Dracena: Real-Time Digital Twin Platform

Tatsuro Matsumoto   Kazuki Matsui   Eishi Morimatsu   Miwa Ueki

Takashi Ohno   Yasuhiko Kanemasa

Advances in IoT and 5G technologies have led to growing interest in "digital twins" that digitalize the real world in real time so that this data can be transformed into value by storing it on the cloud where it can be used without needing to know the devices it came from. Key requirements for building a digital twin are the ability to process large amounts of stream data at high speed and to make additions or changes to the processing without interrupting services. In response, Fujitsu Laboratories has developed the Dynamically-Reconfigurable Asynchronous Consistent EveNt processing Architecture (Dracena) as a reactive platform that can be used to develop and deploy services speedily and flexibly by creating cloud-based digital twins of people, things, and contexts from the real world. The platform features the ability to process large amounts of stream data in real time and to make additions or changes to this processing without shutting down the system. This article describes the architecture of Dracena, the technologies developed for it, and the applications where it is used.

1. Introduction

Advances in IoT and 5G technologies over recent years are raising expectations for services that can make use, in real time, of the large flows of data arriving in the cloud from the many different types of devices located in factories, social infrastructure, and homes, thereby transforming this data into value. These services have already entered commercial use, with real-world data being collected in the cloud in large quantities and utilized for visualization, analysis, and prediction. It is hoped that they will also proceed to the next step in which data and services from various different industries are combined and developed into ways of controlling the real world in real time. One example from the field of connected cars that represents a step toward the practical implementation of autonomous driving is the work being done on how to analyze large amounts of data collected from vehicles, such as their speed and location, and how to provide drivers with information about hazards and other feedback.

One technology that will play an important role in enabling services like this is that of digital twins [1] (**Figure 1**) that convert the real world into digital data and link to digital spaces. Digital twins use the cloud to collect data on the real world (people and things) and provide the results of analyzing this data to services and applications. In addition, these digital twins represent this raw real-world data in contexts that are meaningful in actual business and life. Doing it this way makes it possible to develop services and applications efficiently, without needing to take account of the specific IoT devices that collect the real-world data they use.

Conventional data processing was performed on batches of data that were stored in a database. Dracena, on the other hand, performs sequential stream data processing of this large volume of data, enabling real-time analysis and prediction in the real world.

Fujitsu Laboratories has developed the Dynamically-Reconfigurable Asynchronous Consistent EveNt processing Architecture (Dracena) as a platform that can create cloud-based digital twins of people, things, and contexts and process large amounts of stream data in real time. Dracena can be used to develop and deploy services speedily and flexibly while allowing additions or changes to be made without shutting down the system.
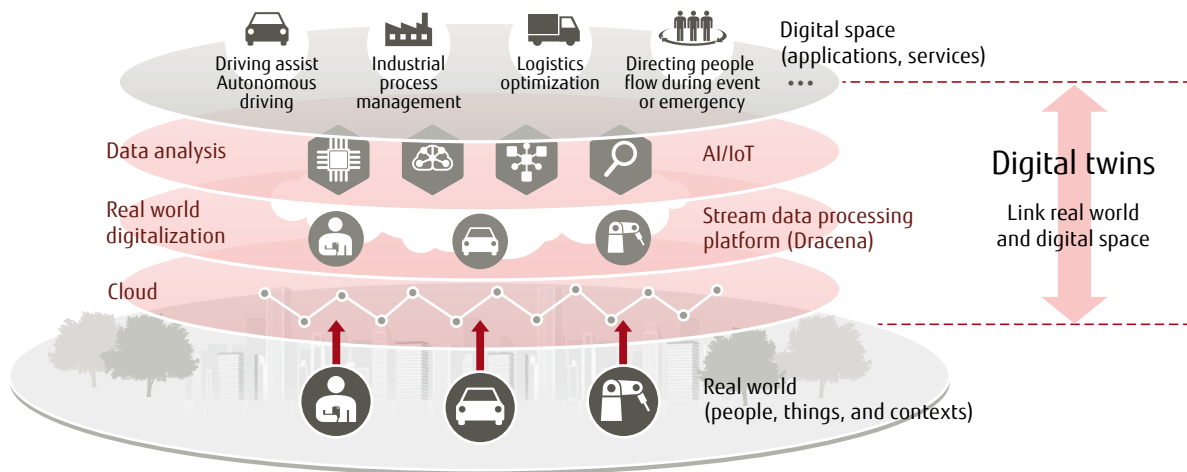
Figure 1
Digital twins.

