Send Orders of Reprints at bspsaif@emirates.net.ae

The Open Medical Informatics Journal, 2013, *7*, 1-7

Open Access

Distributed Storage Healthcare — The Basis of a Planet-Wide Public Health Care Network

Nikolaos Kakouros*

Johns Hopkins University, Baltimore, MD, USA

Abstract: *Background*: As health providers move towards higher levels of information technology (IT) integration, they become increasingly dependent on the availability of the electronic health record (EHR). Current solutions of individually managed storage by each healthcare provider focus on efforts to ensure data security, availability and redundancy. Such models, however, scale poorly to a future of a planet-wide public health-care network (PWPHN). Our aim was to review the research literature on distributed storage systems and propose methods that may aid the implementation of a PWPHN.

Methods: A systematic review was carried out of the research dealing with distributed storage systems and EHR. A literature search was conducted on five electronic databases: Pubmed/Medline, Cinalh, EMBASE, Web of Science (ISI) and Google Scholar and then expanded to include non-authoritative sources.

Results: The English National Health Service Spine represents the most established country-wide PHN but is limited in deployment and remains underused. Other, literature identified and established distributed EHR attempts are more limited in scope. We discuss the currently available distributed file storage solutions and propose a schema of how one of these technologies can be used to deploy a distributed storage of EHR with benefits in terms of enhanced fault tolerance and global availability within the PWPHN.

We conclude that a PWPHN distributed health care record storage system is technically feasible over current Internet infrastructure. Nonetheless, the socioeconomic viability of PWPHN implementations remains to be determined.

Keywords: Electronic health record, distributed storage healthcare, public health care network, peer-to-peer networking.

INTRODUCTION

Modern information technology (IT) can transform manual paper record keeping with the potential to reduce healthcare costs and improve patient safety and outcomes. The electronic health record (EHR), currently generated and maintained by and large within an institution, allows patient health information to be readily shared by standard electronic transactions with various entities within health information exchange networks. Such a network may include ambulatory clinics, sub-acute care environments, other hospitals as well as payers and even patients through tethered patient health care record (PHR) interfaces. There are many different issues to consider when deploying EHRs such as language, clinical vocabularies, ontologies, policies, and technology but unifying concepts are those of achieving high availability and security of the sensitive EHR data.

BACKGROUND

In 2005, the Healthcare Information and Management Systems Society (HIMMS), in an effort to encourage high levels of IT adoption to improve patient care quality and safety, introduced the HIMMS Analytics EMR adoption model (EMRAM) to indicate how individual hospitals and integrated delivery systems in the USA and Canada adopted IT. EMRAR assesses up to 250 potential health IT applications and scores the hospital from 0 to 7.0 [1]. There are eight stages, each witha higher level of health IT integration towards an entirely paperless environment. Since spring 2010, EMRAM is being applied to European hospitals on a trial basis, as the model is being adapted to better suit the different health care market. As of April 2011, there already are 56 hospitals in the USA that have achieved EMRAR Stage 7,which has been described as the "pinnacle of an environment where paper charts are no longer used to deliver patient care" [2].

With such high levels of IT healthcare integration come new considerations. Traditional media, such as magnetic tape and optical discs, used for archival storage of fixed content components of EHR (such as imaging, old reports etc) do not allow fast online access to the stored data. The cost advantage of magnetic tape as a storage medium has been greatly eroded over the recent years due to rapidly falling prices of hard disk drives. The latter allow for high speed and random access to data and are already being used in non-archival storage systems, such as "storage area networks" (SAN) and "network attached storage" (NAS). Digital disk drive storage will likely replace paper and the older digital archiving media.

Nonetheless, with no paper records, medical care can grind to a halt if the digital EHR data becomes unavailable. Natural disasters can seriously damage localized data centers and one can envisage how these may prove targets for terrorist attacks. A relatively small electromagnetic pulse

*Address correspondence to this author at the Johns Hopkins University, Baltimore, MD, USA; Tel: 774-442-5043; Fax 774-441-6303;
E-mail: nikos@doctors.org.uk

1874-4311/13

(EMP) flux compression generator device can be built from off-the shelf components or dual use technologies, and HERF (High Energy Radio frequency) guns can also disrupt electronics and communication equipment over a wide area. Information warfare has been identified as the "perfect terrorist weapon" by the International Institute for Counter-terrorism [3]. In times of war, military attacks to an enemy's EHR datacenters could cripple the country's healthcare from within.

