Storage-centric Trends and STRAGEX

Takeshi Tachi and Shunichi Amemiya†

Abstract

This article describes case examples of implementing *storage-centric* technology with a focus on the STRAGEX system developed by NTT Laboratories. A storage-centric solution offers an optimal means of managing information for organizations by centralizing the servers and the many local hard disks on personal computers that are normally found in offices.

1. Recent issues in information management

In Japan, the recent enforcement of the Personal Information Protection Law and Japan's version of the Sarbanes Oxley Act (J-SOX) has forced organizations to adopt stricter rules and operations for handling information. At the same time, advances in office automation have resulted in the generation of huge amounts of data on a daily basis, involving not only documents and forms but also email, image files, and other types of information. To examine this problem, we classify information assets to be managed into three categories (**Fig. 1**).

The first category is structured data used in core business systems such as customer and business data managed at the company level. Such data has been managed in a relatively thorough manner by information technology (IT) systems for many years. Information assets of this type have traditionally been protected within the company, but it has recently become important to enable users in outside environments or partner enterprises to access this type of data in a secure manner to raise business efficiency.

The second category is non-structured data such as email and shared files in the file server. Systems for handling this type of data have often been constructed and managed on a site-by-site basis, but the volume of non-structured data has been increasing dramatically and there seems to be no end to the problem of information leaks. Mail archiving and management of file server access logs are examples of solutions

that enable non-structured data to be managed in a more systematic manner.

The third category is local data and software assets used on office personal computers (PCs), the management of which has often been the responsibility of individual employees. The rapid evolution of so-called end-user computing environments, including mobile terminals, has made the management of information on these terminals a major problem. There is also a lot of variety in these environments from site to site, which makes it especially difficult to establish uniform systems and rules to manage data and assets of this type.

In response to these complicated information management problems, the *storage-centric* approach is a solution that promises an optimal means of managing information by centralizing storage—the medium for managing information assets—in a server room.

2. Storage-centric technology trends and implementation examples

We first describe technology trends in the IT infrastructure of organizations as background to storage-centric technology and then give some actual examples of storage-centric solutions (**Fig. 2**).

2.1 Storage technology trends

Network storage is one of the most important technologies for achieving storage-centric solutions. Typical of network storage are relatively expensive products like a storage area network (SAN) that have come to be used mainly by core business systems. These products generally use a specialized method called fiber channel. However, when one attempts to

---

† NTT Research and Development Planning Department
  Chiyoda-ku, Tokyo 100-8116 Japan
  Email: s.amemiya@hco.ntt.co.jp

Terminal Environment

Fig. 1.   Information assets managed at organizations and associated problems.

Fig. 2.   Storage-centric technology trends and actual examples.

apply them to mail and file servers, a problem arises, in addition to the high cost: there are few engineers in the user's organization with the skill and know-how to deal with such special products. Consequently, attention has come to focus on a relatively inexpensive system called the Internet protocol storage area network (IP-SAN) using the iSCSI (Internet small computer system interface) protocol that can be used on an IP network. Products incorporating this system are plentiful, and server vendors and other suppliers are releasing products equipped with easy-to-use interfaces and diverse tools, making it easier than

ever to implement network storage. In addition, information lifecycle management solutions have recently attracted attention as a means of achieving compliance in document management, mail archiving, etc., and the need for such solutions has motivated organizations to actively invest in storage systems.

2.2   Network technology trends

When storage is centralized, it is important to upgrade local area networks (LANs) and wide area networks (WANs) to broadband speeds. The current growth in the use of 100-Mbit/s Ethernet and Gigabit

Ethernet LANs will help drive the implementation of storage-centric solutions. On the other hand, centralizing storage via WANs is still fraught with many performance-related issues. Specifically, in communications between remote sites with a relatively large transmission delay, the throughput of TCP/IP (transmission control protocol, Internet protocol) communications will drop dramatically and the level of performance that one might expect in accessing servers and storage from remote locations will not be achieved.

Telecommunication equipment vendors have begun to offer WAN acceleration equipment with protocol conversion, data caching, and other advanced functions with the aim of improving throughput. This is an effective approach for raising data-transfer efficiency when accessing file and mail servers via a WAN. However, it must be kept in mind that this approach can be quite expensive depending on the bit rate of transmission lines and that some protocols are not compatible with such functions. As a result, scenarios to which this approach can be applied are still limited.

2.3 Implementation example by servers

The most basic approach to a storage-centric format is to centralize the placement and management of physical servers in a data center or other location. The use of an integrated operations management tool in this case can significantly reduce the time and labor required for maintaining and operating servers. As a next step, multiple servers can be grouped on a single physical server by using virtualization techniques. In this way, data and software assets on servers can be physically grouped on a single storage device, thereby simplifying backup measures. Furthermore, as a more advanced approach, servers themselves can be made *diskless* and initialized via the network (network boot system). This has several advantages. For example, it enables the system to restore a server immediately in the event of a hardware failure (by switching to an alternate machine) and makes it easier to add more servers when higher processing performance is needed. In any case, a storage-centric format is an effective means of reducing the total cost of ownership for an organization or data center that needs to manage a large number of servers.