This article describes the architecture of Dracena, the technologies developed for it, and the applications where it is used.

2. Existing technologies and associated problems

This section describes three of the problems with past techniques for stream data processing.

1) Difficulty of maintaining state information

Identifying when events occur in the real world requires that information be held on the current state so that changes can be detected in real time. Using the stream of data generated by vehicles to determine when a road hazard has arisen, for example, requires the ability to detect when the data collected from a number of vehicles driving on the same road changes from its normal state to a state indicative of a hazard. Unfortunately, past techniques have achieved high speed by only storing limited amounts of state information, or none at all (stateless operation) in memory. This means that the database needs to be accessed when information about the previous state is required, inevitably resulting in significant degradation of processing speed.

2) Difficulty of data sharing and interoperation

Past practice has been for systems to be vertically integrated into silos, with tight coupling all the way from devices to services. This has made interoperation with other services difficult. Furthermore, the way data streams are formatted is dependent on each service, and in order to use these data for general purposes, they need to be converted to an easy-to-use data format, which requires not only development costs but also machine resources to do the conversion.

3) Need to halt services when making additions or changes

Given the difficulty of anticipating what might happen in the real world, real-world services regularly need to make additions and changes to how they work. Such outages, however, are not acceptable even for a very short time for services that provide information for applications such as autonomous driving. The way this was dealt with in the past, by operating two systems and switching between them, only added to costs.

3. Dracena architecture and implementation

Overcoming the problems described above requires the ability to handle the large volumes of data from the real world in real time while also combining non-stop operation with the agile development and enhancement of different types of services that work with the people, things, and contexts of the real world. This section describes the architecture and implementation of the Dracena platform developed by Fujitsu Laboratories for this purpose.

3.1 Architecture

Dracena is a platform for the provision of timely services through the real-time processing of stream data transmitted to data centers at high frequency from real-world devices numbered in the millions or tens of millions, such as connected cars or smartphones. **Figure 2** shows the architecture of Dracena.

Dracena uses large amounts of stream data from the real world to maintain in-memory state information for "objects" that represent people, things, and contexts. The platform also has the facility for additional plug-in programs that can be invoked by changes to these in-memory object states. The capabilities of these plug-ins include being able to update their own states based on the program results and the passing of new event data to other objects to enable the chaining of data updating. These features give Dracena a reactive architecture in which processing is asynchronous and event-driven. They also make Dracena a new type of real-time processing platform for the cloud era that is scalable with high throughput and low latency. Two key features of Dracena are as follows.

1) Real-world objects and service objects

To prevent IoT systems from becoming silos, Dracena abstracts real-world data in real time on the cloud to provide flexible digital twins that represent people, things, and contexts, thereby allowing a data processing operation to be added that is specific to particular services [2]. It is also able to detect when events occur in the real world by storing the data received from devices as state information in the real-world objects, keeping it in both raw form and pre-processed for use by services.

Those objects that hold state information about real-world people and things that is comparatively generic and not service-specific are called "real-world objects" and those that perform service-related operations are called "service objects." Real-world objects can be shared by different departments within and between companies. In addition, common data processing among services can be made into microservices as service objects, which can be connected and linked together. Thus, efficient development and event processing are realized.

