Optimizing cloud environments over a wide-area network
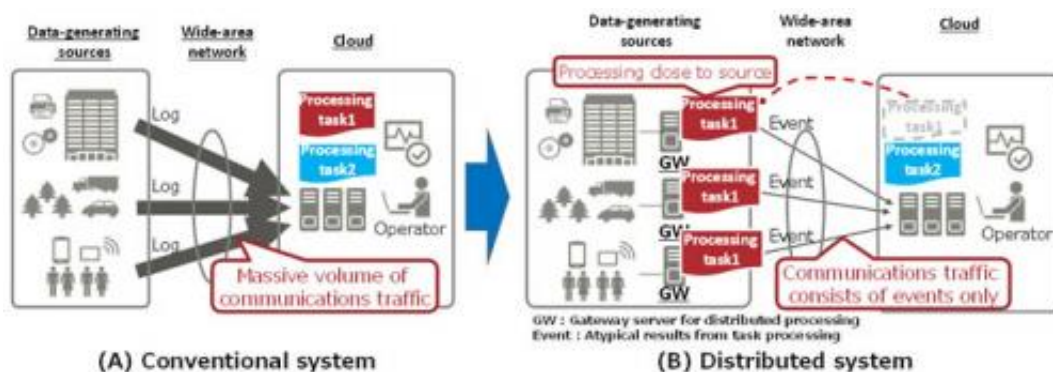
March 14 2014

Figure 1: Approach used in a distributed system

Fujitsu Laboratories Ltd. today announced, in an industry first, it has developed distributed service platform technology that automates system architecture and operations in line with changes in service conditions, and conducts distributing deployment of cloud-based processing and data to servers within a wide-area network.

With recent advances in "machine to machine" (M2M) technology, in which machines communicate among themselves without human intervention, and with the widespread use of mobile devices, there has been a surge in the amount of data exchanged between cloud services and devices. As a result, to deliver services in real time given limited network bandwidth, it is necessary to distribute processing to intermediate servers situated either near where the data are generated or

over a wide-area network.

To facilitate this, Fujitsu Laboratories has now developed a distributed service platform for which, in accordance with varying service requirements, such as changes in the number of devices that can be connected, the optimal location to deploy is determined, with deployment conducted automatically. This cuts the construction time required for redeployment of a distributed system from several days to several minutes. And because it can also accommodate the operations of systems comprised of several hundreds of thousands of devices and servers, which are impossible to accommodate manually, cloud services that start on a small scale can be automatically scaled up. Such scaling up can be done as the need to spend time on such tasks as revising applications becomes redundant.

Background

In recent years, as sensors and mobile devices have become smaller and communications technology has become more sophisticated, increasing attention is being focused on M2M services, in which information from various equipment or sensors connected to a network is transmitted to cloud services to enable, for example, energy management or remote maintenance of equipment. As the number of sensors or devices connected to a network expands, the volume of communications traffic increases. Accordingly, to maintain a service's response performance, there is a need to strengthen the cloud system and communication network equipment, but the added cost entailed is becoming a major problem.

In response, distributed systems are garnering increasing attention. With distributed systems, a portion of the cloud-based processing or data is distributed for processing to intermediate servers known as "gateway servers," which are connected to the network and situated close to where

the sensors or devices are located. This is designed to reduce the volume of communications traffic and improve response performance.

Technological Issues

In distributing a portion of the cloud-based processing or applications, to maximize the reduction in communications traffic, it is important to design in advance which gateway servers, among the multiple gateway servers along the network, are most effective for executing the processing tasks. Determining the most effective gateway servers to deploy requires a comprehensive evaluation of a number of variables, such as the volume of data generated by the devices, the location of the gateway servers on the network, the nature of the processing tasks, and the cost of communications on the network. Using current manual methods, there have been several problems, including higher costs associated with longer design times, and the difficulty of responding rapidly to sudden changes in data volumes. Another problem that has emerged is that manual design methods have not been able to keep up with increases in the number of devices or gateway servers.
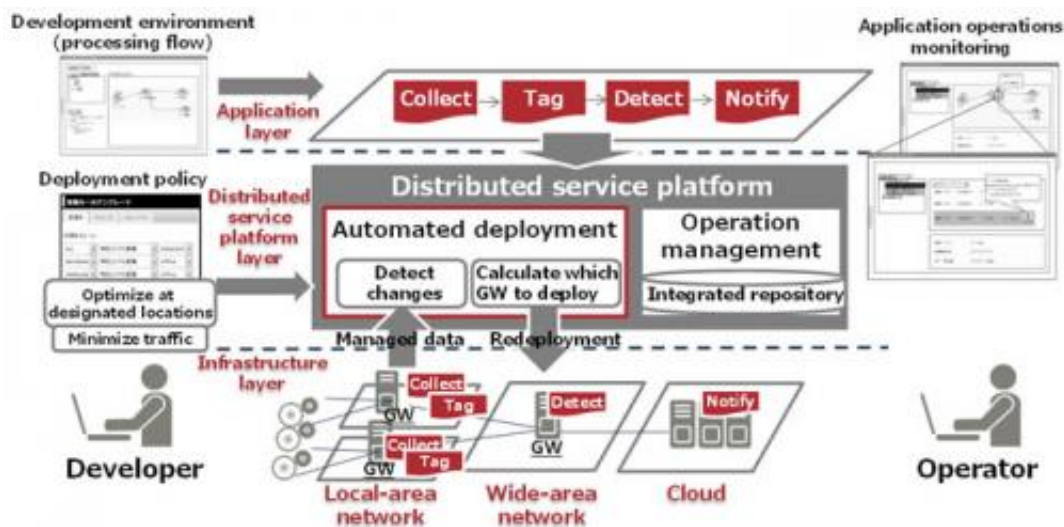

