6] A. Kishida, Y. Asai, K. Nagata, R. Nagatsu, and Y. Takatori, "Use Cases for Wi-Fi Business Solutions in UHR," IEEE 802.11-22/1493r0, Sept. 2022.

Akira Kishida
Senior Research Engineer, NTT Access Network Service Systems Laboratories.
He received a B.E. and M.E. from Tokyo University of Agriculture and Technology in 2005 and 2007 and Ph.D. in engineering from the University of Tsukuba, Ibaraki, in 2015. In 2007, he joined NTT Access Network Service Systems Laboratories then joined NTT DOCOMO in 2015. He is currently engaged in the standardization of the IEEE 802.11be and IEEE 802.11bn as the delegate of NTT Access Network Service Systems Laboratories. He received the IEICE (Institute of Electronics, Information and Communication Engineers) Young Researcher's Award in 2010.

Optical Loss Prediction Tool 1.0 toward Enabling Preventive Maintenance of Optical Fiber Cables

Technical Assistance and Support Center, NTT EAST

Abstract

Technical Assistance and Support Center, NTT EAST developed the Optical Loss Prediction Tool 1.0 for predicting transmission loss of optical fibers. This article presents an overview and the functions of this tool. This is the eightieth article in a series on telecommunication technologies.

Keywords: OTDR, transmission loss, loss prediction

1. Introduction

Optical fiber cables are deployed in diverse environments, such as urban and mountainous areas, and play a role of a social infrastructure supporting not just information and communication services but also various social activities. Transmission loss of optical fibers in the cable may increase due to age-related deterioration caused by the environmental conditions under which the cable is deployed. To maintain communication quality through optical fiber cables as a social infrastructure, it is important to regularly monitor transmission loss of the optical fibers and preventively take necessary maintenance, such as cable replacement, before a problem occurs.

Technical Assistance and Support Center (TASC), NTT EAST developed the Optical Loss Prediction Tool 1.0 for predicting the transmission loss of optical fibers from periodically conducted and stored optical measurement data. This article presents an overview and the functions of this tool.

2. Development background

Optical fiber cables may deteriorate over time due to the environmental conditions, which can cause an increase in transmission loss [1]. For example, in underground optical fiber cables that have been laid in a humid environment for an extended period, blisters may form inside the coating of the optical fibers, as shown in **Fig. 1**. In such optical fibers, the continuous application of multiple invisibly small bending (microbending) causes excessive transmission loss. Even small locally, the transmission loss accumulated over long distance (over hundreds of meters to several kilometers) can increase so much that service cannot be provided, when microbending occurs. To maintain optical fiber cables in good condition, it is important to periodically conduct tests to determine the state of transmission loss and take preventive measures, such as replacing cables, before a problem occurs.

An optical time domain reflectometry (OTDR) is generally used to determine the condition of optical fiber cables. An OTDR uses pulsed light as a test light injected into an optical fiber and measures the backscattered light and reflected light. The round-trip delay of the backscattered or reflected light is proportional to the distance to the scattering or reflection point in the optical fiber; therefore, the transmission state along the length of the optical fiber cable can be estimated by acquiring the intensity of the backscattered

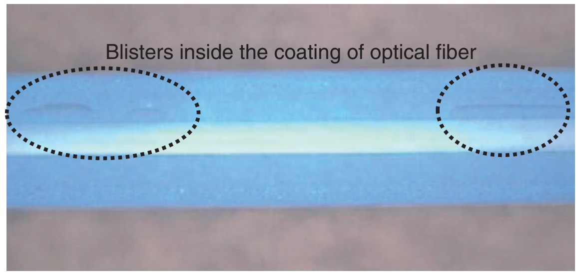


Fig. 1. Magnified image of an optical fiber with microbending.