2) Seamless upgrading without service interruption

Dracena allows additions or changes to data processing programs to be made dynamically, without having to shut down online services and without compromising the scalable processing of real-time data that number in the millions. **Figure 3** shows a diagram of how service updates are performed without halting operation [3]. First, when an upgrade or modification is made to the data processing program, the Java class of the program is distributed to objects as a message in the same way as an event, shown as (1) in the figure. Next, the Java reflection mechanism is used to dynamically incorporate the new processing program, shown as (2). Finally, the service update is enabled by
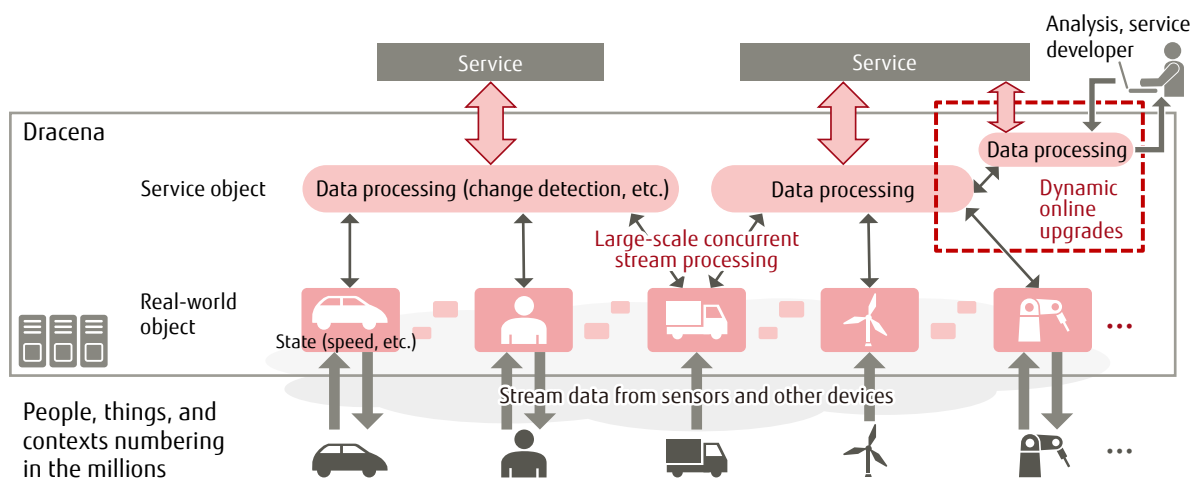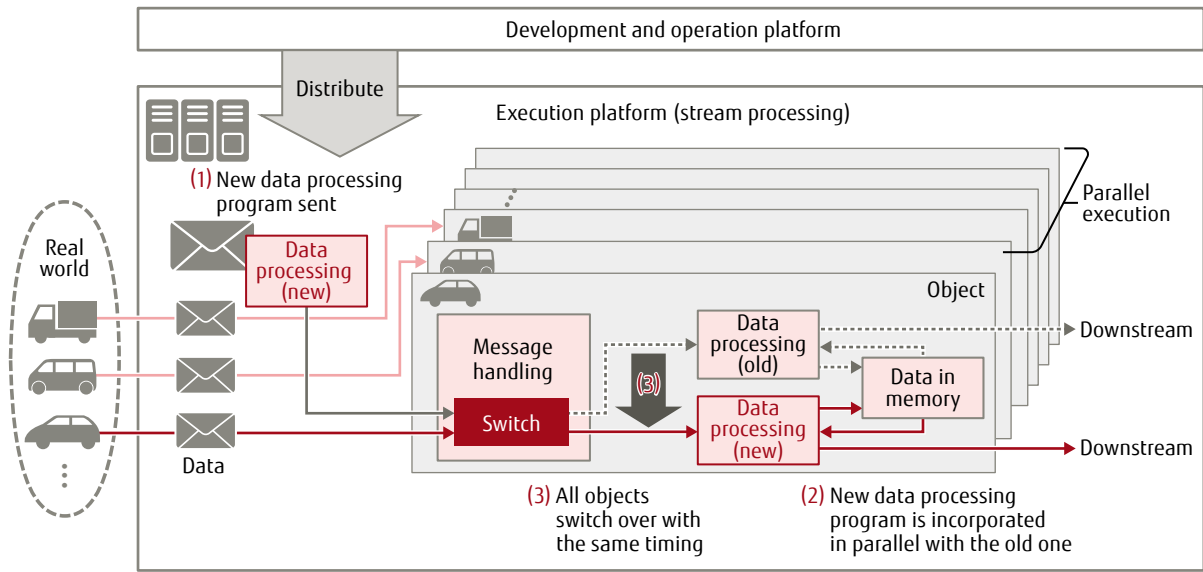

Figure 2
Dracena architecture.

