Grid Technology Applied to Disaster Recovery

KATO Masayuki

Abstract

Business Grid Middleware (GMW) was developed as part of the national project led by the Ministry of Economy, Trade and Industry. A functional summary and features of the middleware are described in this paper, as well as an introduction of a business implementation example for disaster recovery control utilizing GMW at Mazda Motor Corporation, with results of validation experiments cited. Furthermore, the WebSAM GlobalGridOrganizer, based on GMW currently being developed at NEC, will also be introduced in this paper.

Keywords

grid, disaster recovery, Recovery Time Objective (RTO) ,Recovery Point Objective (RPO) autonomy, virtualization, Zero Administration Archive (ZAR)

1. Introduction

Major natural disasters have been occurring throughout the world in recent years. Some of the more recent events still fresh in our minds include the Sumatra-Andaman Earthquake (commonly known as the 2004 Indian Ocean Earthquake) in December 2004, Hurricane Katrina in the southern states of the United States in August 2005 and the Java Earthquake that occurred in May this year. Major disasters that have in the past occurred only once in several to several tens of years have occurred in almost successive years, raising the social interest for disaster recovery action to an extremely high level.

Shifting our sights to organized activities by businesses and public institutions on the other hand, we see that many organizations are taking advantage of IT not merely as a tool but also as a foundation that supports the activities of such organizations, due to the permeation of the Internet and dramatic reductions in the cost of telecommunication lines. This means that in comparison with the past the social roles and responsibilities of IT have increased significantly, with disaster recovery strategies involving IT systems becoming an issue of greater interest for all those who use them, regardless of the scale of their organization.

Research on the grid technology, particularly the business grid technology, had been conducted with an objective for improving the efficiency of resource utilization and operability through the virtualization of IT resources as well as the implementation of autonomous controls. This paper introduces business grid middleware as an example of a technological application for disaster recovery strategies.

2. Summary of Business Grid Middleware and ZAR

The business grid middleware (hereinafter referred to as "GMW") is a middleware developed between the years of fiscal 2003 and 2005 in a national project known as the Business Grid Computing Project. A summary of GMW is shown in **Fig. 1.**

The functional summary of GMW, as well as a definition of business operations (Zero Administration Archive: ZAR),

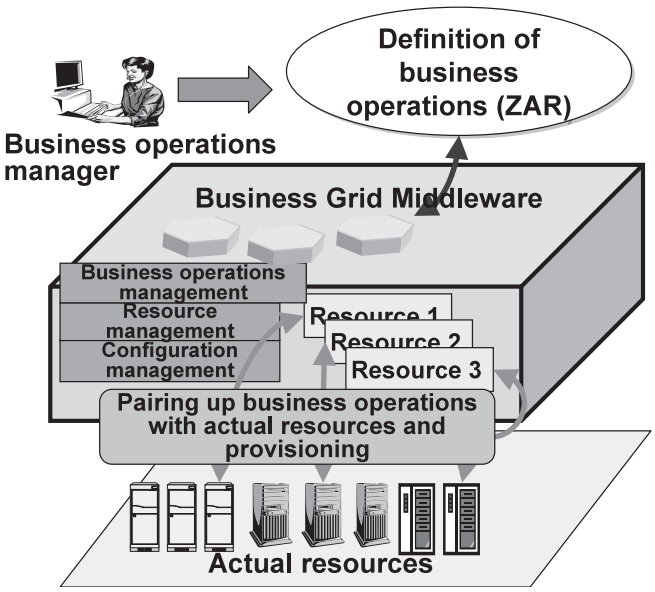

Fig.1 Summary of GMW.

which is an extremely critical concept for explaining the GMW, are described in this paper.

2.1 Summary of GMW

GMW was developed for the purpose of creating autonomous controls in data center environments, in order to allocate resources to users that are necessary for business operations, to carry out provisioning, as well as for responding to disorders and load fluctuations. Users prepare a definition for business operations (ZAR), which is described later in this paper. ZAR is entered in the GMW, which in turn allocates the necessary resources for the defined business operations and performs the provisioning. It is, furthermore, also possible to add or delete autonomous resources for dealing with server disorders or load fluctuations. Users must describe the relevant policies and actions when implementing autonomous controls, however, there is no need for users to be aware of the physical locations or attributes of the resources. Users only need to describe the types and quantities of the necessary resources, as well as the procedures for distribution and redistribution in a virtualized form, using ordinary expressions. This feature makes it possible to use an identical definition of the business operations and perform operations anywhere, as long as such locations are data centers managed by the GMW.

