A Fast and Exact Verification of Inter-Domain Data Transfer based on PKI

Im Y. Jung* · Hyeonsang Eom** · Heon Y. Yeom***

Abstract

Trust for the data created, processed and transferred on e-Science environments can be estimated with provenance. The information to form provenance, which says how the data was created and reached its current state, increases as data evolves. It is a heavy burden to trace and verify the massive provenance in order to trust data. On the other hand, it is another issue how to trust the verification of data with provenance. This paper proposes a fast and exact verification of inter-domain data transfer and data origin for e-Science environment based on PKI. The verification, which is called two-way verification, cuts down the tracking overhead of the data along the causality presented on Open Provenance Model with the domain specialty of e-Science environment supported by Grid Security Infrastructure (GSI). The proposed scheme is easy-applicable without an extra infrastructure, scalable irrespective of the number of provenance records, transparent and secure with cryptography as well as low-overhead.

Keywords : Domain-based Trust Reasoning; Open Provenance Model; e-Science Environment; PKI

Received : 2011. 07. 28. Final Acceptance : 2011. 09. 05.

※ This research was supported by the MKE(The Ministry of Knowledge Economy), Korea, under the ITRC(Information Technology Research Center) support program supervised by the NIPA(National IT Industry Promotion Agency)(NIPA-2011-(C1090-1111-0004)).

* School of Computer Science & Engineering, Seoul National University, Post-Doc Researcher), e-mail : iyjung@dcslab.snu.ac.kr,

** Assistant Professor, School of Computer Science and Engineering, Seoul National University, e-mail : hseom@cse.snu.ac.kr,

*** Professor, School of Computer Science and Engineering, Seoul National University, e-mail : yeom@snu.ac.kr

1. Introduction

Provenance states how data reaches its current state [Simmhan et al., 2005; Buneman et al., 2001; Cui et al., 2000]. The information to form provenance includes the processing history of data, the related memos, the transfer history of data and so on. Open Provenance Model (OPM) [Moreau et al., 2010; Open provenance model] allows users to grasp what provenance says easily. OPM represents provenance with artifact, arc, process and agent. As immutable piece of state, the artifact on OPM includes a physical object and a digital representation such as data. Process means an action or the series of actions performed on or caused by artifacts, resulting in new artifacts. Arc on OPM shows "some change" between two artifacts or one artifact and one process or two processes. Agent is a contextual entity acting as a catalyst of a process. But, OPM does not prove or show the trust of process, artifact and arc. Because the information to form provenance is accumulated as data evolves and there is no standard for the format or the structure to record provenance, it is hard to trace and verify data with provenance as provenance increases. As electronic record, data is easy to transfer, copy and modify. Provenance which records the processing history of data is not created and managed by one system or a domain generally [Tilmes and Fleig, 2008]; e-Science Grid is one example of the domain. Along the transfer path of data, provenance is only added by different agents or provenance recording systems. As provenance increases, it gets a heavy burden to trace all the information provenance provides along the domains data transferred.

On the other hand, even though provenance is an important ground to trust data, it is another data to be verified. OPM does not represent the trust of data. Therefore, it is necessary to verify whether the artifact described by provenance is the exact data referenced and the processing described causes the state change of data really.

The domain such as e-Science environment, where data as research product is important, needs provenance to prove the trust of data [Bose, 2002]. It is emphasized that origin and processing history of data is crucial for encouraging effective sharing of science research data among collaborative scientific communities [Jagadish et al., 2004].

2. Related Works

The issues of trust reasoning with provenance are the requirement of expert system or the credential records from reputation system [Hartig et al., 2009; Prat et al., 2008; Keijzer et al., 2007; Gil and Artz, 2007]. The fast and exact verification scheme proposed in this paper tracks data transfer and data origin only with provenance as [Chapman et al., 2010].

Many researches track the history of data along the causality of the data states from the current version of data to the origin of the data [Hasan et al., 2009; Braun et al., 2008]. And, many researches assume that provenance is believable [Tilmes and Fleig, 2008; Barkstrom, 2010]. A scheme of provenance tracking using

bit vector [Gadang et al., 2008] was proposed to overcome the weaknesses of the aforementioned approaches; heavy load to track the causality of the data versions. In the paper, it was possible to know the data origin using several bits. However, it was not possible to check exactly which version of data was referenced. The information for data source is somewhat inexact without considering the data version, because the data evolves with one identity. Moreover, the paper does not verify whether the bit vector is trusted. The overhead of the approach includes the requirement that all domains must know each bit position at the bit vector by which they are represented in advance. Any additional domain causes the bit vector to be updated and requires the consent of all domains.

