Research and Development Directions for Future Storage Solutions

Manabu Kitamura
Akira Fujibayashi
Masayuki Yamamoto
Nobuyuki Osaki

OVERVIEW: Along with diversification of data handled by IT systems and changes in the social environment, needs unlike those conventionally associated with storage systems are arising. As regards the primary needs that have been growing over recent years, the rapid growth in data targeted for long-term storage is often quoted. Consequently, it is becoming necessary to suppress power consumption of IT equipment, ensure data management spanning a long period exceeding equipment lifetimes, provide a means to access that data, and reduce the risk of data leakage by internal wrongdoings. In response to these needs, Hitachi is actively involved in standardization of technologies required for long-term data storage. These activities include research on storage-control technology that combines reduction of power consumption of IT equipment and high performance, a content-reference interface for making migration between long-term-storage systems easy, and stored-data encryption technology.

INTRODUCTION

AS data volumes administered by corporate information systems have grown explosively, the importance of storage systems has become paramount. In the existing storage market, emphasis is being placed on operation and maintenance of "mission critical" information (such as on-line transaction services dealing with major companies), where access to data at high speed and high availability of data are demanded. In recent years, however, data for handling IT (information technology) systems is continuing to diversify (for example, as analysis data for supporting decision making of business operations and individual information like medical charts), and the amount of such data is on track to surpass the amount of mission-critical information. In addition, to comply with regulations such as the SOX (Sarbanes-Oxley) Act (Public Company Accounting Reform and Investor Protection Act of 2002), it is becoming necessary to further expand the amount of stored data related to business activities and store it over long time periods.

In regard to issues concerning long-term storage of diversified data in increased volumes, the rising cost of operation and maintenance of this data is the primary concern. Accompanying the growth in data volumes has been an increase in the number of servers and storage devices, leading to a major concern regarding operation of IT systems, namely, increased power consumption. Moreover, linked to concerns for environmental problems over recent years, expectations to cut power consumption of IT equipment have been getting stronger.

The secondary issue is maintaining data-access method. According to the type of data stored on IT systems, some data must be continually stored for longer than the lifetimes of equipment and applications. Along with avoiding losing access to stored data on renewal of devices and applications and storing and maintaining data over long periods, it is necessary to come up with a scheme for allowing instant access when needed. Furthermore, preventing information leakage to third persons must go hand in hand with the scheme allowing instant access when needed. Since the trend towards long-term storage of personal information is strengthening along with the trend towards diversification of data, the risk of data leakage by internal crime is raising apprehension.

With these issues in mind, we at Hitachi are pursuing research and development on storage systems and solutions suitable for low-power-consumption, long-term data storage in addition to promoting

standardization activities for the technologies required for long-term data storage.

In the rest of this report, as efforts to create storage technologies that address the issues described above, power-consumption reduction of storage systems, long-term content-storage technology, and storage-security technology are overviewed in terms of recent technical trends.

TRENDS REGARDING LOW-POWER-CONSUMPTION TECHNOLOGY

In recent years, with the development of information-processing technology and communications technology, there has been rapid growth in IT equipment such as servers and storage devices, and the power consumption of these devices is forecast to become enormous from now onwards (see Fig. 1). On top of this concern for the IT-equipment power consumption, concerns for environmental issues are also growing.

For reducing IT-equipment power consumption, there are two main approaches: (1) utilize technology for reducing power-consumption at the component and device levels, and (2) utilize technology for operation and management that targets low-power consumption.

As a way of reducing power consumption of storage devices, MAID (massive array of idle disks) technology has been introduced[(2)]. The basic principle of MAID technology is to reduce power consumption by stopping a disk when it is in the "un-accessed" state. By applying this technology, as the number of disk drives increases or as the uninterrupted periods without disk access get longer, the power-saving effect gets stronger. New product applications for uses like archiving and backup (which have the potential for the above-described usage patterns) are appearing all the time. On top of that, to make the time that "stoppable" components in business systems are stopped as long as possible, research on operation and management (by means of coordination with system middleware) for automatically and finely managing MAID control in response to conditions of loads such as services has caught on.