On a more mundane note, random hardware failures can lead to significant loss of data and functionality. For example, the EHR at three different Karolinska hospitals was unavailable on the 11th and 12th of December 2008 due to failure of one disk on one machine, leading to severe disruptions including the cancellation of a numbered of scheduled surgical operations, as the caregivers could not access patients' medical records [4]. Securing reliability of access to the data is, therefore, essential. Decentralizing data archiving from clustered data centers to a distributed storage system would help ease such concerns.

In an optimistic view of the future one can foresee a planet wide public health network (PWPHN). This network would unify healthcare data throughout participating countries and make a patient's health record available where and when it is needed. This would require the exchange of vast amounts of highly sensitive and important information. In this paper, we will review the current literature on distributed storage and assess the suitability of a worldwide distributed storage EHR solution implementation for a PWPHN.

METHODS

Literature considered in this project includes national and international journal articles, professional or academic and accessible *via* MEDLINE, CINAHL, EMBASE and Web of Science (ISI), and Google Scholar. The search was expanded to include e-papers in repositories (e.g. arXiv.org, Association for Computing Machinery digital library, CIteSeerx) and by using the Google search engine to locate other resources such as online recorded presentations, and non-authoritative news articles.

RESULTS

Our search of the literature revealed that most distributed EHR attempts tend to be of relatively small scope. An example includes early work with a web-enabled hematological system with distributed storage of the large histopathology slide imaging but still retaining a centralized master node database [5]. One of the forerunning country-wide health networks is the ambitious United Kingdom National Health Service (NHS) Spine. In its full implementation the NHS Spine would make every UK patient's "summary care record" (viz.EHR) available throughout the country, linked to a personal demographics service holding the details of all NHS patients. The Spine also supports clinic visit scheduling ("choose and book") to enable hospital staff to arrange clinic visits at a convenient time for the patient. The Spines' electronic prescription service links physicians to community pharmacies and adds prescription information directly to the patient's EHR. A Secondary uses service (SUS) employs the Spine for data exchange that enables a range of audit, reporting, analysis, research, planning and billing functions.

The Spine utilizes an Access Control framework that provides a single log-in for each healthcare professional accessing a patient's NHS care record, providing information on a need-to-know basis depending on the user's role and "legitimate relationship" with the patient. This amounts to 50 million patients and 200 million clinical records per year.

The storage infrastructure of the Spine is very complex. It uses Oracle's Real Application Clusters (RAC) dependent on Oracle Clusterware to bind three nodes together into a single logical server for high availability. 122 such 3-node Oracle databases, interconnected in storage area networks provide storage for the Spine. Unfortunately, the Spine has failed to as yet reach its full implementation due to public mistrust particularly regarding privacy concerns, opt-outs by multiple healthcare providers and serious setbacks with implementation of shared imaging. Initial deployment of the Spine was over the NHSnet private wide area network that was replaced half-way through the project in 2006 by the N3 network. N3 connects all NHS locations and 1.3million NHS employees across England as a high availability, fast broadband network including secure VOIP. Despite the multi-billion dollar outlay, the system remains currently underused. As of November 2010, only 0.13% of the UK population have opted to join Healthspace, the internet accessible personal health record (PHR) component of the system [1].