On the other hand, new services or new system structures for the trust reasoning with provenance were proposed on e-Science environment [Anand et al., 2010; Szomszor and Moreau, 2003; Tilmes and Fleig, 2008; Barkstrom, 2010]. But, there is little mention of the tracking overhead for the trust reasoning of data with provenance especially for the situation that the provenance records are extended several domains [Hasan et al., 2009]. [Barkstrom, 2010] proposed BFS algorithm to track the history of data processing with provenance; the tracking cost is proportional to the number of the events which cause provenance to be recorded. [Tilmes and Fleig, 2008] pointed out the difficulty of tracking data processing with provenance especially when the data is distributed or transferred among multiple organizations. It proposed two types of meta data to reproduce the data rapidly; it focused on how to reproduce a specific version of data fast. But, it did not mention the verification of the data reproduced by provenance or the tracking overhead of provenance to prove the trust of the data reproduced.

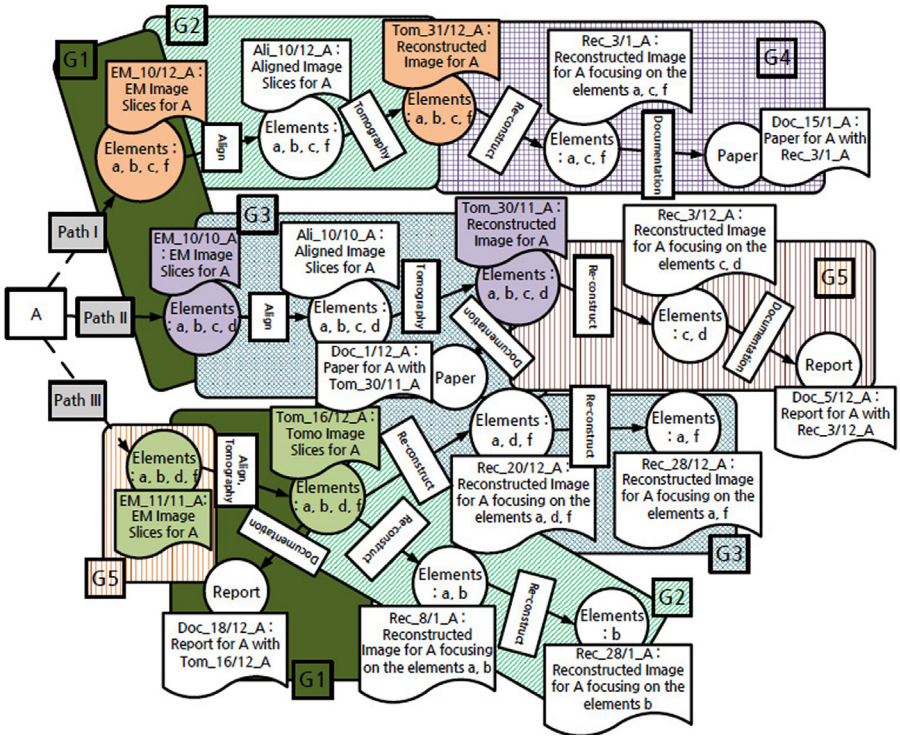
3. e-Science Domain and Problem Specification

e-Science environment, which is also called e-Science Grid or e-Science domain, provides an improved open infrastructure for science research by mitigating the limitation of time and space which is an obstacle of offline laboratory [Sahoo et al., 2008]. It enables scientific experiments by remote control of the expensive equipments. The enormous data produced from the experiments and processing can be stored at e-Science domain safely and retrieved by any user authorized at any time at any place via Internet. High performance computers, which can process large data in a rather short time, can be shared on e-Science environment. Each e-Science Grid assorts intra-domain security supported by Grid Security Infrastructure (GSI) [Foster et al., 1998; Welch et al., 2003; The Globus Security Team, 2005] to protect its expensive equipments and its important data. However, the data transferred between the domains is not protected [Lang et al., 2006]. Provenance is needed by the domain where the trust of data is important. Because e-Science environment is constructed to support science research, it is important to secure the data obtained during the research and to guarantee the verification of the history of creation and proc-

essing of the data. Using provenance, a user can trace “process” or the “data” that led to the aggregation of services producing a particular output on e-Science domain [Bose, 2002].

