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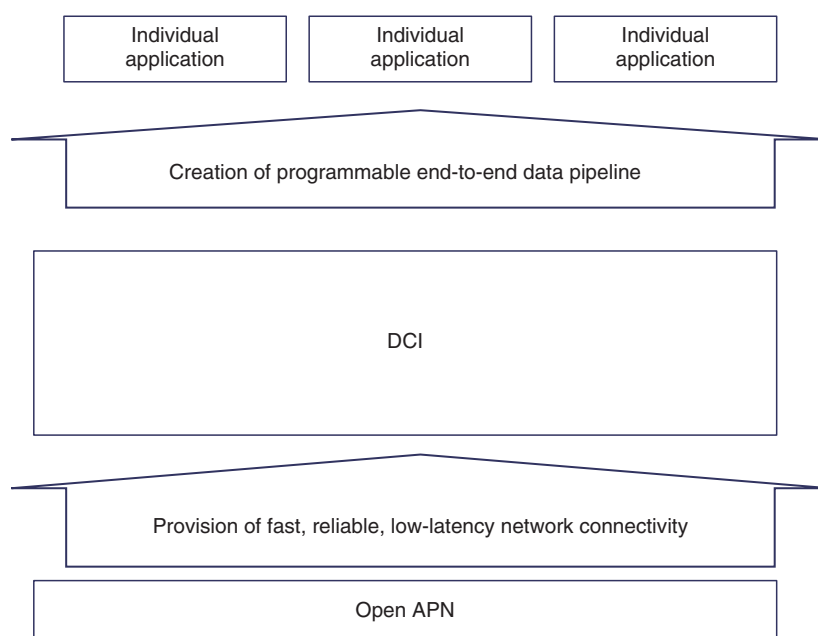
Fig. 1. Overview of ICT infrastructure centered on DCI.

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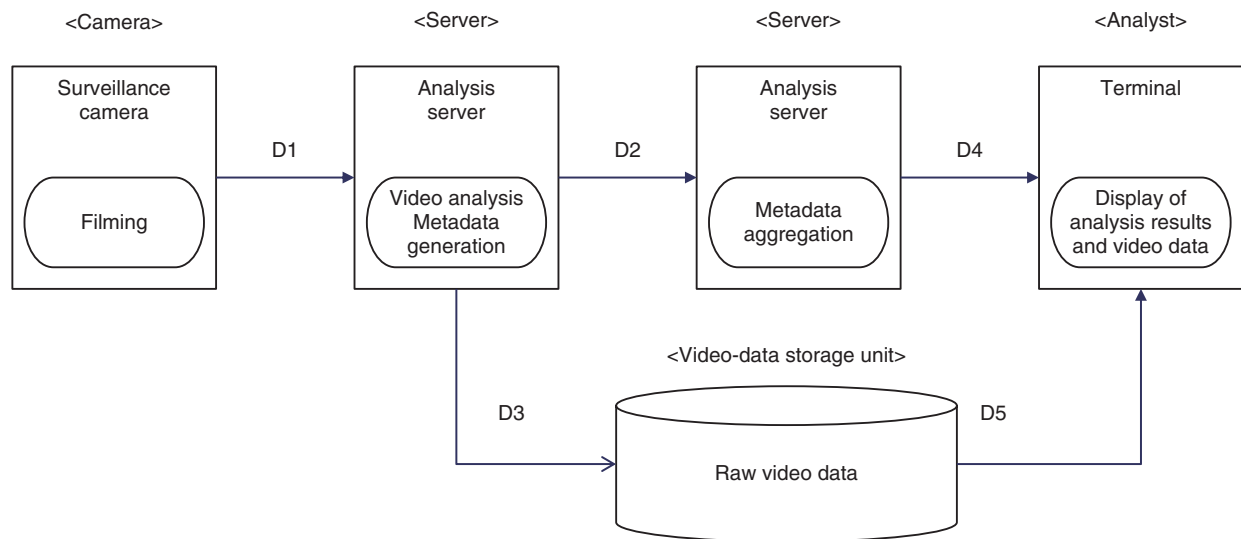
Fig. 2. Example of a data pipeline of a system that aggregates and analyzes surveillance-camera footage on the basis of the metadata.