Distributed Health Information Storage

Although, based on our search, distributed EHR solutions have been very limited in scope and uptake, we believe a distributed file system (DFS) for the PWPHN is feasible using current technology. In brief, each healthcare provider network could maintain its own localized or wide-area EHR system that would asynchronously synchronize with the cloud for archival purposes. The information on the cloud would provide data backup security but also a global view of EHR information on a patient. In many healthcare environments there already exist independently developed solutions that do not easily interoperate with each other and follow their own convention of creating, maintaining and storing patients' EHR. Even within a single EHR vendor there may be limited interaction between "charts" maintained by different specialties or provider locales due to the use of different sets of templates. The first hurdle to the PWPHN will be the extraction of relevant information that can be shared in a standardized format within the global EHR. Text based information, for example, may be standardized by using accepted HL7 (Health Level 7) protocols, while the DICOM international image communications protocol standard may be used for imaging. Linkage to outside data sources, such as laboratories and medication-management systems (e-prescribing) can be handled as a single point of integration, becoming available to all providers. There will effectively, therefore, be the provision of distributed EHR storage as a service. The principal considerations in such a repository are the method of storage and most importantly its scalability, high availability even in the presence of network or hardware faults, low latency even in situations of high system load and, importantly, security features to protect data from unauthorized access and health data corruption attacks.

Storage

Over the last couple of years we have seen the emergence of 'cloud EHR' systems based on proprietary network solutions. The best established of these is the Amazon Elastic Compute Cloud (EC2) that provides abstracted computer resources on which an EHR system can be run, with HIPAA-compliant security. Storage for EHRs running on an EC2 instance is provided by the Simple Storage Service (S3). The details of this proprietary system are unknown, but it can store objects up to 5 terabytes in size with claimed 99.99% availability over a given year. Nonetheless, there have been questions regarding HIPAA compliance and data security due to single key authentication of Amazon Web Services (AWS) and IP unrestricted access to the interfaces [6]. AWS does offer a service for more sensitive workloads (GovCloud) but this is only available to government agencies.

A different approach to distributed long-term EHR archiving comes from the DIGHT (Distributed Information store for Global Healthcare Technology) project that is being deployed as a nationalized EHR for the citizens of India. The Indian Centre for Development of Advanced Computing (C-DAC) is responsible for the front-end interface and EHR data standardization whereas the Swedish Institute of computer Science is responsible for the distributed storage aspects of the project using open source technology and open standards. The DIGHT project proposes to develop data replication algorithms that ensure the security, availability and low latency of the Indian EHR data. The proposal centers on design of partially synchronous networks where different health data are assured varying levels of consistency across the network depending on their nature and immediacy. The storage is based on multiple data clusters with use of Distributed Hash Tables for data retrieval. The data will be encrypted and an audit trail maintained. There is so far, no available information on the real world deployment and performance of the proposed DIGHT system.

We herein propose a further still approach to distributed EHR storage. Briefly, each healthcare provider would act as a node by providing storage for use in a peer-to-peer (P2P) distributed file storage "cloud" that is at least equal capacity to the health data it owns and needs to archive. The resulting network can be linked together with technology similar to that employed in existing P2P DFSs. One such example is the University of California, Berkeley OceanStore project that provides a highly available and durable storage over untrusted servers using promiscuous data caching and a Byzantine-fault tolerant commit protocol (that is, a system tolerant to components that fail in arbitrary ways). Although it includes versioning and is claimed to be able to survive "all but a planet-wide failure event" the system is best suited to keep a permanent, read-only form of a data object and, therefore, not well suited for the PWPHN EHR (Table **1**). Similarly, the P2PCooperative File System (CFS) by the Massachusetts Institute of Technology (MIT), although equally robust, is a read-only distributed file system [7].

On the other hand, a system similar to Wuala, developed at the Swiss Federal Institute of Technology (ETH) may be better suited to the task and will be briefly described as a prototype application. The Wuala network can accept any size of file and includes public but also secured and shared data objects. It is deployed over an untrusted P2P network of users and thus a primary concern is maintaining data in case a P2P node goes offline. One approach is to ensure that the data always remains in the network, i.e. a node uploads all its data prior to going offline, but this is ofcourse time consuming during planned shut-downs and impossible in case of unexpected failures. Another approach is to introduce redundancy in the storage by replicating data across nodes. If a node fails, other nodes have to take over the failed node's key range and copy the data they become responsible for in order to maintain the replication factor. Simple replication, however, requires very high copy number to ensure high levels of availability. In Wuala, the data is encrypted (128-bit AES) and fragmented, using cipher-block chaining, prior to storing in the network. The data is not replicated across nodes, but rather secured by the use of forward error correction (FEC) or "erasure codes" that achieve orders of magnitude higher reliability for the same level of redundancy compared to replication. The Reed-Solomon FEC coding family is used in Wuala. Queries in case of temporary unavailability of nodes are tolerated as long as a certain number of nodes involved in the algorithms are alive.

