IT Resource Management Technology for Reducing Operating Costs of Large Cloud Data Centers

Yukinori Sakashita
Yutaka Kudo, Ph.D.
Masataka Nagura, Ph.D.
Takato Kusama

OVERVIEW: Growing use of the cloud is making data centers larger and more complex. This is increasing operating costs and placing a greater burden on the administrators charged with managing data center operations. In response, Hitachi is taking steps to deal with the three key challenges faced by administrators: the difficulty of centralized management of increasingly large data centers, the growing workload being placed on individual administrators due to the specialized nature of advanced management skills, and the increasing amount of time spent on coordinating administrators of different layer. To overcome these challenges, Hitachi has implemented management repository technology that allows centralized administration of large-scale IT resources, root cause analysis technology that formalizes the advanced know-how of administrators in the form of structured knowledge, and virtual server and storage administration coordination technology that automates storage configuration and reduces the workload associated with coordinating different administrators.

INTRODUCTION

THE quantity of digital data has grown explosively in recent years as mobile devices such as smartphones and tablet personal computers (PCs) have proliferated. The total quantity of digital data in the world is forecast to reach 73 Zbyte (Z: 10^{21}) by 2020, of which it is estimated 44.4% will be stored in the cloud[(1)]. This is approximately 10 times larger than the quantity in 2012. The emergence of the cloud is also driving a shift away from ownership and toward sharing of information technology (IT) resources, with the aim of cutting costs by consolidating these resources in data centers rather than having them dispersed as they have been in the past.

Recognizing this growth in use of the cloud, Hitachi supplies a wide range of products extending from the servers, storage, and other IT resources used at data centers through to Hitachi cloud computing solutions.

This article identifies three challenges faced by administrators seeking to cut the operating costs of data centers, which are becoming larger due to growing use of the cloud, and describes IT resource management techniques for overcoming them.

CHALLENGES FOR IT RESOURCE MANAGEMENT

While data centers are getting larger, the number of administrators available to manage them remains roughly constant. This has created a need to reduce the workload imposed on administrators so as to cut operating costs.

In response, Hitachi has identified three key challenges faced by the administrators of these increasingly large data centers (see Fig. 1).

The first challenge is that the centralized management of IT resources has become difficult. The growing size of data centers has led to an explosion in the quantity of IT resources that need to be managed, making it difficult to continue with the centralized management practices used in the past to deal with configuration, performance, and other administrative information. By impeding a clear understanding of the overall situation at large data centers, this makes timely administration impossible.

The second challenge is the specialized nature of advanced management skills. In addition to data centers becoming larger, advances in virtualization technology for servers and storage are making system configurations more complex. Even more so than in the past, this demands that administrators acquire advanced know-how and extensive experience. In many cases, however, companies are unable to obtain enough administrators with adequate experience. This is placing a growing workload on individual administrators with specific advanced management skills, making it difficult to get many tasks completed quickly.

