

Trend of Long-term Preservation of Digital Assets and New Content Archive Solutions

Yoshitaka Kawamura
Masayuki Kitamura
Yoshihiko Miyazawa
Kenichi Soejima

OVERVIEW: Given the recent emphasis on handling compliance for internal controls and business risks, the need for content archives that store electronic data (such as e-mails and contracts, public documents, and drawings) for long periods in a form that cannot be over-written is growing rapidly. The “Hitachi Content Archive Platform” — which is a storage appliance for content archiving — is an “appliance product” that unifies management-software-installed servers and storage in such a way as to allow customers to easily integrate it with existing systems. Featuring data-protection functions (such as prevention of falsification and premature deletion) and high-efficiency disk-utilization functions, the Hitachi Content Archive Platform integrates with ISV software (which has functions for automatically archiving data from middleware linked with various operation systems used in businesses), thereby providing archive solutions for e-mail data, ERP data, file data, and log data.

INTRODUCTION

AS a consequence of the evolution of the information society, all information used in business organizations is being digitized, and the amount of content such as e-mails and contracts, public documents, technical

drawings is growing rapidly. Furthermore, from now onwards, in accordance with legal regulations, it is forecast that the amount of content data in long-term storage will continue to grow. Accordingly, how this data will continue to grow and how to reduce and

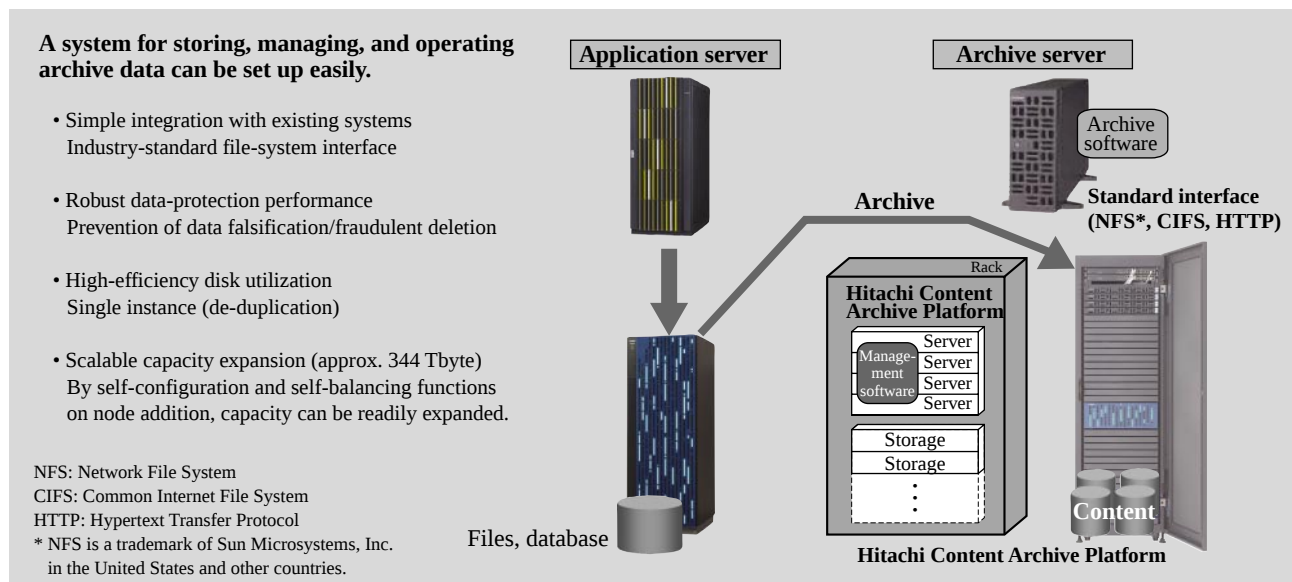


Fig. 1—“Hitachi Content Archive Platform” Storage Platform for Content Archive.
Hitachi Content Archive Platform — a content-storage archive for long-term storage of business e-mails — is a hard-disk-based storage for archiving content with a robust data-protection function. Compared to performance of archives using conventional media, scalability, reliability, and efficiency of searching and operation are dramatically improved. Moreover, it is an appliance product that unifies management software for servers, storage, and content archives.