Figure 3
How service updates are performed without halting operation.

coordinating the timing for switching over to the new processing program, shown as (3).

3.2 Implementation

The Dracena implementation involves an optimal combination of the latest open-source software, including Apache Kafka [4] for distributed message queuing and the Apache Flink [5] distributed stream processing engine. These serve as a base on which the objects and plugins that comprise real-time digital twins are uniquely implemented. This approach to implementation is in line with international trends in this era of digital transformation that takes maximum advantage of open source. Dracena achieves maximum performance by optimizing resource use through the detailed analysis of the performance characteristics of the overall system made up of these various different elements.

4. Potential applications

This section looks at typical applications where Dracena is used.

4.1 Mobility

Figure 4 shows a Dracena application in the mobility sector. By the collection and analysis by Dracena of sensor data from large numbers of network-connected cars, a next-generation traffic information service can

be provided that obtains real-time information not only about vehicles and drivers but also on road obstructions and other aspects of the ever-changing outdoor environment over a wide area.

One example would be to detect when frequent instances of sudden braking occur on a particular section of road and use this to warn approaching vehicles. It is also possible to quickly identify and address the cause of the problem by analyzing image data collected from passing vehicles. Another possibility would be the accurate prediction of potential flooding based on windscreen wiper usage and weather data and the use of this information to provide safe route guidance.

4.2 Logistics

The spread of e-commerce and the globalization of manufacturing are increasing the complexity of logistics all of the way from upstream to downstream of the distribution chain. Obtaining real-time information on ever-changing real-world conditions over a wide area can help improve work safety and efficiency in industries like this, including such information as the status of freight at all steps along the supply chain and the progress of work at the distribution centers and warehouses along the way.

By collecting information on the locations of delivery trucks and performing analyses that combine this
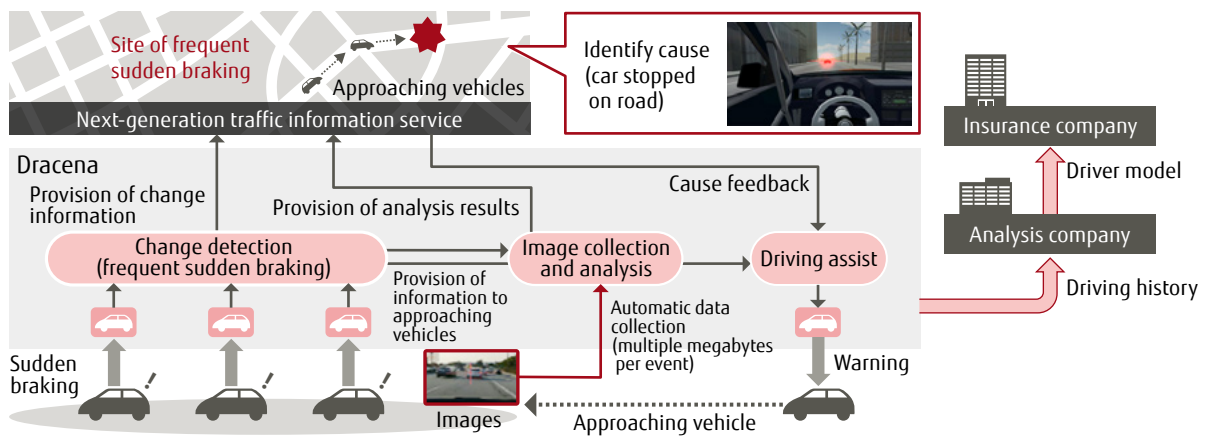
Figure 4
Mobility application.

with other data, such as continually changing traffic conditions and the availability of space for unloading goods at the destination warehouse, it is possible to update and improve logistics planning in real time. Real-time analysis of the status of warehouse forklifts and staff can also be used to improve safety or to make work more efficient.

Against a background of urbanization and aging demographics, the logistics industry has been trailing practices such as the use of robots in warehouses and autonomous vehicles or drones for last-mile delivery. This indicates how Dracena will become increasingly important in the future as a platform for digital twins at workplaces where reforms are being undertaken in ways that involve the coexistence of people and systems.