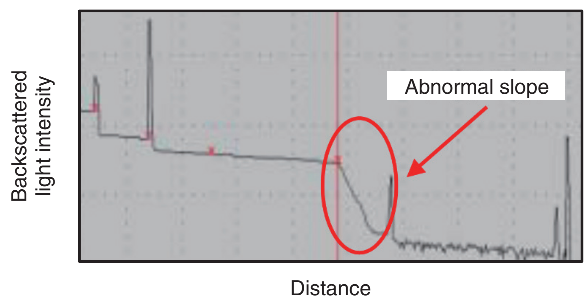


Fig. 2. Waveform of optical fiber cable with increasing transmission loss measured using an OTDR.

or reflected light as a function of time [2]. **Figure 2** shows that in a cable section where transmission loss has increased due to cable aging, etc., the slope of the waveform measured with an OTDR, which is the backscattered light intensity change with respect to distance, increases abnormally. Therefore, by analyzing the slope of the waveform with respect to distance, it is possible to determine the condition of transmission loss.

NTT EAST Group conducts monthly optical measurements on optical fiber cables in optical access networks using an OTDR. However, these periodic OTDR measurements only confirm the presence or absence of abnormalities at the time of measurement, and there is no method for detecting abnormalities before they become apparent. If an abnormality is detected during periodic OTDR measurements and the abnormal optical fiber cable must be renewed, it can take several months to a year from design to completion of installation. Therefore, it might be difficult to secure assignable fibers if the abnormality is addressed after it becomes apparent.

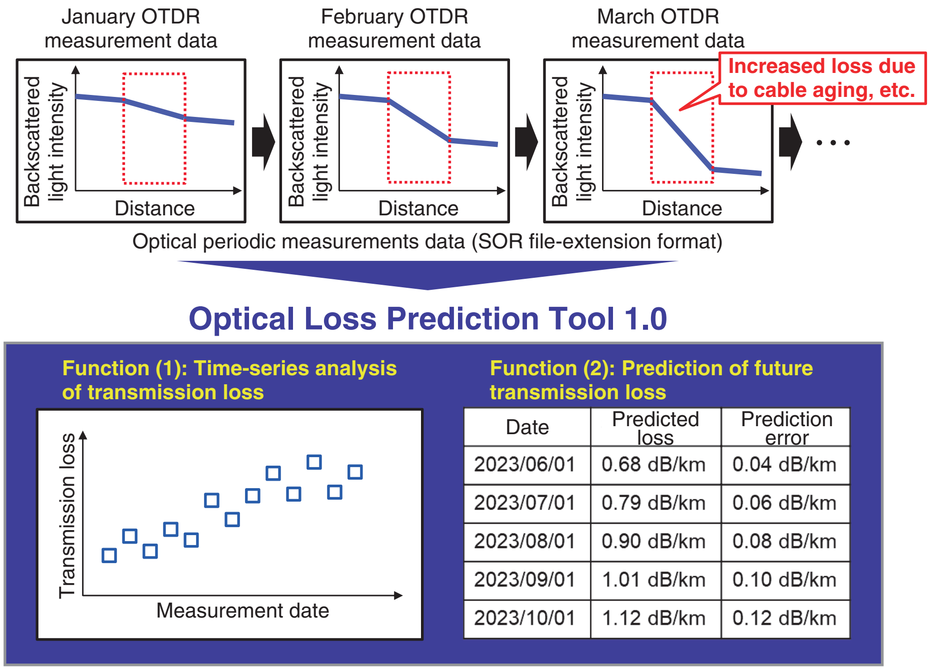
Responding to the above issues, TASC developed the Optical Loss Prediction Tool 1.0 to identify transmission loss abnormalities in advance. This tool uses periodically conducted and stored optical measurement data to analyze time-series changes in past transmission loss and enables prediction of future transmission loss.

3. Overview of Optical Loss Prediction Tool 1.0

Figure 3 shows an overview of the Optical Loss Prediction Tool 1.0. The tool has two main functions: (1) time-series analysis of transmission loss and (2) prediction of future transmission loss. The following sections explain each function.