Figure 2: Distributed service platform

About the New Technology

For a distributed service platform that enables a distributed system, by detecting changes relating to the infrastructure, such as increases or decreases in the number of devices or volume of data, and repeatedly calculating the optimal deployment configuration and redeploying equipment accordingly, it is possible to achieve an optimized distributed deployment that keeps pace with changes in the execution environment. Accordingly, for large-scale systems applications, it is important to increase the efficiency of automated deployment to increase, for example, the number of devices collecting data or the number of gateway servers.

Now Fujitsu Laboratories has developed a distributed service platform that automates distributed deployment for application processing. This platform can also be applied to large-scale systems.

The technology that increases the efficiency of automated deployment has the following features.

1. Technology that accelerates the calculation of deployment configurations

Ordinarily, the calculations required to determine which gateway server is to conduct processing to achieve the optimal combination are very time-consuming. Now Fujitsu Laboratories has developed an algorithm that performs a combined search to determine whether it is best to prioritize the gateway server with the shortest network path depending on the characteristics of the processing task to be assigned, or whether it is best to prioritize the gateway server that is closest to the device (see

section A in figure 3). As a result, the time required for calculation was cut to approximately 1/500th of current levels.

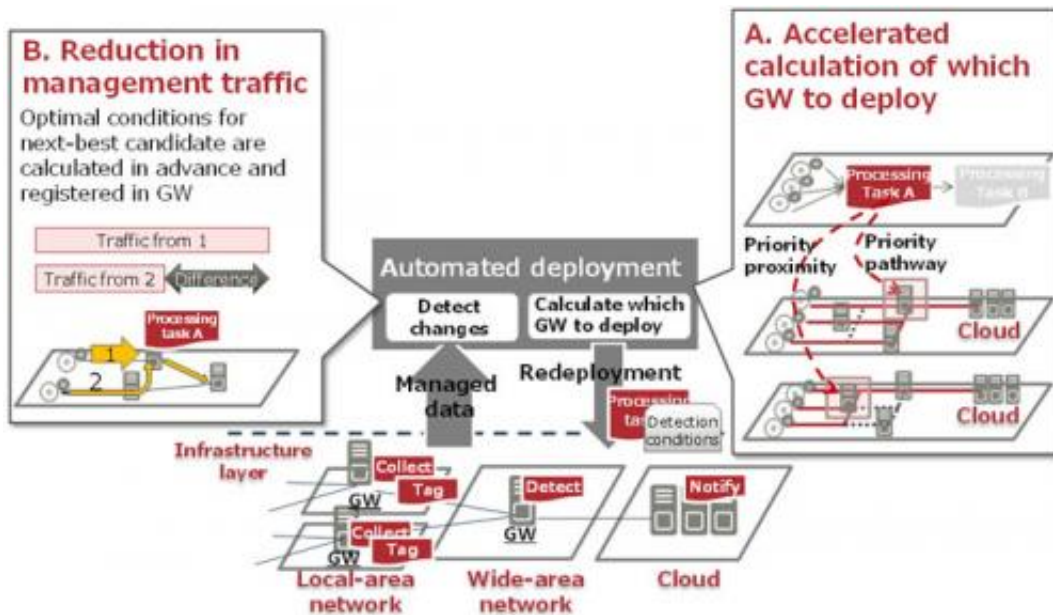## 2. Technology for reducing the volume of traffic managed

Figure 3: Technology that increases efficiency of automated deployment

For redeployments to keep immediate pace with infrastructure changes, it is necessary to frequently send management information to a server monitoring the system (the monitoring server) that indicates the status of the infrastructure. The increase in management traffic this entails, however, is a problem, particularly for large-scale systems. Fujitsu Laboratories has developed technology that, in addition to calculating the optimal candidate for deployment, also calculates the next-best candidate, and even calculates the change in traffic conditions that are optimal for the next-best candidate. By deploying the calculation results

and the processing task to the gateway server, only changes necessary for redeployment are transmitted to the monitoring server (see section B in figure 3). As a result, because minor changes that do not affect the redeployment are not sent to the monitoring server, management traffic was able to be reduced to approximately 1/700th the level achieved using the previous approach, which required periodic transmissions.

Results

In this distributed service platform, the automated deployment structure and the operations management environment are unified. The platform has been designed so even a tool for defining the application processing flow can be used. By using this platform to respond to changes in the infrastructure environment, the time required for the redesign, configuration, and operations work, which used to take several days using a manual approach, can now be cut to just several minutes. Moreover, the redeployment of systems comprised of several hundreds of thousands of devices and servers, which are impossible to accommodate manually, can also be performed in a few minutes. As a result, service providers can swiftly provision and cycle through the operations lifecycles of distributed systems, and, by optimizing distributed processing, can reduce by approximately 1/100th the overall volume of communications traffic.

Fujitsu Laboratories plans to adapt the technology for multitenant applications, in which a cloud system houses multiple services, and tie it into open standard technologies, such as software-defined networking, which enables dynamic changes to network structures. The aim is to commercialize the technology within fiscal 2014.

Provided by Fujitsu