Software-Defined Container-Based Data Centers

3.1 Overview

Here we will give an outline of the co-IZmoSD software-defined container-based data center we are developing and testing, as well as associated research, with the aim of realizing both high efficiency and reliability by exerting precise control over data center facilities using software.

3.2 Data Center Requirements

The minimum requirements for enabling an Internet service are an Internet connection and an IT server to run the service on. The fact that this is all you need to provide services to every Internet-connected user is one of the Internet's great advantages. However, more consideration will be required to continuously provide services to a large number of users. There are a range of issues that are almost impossible to deal with in an ordinary home or office, such as redundancy in anticipation of a failure occurring (a power outage, interruption of Internet access, or fault in equipment accommodating access lines), securing the floor load of the location where racks housing large numbers of IT servers will be installed, and increasing the cooling capacity for processing the exhaust heat of equipment.

Data centers provide features for enabling stable Internet services*1. They are facilities specifically designed to run massive numbers of IT servers stably and continuously, resolving the aforementioned issues all at once. Elements such as the cooling systems, power systems, and external connectivity have redundant configurations, enabling operation to continue even if a fault occurs in part of the data center. Additional services are also provided, such as physical security that includes strict entry management and device status monitoring, as well as device protection through fire extinguishing equipment specifically designed for IT devices.

Data centers need to provide stable operation, but leading-edge data centers are also expected to offer other features. "Optimization" and "massive integration" have been keywords for the technological development of Internet-oriented data centers over the past ten years or so. There is also a strong connection between these two features and the growth of cloud computing.

The terms "hosting" and "housing/colocation" are used as data center usage categories. Although the ownership of equipment differs, in both instances the role of the data center is to install the equipment that customers will use. In these cases, the main service of the data center is to provide the surrounding environment for the installed equipment. As long as an agreement is in place, it is necessary to continue providing services such as cooling and power regardless of equipment operating status, even in extreme cases such as when no actual equipment exists.

In contrast, at data centers whose main purpose is cloud computing, in most cases either entire floors or the entire data center itself is used for a single cloud service. Because the data center is occupied by a single service or operator, it is characteristic for the operation of the data center and equipment to be unified. By unifying these, it is possible to apply operating and optimization techniques not used in traditional data centers. For example, you can install IT servers that operate in higher temperatures than normal and change various design criteria such as the building and cooling equipment to create data centers that don't need as much cooling. There have also been aggressive trials involving IT servers with small integrated backup power supplies combined with power management software to eliminate the expensive backup power supply equipment (massive batteries or power generators) usually required at data centers. This demonstrates that it is becoming more appropriate to think of modern data centers targeted at cloud computing as organic systems that are integrated with the servers they accommodate, rather than facilities merely for housing IT servers.

*1 In a broad sense "data center" can refer to any building that houses computer equipment, but here we use the term to mean only data centers for Internet services (IDC).

3.3 Data Center Optimization

Data centers provide economies of scale due to the fact that large numbers of servers are housed in the same place. This is because ancillary equipment can be shared, and also operated at an efficient scale, making it possible to maintain data centers in a more efficient state than other environments. An indicator called PUE (Power Usage Effectiveness) is frequently used to show data center efficiency. PUE is defined using the following formula.

$$\text{PUE} = \frac{\text{Total data center power consumption}}{\text{IT equipment power consumption}} = \frac{\text{IT equipment power consumption + facility power consumption (+ power loss)}}{\text{IT equipment power consumption}}$$

For example:

- IT equipment power consumption = 10 kW
- Power consumption of any other equipment in the data center = 12 kW

In this case, the resulting PUE would be 2.2 as shown below.

$$\text{PUE} = \frac{10\text{ kW} + 12\text{ kW}}{10\text{ kW}} = 2.2$$

PUE indicates how much facility equipment (power supply, cooling, and lighting equipment, etc.) is required to operate IT equipment, so the theoretical best value for PUE is 1.0, which means no additional equipment is required to operate IT equipment. The PUE of typical data centers in Japan is around 2.0. Data center facilities consume the same amount of power as IT equipment, but this is still substantially more efficient than operating servers in a standard office building. However, due to factors such as the overall increase in IT equipment numbers due to the growth of Internet services, we need to pursue a higher level of efficiency for data centers. In particular, the stagnation of improvements to scale efficiency is a make-or-break issue at cloud data centers where systems consisting of vast quantities of servers are constructed. A PUE of about 1.2 is given as a target value for all advanced, large-scale data centers designed and built since 2010, and apparently actual PUE performance is in the range of about 1.1 to 1.3.