3.1 Time-series analysis of transmission loss

The tool's first function is to visualize time-series changes in transmission loss in any cable section. To use this function, the folder containing optical periodic measurement data is specified on the input screen. The tool then loads all optical periodic measurement data from the specified folder at once. The loaded measurement data, which is compliant with the Standard OTDR Record (SOR) data format [3], includes the backscattered light intensity and distance information measured using an OTDR, as well as the measurement date. When analyzing transmission



Date	Predicted loss	Prediction error
2023/06/01	0.68 dB/km	0.04 dB/km
2023/07/01	0.79 dB/km	0.06 dB/km
2023/08/01	0.90 dB/km	0.08 dB/km
2023/09/01	1.01 dB/km	0.10 dB/km
2023/10/01	1.12 dB/km	0.12 dB/km

Fig. 3. Overview of Optical Loss Prediction Tool 1.0.

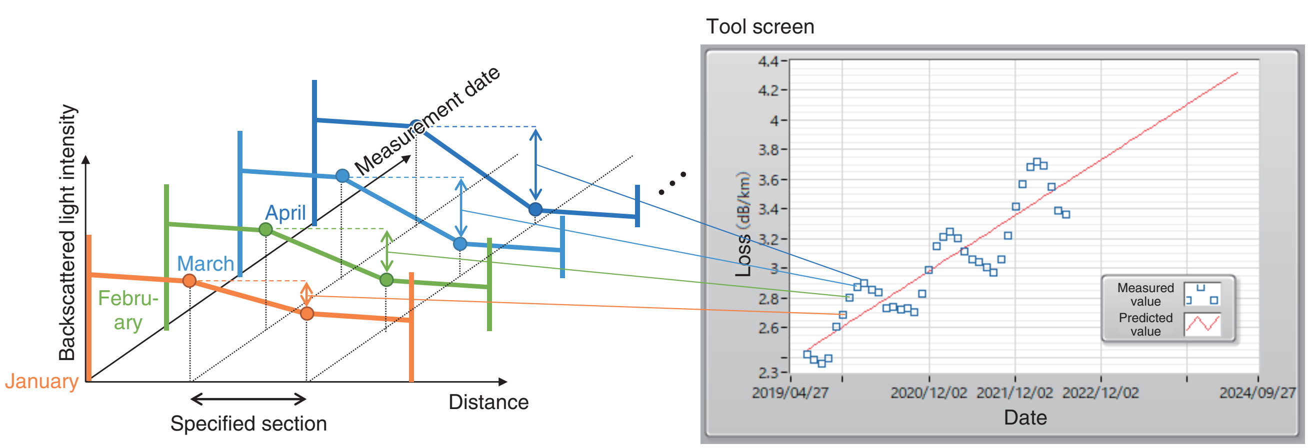


Fig. 4. Schematic diagram of time-series analysis of transmission loss.

loss, the distance information for the cable section is specified. The transmission loss values are then automatically analyzed for all loaded SOR extension data by calculating the backscattered-light-intensity changes in the cable section on the basis of the specified distance information. These values are associated with the date of measurement recorded in each SOR extension data, and displayed in a graph arranged in order of measurement date, as shown in **Fig. 4**. This graphical visualization of time-series changes in transmission loss enables us to see the degradation trend of transmission loss in the specified section.

3.2 Prediction of future transmission loss

The tool's second function is to predict future transmission loss in the specified cable section. This function uses the analyzed time-series data mentioned in

Measurement date	Predicted loss (dB/km)	Prediction error (dB/km)
2022/07/31	3.604208	0.244726
2022/08/30	3.635487	0.256302
2022/09/30	3.666767	0.268041
2022/10/30	3.698047	0.279942
2022/11/29	3.729326	0.292006
2022/12/30	3.760606	0.304233
2023/01/29	3.791886	0.316622

Fig. 5.   Example results of output by transmission loss prediction.

Section 3.1 as training data for transmission loss prediction. Transmission loss prediction is carried out by applying a regression analysis to the time-series data and extrapolating the obtained regression line to the specified prediction date, as shown with the red line in Fig. 4. The predicted losses are listed next to the specified prediction date, as shown in **Fig. 5**, in a manner that makes it possible to quantitatively understand the predicted increase in transmission loss. As also shown in Fig. 5, the predicted values of transmission losses as well as estimated prediction errors are displayed as a list up to the specified prediction month, enabling a quantitative understanding of transmission loss deterioration prediction. The prediction errors, which are automatically estimated on the basis of the specified confidence coefficient (probability that the prediction value will fall within the error range) and prediction date, are displayed together with the prediction values in this list box. This function makes it possible to determine the prediction value of future transmission loss and error range, which can be useful in planning cable renewals.