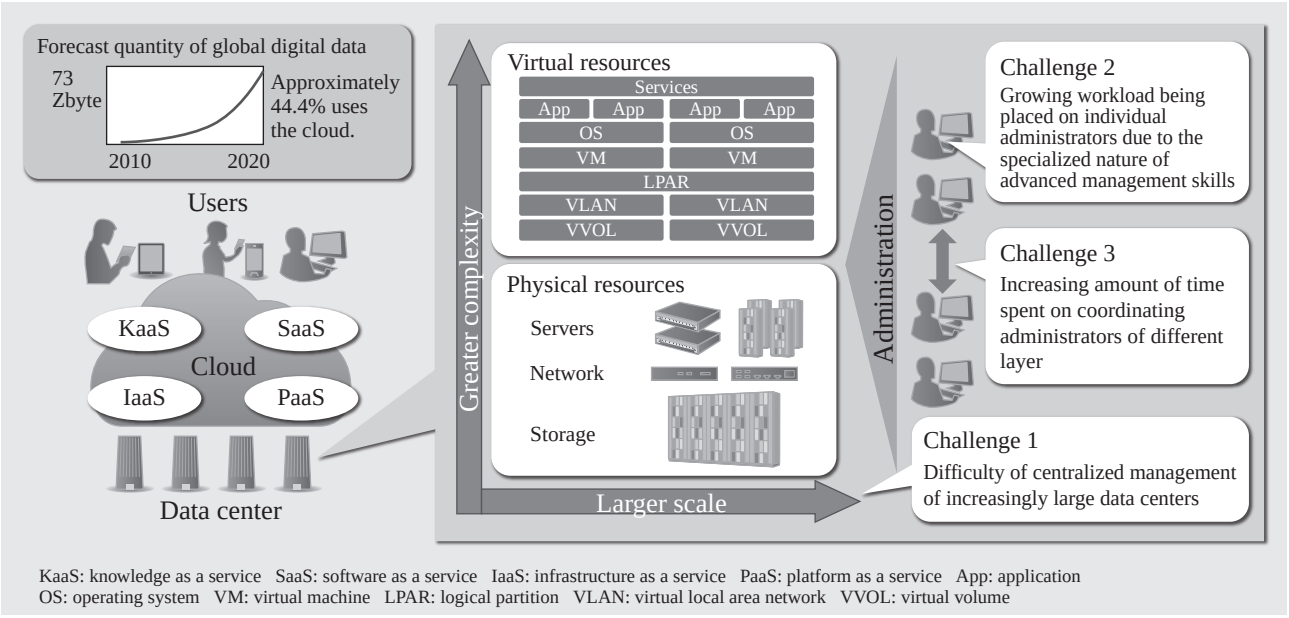

Fig. 1—Challenges for Administration of Large Cloud Data Centers.
As use of large cloud data centers grows, new challenges are arising in their administration due to the growing scale and complexity of both physical and virtual resources.

The third challenge is the increasing amount of time spent on coordinating administrators of different layer. As data centers become larger, administrators' working practices are becoming increasingly compartmentalized based on factors such as the type of resources being administered or the services provided. This creates situations where the work of a number of administrators with different roles must be coordinated. Not only does this human involvement slow the time taken for work to be completed, it is also sometimes a cause of faults as their differing responsibilities result in administrators' know-how becoming more specialized, impeding the exchange of information between them.

REPOSITORY TECHNOLOGY FOR MANAGEMENT OF LARGE-SCALE IT RESOURCES

This section describes the use of repository technology for management of large-scale IT resources as a way of dealing with the first challenge.

As data centers become larger and more complex, this is increasing the quantities and types of IT resources being managed, and also the amount of data handled by the management software that administrators use in their daily work. This has led to problems such as insufficient memory or slow execution for the management software that has been used in the past for centralized management, making this approach to administration increasingly impractical. The main cause of this lies in the databases (DBs) used for storing and searching of configuration information for IT resources. Typically, management software stores configuration information acquired from a particular managed IT resource in a corresponding table in the DB (so that server configuration information is stored in a server table and storage configuration information is stored in a storage table, for example). However, as the quantity of data being handled by management software grows, this increases the number of tables that must be dealt with for a single operation, such as displaying a report screen that gives an overview of the entire data center. It is this that results in the DB running out of memory or being slow to execute.

A common solution to this problem is to index the DB. However, as the quantity of data and number of tables increases, the number of indexes that need to be specified also rises, and this causes its own problems with the execution time required to generate the indexes and with the quantity of data they contain. An estimate based on the quantity of IT resources at a large data center puts the size of the required indexes at more than 100 Gbyte. In response, Hitachi has developed repository technology for the management of large-scale IT resources that minimizes memory use and supports fast searching (see Fig. 2). This technology analyzes the main use cases applicable to data center administration to identify which data is actually needed, then generates consolidated tables