A shortfall of Reed-Solomon coding, however, is that the entire data needs to be downloaded from the network to repair a failed node [8]. More recent work on this field has led to the description of "regenerating codes" that retain the FEC reconstruction property but minimize repair bandwidth [9]. Each node periodically checks the data on the cloud and creates new fragments if any nodes storing data appear to have permanently left the system. Parallel downloading from multiple nodes can give fast download speeds of data off the EHR cloud.

In case of EHR systems, legitimate access requests to the cloud EHR are likely to be geographically nestled and query the same nodes. Locality keeps routing distance to the objects as short as possible and improves latency, reduces the chance of loss or corruption, increases the chance of the data remaining accessible in case of near-catastrophic loss of interconnectivity and reduces bandwidth across the network. Such object location systems include Tapestry, as used by the OceanStore project [10]. Structured systems such as these, however, leave the cloud subject to clever attacks whereby an attacker (or enemy) could pull out a number of nodes simultaneously off the system and cause the PWPHN to be partitioned. The addition of random links to more distant nodes can speed up routing and potentially avoid partitioning problems taking benefit of "small world effects" [11-13]. The previously mentioned Amazon S3 is similarly partitioned into 7 geographical compartments, plus an additional network for the GovCloud. The partitioning is so strict that data objects stored in a specific Region never leave the Region unless specifically transferred out.

The paradigm used by Wuala is based on the Chord overlay scheme that defines three node classes: Super Nodes, Storage Nodes and Client Nodes. The Super Nodes, that are responsible for network message routing, must be constantly online and should be managed by the organization that provides the distributed storage service. In the case of a PWPHN, a suitable organization such as the WHO (World

Health Organization) could serve as the service guarantor and SuperNode manager.

To promote equity amongst nodes (viz. healthcare providers) each node will be allowed to write to the EHR cloud the same amount of data as they are providing in storage. If the provider wishes to store health record files in the cloud that are not part of the main EHR dataset, these could be accommodated, potentially for a higher storage provided to the cloud (e.g. providing 1.25TB to the cloud to permit storage of 1TB of additional non-core EHR data) [14]. Data for the provider's patients will reside locally at each node for immediacy of access with asynchronous synchronization to the cloud EHR.

The cloud EHR, therefore, will effectively not only create the online health record information repository for the PWPHN but also serve as an archiving solution for each provider. Given the redundancy of the data storage in the cloud, the risk of loss is lower than other archiving solutions, whilst the setup cost would be less, as the fault-tolerance of each storage node can be quite high. This cost-saving possibility is likely to entice many providers to join the PWPHN, the success of which would depend on a high percentage of providers opting in.

Scalability

According to the CAP theorem, three desired properties of distributed databases are Consistency of data, Availability and Partition tolerance. In 2000, Brewer made the conjecture that in an asynchronous network it is impossible to achieve all three [15, 16]. Nonetheless, using a number of techniques such as distributed hash tables the system can have high availability, scalability and security [17]. An EHR system will tend to store a very large number of small records (such as textual summary care records) and a smaller number of very large records (such as imaging). Similarly, data access will be small read/writes (text updates) and large streaming read/writes (image viewing).

Shared nothing architectures (SNA) are becoming more prevalent in data warehousing. With SNA,each node being independent and self-sufficient gives it almost infinite scalability as more healthcare nodes will get added but with no single bottleneck to slow the system down. Google has demonstrated this feature of a pure SNA system very well and calls it “sharding”. In terms of the PWPHN database architecture 'sharding' or horizontal partitioning could occur in many levels. By splitting the database by rows, rather than columns (as in the process of database normalization) the index size is generally reduced. Furthermore, a database “shard” containing a patient's record can be placed on specific nodes in the PWPHN, geographically near where the patient may seek health care and, therefore, from where an access request may originate. Real-world segmentation (e.g. trans- Atlantic data separation) may be appropriate but not necessarily enforced. The additional advantage of sharding the PWPHN database would be that large partitionable tables can be split across the server/nodes whilst smaller tables can be replicated across them *en masse*.