2.4 Implementation example by office PCs and mobile terminals

A storage-centric approach using *thin clients* is an effective means of centralizing non-structured data on office PCs that has traditionally been the most difficult to manage. In particular, the recent spread of telecommuting has made the use of thin clients an effective means of using information in a remote environment by accessing it from a mobile terminal not only from the viewpoint of enhanced security, but also in terms of minimizing the amount of data to be transferred. Moreover, the recent growth of broadband access environments and Internet virtual private networks (VPNs) has significantly reduced the barriers to implementing thin clients. At the same time, it must be kept in mind that the office PC provides a vital user interface that employees experience on a daily basis. The problem then is how to achieve a storage-centric environment that is secure without losing any of the user-friendly and convenient features that users have become accustomed to on their PCs.

3. STRAGEX

NTT Laboratories developed STRAGEX as part of these worldwide trends in storage-centric solutions. It is a management technology that performs network boots of office PCs from centralized storage by using the iSCSI protocol. In other words, STRAGEX enables all local data and software assets used by office PCs to be maintained in centralized storage. An example of implementing STRAGEX in an actual office environment, namely, the Teishin Building of NTT Holding Company, is shown in **Fig. 3**.

3.1 Difference from screen-transfer thin clients

In thin clients using the so-called screen-transfer system, it is true that most data and applications are centrally stored on the server side, but this approach has limitations. For example, (1) all applications and peripheral devices used by conventional PCs may not necessarily work, (2) applications like three-dimensional computer-aided design (3D-CAD) and video playback requiring a high-performance central processing unit may not operate at the expected levels of performance, and (3) some software assets must be left on the PC side such as a thin-client operating system.

In contrast, STRAGEX uses the network boot system that exploits the high performance of the iSCSI protocol to achieve (1) usability and performance in no way inferior to a conventional PC, (2) scalability in which performance is not easily degraded even when booting and using a large number of PCs simultaneously, and (3) completely diskless PCs.
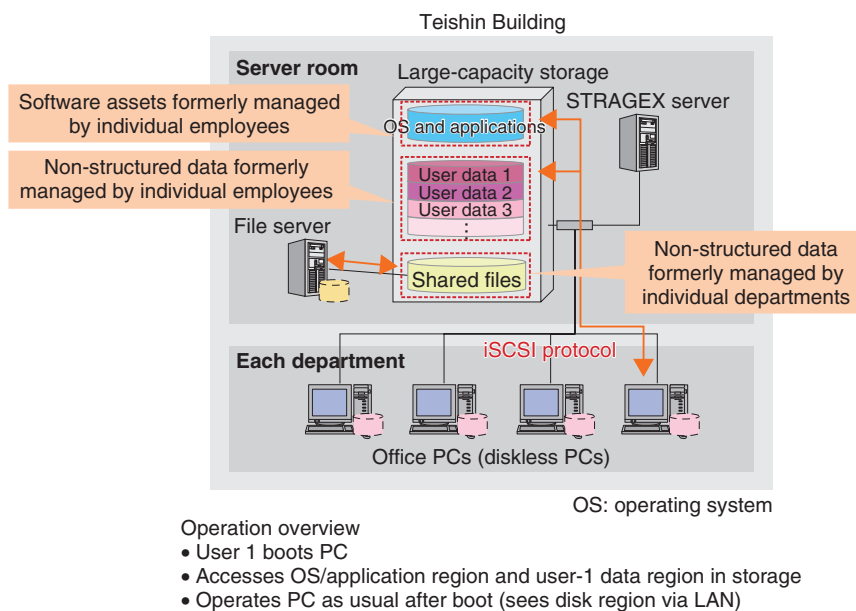
Fig. 3.   Implementation of STRAGEX in the Teishin Building.

3.2   Ease of achieving a uniform PC environment

Another important feature of STRAGEX is the ease of uniform PC environment management. Releasing office PCs for use after converting them to a diskless format makes it easy to establish a thorough rule-based PC environment beforehand (in terms of both hardware and software). In addition, the use of a high-speed disk-copy function here can significantly speed up the process of installing software on a computer prior to its release. For example, when one is setting up a large number of PCs in the conventional manner with a wide variety of office applications that must be combined in a complicated manner, the work involved in booting up each PC and launching every application takes a considerable amount of time and labor. The advantage of centrally managing PC storage should therefore be obvious in this regard.