HVEM Grid [Jung et al., 2007; Han et al., 2006] is a concrete example of e-Science environment. The experiments using High Voltage Electronic Microscope (HVEM) on HVEM Grid are executed with the purpose to know the elements and the microscopic structure of the sample to be studied. When the EM images obtained from HVEM and the images processed of the EM images are kept in HVEM Grid as experimental results, the information for the images and their processing is recorded as provenance. The provenance states by what processing condition and on what environment the images come to exist and change. <Figure 1> shows an OPM for the data produced and processed in HVEM Grid. Process which causes the changes on artifact is expressed as rectangle, artifact such as image is drawn as circle. The detail information recorded in provenance is written in the document symbol. A in <Figure 1> is a material for the experiment with HVEM. HVEM Grid has the EM images of A from HVEM and the processed images of the EM images through alignment, tomography and reconstruction.

One day, a researcher who read the paper, Doc_15/1_A, gets to have one question. He, who knows A, is doubtful whether Rec_3/1_A originated from A.



〈Figure 1〉 Data Evolution on OPM for HVEM Grid

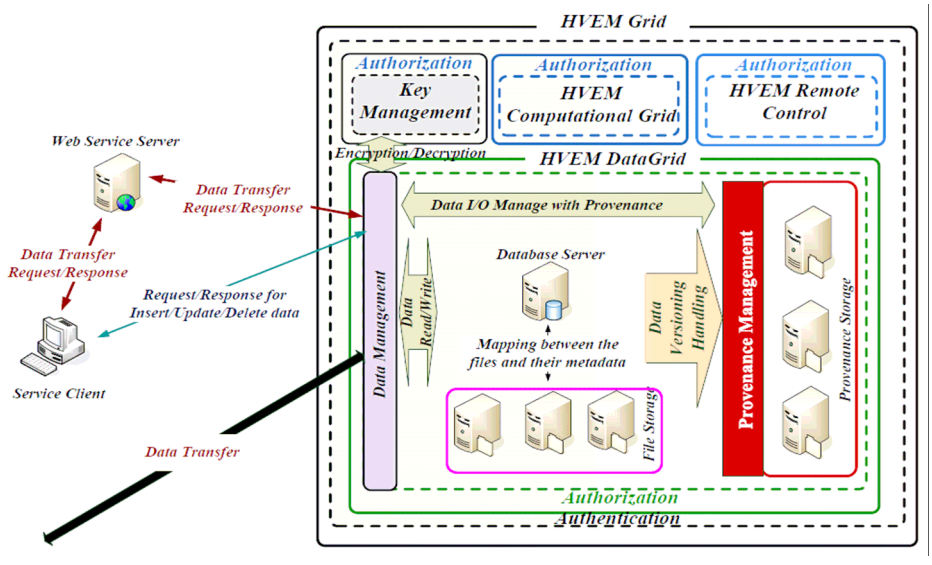
4. A Fast and Exact Verification of Inter-Domain Data Transfer and Data Origin on Open Provenance Model

This paper solves the question raised at section 3 with the two verifications : the verification of inter-domain data transfer and the verification of the real origin of data. The fast and exact verification of data transfer and data origin was proposed as the first step to track the entire history of data with provenance. The next step would be the tracking of data evolution at each domain as needed. The two steps compose the two-pass verification.