Against such a backdrop, IIJ has been performing a range of data center optimizations. One example of improvements to the efficiency of data center equipment is the Matsue Data Center Park, which we have operated in Matsue, Shimane Prefecture since 2011. Matsue Data Center Park optimizes initial investment costs and operating costs using a standard configuration that consists of modular data center containers and cooling modules. The year-round PUE performance is about 1.2, making it an extremely efficient data center facility. Refer to the website*2 for more information about Matsue Data Center Park.

3.4 Integrated Data Centers

At the latest generation of data centers, where component improvement optimizations approaching the theoretical threshold PUE value of 1.0 are on the verge of being realized, a paradigm shift that includes rethinking the definition of data centers is needed to achieve higher efficiency. Instead of treating data centers as facilities for simply housing IT equipment, it is necessary to search for technology that can be applied to systems where data centers and IT equipment are integrated.

Because forms of data centers utilization that enable a shift to systems that integrate IT equipment to be considered are limited to cases where the data center operator and IT system operator are closely related or the same entity, we will refer to these as "integrated data centers" here. Table 1 shows the differences between existing data centers and integrated data centers.

Item	Existing Data Centers	Integrated Data Centers
Operator	Data center and internal systems have different operators	Data center and internal systems have the same operator
Data center environment	The data center operator cannot change environment requirements without users' permission	Configurable according to requirements due to the operators being the same
Housed equipment	Selection is not possible because housed equipment is either brought in or specified by the customer	Selection of housed equipment according to design and operating requirements is possible
Operating requirements	Static operation is carried out to constantly fulfill the service level specified in advance (with specifications for temperature and humidity, power supply, connectivity, etc.)	Operation is carried out based on criteria that are determined dynamically taking into account housed equipment and service operating requirements

Table 1: Differences Between Existing Data Centers and Integrated Data Centers

*2 Matsue Data Center Park (http://www.iij.ad.jp/DC/en/about/index.html)

At integrated data centers it is possible to control the IT equipment environment at your discretion, which is something data center operators could not get involved in up until now. IT system operators can also control data center behavior. At existing data centers you could only specify static operating requirements such as an IT rack exhaust temperature of 22 - 23 degrees and humidity of 40% RH. Integrated data centers enable reciprocal interactions, such as changing data center operating conditions based on the operating status of IT systems, or controlling the availability of IT systems according to data center operating restrictions. System integration provides the following benefits:

- Dynamic, highly responsive control
- Interaction between the data center facility and IT equipment

3.5 Interaction Between Data Centers and IT Systems

A software interface for handling both the data center and IT systems is necessary to achieve a high degree of efficiency through interaction between them.

In most cases a standard API or protocol is specified for IT systems, so it is possible to use these to obtain or control various information as required. Also, progress has been made toward the examination, evaluation, and standardization of systems for the automatic control of IT servers and network switches, as well as techniques for constructing systems that consolidate and manage the racks as a whole, so the scope of control possible is sufficiently broad.

On the other hand, there are issues related to the control of data center equipment. The largest issue is that data centers are a concept representing the convergence of things and facilities, so there is no sense of the reality of the situation as with information systems. Because no end point for control has been defined, when managing a data center it is actually necessary to handle all its components individually. There are also other issues regarding these individual components. Nowadays, there are sufficient possibilities to choose components used in the data center facility that have networking functionality to communicate with each other. However, even though these components have such functionality, they still tend to use dedicated network protocols and vendor specific (non standardized) APIs. Some special care would be required to integrate such components with existing IT systems.