2.2 Structure and Roles for ZAR

Descriptions for the definition of business operations (ZAR), which have been mentioned a number of times in this paper, are provided next. The abbreviation "ZAR" stands for "Zero Administration Archive," thus named with the expectation of reducing operational management costs, as infinitesimally close to zero as possible. A ZAR is a file comprised of a collection of several components, which are compressed as a file in the ZIP format. The internal configuration of ZAR is shown in **Fig. 2**.

Many components in ZAR are text files in the XML format. Some of these components are singularly unique in the overall ZAR, while others are components that exist in each scope. A scope is a range in which the components that belong to an applicable scope are valid. For example, when a business operation is performed by its distribution over multiple sites, and if the executed applications are to be different for each of these sites, then a separate scope is described for each site. Disaster recovery also takes advantage of such flexibility in the descriptions of ZAR.

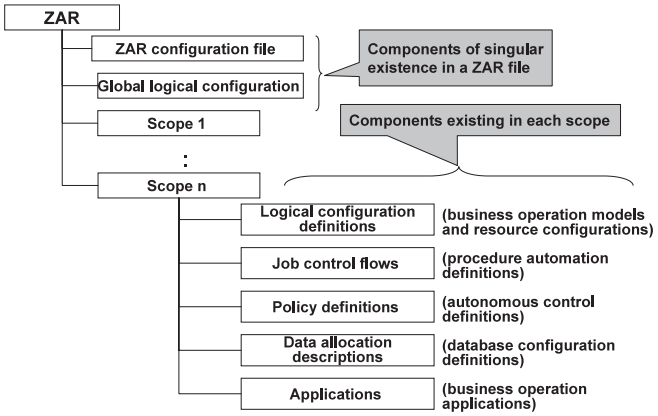

Fig.2 Configuration for definition of business operations(ZAR).

3. Features of Disaster Recovery with GMW

Features of disaster recovery control with GMW include (1) improvement of operability through the virtualization of resources, (2) resource adaptations provided by prioritized controls, as well as (3) acceleration of processing to high speeds through linkups with server synthesized functions.

With GMW, it is possible to manage resources from a perspective of business operations, without users being aware of the physical resources, through the virtualization of resources. For example, data replications are performed between a system in operation and a system in standby during disaster recovery operations, but in order to perform such data replications it is ordinarily necessary to have an operation design that depends on replication methods. With GMW, however, operational procedures are identical regardless of the particular replication method involved, freeing users from complex operational designs.

Resource adaptations with prioritized controls are functions that change the allocation of resources based on the priority of business operations. It is not necessary to prepare resources dedicated for standby purposes on secondary systems when this function is used. With ordinary disaster recovery systems it is necessary to have resources available for standby systems that are equivalent to those for the system in operation. Using GMW, however, resources can be adapted from business operations with a lower priority (for example from a system for development) when a disaster occurs, making it possible to use resources in an efficient manner.

Finally, conventional technologies incorporated into infrastructures, such as servers and storage devices, are linked up with the GMW out of consideration for their use in actual user

environments. Conditions for RTO (Recovery Time Objective) are extremely strict for mission critical business operation systems. Distribution of applications, on servers when a disaster occurs, results in a recovery time that lags too far behind the targeted recovery time. On the other hand, a technology that involves the preparation of a disk image that will be provisioned in advance and a server synthesized at high speeds by using such a disk image are recently becoming common practice. Such a technology is incorporated into products from NEC, such as WebSAM SigmaSystemCenter as well and GMW, which also offers a close linkup with such functions.

4. Validation Experiments

In the Business Grid Computing Project, validation experiments were held in collaboration with actual user businesses during middleware development. Such efforts were made in order to evaluate the GMW using the business operational systems of businesses, to validate the middleware for it to be adequate for use in business environments.

Such validation experiments were conducted with Mazda Motor Corporation and Nihon Keizai Shimbun, Inc. Details of validations conducted at Mazda Motor Corporation are described in this paper. A summary of these validations is shown in **Fig. 3**.

A validation experiment, conducted at Mazda, took into consideration two business operations, "domestic sales company dealership sales systems" and "universal data collection and distribution systems," as business operations subject to validation. Evaluations included verification that the system is able to achieve RTO and RPO comparable with the costs involved for business operations with both high and low service level agreement (SLA), as well as functions for reducing the scale for business operations with a lower priority and reallocating resources for business operations with a higher priority. In all cases, recovery was completed within the targeted time frame, enabling us to confirm that the features of the aforementioned GMW is effective for the business operational environment of actual users. A part of the evaluation results is shown below.