4. Conclusion

This article introduced a tool for predicting future transmission loss in optical fiber cables. We believe that this tool will help prevent failures caused by changes in transmission loss due to age-related deterioration, etc., and contribute to the planned renewal of cables. Toward practical use for cable maintenance, we are currently discussing the utilization method of this tool with the relevant departments of NTT EAST and NTT WEST. We will also improve our prediction accuracy by incorporating machine learning and other methods into this tool.

The Access Network Engineering Group in TASC offers technical support to resolve challenging access facility failures that occur in the field throughout Japan. We will continue to contribute to field troubleshooting and develop tools that enhance technical skills and efficiency in the field by leveraging the expertise acquired from technical support and the knowledge gained through the investigation of the causes of failures.

References

[1] K. Sohma, N. Iwaguchi, T. Kawano, T. Fujii, and Y. Koyano, "Estimation of Long-term Change in Physical Property of Optical Fiber Coating Considering Effect of Humidity," SEI Technical Review, Vol. 187, pp. 12–17, 2015.

[2] NTT EAST, "Methods for Finding Failures in Optical Fibers," NTT Technical Journal, Vol. 18, No. 10, pp. 53–54, 2006 (in Japanese).

[3] Telcordia SR-4731, "Optical Time Domain Reflectometer (OTDR) Data Format," Issue 2, 2011.

External Awards

Technical Development Award

Winners: Tatsuya Kako, NTT Computer and Data Science Laboratories; Hironobu Chiba, NTT Computer and Data Science Laboratories
Date: May 27, 2023
Organization: Acoustical Society of Japan

For development of open-ear earphones that minimize sound leakage by radiating antiphase sound waves.

SCIS Innovation Paper Award

Winners: Junko Takahashi, NTT Social Informatics Laboratories; Shota Kita, NTT Basic Research Laboratories; Akihiko Shinya, NTT Basic Research Laboratories; Kazumaro Aoki, Bunkyo University; Koji Chida, Gunma University; Fumitaka Hoshino, University of Nagasaki
Date: July 20, 2023
Organization: The Institute of Electronics, Information and Communication Engineers (IEICE) Technical Committee on Information Security (ISEC)

For "Photonic Cryptographic Circuit Technology and Operation Verification."
Published as: J. Takahashi, S. Kita, A. Shinya, K. Aoki, K. Chida, and F. Hoshino, "Photonic Cryptographic Circuit Technology and Operation Verification," Proc. of the 40th Symposium on Cryptography and Information Security (SCIS 2023), 1E2-2, Kitakyushu, Fukuoka, Japan, Jan. 2023.

Young Scientist Award

Winner: Hiroyuki Fukumoto, NTT Network Innovation Laboratories
Date: August 21, 2023
Organization: 35th Union Radio-Scientifique Internationale General Assembly and Scientific Symposium (URSI GASS 2023)

For "Implementation of Bidirectional High-rate Underwater Acoustic Communication Systems for Fully Wireless-controlled Remotely Operated Vehicles."
Published as: H. Fukumoto, R. Okumura, Y. Fujino, S. Ohmori, Y. Ito, T. Ishihara, and Y. Tabata, "Implementation of Bidirectional High-rate Underwater Acoustic Communication Systems for Fully Wireless-controlled Remotely Operated Vehicles," URSI GASS 2023, Sapporo, Hokkaido, Japan, Aug. 2023.

Engineering Sciences Society Service Award

Winner: Yukihiro Bandoh, NTT Computer and Data Science Laboratories
Date: September 13, 2023
Organization: IEICE

For contribution to the management and activities of the Technical Committee on Smart Info-Media Systems.

Signal Processing Young Researcher's Award

Winner: Takayuki Sasaki, NTT Computer and Data Science Laboratories
Date: November 7, 2023
Organization: IEICE Technical Group on Signal Processing

For "Fast Graph Fourier Transform based Graph Signal Denoising."