Nonetheless, transactions on a DHT are traditionally relatively slow because write operations have to establish membership for the operation and then perform the write using some kind of consensus algorithm. Recently published work may prove a solution for the lookup inconsistency problem with faster writes based on a simple quorum algorithm though this approach has yet to be proven in practice [18]. Alternatives, such as Bigtable used by Google for web indexing, Google Earth and Google Finance, have also proven to be scalable to petabytes of data and thousands of machines whilst meeting the latency requirements [19].

Security

In order to attract and retain users, the cloud EHR PWPHN must be trusted, as unauthorized disclosure or adulteration of a patient's EHR may impact their health, employment prospects and social standing [20]. Security against unauthorized access is one of the greatest concerns in a PWPHN accessible from multiple nodes across the planet. A trusted PWPHN can only be achieved if appropriate measures are taken to secure the information; efforts in this direction recently rely on public-key cryptography and digital certificates.

The principal concern is not the safe transmission of information across the public network, as sufficiently strong encryption algorithms can be applied efficiently. The risk arises by the high number of users requiring access and the difficulty in assessing their clearance and the trustworthiness of their sites [21]. In the PWPHN, a number of independent and geographically distant providers have authority to administer access to their resources. Therefore, rather than maintaining a centralized agency, authority can be delegated to regional security administrators (see Fig. **1**). Control over the regional administrators can be centrally administered (e.g. by the World Health Organization), but they can have considerable autonomy within their regions. The delegation can be repeated to set-up sub-regions down to the level of a specific healthcare provider authorizing each of its staff for access [22]. The DIMEDAC (DIstributed Medical Database Access Control) security policy has been previously proposed for this purpose [21]. This includes location controls (based on site, domain and living space - i.e. established relationship between doctor and patient) as well as access control mechanisms. The later define sensitivity levels in the EHR data that can be accessed depending on the user's role (role-based access control, RBAC) [23]. It also takes into account location hierarchy, such that, for example, a user accessing data from beyond his administrative domain receives reduced privileges for his role. A multi-dimensional access matrix can then be used to define how a user's role and location affects authority to access a data set within a certain node. The system, with an example DIMEDAC certificate-based implementation is described by Mavrides *et al.* [21]. The security would be complemented by a maintained audit trail of EHR accesses and all data cryptographically attributed to an authorized entity (such as a hospital).

Blaze, Bleumer and Strauss proposed in 1998 an application called *atomic proxy re-encryption*, where a semi-trusted proxy can convert a cipher text for Alice into a cipher text for Bob without seeing the underlying plaintext [24]. Such a system of a fast and secure re-encryption has been predicted to become popular for managing encrypted file systems [25]. The system could be incorporated to the security of a PWPHN. The security server of a healthcare