1) Data Replication Comparable with SLA
An RTO of 87 minutes and 32 seconds, as well as a RPO of 51 minutes and 40 seconds were achieved for a business operation with high SLA.

2) Resource Adaptations According to Priority
We confirmed that business operations can be restored through the adaptation of resources from business operations with a lower priority in backup sites (operating time of ten minutes).

5. Implementations to Product Businesses

The GMW, which was developed by the national project, is to be commercialized as the WebSAMGlobalGridOrganizer (hereinafter referred to as the "GGO") at NEC. GMW offers various features and advantages described in this paper, however, a number of items need to be improved in order to pro-

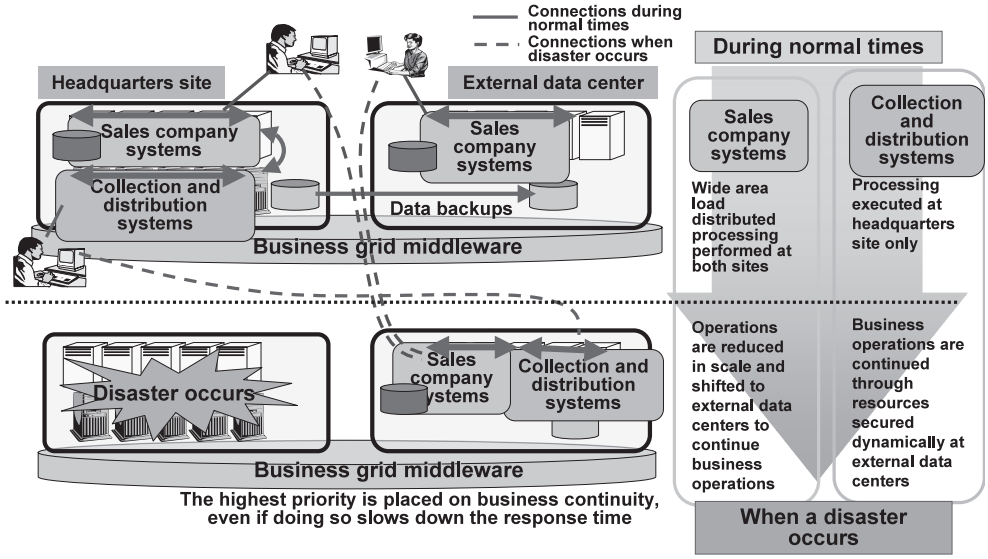

Fig. 3 Summary of validation experiments.

vide a product for use by our customers. Representative examples of such items are introduced below.

(1) Improvement of User Interface

An emphasis was placed on faithfully realizing the intended functions with the user interface function for the GMW and, as a result, the usability and appearance are not adequate. Based on such reflections a user interface, which is user friendly and intuitive in its operations, will be the aim for GGO.

(2) Virtualization of Storage Devices

Development of the GGO is ongoing to make it possible to handle storage devices other than servers as independent IT resources. As the first step toward this goal, development is to include iStorage, which is a storage device manufactured by NEC, as a resource subject to control.

(3) Improvement of Quality and Other Issues

In order to ensure that the product can be used by our customers with peace of mind, quality improvements, mainly on the processing of irregularities will be conducted, along with activities relating to cluster configurations and a raised level of reliability for the GGO itself.

6. Conclusion

Descriptions on disaster recovery with business grid middleware and the activities relating to commercialization have been provided. As part of the WebSAM product family, the GGO is expected to become a core product that can achieve autonomous controls and virtualizations in the future and sequential enhancements are expected over time for it to be a central middleware for disaster recovery solutions, the needs for which are likely to increase in the future.

References

1) Information-technology Promotion Agency, Japan: "Second Promotion Committee Meeting material of Business Grid Computing Research and Development Work for Fiscal 2005", March 2006, http://www.ipa.go.jp/software/bgrid/05iinkai2nd.html

2) Information-technology Promotion Agency, Japan: "IT Foundation of the Next Generation Pioneered by Business Grid", ASCII, April 2006.

Author's Profile

KATO Masayuki
Senior Manager, Grid Promotion Center,
First System Software Division,
Systems Software Operations Unit,
NEC Corporation