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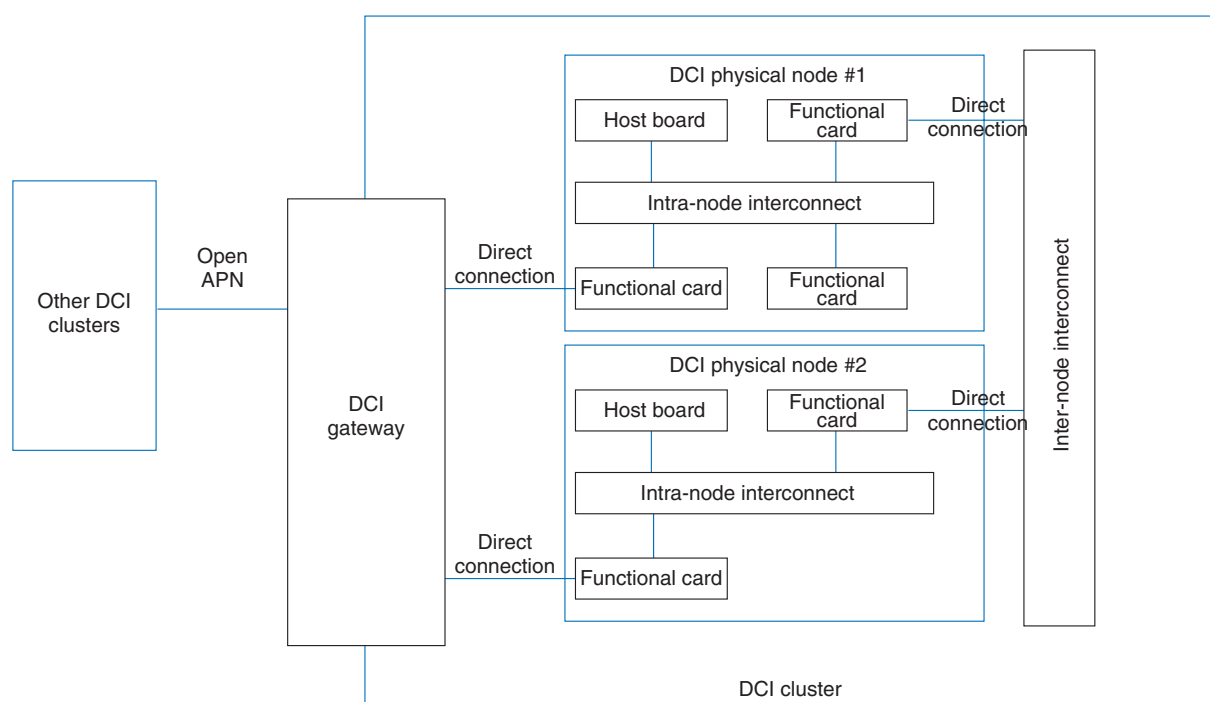
Fig. 3.   Schematic of various resources managed by the DCI.

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
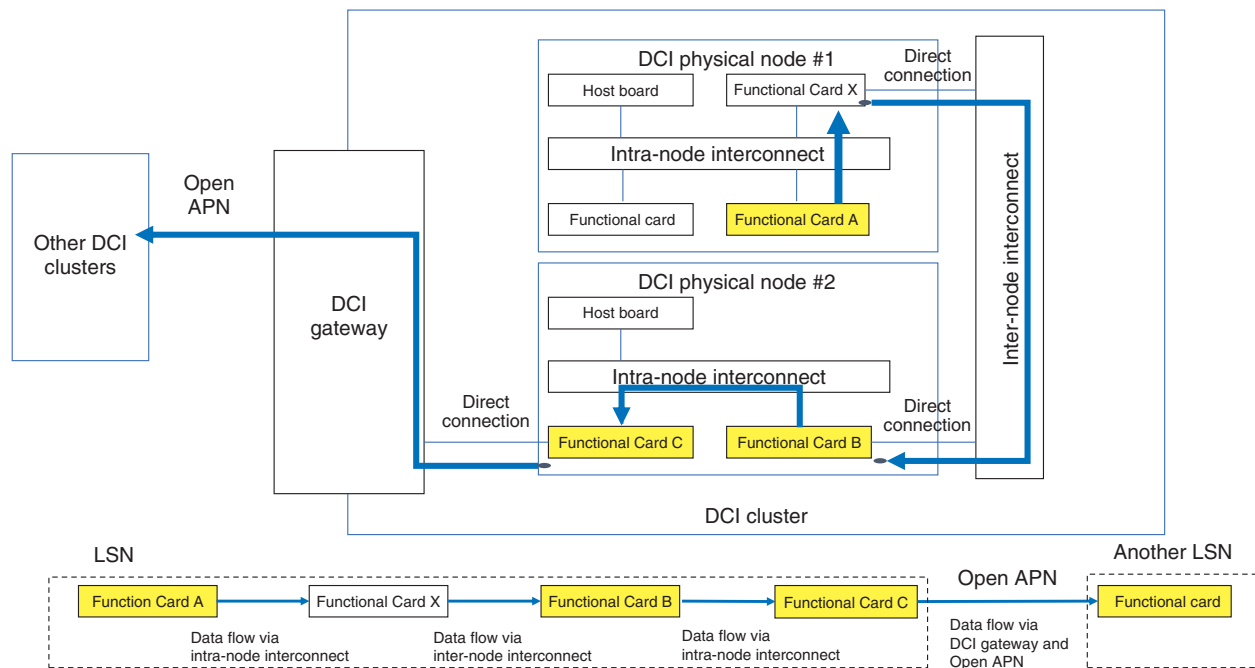
Fig. 4.   Example of data-pipeline configuration.

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.