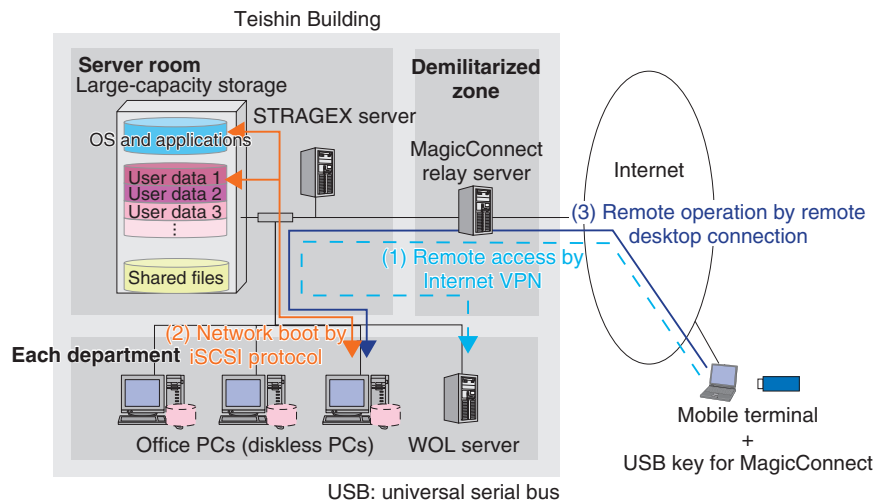### 4.   STRAGEX and thin-client hybrid system

Terminals that use a network boot system like STRAGEX can be thought of as a secure client system that does not sacrifice any of the convenience or processing power of conventional *fat* clients with local hard disks and that presents few barriers to adoption from the user's point of view. In a remote access environment, however, the current state of affairs is such that the large volume of traffic present at boot-up time produces a bottleneck that prevents

the terminals from being used. It is therefore necessary for the system to adaptively use both the network boot system and the screen transfer system, which can be used in slower network environments, as the situation demands. More specifically, a hybrid system that combines both of these systems is required.

The configuration of a hybrid system now on trial at the Teishin Building is shown in **Fig. 4**. Here, a network boot system using STRAGEX is used for access by an in-house LAN, while MagicConnect 2.0 [1] from NTT IT Corporation is used as a thin-client system for access in a remote environment.

4.1 Operation overview

First, a mobile terminal needs an Internet access environment, which can be just a Web browser. The user now performs a remote boot of a STRAGEX PC by using the Wake up On LAN (WOL) standard via the MagicConnect relay server. Next, the user accesses that STRAGEX PC by the screen transfer system using a remote desktop connection. In this way, a user in a remote environment can use an application-and-data-access environment completely the same as that of an in-house office PC's desktop environment. This hybrid system supports security requirements such as no local downloading of accessed data onto a mobile terminal.

Fig. 4.   Hybrid system combining thin clients and STRAGEX.

4.2   Features of hybrid system

Unlike systems that use SSL-VPN (SSL: secure sockets layer), there are no limitations on the applications and in-house servers that can be accessed, which makes the hybrid system optimal for telecommuting. Strictly speaking, however, it is not ideal for applications that involve the transfer of huge amounts of screen data like video playback and 3D-CAD. Furthermore, in contrast to systems that use IPsec-VPN, our hybrid system can be set up so that absolutely no data is left on a mobile terminal, which makes it superior in terms of security and uniform information management. In short, unfinished materials, personal profiles, and other data left on a desktop environment can be used just as they were next time the desktop is accessed from whatever environment. That is, a desktop PC is essentially an employee's *work-environment-preservation server*.

Since data loss due to local hard-disk problems or operational errors is not an uncommon occurrence in ordinary PCs, a storage-centric approach like STRAGEX, which centrally manages desktop envi-ronments and backup measures, is a significant step forward.

5.   Conclusion

This article described recent technical trends in storage-centric technology and the use of the STRAGEX system developed by NTT Laboratories in an actual office environment. At present, STRAGEX technology is being commercialized as CoreBoot^{TM} after having been given functional enhancements such as a terminal inventory management function and enhanced operability by NTT DATA Corporation and NTT DATA Engineering Systems Corporation. The other Special Feature articles in this issue provide technical explanations of STRAGEX, introduce CoreBoot^{TM}, and discuss its applicability to design sites in the manufacturing industry.

Reference

[1]   http://www.magicconnect.net/

Takeshi Tachi

Senior Manager, Open Source Software Center, NTT.

He received the B.E. and M.E. degrees in electrical engineering from Osaka University, Osaka, in 1987 and 1989, respectively. He joined NTT Telecommunication Networks Laboratories in 1989. He received the M.S. degree in industrial engineering and operations research from the University of California, Berkeley, in 1997. He has been engaged in the development of IP network systems and banking systems and in the planning of R&D strategies in the field of information sharing platforms.

Shunichi Amemiya

Producer, Research and Development Planning Department, NTT.

He received the B.E. and M.E. degrees in electrical engineering from Waseda University, Tokyo, in 1993 and 1995, respectively. He joined NTT DATA in 1995. He is a member of the Institute of Electronics, Information and Communication Engineers of Japan.