Kekicheff Award

Winner: Eikazu Niwano, NTT Social Informatics Laboratories
Date: November 10, 2023
Organization: GlobalPlatform

For his outstanding dedication and commitment to GlobalPlatform. He has been a key part of its organization since 2005, serving on the GlobalPlatform Board of Directors and leading the Japan Task Force. This award recognizes his extraordinary, impactful efforts.

MNC 2022 Most Impressive Presentation Award

Winners: Katsuhiko Nishiguchi, NTT Basic Research Laboratories; Akira Fujiwara, NTT Basic Research Laboratories; Kensaku Chida, NTT Basic Research Laboratories
Date: November 14, 2023
Organization: 35th International Microprocesses and Nanotechnology Conference (MNC 2022)

For "Room-temperature Single-electron Sensor for AC Signals beyond an RC Time Constant ~ Energy Transfer via a Non-leakage DRAM in Non-equilibrium ~."
Published as: K. Nishiguchi, A. Fujiwara, and K. Chida, "Room-temperature Single-electron Sensor for AC Signals beyond an RC Time Constant ~ Energy Transfer via a Non-leakage DRAM in Non-equilibrium ~," MNC 2022, 10B-2-1, Tokushima, Japan, Aug. 2023.

MNC 2022 Most Impressive Poster

Winners: Megumi Kurosu, NTT Basic Research Laboratories; Daiki Hatanaka, NTT Basic Research Laboratories; Ryuichi Ohta, NTT Basic Research Laboratories; Keiko Takase, NTT Basic Research Laboratories; Hiroshi Yamaguchi, NTT Basic Research Laboratories; Yoshitaka Taniyasu, NTT Basic Research Laboratories; Hajime Okamoto, NTT Basic Research Laboratories
Date: November 14, 2023
Organization: MNC 2022

For "High-overtone Bulk Acoustic Resonator with Epitaxial AlN Layer Grown on n-SiC."
Published as: M. Kurosu, D. Hatanaka, R. Ohta, K. Takase, H. Yamaguchi, Y. Taniyasu, and H. Okamoto, "High-overtone Bulk Acoustic Resonator with Epitaxial AlN Layer Grown on n-SiC," MNC 2022, 11P-4-31, Tokushima, Japan, Aug. 2023.

Best Paper Award

Winner: Shin'ya Yamaguchi, NTT Computer and Data Science Laboratories
Date: November 15, 2023
Organization: The 15th Asian Conference on Machine Learning (ACML 2023)

For "Generative Semi-supervised Learning with Meta-Optimized Synthetic Samples."
Published as: S. Yamaguchi, "Generative Semi-supervised Learning with Meta-Optimized Synthetic Samples," Proc. of ACML 2023, Istanbul, Turkey, Nov. 2023.

Runner-up rank

Winners: Takumi Fukami, NTT Social Informatics Laboratories; Yuusuke Yamasaki, NTT Social Informatics Laboratories; Iifan Tyou, NTT Social Informatics Laboratories; Kenta Niwa, NTT Communication Science Laboratories
Date: November 15, 2023
Organization: NeurIPS 2023 competition on "Privacy Preserving Federated Learning Document VQA" (PFL-DocVQA)

For "Differential Privacy Analysis for DP-CLGECL."

IDSA Ambassador

Winner: Koki Mitani, NTT Social Informatics Laboratories
Date: November 22, 2023
Organization: International Data Spaces Association (IDSA)

For outstanding commitment to promoting international development and cooperation of data spaces driven by IDSA.