Over recent years, in regard to high-speed computers for large-scale scientific calculations, we are trying to make power consumption and high performance of large-scale, large-capacity storage systems compatible. Accordingly, a research-and-development project on adaptive data-allocation control with data-access prediction and storage tiering technology for high-performance storage and low-cost/

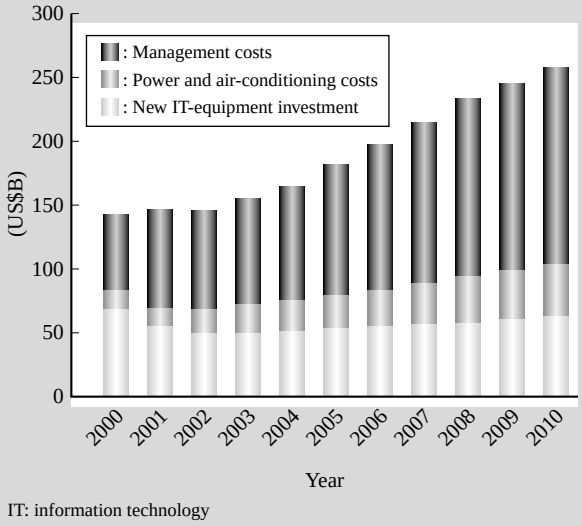

Fig. 1—Forecasted Transition of Cost Breakdown of Data-center Operating Costs.
It is forecasted that power and air-conditioning costs for a data center will close in on investment in new IT-equipment by 2010.

low-power-consumption storage is currently being implemented by Hitachi, Ltd. and Tohoku University.

On the other hand, in contrast to MAID technology (which is effective for usage patterns in which access has prior locality), research and development on technologies for power saving in components and devices for accomplishing power saving with more general utilization patterns is continuing. One example is research applying new low-power-consumption media such as non-volatile storage media like flash memory and phase-change memory (which consume extremely low power during operation compared to hard-disk devices) in servers and storage devices. Another example is research on trying to reduce unnecessary power consumption and reduce the data volume itself held by storage systems (by eliminating duplicated data by means of so-called "de-duplication").

TRENDS REGARDING LONG-TERM-DATA-STORAGE TECHNOLOGY

Functional Requirements for Long-term Storage of Content

In the case of the health-care field in the USA, the HIPAA (Health Insurance Portability and Accountability Act) — which was established in regard to the promotion of the digitization of medical information and the accompanying issues concerning

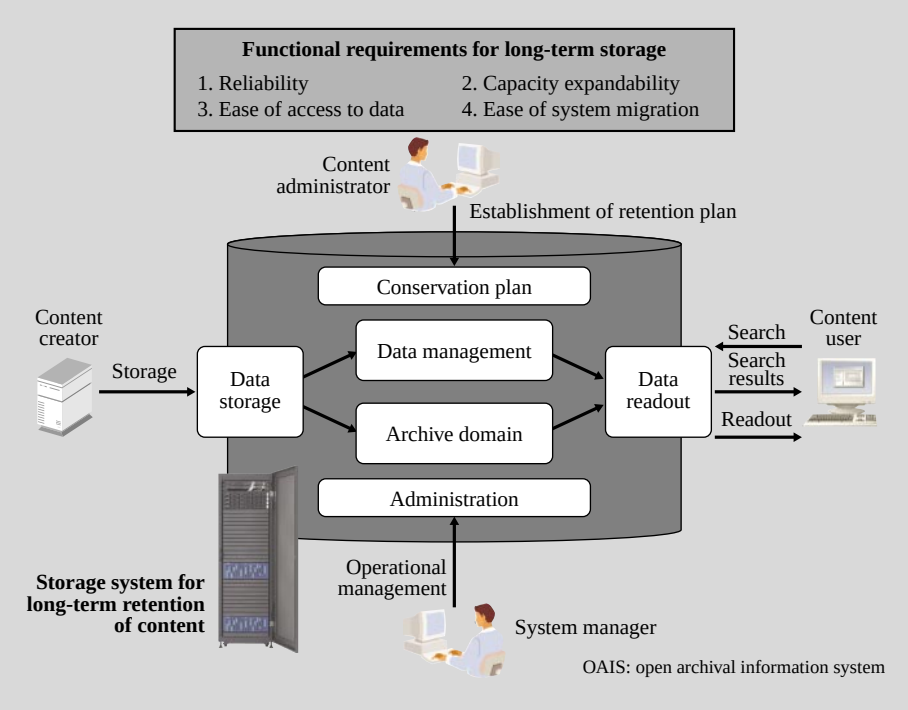

Fig. 2—Functional Requirements for Long-term Storage Based on OAIS Model.
In regard to long-term storage systems for digital information, the OAIS (open archival information system) model is the guiding principle specified by the International Organization for Standardization (ISO). For long-term storage of content with a lifetime longer than that of applications and devises, ease of migration of systems is prescribed as a key requirement.

privacy protection and security assurance — stipulates that medical records of an adult patient must be kept for two years after the death of that patient. In regard to corporate business systems (which are bearing the digitization of technology), to comply with HIPAA, it is necessary to store data for periods exceeding the lifetime of equipment and applications. Functional requirements regarding storage systems for long-term storage of content — including "Hitachi Content Archive Platform" — based on an OAIS (open archival information system)[3] model are described in the rest of this report (see Fig. 2). The OAIS model is a guideline in regard to establishing long-term storage of digital information, and is being specified by the International Organization for Standardization (ISO).