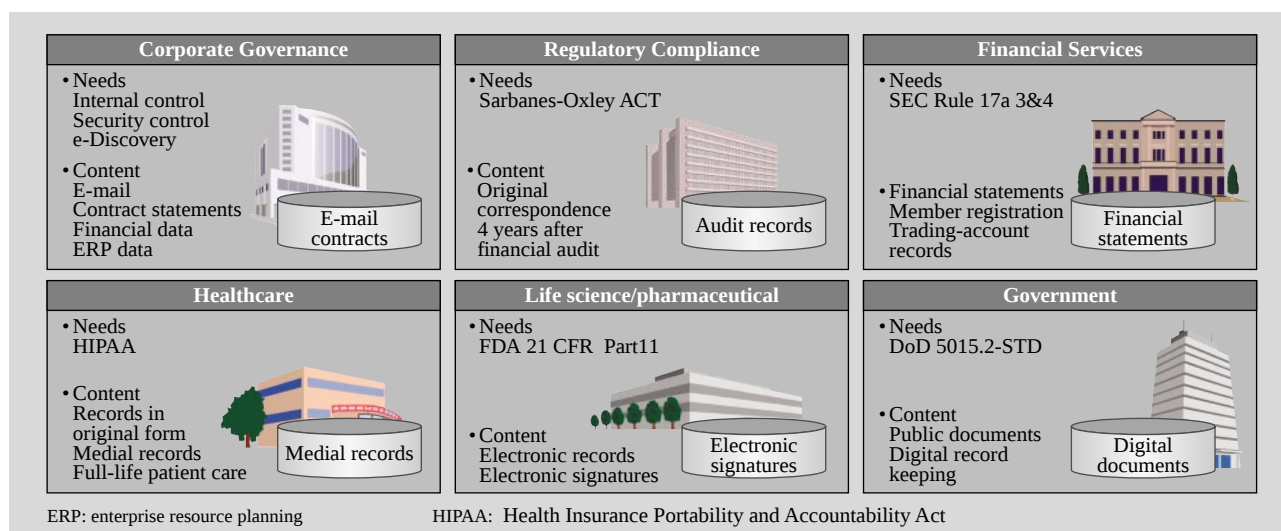


Fig. 2—Needs of Content Archive.

Various kinds of e-mail data used in business must be stored for long periods in accordance with legal requirements.

efficiently manage costs are becoming big challenges for businesses.

In response to these needs, Hitachi, Ltd. announced the “Hitachi Content Archive Platform” — a storage appliance for content archiving — in May 2007. After that in October 2007, we upgraded its “solution services” menu and strengthened its “content archive solutions” (see Fig. 1).

In the rest of this report, our latest content archive solutions — centered on the Hitachi Content Archive Platform — are described.

NEEDS OF THE CONTENT-ARCHIVE MARKET

In accordance with legal regulations and acts, in addition to the e-mail, contract documents, accounting data, and so on required for business activities, it has become necessary to store various kinds of electronic data used in each business category over a long period (see Fig. 2).

In regard to recent legal regulations, such as internal control and compliance, data is being accessed rapidly for auditing purposes, and “e-Discovery” systems (which seek to disclose data as evidence) are being legalized in the USA. Furthermore, as a result of the revision to the “Federal Rules of Civil Procedure” (FRCP) in December 2006 in the USA, the effects of these regulations on global businesses have started to be felt.

While being protected from falsification, this data also requires protection from premature deletion or

erasure by incorrect operation. Storage periods are lengthening and the amount of target data is growing year on year; consequently, data management will become a bigger challenge from now onwards.

