MEDINFO 2007
HISA Ltd Australia
© 2007 The authors. All rights reserved.

Data Integrity Management for Medical Image Data Archive Using Digital Signatures and Time-Stamp Certificates

Raimund Vogl [a], Michael Berreck [b], Thomas Brauns[b], Gerald Geiger[b], Oliver Haid[b], Robert Penz[a], Christoph Pirchl[b], Dietmar Reiter[b], Manfred Wallinger[b], Joachim Zaers[c], Georg Lechleitner[b]

[a] *HITT – health information technology tirol, Austria*
[b] *Tiroler Landeskrankenanstalten, Information Management Department, Austria*
[c] *JoSoft Consulting, Germany*

Abstract and objective

Medical image data archives have been introduced at Innsbruck University Hospital in 1997 using write-once media (CD-R) for long term storage. Due to increasing data volumes, a migration of the archived data to high capacity magnetic tape storage media was necessary. To fulfil the requirements for guaranteed integrity and non-repudiation of the archived data, an additional data integrity layer was added to the commercial PACS software product to add digital signatures to the archived data and time stamp certificates. A time stamp server software suit had been developed by the hospital IT team for that purpose in 1999, issuing more than 2.4 million certificates since. The data integrity layer proved essential in early detection of a severe hardware malfunction in the tape archive system, which resulted in corruption of archived data in December 2000.

Keywords:

PACS, medical archives, time-stamping, data integrity

Introduction

The guarantee of data integrity in archiving medical data is a prime concern. Usually, this is resolved by using unalterable (write-once) media like WORM, CD-R, DVD-R etc. The first choice of archive media for the PACS introduced at Innsbruck University Hospital in 1997 was CD-R jukeboxes, and image data was only stored in one copy. Due to the fast growing volume of data archived, a decision to change to higher capacity magnetic tape media had to be taken in 1999, also to allow for duplicate archive copies of data to prevent data loss in case of media corruption (actually, in the course of data migration of 1.8 TB of compressed image data in 1999, one CD-R medium - from a total of 4.800 - was found to be corrupt, resulting in the loss of 157 imaging studies). To fulfil the requirements for guaranteed data integrity and eliminate the potential for subsequent alteration of archived data (non-repudiation) when using rewriteable media like magnetic tapes, an additional data integrity layer was added by the hospital IT team to the commercial PACS software product. This data integrity layer is a set of scripts that scan the PACS database for trigger events for archiving and retrieving imaging studies and provide the corresponding fundctionalities by interfacing to the archive management software of the magnetic tape archive. The archiving script scans the database for imaging studies due to be archived, then one by one collects together all image files of this study, and adds a digitally signed "study summary files" to the data set to be archived that allows to validate the integrity of the archive data (in subsequent retrievals) and also requests a time-stamp certificate from a time-stamp server for this archive data set, as a guarantee that the given data set was present at the specified point in time (and not foisted later in a possible fraud attempt).

Methods

To enable the commercial PACS product to use a tape archive system for long term archiving (instead of the CD-R jukeboxes), a software layer was developed by the hospital IT team to connect to the archive management software system. It was decided to archive image data in packages containing one imaging study each (ranging up to Gigabyte size for large CT exams). A study summary file is created for each study, representing an enhanced directory listing of all image files in the study, containing filename, creation date, file size and an MD5 hash of the image file itself (the MD5 hash algorithm was chosen for its fast execution time essential for the hardware available back then). This study summary file is digitally signed with a server specific private key (controlled by hospital IT staff). On retrieval of the archived data set at a later point in time, the signature of the study summary file is verified and the content of the study summary file is cross checked against the retrieved image data – allowing to detect any alteration in the image data files.

This still leaves open the possibility for manipulation by the hospital IT team in control of the signature keys. To eliminate this potential security breach, a digital time-stamping service was developed by a member of the hospi-

tal IT team in early 1999 as a Java application running under Linux. This service has been continuously operational since, issuing more than 2.4 million time-stamp certificates, one for each imaging study archived.

This time-stamping server receives request for issuing a time-stamp certificate via a proprietary TCP/IP socket based network protocol (standards for time-stamp certificate formats and request protocols have only been defined years later [1]). This usually contains an SHA1 hash of the actual data that is to be time-stamped (in our case, an SHA1 hash of the study summary file). The certificate issued is a text data set that is signed by the time-stamping system with its own private key. It is comprised of the data to be time-stamped (the SHA1 hash), the UTC time-stamp, a sequence number, and a hash of the previously issued certificate. This linear linking scheme is the simplest way to guarantee a tamper-prove sequence of time-stamp certificates. It is impossible to alter one of the certificates at a later point in time (even when in control of the signature key), because this would require changing all subsequently issued time-stamp certificates as well. To render even this impossible, once a month, the most recently issued time-stamp is send to a broader group of people to make it publicly known (publishing it in a printed periodical with wider circulation was also considered).