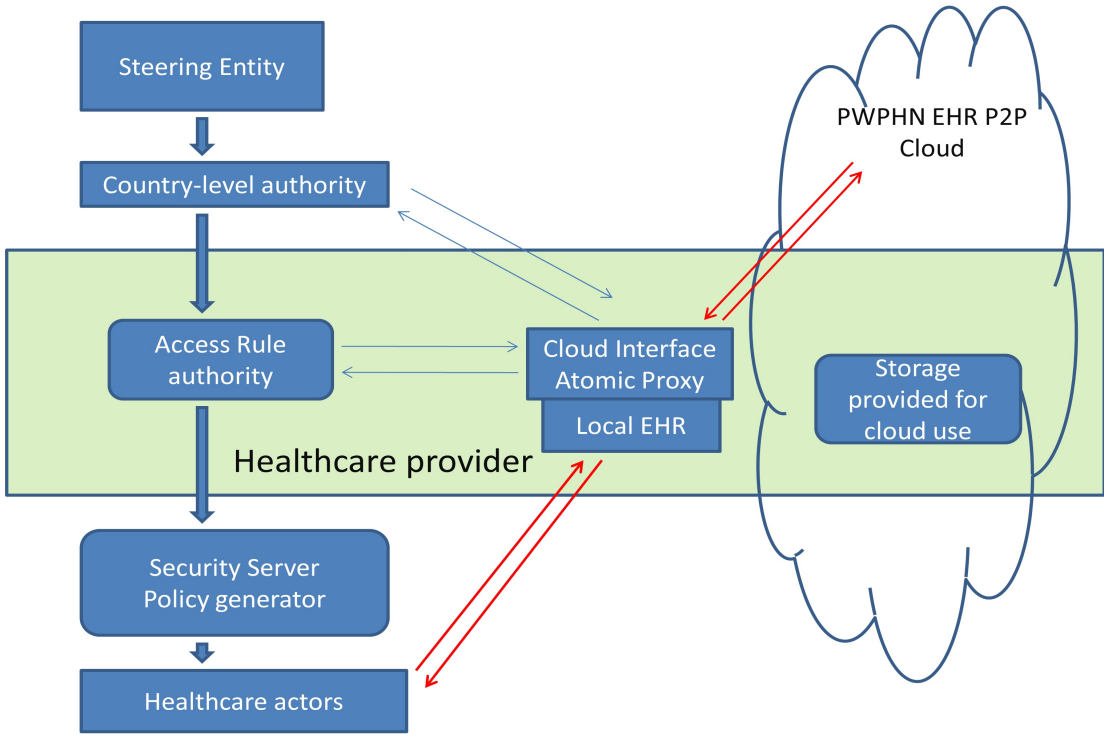

Fig. (1). The PWPHN Steering entity delegates authority to regions and sub-regions down to the level of the healthcare provider (e.g. hospital). Policy propagation occurs down to the security server policy generator that authenticates and authorizes health workers reading and writing on EHRs based on user-role, dataset, and user-location hierarchy employing a three-dimensional access matrix to define final user permissions (see text for details). Healthcare actors interact with the local EHR. The healthcare provider provides storage (at least equal in size to the local EHR requirements) for cloud usage. In exchange, the local EHR data is archived to the P2P cloud. Some of this can remain private and specific to the healthcare provider, but most becomes part of the PWPHN EHR and available to other healthcare providers planet-wide.

provider, considered semi-trusted by the planet-wide steering entity, could use atomic proxy re-encryption to pull data off the EHR cloud and transform the cyphertext so that it may be read by the healthcare provider's own secret key. Any data contributed by the healthcare provider to the cloud would be similarly encrypted using the master steering authority public key so that atomic proxy re-encryption can be implemented at a remote site.

Such complicated yet robust schemes are essential to establish trust in the PWPHN system. In a recent study, although the majority of patients and physicians believe that the benefits of computerization outweigh the risks of potential loss of confidentiality, there remain a significant proportion (10% in the study) who oppose the computerized sharing of health information [26]. Significant opt-outs from the PWPHN enforced by proponents of such notions may jeopardize the universality of the PWPHN cloud EHR. High security is essential to avoid such catastrophic opt-outs.

Human Nature

A remaining concern, despite any security features that may make the PWPHN effectively inaccessible by outsiders, is the danger from within. In the recent Websense Security Labs™ 2012 Threat Report, the human element remains identified as the weakest link in networked systems. Large database systems appear to conform to Cambridge professor Ross Anderson's rule that “a large, functional database can never be entirely secure, while a completely secure database can never be functional”. To maintain functionality, the database must remain accessible to certain authorized individuals. Yet, due to human nature, the temptation will always be there for authorized individuals to use their access rights to inappropriately access patients' records - sometimes simply out of curiosity.

Such incidents occurred soon after the NHS Spine became available. In January 2010, Dr. Andrew Jamieson, a doctor in Scotland UK, looked up “out of interest” the emergency care summary records of the UK prime minister as well as other famous people including journalists, footballers and politicians. Interestingly, a decision was made not to prosecute him, possibly as this would require bringing the victims’ EHR information to court as evidence.