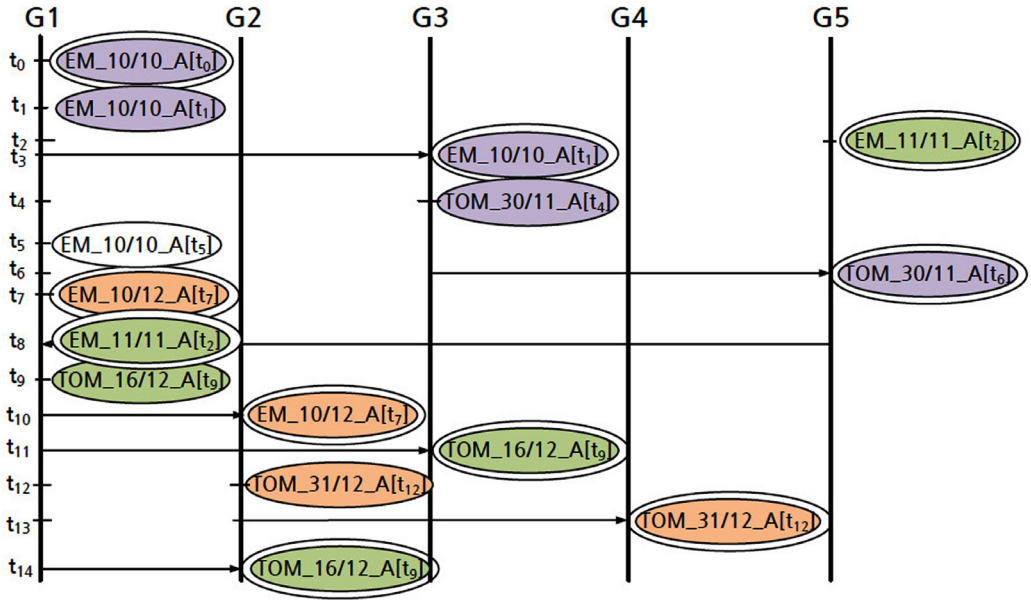
This paper assumes that each e-Science Grid is believable. It does not consider camouflage of the Grids and collusions between the Grids.

4.1 Preparation to Verify Inter-Domain Data Transfer and Data Origin

e-Science Grid is the domain where expensive experimental instruments and high performance computing devices are shared. Researchers save their significant data to the domain, too. Because the resources and the data which e-Science Grids manage are valuable and important, e-Science environment has its logs for the access to the resources and the data for the purpose of defense against attacks or system failures. The logs are recorded and managed by e-Science domain directly and cannot be touched by any individual who is not authenticated and authorized. Therefore, each e-Science environment is responsible for the truth of provenance and the data related to the provenance as shown in <Figure 2>. And, the trust of the data produced on e-Science environment can be estimated by the trust of the domain; this paper assumes that each e-Science Grid is believable. The processing history of data in provenance can be tracked and verified with the logs on the domain for the usage of instruments and com-



〈Figure 2〉 e-Science Grid : HVEM Grid



〈Figure 3〉 Data Flow among e-Science Domains

puting infrastructure, the storage access.

<Figure 1> represents the domains data has been transferred as data evolves on OPM : G1, G2, G3, G4 and G5. Each state of data, which is shared between the sender domain and the receiver domain at a data transfer, is indicated, too. For example, *EM_10/10_A* is shared by *G1* and *G3*. The problem is how to verify the data transfer and the data version transferred represented on the OPM of HVEM Grid.

<Figure 3> shows what to check in order to verify the data transfer between domains. At <Figure 3>, *EM_10/12_A* should be checked as follows. Both *G1* and *G2* should confirm that the data version *EM_10/12_A*[t_7] was transferred at t_{10} from G1 to *G2*. Both *G2* and *G4* should confirm that the data version *TOM_31/12_A*[t_{12}] was transferred at t_{13} from *G2* to *G4* in the same context. As data evolves, the states of data change. There is no guarantee for all the states of data to be kept and to be retrieved at any time. Therefore, each domain should add the following information to provenance.

- Each state of data and The effective time interval of the state

The effective data version of *EM_10/10_A* from t_0 is represented as *EM_10/10_A*[t_0]. The hash value of *EM_10/10_A*[t_0] and its effective time interval is kept as the data state of *EM_10/10_A* from t_0. The effective time interval is the duration each data version exists along the causality of data evolution to the current state of data. It becomes an important clue to search the data origin of the current state of data in a domain. At <Figure 3>, the effective duration of *EM_10/10_A*[t_0] is [t_0, t_1). At t_1, *EM_10/10_A*[t_0] changes its state to *EM_10/10_A*[t_1].

- The data origin at a domain should be tracked by considering the two cases as follows. Each domain should add the proper information to provenance.
- The data origin transferred from other domain.