At the IIJ Research Laboratory, we are attempting to resolve these issues by implementing a new software layer that provides abstraction to enable data centers to be treated as part of an information system. In addition to abstraction, we are moving ahead with research into improving compatibility between data centers and IT systems by combining the following:

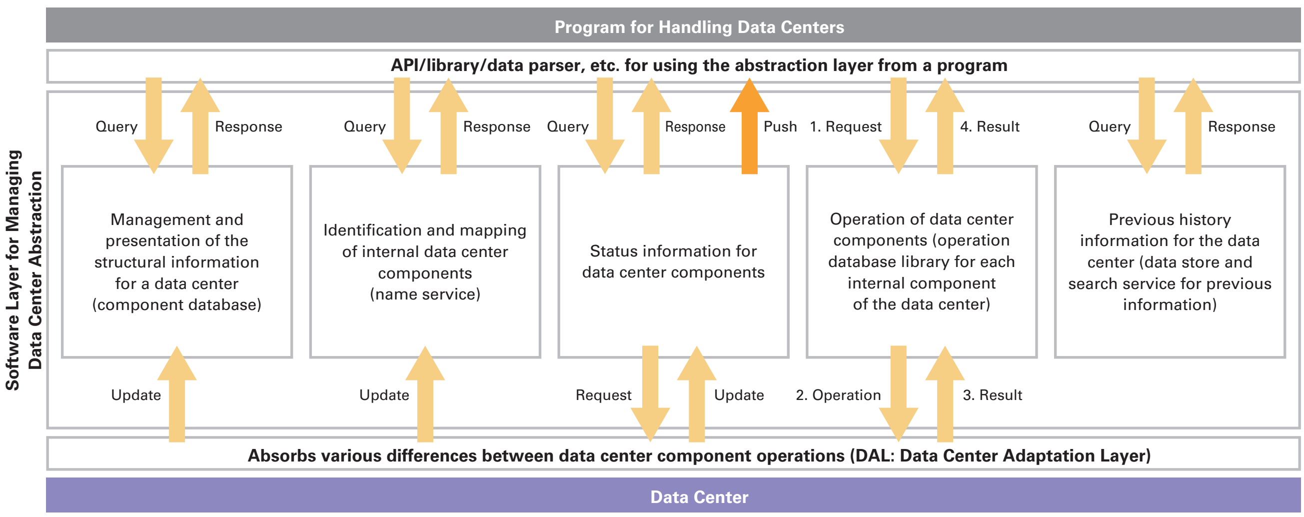

Figure 1: Software Abstraction of Data Centers

- Establishment of data center structure description techniques
- Standardization of a data format for internal information at data centers
- Preparation of procedures (APIs, etc.) for controlling the internal components of data centers
- Preparation of standard libraries for data center control programs

Figure 1 shows the components of a software group for handling data centers.

3.6 co-IZmoSD Container-Based Data Centers for Research and Development

Co-IZmoSD (Figure 2) is a data center facility we operate to perform research and development regarding data centers that are highly compatible with IT systems. It is based on co-IZmo*3, which was developed at IIJ as an integrated, compact, container-based data center in 2013. After adding functions required for research and development, co-IZmoSD was implemented in 2014. The SD part of co-IZmoSD stands for "software-defined," and indicates this experimental facility is aimed at maximizing the components in existing data center facilities that can be controlled using software.

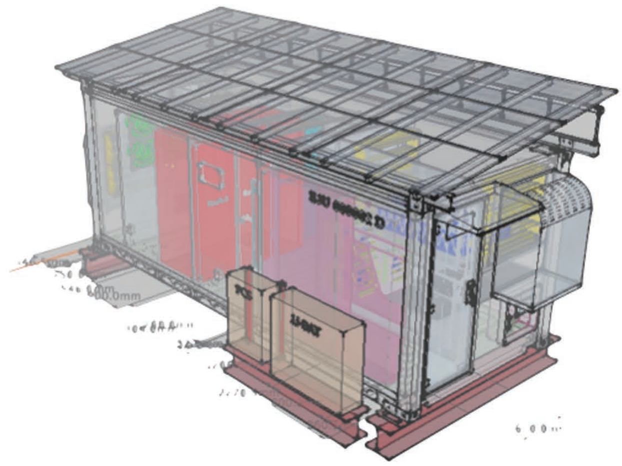

Figure 2: External View (Left) and Perspective View (Right) of co-IZmoSD

Item	co-IZmo	co-IZmoSD
Architecture	Container-based data center with integrated IT/cooling modules	←
Size	ISO 20 ft container, High Cube (9.6 ft)	ISO 20 ft container, standard (8.6 ft)
Footprint	8 m x 3 m = 24 m2	←
Cooling System	Direct outside-air free cooling, chiller-less	←
Racks	3 racks (46U)	3 racks (42U)
Rack Configuration	Cold/hot aisle containment	←
Cooling Capacity	Max 30 kw (depends on outside air conditions)	←
Internal Environment Provisions	Independent standard based on Ashrae 2011 allowable 3	←
Maintenance System	Fire detection sensor, condensation sensor, etc.	←
Auxiliary Battery	None	12 KWh Li-ion battery unit
PV + PCS	None	System equipment rated for 10 KWh
Control System	Stand-alone	Stand-alone + external control hybrid
Power Supply System	Uniform	Prioritized hierarchical power supply system