Prosecution for such breaches of security may be even more difficult if the PWPHN was accessed from a different legislative area, country or continent. Centralized policing of the system would be near impossible and a more practical solution would be the delegation of responsibility to each of the trusted nodes, similar to the delegation of authority described above.

Table 1. Application of Distributed Storage Systems in Healthcare Application

System	Storage Infrastructure	Comments
Berkeley Ocean Store, UCLA	Tapestry overlay for resource virtualization, FEC, location-independent routing abstraction	Pro: Best suited for read-only archiving. Con: prone to partitioning
Cooperative File System (CFS), MIT	DHT DHash and Chord location protocol	Pro: Decentralized, scalability Con: File system layer only provides read-only file system semantics, no quota
Wuala, Swiss FIT	DHT cipher-block chaining, Reed-Solomon FEC, Chord P2P overlay (super nodes, storage nodes and client nodes)	Pro: provider archiving purposes, scalable, peers are both suppliers and consumers of resources, quota
BigTable, Google	Google File System, DFS by SNA sharding	Pro: scalable Con: single master server, not distributed outside Google
Simple Storage Service, Amazon	Zero-hop DHT, details unknown	Pro: Excellent latency and scalability. Con: Closed proprietary design

MIT, Massachusetts Institute of Technology; FIT, Federal Institute of Technology; UCLA, University of California, Los Angeles, SNA, shared nothing architecture, DHT, distributed hash table; FEC, forward error correction; DFS, distributed file system.

CONCLUSIONS

The Internet provides previously almost unimaginable opportunities for data sharing and interaction between doctors, patients and researchers. After over three decades of struggling with the limitations and vagaries of proprietary solutions, health service organizations worldwide now have the means towards establishing a ubiquitous and seamless system using the Internet's attractive infrastructure. Andrew Tanenbaum, the famous professor and writer of the standard computer science textbooks, humorously wrote in his book "Distributed Operating Systems" that "The design of a world-wide, fully transparent distributed file system for simultaneous use by millions of mobile and frequently disconnected users is left as an exercise for the reader" [27]. Fifteen years later, we appear to be very close to completing the exercise left to us by Professor Tanenbaum and may consider applying the results to healthcare storage.

Such a system of distributed health care record storage would be an integral part of a cloud EHR but significantly raises the element of risk regarding the integrity and confidentiality of the information. Our review of the currently available technologies suggests that such a system is technically feasible. Nonetheless, the socioeconomic and political acceptability of such a planet-wide EHR system remains to be established.

CONFLICT OF INTEREST

The author confirms that this article content has no conflicts of interest.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

[1] Collins T. The DH "Healthspace" IT project few patients want or use. ComputerWorldUK; 2010 Nov; [cited 2011 April 20]; Available from: http: //blogs.computerworlduk.com/the-tony-collins-blog/2010/11/the-dh-healthspace-it-project-few-patients-want-or-use/index.htm.

[2] HIMMS analytics. HIMSS Analytics Stage 7 Award – Paperless and Proud of IT! 2011 Mar [cited 2011 April 25]; Available from: http: //www.himssanalytics.org/hc_providers/stage7Award.asp.

[3] Shahar Y. Information Warfare. Institute for Counter-Terrorism; 1997 Feb [cited 2011 April 25]; Available from: http: //www.ict.org.il/Articles/tabid/66/Articlsid/715/currentpage/39/Default.aspx.

[4] Jerräng M. It-haveri på Karolinska universitetssjukhuset (IT crash at the Karolinska University Hospital). IDG; 2008 Dec [cited 2011 April 24]; Available from: http: //itivarden.idg.se/2.2898/1.202051.

[5] Gortzis L, Koubias S, Nikiforidis G. Design and implementation of a web-enabled haematological system. Comput Methods Prog Biomed 2004; 75(3): 221-34.

[6] Balding C. Is Amazon AWS Really HIPAA Compliant Today? : CloudSecurity; 2009 Apr [updated 4/8/2009; cited 2012 July 25]; Available from: http: //cloudsecurity.org/blog/2009/04/08/is-amazon-aws-really-hipaa-compliant-today.html.