4.3 People flow

The spread of smartphones in recent years has made it easier to track people flow. This information can provide the basis for applications such as reducing crowding at stadiums or other event venues by tracking the flow of large numbers of people and guiding them toward routes with fewer people. Another potential application for tourist sites would be to determine where people currently are, where they have come from, and where they stopped along the way, and to use this to recommend other places to visit.

One consideration in applications that deal with the flow of people is the difficulty of coming up with recommendations that suit particular individuals.

Unlike mobility and logistics, providing people with suggestions chosen to maximize convenience may not be what they want.

This is where the ability of digital twins to provide real world feedback comes into play. The characteristics of individuals can be learned by tracking how they act in response to feedback, and algorithms [6] can be developed and implemented that offer recommendations that better suit them.

5. Future challenges and opportunities

Fujitsu Laboratories intends to continue with technical development aimed at overcoming the following two challenges.
1) Provision of an efficient development environment

To boost the efficiency of development for services that interact with the real world, Fujitsu Laboratories plans to offer a framework that developers can use for the simple GUI-based design and development of real-world and service objects.
2) Rapid acquisition of real-world models

The building of digital twins requires models that use the collected data to detect the status of people, things, and contexts as well as changes of state. Most current AI and machine learning techniques for model acquisition work by batch processing, posing the risk that models, once acquired, will soon be left behind by rapidly changing external circumstances. Accordingly, Fujitsu Laboratories is developing techniques for dynamic model building from real-world stream data so

as to progressively keep models in step with what is actually happening.

6. Conclusion

This article described the Dracena reactive platform for the real-time processing of large amounts of real-world stream data. Dracena enables the flexible and timely development and enhancement of services by allowing additions and changes to be made to stream data processing without the need for a temporary shutdown of the entire service.

In the future, Fujitsu Laboratories intends to expand the scope of application of Dracena while also looking toward the next generation of digital twins that incorporate features such as automatic learning for models of the real world.

All company and product names mentioned herein are trademarks or registered trademarks of their respective owners.

References and Notes

[1] M. Grieves and J. Vickers: "Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems." Trans-Disciplinary Perspectives on System Complexity, 2016, Springer: Switzerland, pp. 85–114.
[2] H. Yamaoka et al.: "Dracena: A Real-Time IoT Service Platform Based on Flexible Composition of Data Streams." 2019 IEEE/SICE International Symposium on System Integrations (SII 2019).
[3] G. Nakagawa et al.: IoT event processing platform using Flink for dynamic updating. DEIM2018-C7-2, 2018 (in Japanese).
[4] Apache Kafka.
    https://kafka.apache.org/
[5] Apache Flink.
    https://flink.apache.org/
[6] T. Imai et al.: Ubiquitous Recommender System with Models of User Awareness. REVEAL '19, Sept. 2019.

Tatsuro Matsumoto
Fujitsu Laboratories Ltd., Super Middleware Unit
Mr. Matsumoto is currently engaged in the development of application technologies for real-time digital twins.

Kazuki Matsui
Fujitsu Laboratories Ltd., Super Middleware Unit
Mr. Matsui is currently engaged in the development of real-time digital twins.

Eishi Morimatsu
Fujitsu Laboratories Ltd., Super Middleware Unit
Mr. Morimatsu is currently engaged in the development of application technologies for real-time digital twins.

Miwa Ueki
Fujitsu Laboratories Ltd., Super Middleware Unit
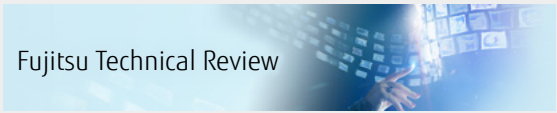Ms. Ueki is currently engaged in the development of platform technologies for real-time digital twins.

Takashi Ohno
Fujitsu Laboratories Ltd., Super Middleware Unit
Mr. Ohno is currently engaged in the development of application technologies for real-time digital twins.

Yasuhiko Kanemasa
Fujitsu Laboratories Ltd., Super Middleware Unit
Mr. Kanemasa is currently engaged in the development of platform technologies for real-time digital twins.

This article first appeared in Fujitsu Technical Review, one of Fujitsu's technical information media. Please check out the other articles.

Fujitsu Technical Review

https://www.fujitsu.com/global/technicalreview/