Results

The time-stamping system started operation in May 1999 and was first tested, when 1.8 Terabyte of losslessly compressed image data were migrated from 9 CD-R jukeboxes to the tape archive system in a 2 month effort in the summer of 1999. The dual archiving of data and the digital signature based data integrity layer proved enormously valuable, when a sporadic hardware defect occurred in the tape drives (or rather a SAN fibre channel gateway device) used in the tape archive system, corrupting part of the image data files.

In December 2000, when retrieving imaging studies from the tape archive the data integrity checks comparing data retrieved with the MD5 hash information from the data summary files stored along with the actual image data showed signs of data corruption and caused the hospital IT staff to initiate a problem analysis with the hardware vendor, which proved to be very difficult due to the erratic nature of the hardware defect. After lengthy analysis and due to the fact, that corrupted data sets could be easily identified, the problem could be tracked down and data recovery procedures were initiated.

The time-stamping service has been operating flawlessly since 1999, issuing a total of 2.4 million time-stamp certificates until now. Apart from the incident with data corruption, no data integrity problems ever occurred.

Conclusion

Ensuring data integrity and guaranteeing non-repudiation for magnetic tapes as long term archive medium for large volume medical image data using digital signatures and time-stamp certificates has proved to be a viable and cost effective solution and has been operational at Innsbruck University Hospital from 1999 on. More than 2.4 million imaging studies have been archived and digitally time-stamped. The inhouse development for the time-stamp server proved to be very stable and has been continuous in operation for 8 years. Nevertheless, the linear linking scheme used to generate a tamper-prove sequence of time-stamps has deficits in the verification process [2] and more efficient schemes have been proposed [3]. Due to the absence of standards available at the time of implementation of the time-stamp service, the protocol for requesting time-stamp certificates had been defined in a proprietary way. Subsequently, standards have become available [1]. And open source implementations for time-stamping services are now also available drawing on these standards (see [4] and [5]), which have the additional advantage to be able to utilize commercial Hardware Security Modules (HSM), which guarantee the tamper-prove generation of time-stamp certificates in hardware. Even though the utilization of time-stamps is undisputed, a new standard compliant system with improved security (eg. by using HSM and new secure hash algorithms) can be seen as a requirement for the new medical image and documents archive system, to which the PACS image data has recently been migrated. Anyhow, digital signature legislation limits the validity of digital certificates (the Austrian "Signaturgesetz" sets a maximum of 5 years for the validity for qualified certificates [6]), and thus a renewal of the signatures on the archived data sets has to be considered, best performed along with a periodic active rearchiving on new media to protect against data corruption due to media ageing.

References

[1] Adams C, Cain P, Pinkas D, Zuccherato R. Internet X.509 Public Key Infrastructure Time-Stamp Protocol, RFC3161 2001. Available from: http://www.ietf.org/rfc/rfc3161.txt (accessed 2006-11-30)

[2] Lipmaa H. Secure and Efficient Time-Stamping Systems. PhD dissertation. Tartu University, 1999. Available from: http://www.adastral.ucl.ac.uk/~helger/papers/thesis/thesis.pdf (accessed 2006-11-30)

[2] Maniatis P, Giuli T, Baker M. Enabling the Long-Term Archival of Signed Documents through Time Stamping. arXiv:cs.DC/0106058 2001. Available from: http://arxiv.org/PS_cache/cs/pdf/0106/0106058.pdf (accessed 2006-11-30)

[4] Open Evidence. Available from: www.openevidence.org (accessed 2006-11-30)
[5] Open Time-stamping Authority tool. Available from: www.opentsa.org (accessed 2006-11-30)
[6] Austrian Digital Signature Law (in German: Bundesgesetz über elektronische Signaturen (Signaturgesetz – SigG)). Available from: http://www.a-sit.at/pdfs/SigG_incl_Novelle2000.pdf and http://www.ta-sit.at/pdfs/SigV.pdf (accessed 2006-11-30).

Address for correspondence

Raimund Vogl, Phd
HITT health information technology tirol,
Leopoldstrasse 1, A-6020 Innsbruck, Austria. rvogl@hitt.at