The first requirement is reliability of systems over the storage period. The second requirement is capacity expandability of the area for archival storage because, as a characteristic of archive operation, stockpiles of stored digital information are increasing monotonically. The third requirement is ease of access to data; that is, data must be easy to store in ("ingest") and readout ("access") from long-term storage by using standard data-access protocols like NFS* (Network File System). The fourth requirement is ease of system migration.

* NFS is a trademark of Sun Microsystems, Inc. in the United States and other countries.

Trends Related to Ease of System Migration

Activities regarding ease of system migration are described as follows. The trade organization of the storage industry, namely, the Storage Networking Industry Association (SNIA), is promoting standardization of the XAM (extensible access method) as a new specification for prescribing the reference interface for content. XAM functions as an abstract layer between OSs (operating systems) and various content applications (such as archiving software for e-mail, files, and databases) and storage equipment, and it facilitates searching for data and its accompanying metadata (i.e. reference information like retention period and access authority) by using the same procedure — even when searching between different applications and different storage media. As a result of adopting XAM, end-users are unconstrained by system configurations; they can thus pay undivided attention to a data-storage strategy, even in the case of different-generation archiving software and storage equipment. Moreover, archive-system migration is made easy by XAM (see Fig. 3). As part of the XAM Initiative[4], a project that started in October 2004, 45 vendor companies, including Hitachi Group [represented by Hitachi Data Systems (HDS); head office, Santa Clara, California, USA], are currently participating. At the start of 2008, the proposed standardization for XAM was submitted to the American National Standards Institute (ANSI), where

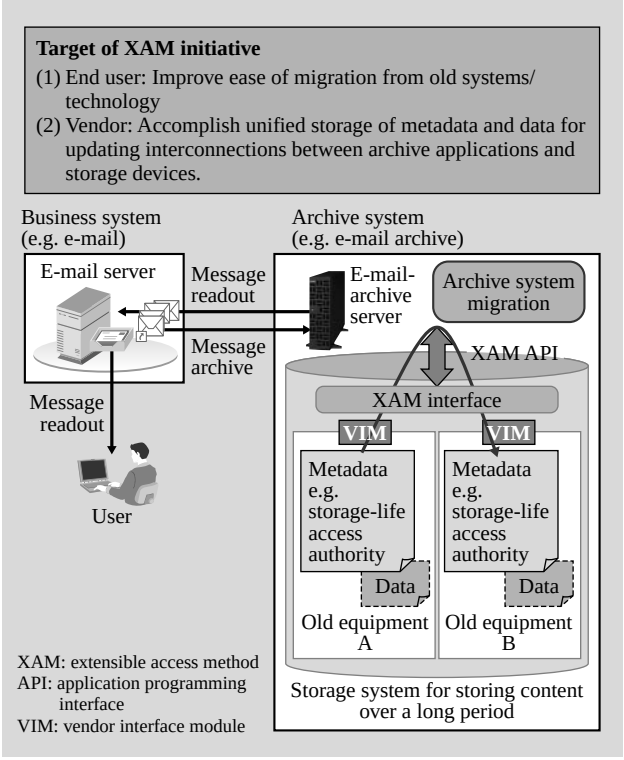

Fig. 3—Overview of XAM Initiative and Image of System Application.
By means of an XAM interface, system migration (including archive-system metadata with no regard to generation or hardware model) becomes easy.

it will be reviewed as an international standard.

TRENDS REGARDING STORAGE-SECURITY TECHNOLOGY

In countries around the world, particularly the USA, it is becoming necessary to strengthen security so as to ensure compliance with regulations like SOX and to prevent information leakage. As an anchor for storing information (which is a valuable business asset), storage plays a key role in assuring security of a business. Technical trends in technologies for supporting information security are described in the following sections.