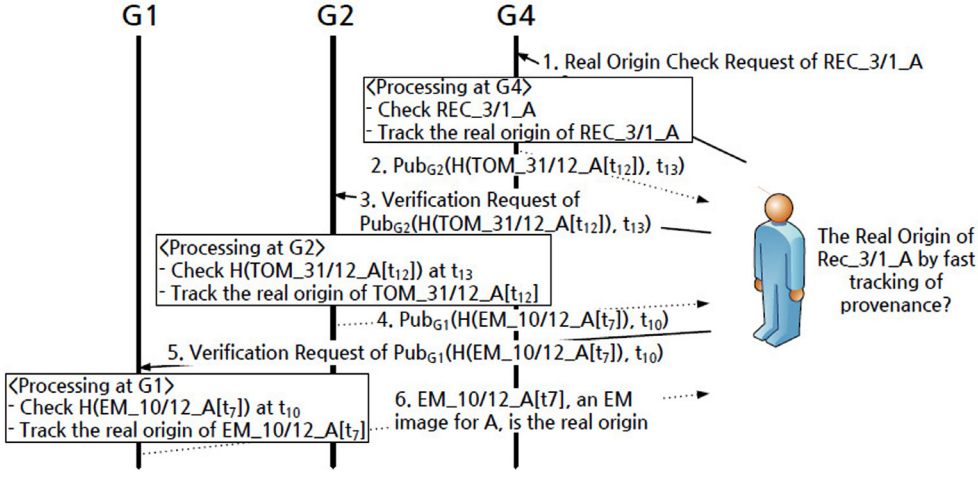
If the data version of *EM_10/10_A* transferred to *G*3 is *EM_10/10_A*[t_1], *G*3 keeps $Pub_{G1}(H(EM\_10/10\_A[t_1]), t_3)$ additionally, which is the encrypted value of $(H(EM\_10/10\_A[t_1]), t_3)$ with the public key of *G*1. The pair of $(H(EM\_10/10\_A[t_1]), t_3)$ shows the existence of $EM\_10/10\_A[t_1]$ from t_3 at *G*3. $Pub_{G1}(H(EM\_10/10\_A[t_1]), t_3)$ indicates that *G*1 will prove the transfer of $EM\_10/10\_A[t_1]$ at t_3.

- The data origin created at the domain without any reference of the data from other domain.

*G*1 holds $Pub_{G1}(H(EM\_10/10\_A[t_0]), t_0)$ additionally. $Pub_{G1}(H(EM\_10/10\_A[t_0]), t_0)$ indicates that only *G*1 can prove the creation of *EM_10/10_A*[t_0] at t_0. If a domain keeps the encrypted information with its public key, it means that the data is created at the domain without any reference of the data on other domains.

4.2 Verification Process of Data Transfer and Data Origin

<Figure 4> shows the tracking process of the real origin of *Rec_3/1_A*. The user to track the data origin asks *G*4 the source of *Rec_3/1_A*. The user receives $Pub_{G2}(H(TOM\_31/12\_A[t_{12}]), t_{13})$, which means *G*4 got $TOM\_31/12\_A[t_{12}]$ from *G*2 at t_{13}, as the answer. When he asks *G*2 the origin of $TOM\_31/12\_A[t_{12}]$ with $Pub_{G2}(H(TOM\_31/12\_A[t_{12}]), t_{13})$, *G*2 verifies $Pub_{G2}(H(TOM\_31/12\_A[t12]), t_{13})$ with its private key, Pri_{G2}. After *G*2 confirms that $TOM\_31/12\_A[t_{12}]$ existed at t_{13}, it informs the user that the source of $TOM\_31/12\_A[t_{12}]$ is $EM\_10/12\_A[t_7]$ and sends $Pub_{G1}(H(EM\_10/12\_A[t_7]), t_{10})$ as the answer. The user asks *G*1 the source of $EM\_10/12\_A[t_7]$ with $Pub_{G1}(H(EM\_10/12\_A\ [t_7]), t_{10})$. After *G*1 verifies $Pub_{G1}(H(EM\_10/12\_A[t_7]), t_{10})$ with its private key and $Pub_{G1}(H(EM\_10/12\_A[t_7]), t_7)$, it informs that $EM\_10/12\_A[t_7]$ is the real origin of *Rec_3/1_A*. Because *G*1 can prove that $EM\_10/12\_A[t_7]$ was produced from



<Figure 4> Verification of Inter-Domain Data Transfer and Data Origin

the experiment with HVEM at t_7 on the domain, $G1$ can confirm $EM\_10/12\_A[t_7]$ to be the real origin.