Best Paper Award

Winners: Atsushi Taniguchi, NTT Network Innovation Laboratories; Yasuhiro Mochida, NTT Network Innovation Laboratories; Sakae Chikara, NTT Social Informatics Laboratories; Yasuyuki Sanari, NTT Network Innovation Laboratories; Keizo Murakami, NTT Social Informatics Laboratories; Momoko Miura, NTT Network Innovation Laboratories; Hirokazu Takahashi, NTT Network Innovation Laboratories; Koichi Takasugi, NTT Network Innovation Laboratories; Hiroki Itoh, NTT Social Informatics Laboratories; Daigoro Yokozeki, NTT Social Informatics Laboratories
Date: November 30, 2023
Organization: International Conference on Emerging Technologies for Communications (ICETC) 2023

For "Disaggregation Architecture of Quantum-safe Security for Optical & Packet Transport Equipment."
Published as: A. Taniguchi, Y. Mochida, S. Chikara, Y. Sanari, K. Murakami, M. Miura, H. Takahashi, K. Takasugi, H. Itoh, and D. Yokozeki, "Disaggregation Architecture of Quantum-safe Security for Optical & Packet Transport Equipment," Proc. of ICETC 2023, Sapporo, Hokkaido, Japan, Nov./Dec. 2023.

IEEE MTT-S Japan Chapter Young Engineer Award

Winner: Masahito Nakamura, NTT Device Technology Laboratories
Date: November 30, 2023
Organization: IEEE Microwave Theorty and Technology Society (MTT-S) Japan Chapter

For "Broadband Dielectric Spectroscopy for Quantitative Analysis of Glucose and Albumin in Multicomponent Aqueous Solution."
Published as: M. Nakamura, T. Tajima, and M. Seyama, "Broadband Dielectric Spectroscopy for Quantitative Analysis of Glucose and Albumin in Multicomponent Aqueous Solution," IEEE JERM, Vol. 6, No. 1, pp. 86–93, Mar. 2022.

IACR Asiacrypt Test-of-Time Award

Winners: Yu Sasaki, NTT Information Sharing Platform Laboratories (At the time of writing the paper); Kazumaro Aoki, NTT Information Sharing Platform Laboratories (At the time of writing the paper)
Date: December 6, 2023
Organization: The International Association for Cryptologic Research (IACR)

For "Preimage Attacks on 3, 4, and 5-Pass HAVAL."
Published as: Y. Sasaki and K. Aoki, "Preimage Attacks on 3, 4, and 5-Pass HAVAL," ASIACRYPT 2008, LNCS, Vol. 5350, pp. 253–271, 2008.

Globecom 2023 Best Paper Award

Winners: Katsuaki Higashimori, NTT Network Innovation Laboratories; Takafumi Tanaka, NTT Network Innovation Laboratories; Takeru Inoue, NTT Network Innovation Laboratories
Date: December 7, 2023
Organization: IEEE Global Communications Conference (GLOBECOM) 2023

For "Efficient Routing Method for Reducing Significant Outages in Optical Networks."
Published as: K. Higashimori, T. Tanaka, and T. Inoue, "Efficient Routing Method for Reducing Significant Outages in Optical Networks," GLOBECOM 2023, Kuala Lumpur, Malaysia, Dec. 2023.

IOP Trusted Reviewer status

Winner: Takuya Hatomura, NTT Basic Research Laboratories
Date: December 12, 2023
Organization: IOP Publishing

2023 IEICE Communications Society OCS Young Researchers Award

Winner: Akira Kawai, NTT Network Innovation Laboratories
Date: December 14, 2023
Organization: IEICE Technical Committee on Optical Communication Systems (OCS)

For "Precise Signal Demodulation in Dynamic Environments Using Partially Frozen MIMO Adaptive Equalizers."
Published as: A. Kawai, M. Nakamura, M. Takahashi, T. Kobayashi, and Y. Miyamoto, "Precise Signal Demodulation in Dynamic Environments Using Partially Frozen MIMO Adaptive Equalizers," IEICE Tech. Rep., Vol. 122, No. 332, OCS2022-69, pp. 51–55, Jan. 2023.

Papers Published in Technical Journals and Conference Proceedings

Northcott Numbers for the Weighted Weil Heights

M. Okazaki and K. Sano

Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl., Vol. 34, No. 1, pp. 127–144, July 2023.

We answer the question of Vidaux and Videla about the distribution of the Northcott numbers for the Weil height. We solve the same problem for the weighted Weil heights. These heights generalize both the absolute and relative Weil height. Our results also refine those of Pazuki, Technau, and Widmer.