[7] Dabek F, Kaashoek MF, Karger D, Morris R, Stoica I. Wide-area cooperative storage with CFS. Proceedings of the eighteenth ACM symposium on Operating systems principles. Banff, Alberta, Canada: ACM 2001; pp. 202-15.

[8] Dimakis AG, Ramchandran K, Wu YN, Suh CH. A Survey on Network Codes for Distributed Storage. Proc IEEE 2011; 99(3): 476-89.

[9] Shah NB, Rashmi KV, Kumar PV, Ramchandran K. Regenerating codes for distributed storage networks. Proceedings of the Third international conference on Arithmetic of finite fields. Istanbul, Turkey; 1893755: Springer-Verlag 2010; pp. 215-23.

[10] Zhao BY, Kubiatowicz JD, Joseph AD. Tapestry: a fault-tolerant wide-area application infrastructure. SIGCOMM Comput Commun Rev 2002; 32(1): 81.

[11] Dai TN, Bach LN, Duy LV. Improving Freenet's performance by adaptive clustering cache replacement. 2009 IEEE-RIVF International Conference on Computing and Communication Technologies: Research, Innovation and Vision for the Future. Vietnam 2009; pp. 195-201.

[12] Korte C, Milgram S. Acquaintance networks between racial groups - application of small world method. J Pers Soc Psychol 1970; 15(2): 101-8.

[13] Milgram S. Small-world problem. Psychol Today 1967; 1(1): 61-7.

[14] Hao LM, Lu SNA, Tang JH, Yang ST. An efficient and robust self-storage P2P reputation system. Int J Distrib Sens Network 2009; 5(1): 40.

[15] Brewer EA. Towards robust distributed systems (abstract). Proceedings of the nineteenth annual ACM symposium on Principles of distributed computing. Portland, Oregon, United States 343502: ACM 2000; p. 7.

[16] Shen X, Yu H, Buford J, Akon ME. Handbook of peer-to-peer networking. 1st ed. Springer: New York 2010.

[17] Risson J, Moors T. Survey of research towards robust peer-to-peer networks: search methods. Comput Networks 2006; 50(17): 3485-521.

[18] Shafaat T, Ghodsi A, Haridi S. Dealing with network partitions in structured overlay networks. Peer-to-Peer Networking Applications 2009; 2(4): 334-47.

[19] Chang F, Dean J, Ghemawat S, *et al*. Bigtable: A distributed storage system for structured data. ACM Transact Comput Syst 2008; 26(2): 4.

[20] Grimson J, Grimson W, Hasselbring W. The SI challenge in health care. Commun ACM 2000; 43(6): 48-55.

[21] Mavridis I, Georgiadis C, Pangalos G. Access-rule certificates for secure distributed healthcare applications over the Internet. Health Inform J 2002; 8(3): 127-37.

[22] Sandhu R, Munawer Q. How to do discretionary access control using roles. Proceedings of the third ACM workshop on Role-based access control; Fairfax, Virginia, United States; 286893: ACM 1998; pp. 47-54.

[23] Lu PY, Song H, He LJ. Design and Implementation of Privilege Management System Based on RBAC. EBM 2010: Int Conf Eng Bus Manage 2010; 1-8: 5078-81.

[24] Blaze M, Bleumer G, Strauss M. Divertible protocols and atomic proxy cryptography. Adv Cryptol - Eurocrypt '98 1998; 1403: 127-44.

[25] Ateniese G, Fu K, Green M, Hohenberger S. Improved proxy re-encryption schemes with applications to secure distributed storage. ACM Trans Inf Syst Secur 2006; 9(1): 1-30.

[26] Perera G, Holbrook A, Thabane L, Foster G, Willison DJ. Views on health information sharing and privacy from primary care practices using electronic medical records. Int J Med Inform 2011; 80(2): 94-101.

[27] Tanenbaum AS. Distributed operating systems. Englewood Cliffs, NJ: Prentice Hall 1995.

Received: July 12, 2012 Revised: September 22, 2012 Accepted: September 25, 2012