Prevention Method of Recuring System Failure with Sharing Know-How in Large-scale Data Center

Hiroyuki Nishino (1210041)

School of Information Science,
Japan Advanced Institute of Science and Technology

February 12, 2014

Keywords: Data Center, Large-scale, Sharing Know-How, Root Cause Analysis, Server Management.

This paper proposes the method of sharing know-how about the system failure in the past for the purpose of preventing recurrence of system failure. The large-scale data centers bigger than $500m^2$ are increasing of late years. The system in such data centers is too big for each system manager to comprehend all of the system by him or her self. That is the reason why decentralized management is general in today's data center. In addition, data centers are suffered from chronic lacking of managers recently. Therefore steady system management is tend to depend on each system manager.

In this situation, the influence of system failure is tend to spread out in every direction. To make matters worth, they often spread over other manager's departments. It is too difficult for system managers to restore through their own resources. Naturally, There are many methods of "Root Cause Analysis (RCA)" for the purpose of supporting repair works. Although those RCA methods are able to output candidate list of Root Cause, just because system managers can specify the root cause does not recessarily mean they can restore the system failure. There is the possibility of existing other element which is relating with the system failure. It calls "Obstacle Outbreak Reason" in this research. Obstacle Outbreak Reason

is hard to detect by using RCA because there is no dependence relationship between The Spot of Root Cause and The Spot of Obstacle Outbreak Reason. For that reason, all of system managers must learn know-how about finding The Spot of Obstacle Outbreak Reason in daily operational services. Even though it is urgent business, there is seldom chance to learn such know-how because repair works are tend to be executed by skilled hand.

As mantioned above, The scale of data centers is getting big, and the systems are decentralized. Thereupon this research aims at preventing recurrence of obstacles which has been happened in the past by sharing know-how about Obstacle Outbreak Reason. In this research, presented data set which is included know-how is defined as " Know-How Information". The proposal technique are constituted following two points. First, accumulation of Know-How Information. Second, presentation of Know-How Information.

In this research, an specific spot of system is expressed as a data set named "Object". Each of objects contains "Object-ID" and "Object-Type-ID". This paper postulates that all of spacific spots of system are expressed by Objects from now on. The Know-How Information includes the result of RCA. In addition, it needs items explaining about the situation and the detail of the past system failure. There are five items in Know-How Information. They are "The Spot of Root Cause", "The Operation of Root Cause", "Similarity Estimation Range", "Obstacle Outbreak Reason" and "The Spot of Reason". the Spot of Root Cause is input natural Object undisturbed. The Operation of Root Cause is input the operation chosen from past records by restoring manager. Obstacle Outbreak Reason is written by restoring manager. Other items are input being chosen Objects by the manager.

These Know-How Information are indicated when the operation which is same as the Operation of Root Cause (hereinafter called "The Target Operation"). However there is a possibility that too much Know-How Information would be indicated if only to compare the target operation with The Operation of Root Cause. Thereupon, the proposal method place Know-How Information in order with using objects in Similarity Estimation Range. In addition, it also browses The Spot of Reason for eliminating

unnecessary information.

System failures relating with Obstacle Outbreak Reason sometimes happen in the Research Center for Advanced Computing Infrastructure (Data center in this educational establishment). The simulations in the following are conducted with being based on the one of those precedents. That system failure had happened when the system manager has been setting the new disk array machine. Although the manipulations had been correct, The system failure has happened by the manipulations cooperating with the configuration of switch between the disk array machine and the virtual disk mounting on the machine.

First simulation validated the difference of output pattern list of Similarity Estimation Range in the two systems which have a same structure. Second simulation increased the number of element pattern of Object-Type-ID, and validated how increase the output pattern list of Similarity Estimation Range. The results shows that this poposal technique is able to indicate accurate Know-How Information in accordance with the system structure. Additionally, The results also show that the proposal method hold good for the system which is mixed various Object-Type-ID.

The precedent in this time treats the disk array machine as one Object. However there are a lot of Objects in the lower layer in the machine. For example, "Redundant Array of Independent Disks" is constituted by plural Disk Objects. Furthermore, These disks can be classified. In addition, The renewed data which is in the RAID disk may be saved in the different disks by "Snapshot". Although the proposal system can dissemble these Objects because they will not affect the system failure, the proposal method mirrors the intention of repaired manager by changing Similarity Estimation Range.

It depends on the system if there is an Obstacle Outbreak Reason there. It is hard for "Expert System" to specify Obstacle Outbreak Reason because there is no dependence relationship. That is the reason why Obstacle Outbreak Reason should be searched by human intellects. However it is also difficult for managers to specify Obstacle Outbreak Reason because of significant amount of information and decentralized management. In this situation, the proposal method is the best method for Preventing system failure happened in the past.