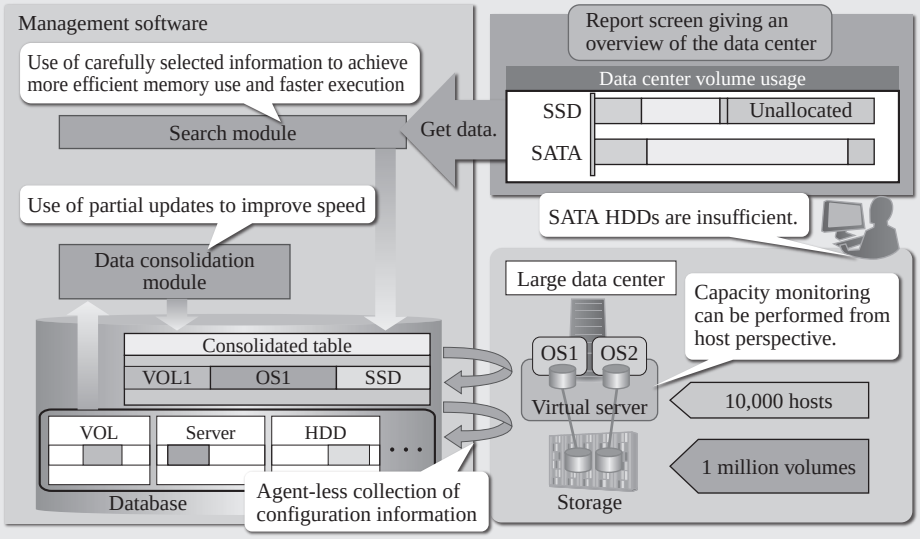

Fig. 2—Platform for Management of Configuration Information Used in Administration of Large-scale IT Resources.
The way data is stored is consolidated based on administrator use cases. Report screens able to provide an overview of the entire data center are made possible by fast search performance.

based on the results. By only defining indexes for these consolidated tables, the repository technology for the management of large-scale IT resources is able to combine efficient use of memory with fast operation. Using this technology, Hitachi has implemented an IT resource management repository with world-class scalability (approximately 40 times that of the previous Hitachi system).

ROOT CAUSE ANALYSIS TECHNOLOGY

This section describes the use of root cause analysis technology as a way of dealing with the second challenge.

Larger and more complex data centers require administrators with more advanced know-how and more extensive experience than in the past. However, companies are often unable to assign enough administrators with adequate experience. Among the different types of administration work, this creates particular difficulties for fault recovery work, which demands a rapid response. To reduce the recovery time, shortening the time between detecting the failures and identifying the root cause is very important. Accordingly, Hitachi has developed root cause analysis (RCA) technology for identifying the cause when a failure occurs.

RCA defines general rules for the patterns of where failures occur and the corresponding root cause. When failures are detected, the general rules are retrieved by looking for matches with the incoming failure events. Using the configuration information for any general rules found by this retrieving, analysis rules that specify the exact relationship between the failure events and the equipment affected are generated and the root cause identified. The technology also calculates a certainty of the root cause based on the rate of failure events actually received from among those failure events contained in the analysis rules that could have been expected to be received when the failure occurred. Together with the certainty factor, the root cause is then reported to the administrator as a candidate for the cause of the problem.

Typically, when a fault occurs on a large system, it only affects a subset of equipment. Accordingly, the generating of analysis rules is only performed for equipment that could potentially be affected by the fault (determined from information on connections between devices). Also, generation of analysis rules is performed on-demand for specific equipment by using the limited general rules associated with received events, including only those rules that are required for analyzing the root cause. This allows root cause analysis to be performed quickly (see Fig. 3).