ISSUES REGARDING ARCHIVES USING CONVENTIONAL MEDIA

Since access frequency decreases with the passage of time, it is often the case that content is migrated from high-speed storage of application servers to optical disks or magnetic tape according to the type of archive. Conventional media such as optical disks and tapes can, however, become unreadable as a result of degradation of the storage medium. Moreover, it is also a problem that it takes time to handle and search media when accessing archives. On top of that, there is the risk of information leakage like loss or theft of media.

CONTENT-ARCHIVE STORAGE

To resolve the issues concerning conventional media like optical disks and magnetic tape, “content archive storage” based on hard disks has recently been gaining attention. In regard to the need in the case that access frequency to archives is pretty low, instead of using high-speed enterprise storage, using storage with low capacity unit price even at low speed makes it possible to manage data more efficiently and reduce the TCO (total cost of ownership). Having WORM (write once read many) media (which cannot be rewritten until a specified storage period is surpassed),

high-efficiency data-management functions (such as de-duplication), and a robust data-protection function (such as periodic checking of data authenticity), the Hitachi Content Archive Platform is a content-archive storage appliance that integrates storage as a content depository and a management server installed with content-archive management software.

CURRENT CONTENT-ARCHIVE SOLUTIONS Solution Systems

Hitachi is providing solutions that apply the Hitachi Content Archive Platform by integrating ISV (independent software vendor) software (with a function for automatically archiving data from middleware cooperating with various application systems to the Hitachi Content Archive Platform) (see Fig. 3). Some example solutions, namely, archiving for e-mail data, ERP (enterprise resource planning) data, file data, and log data, are described in the following sub-sections.

E-mail-data-archive Solution

With the dramatic increase in the use of e-mail, the importance of e-mail within businesses is ever increasing. That is to say, e-mails and their attached files contain “intellectual property,” owned by particular businesses, in abundance. Moreover, to swiftly handle risks such as law suits, business-transaction trouble, and information leakage, the need

to store data securely over long periods and quickly retrieve it for perusal, under a condition in which a massive volume of e-mail is close to “on-line,” is getting stronger. Accompanying the massive growth in volumes of e-mails, from the system-operation aspect, problems related to resources of e-mail server (such as increase in size of in-boxes, disk volume, backup time, restoring time during failures, and decrease in system processing power) are getting worse.

As for the e-mail-client aspect, since the capacity of e-mail in-boxes is limited, it is often the case that e-mails are stored on the local disk of a PC (personal computer). Accordingly, managing e-mails is a troublesome task.

Our “e-mail-data archive solution” provides the solutions to these problems with e-mail management. By coordinating existing e-mail systems and Hitachi Content Archive Platform and automatically archiving e-mails according to a pre-set policy (such as received date, access date, and size of e-mail), this solution prevents falsification and premature deletion and assures secure management of e-mails.

ERP-data-archive Solution

As a consequence of Japan’s “Electronic Storage Act,” it is possible to electronically store bookkeeping materials created on computers. Presently, in the case of storing accounting records, journals, ledgers,