Table 2: Basic Specifications and Differences Between co-IZmo and co-IZmoSD

*3 IIJ, "IIJ to Begin Testing Container-Unit Data Center Modules with Year-round Outside-Air Cooling" (http://www.iij.ad.jp/en/news/pressrelease/2013/0408.html).

Table 2 shows outline specifications for co-IZmo that served as the base, as well as co-IZmoSD. Because co-IZmoSD was designed as a successor to co-IZmo, a variety of improvements have been made based on the implementation and operating experience of co-IZmo. IIJ has been continuing the research and development of container-based data centers beyond co-IZmoSD, and the co-IZmo/I*4 modular data center that uses an indirect outside-air cooling system is the latest incarnation of this architecture.

As shown by the co-IZmoSD specifications in Table 2, the basic architecture is the same as co-IZmo. About the only differences are that co-IZmoSD is based in an ISO container with a lower height (standard size), and it includes various ancillary equipment such as solar cell panels and external power supply units. On the other hand, the internal structure is very different, featuring a control mechanism that offers a high degree of freedom. Figure 3 shows an outline of the internal structure of co-IZmoSD.

As illustrated in Figure 3, the various environmental sensors, actuators (fans and dampers on the air intake and exhaust, etc.) and mechanisms for controlling the power supply equipment (power system toggle switch on the power panel and sensors for the power supply, etc.) in the data center are consolidated via a PLC (programmable logic controller) network. A PLC is programmed with the minimum required logic, and enables normal data center operations or equipment maintenance work by itself. In co-IZmoSD the PLC program is further expanded into a design that enables hybrid operation combining external control with the PLC's internal control logic. PLCs are highly reliable, but it is difficult to increase the scale of the internal program or shorten the development cycle. However, implementing this in tandem with an external system as we have done here enables more advanced control that would not be possible using a regular PLC.