The connection relationships between equipment at large data centers are complex, and a fault occurring at such a data center often results in a very large number of failure events being detected. This makes it very difficult to determine which failure events are related and which failure events to consider when indicating the root cause. To deal with this, Hitachi investigated in advance how failure events occur under the customer's system to generate the general rules that can then be used as the basis for the analysis of failure events when an actual fault occurs. This shortens the time taken from fault detection until the root cause is identified. The technology also analyzes faults on

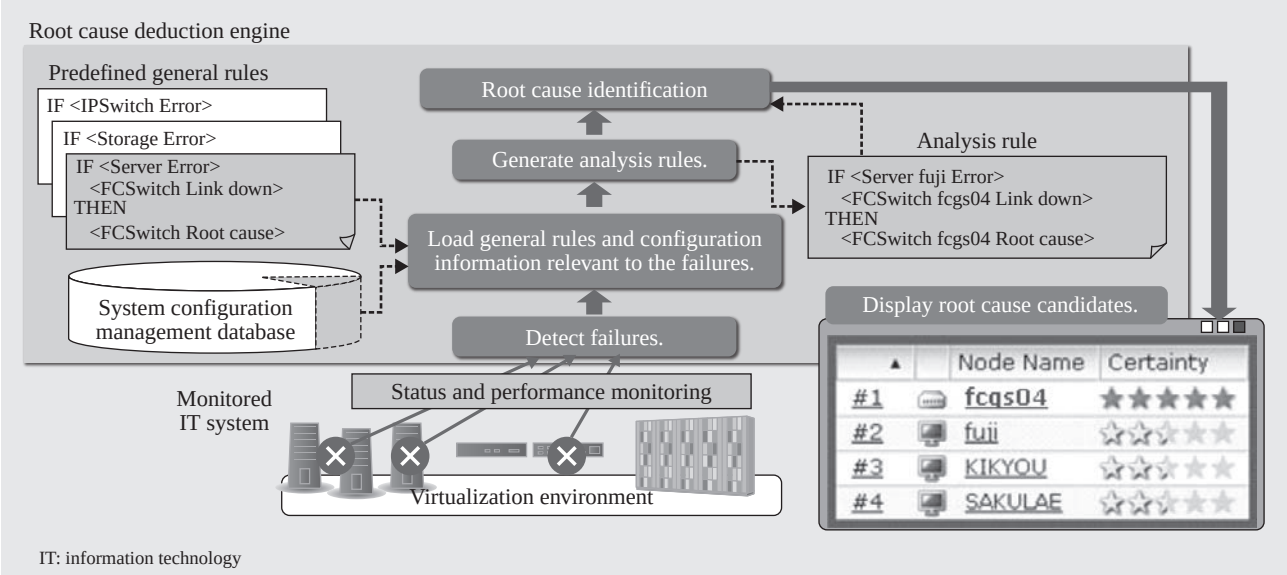

Fig. 3—Procedure for Root Cause Analysis.
This system performs rapid root cause analysis for large cloud data centers by loading the configuration information for the devices that connect to the devices which the failure events are detected, and then analyzing by using analysis rules generated from general rules and this information.

servers, storages, and network equipment. Identifying the root cause becomes even more difficult in cases when a fault on one of these devices affects other types of device. Hitachi's technology can analyze the cause of failure events occurring rapidly, even if the site does not have a separate administrator dedicated to each type of device.

In this way, the cost of operating large cloud data centers can be cut by minimizing the workload associated with responding to faults.

COORDINATION TECHNOLOGY FOR VIRTUAL SERVER AND STORAGE MANAGEMENT

This section describes the use of coordination technology for virtual server and storage management as a way of dealing with the third challenge.

In the cloud, it is necessary to be able to allocate IT resources rapidly when users require them. However, as the administration of large data centers is split between administrators of different layer, some companies adopt practices whereby an in-house workflow must be followed to coordinate this work between the different administrators. In the case of services that take an infrastructure-as-a-service (IaaS) approach to the provision of virtual servers, for example, the administrators responsible for virtual servers and storage respectively need to work together on a system configuration in order to provide both the virtual servers for the system and the storage used to hold its data. This involvement of people in the process makes it difficult to achieve the prompt service expected from the cloud.