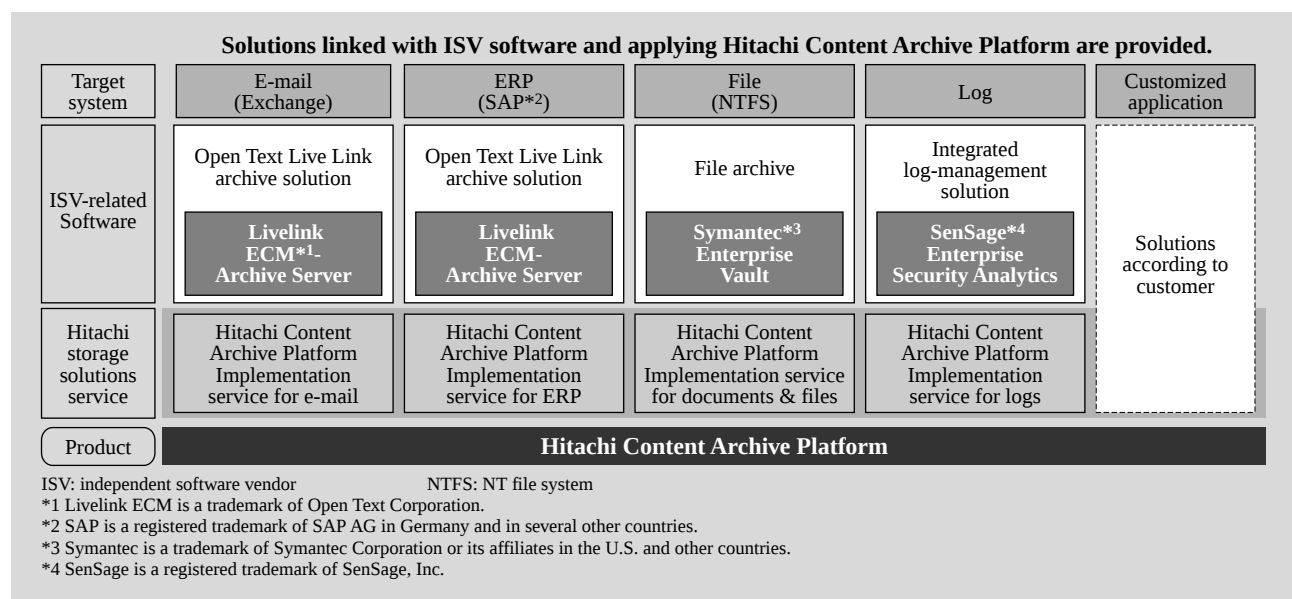


Fig. 3—Content Archive Solutions.

By means of solutions and services linked to the Hitachi Content Archive Platform, each archive solution is realized.

auxiliary books, printout forms, account-related materials, bill statements, etc. on paper, high costs for vast amounts of paper output, transport, and storage are incurred. What's more, in the case of storing account books in ERP systems, along with the massive increase in volume of stored data, response time to user requests gets longer. Furthermore, it is necessary to upgrade hardware due to lack of disk capacity, decrease in server processing performance, and so on. In addition, the problems of increasing backup time and insufficiency of the ERP server's backup window (i.e., the time period in which servers and processing services must be shut down) arise.

"ERP data-archive solution" supports business solutions to these problems concerning ERP systems. By coordinating existing ERP systems and Hitachi Content Archive Platform and periodically archiving accounting records and printout forms according to pre-set policies on a monthly or yearly basis, this solution enables secure storage of ERP data.

File-data-archive Solution

As for general file data that is not managed by applications, various kinds of data (e.g. text data and spreadsheet data) exist. Consequently, when file servers are used over a long period of time, accumulation of old files causes expansion of storage capacity and addition of new file servers becomes necessary.

Our "file-data-archive solution" enables storage, management, and retrieval of file data from various sources under a file-server environment. In addition, by securely, simply, and automatically archiving data requiring long-term storage on Hitachi Content Archive Platform according to a pre-set policy (such as retention period for all files and directories), this solution simplifies management of storage and reduces storage costs. Files that have been archived can also be accessed without user being aware (see Fig. 4).

Log-data-archive Solution

For checking content archives and analyzing security of businesses, it is important to collect, store, analyze, and report vast amounts of logs generated by the multitude of servers used in businesses. Since log-management systems are aimed at security analysis, logs stored by log-management systems need to be stored in secure storage that cannot be falsified or fraudulently deleted. What's more, storage of vast amounts of logs and secure storage in line with requirements of legal regulations are necessary.