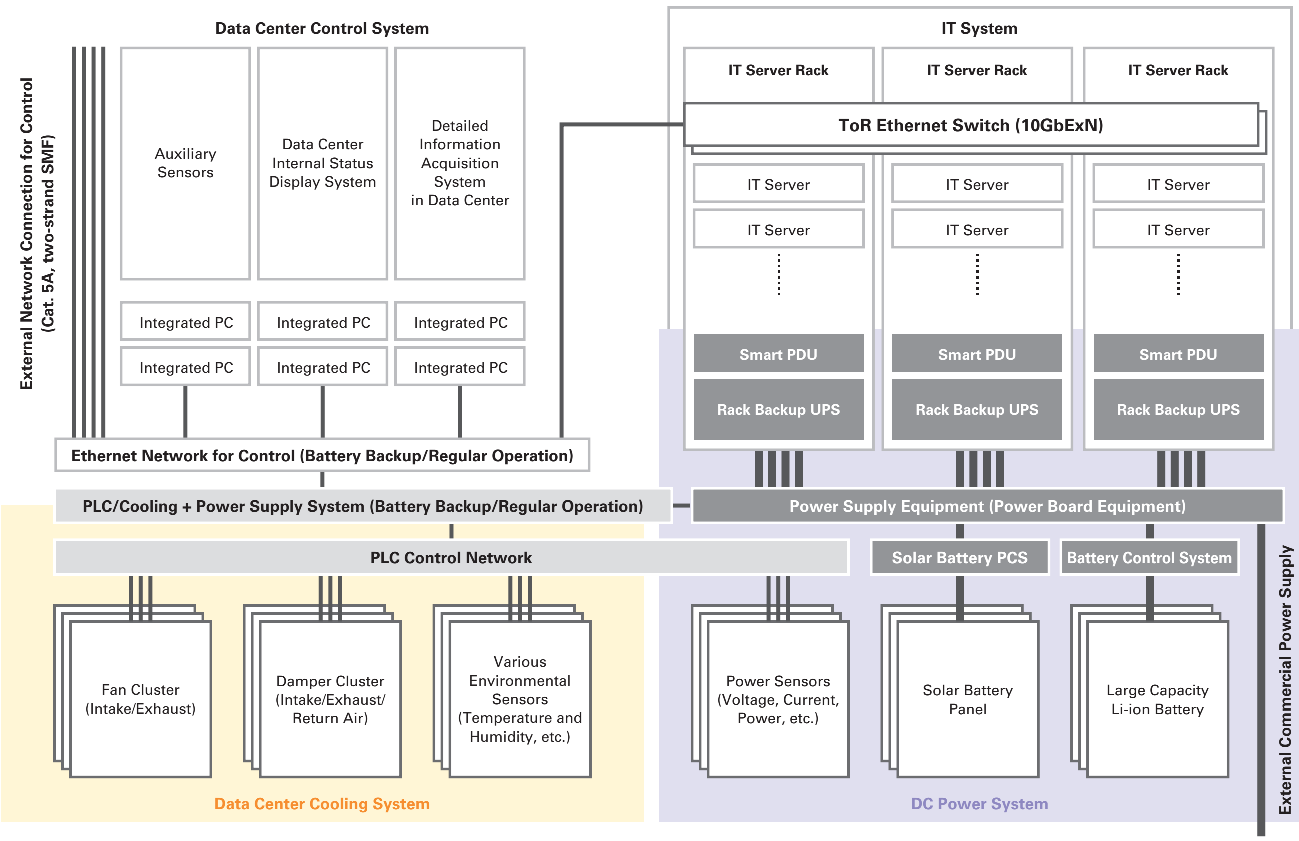

Figure 3: Internal Structure of co-IZmoSD

*4 co-IZmo/I (http://www.iij.ad.jp/DC/en/products/coizmo_i.html).

Components such as ancillary equipment (photovoltaic (PV) panels and PCS, large battery system), the PDU and UPS inside racks, and various sensor units not used for air conditioning control, etc., are also connected to a network. This makes it possible to access almost all components from a network outside the internal IT system or co-IZmoSD. In addition, the power supply system is equipped with features for supplying power from multiple power supply sources, and supporting coordination with each battery unit. You can also execute functions such as the partial shutdown or temporary suspension of the data center via software.

3.7 An Operating System for Data Centers

We designed and built co-IZmoSD as a research-oriented data center for testing the concept of interaction between data centers and IT systems. As co-IZmoSD provides sufficient functionality for interaction, we can verify the concept by combining these features with associated software components.

When you include IT systems, the components that make up a data center are many and varied. Data centers are also non-stop systems that are basically never shut down completely. Viewed the other way, at data centers there is always a possibility that a certain part is being replaced or is shut down for maintenance, meaning they have a fluid structure. It is difficult to control a massive system with a fluid structure from scratch via codes without any support, so there is a need for a software platform that supports the programming and execution of codes while taking these factors into consideration. We are progressing with the research and development of an operating system for data centers that assists software implementation aimed at data centers that enable flexible control through coordination with external systems, such as co-IZmoSD.

We began working on conceptual designs for a data center operating system together with the launch of co-IZmoSD, and we implemented the basic functions as a prototype in fiscal 2015. At this stage we are working on improving the quality of the software while running it on co-IZmoSD.

3.8 The Future of Software-Defined Data Centers

Data centers are already essential facilities for providing stable Internet services, and in the coming years their numbers and the scale of the IT servers they house are expected to continue growing. In light of this, the pursuit of more efficient data centers will be an unavoidable issue for Internet and cloud service providers.

In this article we provided a perspective on the data center research we are conducting at the IIJ Research Laboratory, and discussed the co-IZmoSD data centers we designed and developed and are operating to test the concept. The focus of our research using co-IZmoSD is already shifting to the software portion. In the future we will continue to perform research into IT system design and control techniques based on the premise of data centers that are made more sophisticated utilizing software that handles both data centers and IT systems in an integrated manner.

Author:
Yojiro Uo
Senior Researcher, Research Laboratory, IIJ Innovation Institute Inc.
Dr. Uo specializes in the research of network infrastructure systems (distributed systems, operating systems, data centers, etc.).