In response, Hitachi has developed technology for coordinating administration that eliminates the need for communication between these administrators, and

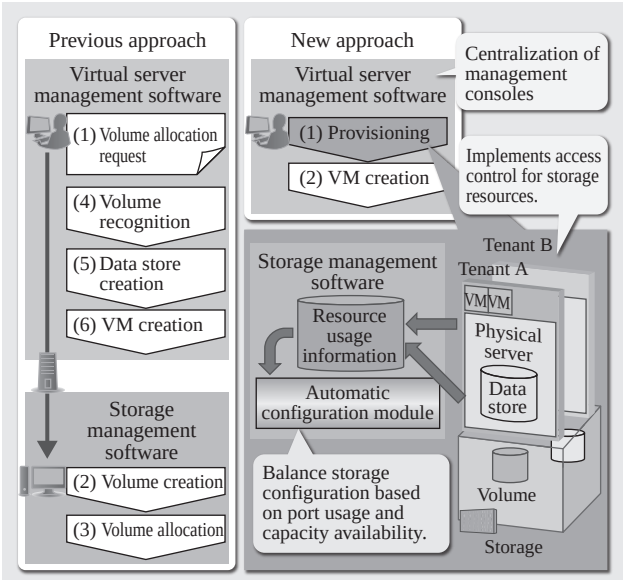

Fig. 4—Storage Administration Technology for Use by Virtual Server Administrators.
The system includes an automatic configuration module for storage that allows even virtual server administrators who lack expertise in storage to perform tasks that in the past would have required the virtual server administrator and storage administrator to work together.

allows the administrator responsible for virtual servers to allocate the virtual server and storage IT resources on his or her own (see Fig. 4). Rather than simply consolidating the administration functions for virtual servers and storage, this technology uses an automatic configuration module to set resource parameters automatically based on a dynamic assessment of storage usage. This means that even a virtual server administrator who lacks storage expertise can configure a system's storage. This shortens the time taken to perform the administration task to a few minutes, instead of the several hours or more required for communication between administrators.

CONCLUSIONS

This article has described Hitachi's repository technology for management of large-scale IT resources, root cause analysis technology, and coordination technology for virtual server and storage management, and how these technologies relate to three specific challenges facing the administration of large cloud data centers.

A common feature of all three technologies is that they reduce the administration workload. Reducing the work associated with tasks frequently performed by administrators is important for reducing the cost of operating increasingly large and complex data centers. Hitachi has plans to reduce these workloads even further in the future. The technologies described in this article have been released as part of the Hitachi Command Suite and Hitachi IT Operations management software.

REFERENCE

(1) Gartner, Inc. Report, "Digital Destination 2020," (Apr. 2009).

ABOUT THE AUTHORS

Yukinori Sakashita

Joined Hitachi, Ltd. in 2003, and now works at the IT Management Research Department, Yokohama Research Laboratory. He is currently engaged in research on IT operations management systems. Mr. Sakashita is a member of the Information Processing Society of Japan (IPSJ) and the chair of Technology Committee of the Storage Networking Industry Association Japan Forum (SNIA-J).

Yutaka Kudo, Ph.D.

Joined Hitachi, Ltd. in 1995, and now works at the IT Management Research Department, Yokohama Research Laboratory. He is currently engaged in research on IT operations management systems. Dr. Kudo is a member of the IPSJ and The Institute of Electrical Engineers of Japan (IEEJ).

Masataka Nagura, Ph.D.

Joined Hitachi, Ltd. in 2009, and now works at the IT Management Research Department, Yokohama Research Laboratory. He is currently engaged in research on IT operations management systems. Dr. Nagura is a member of the IPSJ and the Japan Society for Software Science and Technology (JSSST).

Takato Kusama

Joined Hitachi, Ltd. in 1999, and now works at the Architecture Design, Integrated Platform Systems Development Operation, IT Platform R&D Management Division, IT Platform Division Group, Information & Telecommunication Systems Company. He is currently engaged in the research and development of IT operations management systems.