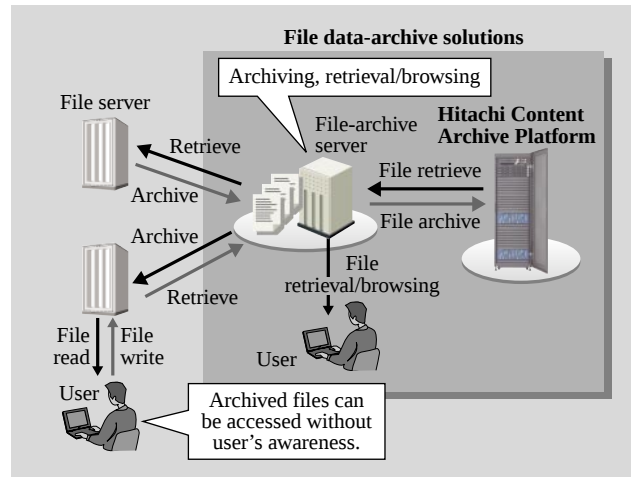


Fig. 4—File-data Archive Solutions.

Applying the Hitachi Content Archive Platform makes it possible to securely and simply store file data.

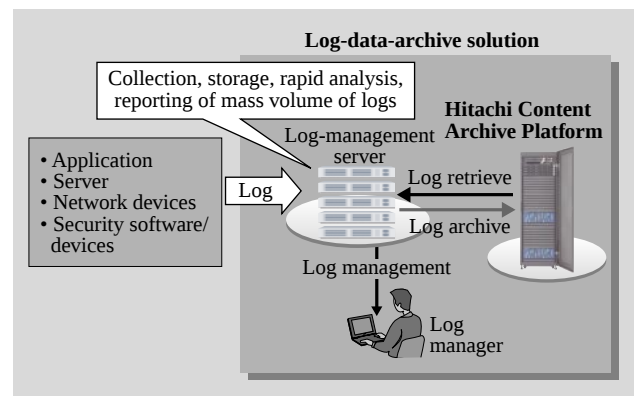


Fig. 5—Log-data-archive Solutions.

Applying the Hitachi Content Archive Platform makes it possible to prevent falsification and fraudulent deletion and securely store log data.

With the "log-data-archive solution," it is possible to securely store log data by means of archiving log data on Hitachi Content Archive Platform — which features robust data-protection functions. In addition, with the (hard-disk-based) Hitachi Content Archive Platform it is possible to analyze and report logs in a short time because log data can be fetched up in a flash as required (see Fig. 5).

CONCLUSIONS

In this report, the latest content-archive solutions — centered on the Hitachi's storage platform called "Hitachi Content Archive Platform" for content archiving that meets the need for long-term storage of

e-mail data used in business — were described. From now onwards, we will continue to expand our diverse solutions that coordinate with various kinds of middleware.

REFERENCES

- (1) Y. Kawamura, “Storage Solutions for Internal Control,” *Hitachi Hyoron* **89**, pp. 718-721 (Sept. 2007) in Japanese.
- (2) Product introduction site, <http://www.hitachi.co.jp/storage/> in Japanese.

ABOUT THE AUTHORS



Yoshitaka Kawamura

Joined Hitachi, Ltd. in 1986, and now works at the Strategic Business Planning Department, the Storage Area Network Systems Solution Division, the Information & Telecommunication Systems. He is currently engaged in planning and promotion of the global business strategy.



Yoshihiko Miyazawa

Joined Hitachi, Ltd. in 1991, and now works at the Storage Area Network Development Department, the Storage Area Network Systems Solution Division, the Information & Telecommunication Systems. He is currently engaged in the ISV integration with Hitachi Content Archive Platform.



Masayuki Kitamura

Joined Hitachi, Ltd. in 1993, and now works at the Strategic Business Planning Department, the Storage Area Network Systems Solution Division, the Information & Telecommunication Systems. He is currently engaged in planning of content-storage products.



Kenichi Soejima

Joined Hitachi, Ltd. in 1993, and now works at the 2nd Group, the Solution Service Development Department, the Storage Area Network Systems Solution Division, the Information & Telecommunication Systems. He is currently engaged in development of storage-solution services.