Trends Regarding Communications Security

The standardization body concerning Fibre Channels, the ANSI work group INCITS (International Committee for Information Technology Standards) T11 is developing protocol specifications for securing fiber-channel communications, namely, FC-SP (Fibre Channel Security Protocols)[5]. Among these protocols, protocol for assuring "integrity" and "confidentiality" of communication data and authentication protocols flowing between host computers, switches and storage devices are being established. To comply with FC-SP, it is necessary at the very least to support one of these authentication protocols, namely, AUTH-A, which is almost equivalent to the so-called CHAP (Challenge Handshake Authentication Protocol) and which is also required to be supported in iSCSI (Internet small computer system interface) protocol. Standardization of FC-SP as ANSI/INCITS 426:2007 was completed in February 2007, and at present, FC-SP-2 (for extending the former) is being discussed by technical group T11. From now onwards, it is expected that even in regard to data centers (which used to be assumed to be operated securely), assuring communication security will become extremely important, in order to stop eavesdropping and impersonation, and that protocols other than AUTH-A specified by FC-SP (for communication encryption, etc.) will become commonly used.

Trends Regarding Security of Stored Data

The repetition of trouble concerning information leakage due to loss of media such as magnetic tapes is being incurred, and in recent years, encryption technology for stored data has been gaining much attention. IEEE P1619 is a standardization project addressing such issues regarding the security of stored data, and Hitachi is actively involved in promoting standardization through participation in this project. IEEE P1619 is separated into several activities, the first of which involved developing a mode of operation suitable for encryption of data on disk storage which does not result in change of data length and designating a key-backup format based on XML (extensible markup language). In addition, it is promoting standardization in regard to management of the key that is used when stored data is encrypted.

Trends Regarding Security Authorization

Whether systems and devices in use have enough security is substantiated by becoming certified by a general third-party evaluation organization. As a standard scheme for this certification, FIPS (Federal Information Processing Standard) 140-2[6] and ISO/IEC (International Electrotechnical Commission) 15408 are available. FIPS 140-2 is an American standard for accrediting whether the encryption module of a device used for encryption [such as VPNs (virtual private networks) and smart cards] satisfies the relevant security requirements. Security requirements

are defined on four levels (i.e. levels 1 to 4), and whether an encryption module has the security to match each level is evaluated and certified. A draft of FIPS 140-3 (which is an improved version of FIPS 140-2) was released in July 2007, and its standardization is moving ahead.

ISO/IEC 15408 evaluates security from the viewpoint of whether products and systems (i.e. not only encryption modules) are properly designed. In recent times, regarding storage-related products (such as Hitachi's disk array subsystems and storage-management software) as well, while products gaining this accreditation are getting onto the market, it is thought that getting such accreditation will become normal.

CONCLUSIONS

Trends in recent storage technologies for handling management of growing volumes of diversified data were described in this report. Storage media used up until now have been developed from a hardware viewpoint focused on operation and management, that is, management of storage systems and storage networks. From now onwards, in addition to making management easier from the "hardware" viewpoint, developing technologies from the "data-management" viewpoint will be required, and close-knit coordination with surrounding services and applications—that is, not just with storage systems—will become necessary. Aiming to provide "storage solutions" that satisfy such demands from society, Hitachi is committed to push its technical development ahead in a positive manner.

REFERENCES

(1) T. Oeda: "Future Trends of Storage System Technology," *Hitachi Hyoron* **85**, pp. 277-282 (Mar. 2003) in Japanese.
(2) D. Colarelli et al., "Massive Arrays of Idle Disks for Storage Archives," University of Colorado at Boulder ACM/IEEE SC 2002 Conference (SC'02) (2002).
(3) CCSDS 650.0-B-1, "Reference Model for an Open Archival Information System (OAIS)," Blue Book, Issue 1 (2002).
(4) "SNIA XAM Initiative," http://www.snia.org/forums/xam/
(5) "FC-SP (Fibre Channel Security Protocols REV 1.6)," http://www.t11.org/
(6) "FIPS Publication 140-2," http://csrc.nist.gov/publications/fips/fips140-2/fips1402.pdf

ABOUT THE AUTHORS

Manabu Kitamura
Joined Hitachi, Ltd. in 1994, and worked at the 8th Research Department, the Systems Development Laboratory. He was engaged in the research and development of enterprise storage systems.

Akira Fujibayashi
Joined Hitachi, Ltd. in 1994, and now works at the 8th Research Department, the Systems Development Laboratory. He is currently engaged in the research and development of enterprise storage systems. Mr. Fujibayashi is a member of the Information Processing Society of Japan (IPSJ).

Masayuki Yamamoto
Joined Hitachi, Ltd. in 1996, and now works at the 8th Research Department, the Systems Development Laboratory. He is currently engaged in the research and development of content archive solutions, storage management software, and storage services. Mr. Yamamoto is a member of the IPSJ.

Nobuyuki Osaki
Joined Hitachi, Ltd. in 1994, and now works at the 3rd Research Department, the Systems Development Laboratory. He is currently engaged in the research and development of storage management software and storage security.