5. Contribution and Evaluation

The contribution of this paper is to construct the two-pass verification of data with provenance instead of the one-pass verification such as the verification according to the causality of provenance, and to propose a scheme to verify data transfer and data origin fast and exactly as the first step of the two-pass verification. If the first pass succeeds without any problem, the second pass can be executed optionally by all the verification of each data version with provenance along the causality of provenance at each domain. The result of the two-pass verification, whose objective is to verify data transfer among e-Science domains and data origin, is the same with that of the verification according to the causality of provenance shown in OPM.

The characteristics of the proposed scheme are as follows.

- It needs no extra infrastructure.

Each e-Science domain has the security schemes for authentication and authorization for the access to the resources in the domain [Welch et al., 2003]. For the scheme proposed, each domain is only to add hashing and encryption/decryption based on PKI to its infrastructure.

- It is scalable.

At the first step of the two-pass verification, the proposed scheme is not affected by the amount of provenance to be verified directly, but by the number of the domains which data is transferred to. The changes on data produce provenance records as many. The overhead to track data origin according to the causality of data states increases as provenance records are produced [Hasan et al., 2009; Braun et al., 2008]. The verification cost of two-pass verification with the first step and the second one is the same with that of the schemes proposed by [Hasan et al., 2009; Braun et al., 2008]. But, with the first step without the second one, users can verify the real origin and the transfer history of data. Therefore, the proposed scheme is scalable. Moreover, if data is transferred to the domains to which the previous versions of the data have been transferred before, the overhead to verify the origin of data is cut down.

- It includes the security scheme to ensure an exact verification.

The proposed scheme confirms the data version effective at certain time interval using hash function of one-way verification function without revealing the content of data. The data state and the time at a data transfer, which are encrypted by the public key of the sender domain.

- The overhead of time and space is low

During the verification, the times for hashing, hash verification and encryption/decryption based on PKI are small as shown in <Table 1> and <Table 2>. The required space to keep hash value is also not large as shown in <Table 3>.[1)]

〈Table 1〉 Hashing Time and Hash Verification Time of the Data and its Provenance

6Data Type	Data Size (MB)	Hashing Time (s)			Hash Verification Time (s)		
		MD5	SHA1	SHA256	MD5	SHA1	SHA256
rec	338	1.572	2.604	7.552	2.98	4.057	9.345
zip	103	1.23	1.214	2.596	0.955	1.285	2.57
pdf	2	0.048	0.038	0.064	0.025	0.037	0.073

〈Table 2〉 Encryption and decryption time of the hash value

Data Type	Hash Algorithm	Encryption Time (s)		Decryption Time (s)	
		RSA1024	RSA2048	RSA1024	RSA2048
rec	MD5	0	0.006	0.181	0.041
	SHA1	0.03	0.003	0.162	0.034
	SHA256	-	0	-	0.034
zip	MD5	0	0	0.141	0.034
	SHA1	0.03	0.003	0.141	0.037
	SHA256	-	0	-	0.038
pdf	MD5	0	0	0.141	0.038
	SHA1	0.03	0	0.144	0.034
	SHA256	-	0	-	0.038

〈Table 3〉 Information size after hashing and encryption with the private key of the e-Science domain

	Hash Algorithm			PKI Algorithm	
	MD5	SHA1	SHA256	RSA1024	RSA2048
Size(B)	16	20	32	256	256

1) The experimental environment was a computer with a Pentium 4 CPU at 3.20GHz with 3G of RAM. The data for experiment was collected from the HVEM Grid. The file format of rec is necessary to convert the 2D images from HVEM to its corresponding 3D image. The performance using crypto++ library 5.5.2 [Crypto++ library] and the QuickHash library [Quickhash library] is given in <Table 1>, <Table 2> and <Table 3>. The provenance information of the data per version was less than 1KB and consisted of the creator, the creation time, the creation environment, and the list of data referenced. The experimental results are the average of 100 trials.

6. Conclusion and Future Work

This paper proposed a fast verification of inter-domain data transfer to track data history from the current state of data to the real origin of the data. The proposed scheme cuts down the tracking overhead of data and provenance along the causality presented on OPM because it checks only the inflow data and the outflow one at each domain along the causality in OPM. The fast tracking is based on the domain specialty such as the intra-domain security of e-Science domain. Moreover, the verification of data transfer is exact and transparent. The proposed scheme requests the sender domain to confirm that the domain possessed the data sent to the receiver domain at the transfer time for the exact verification of data transfer. The hash values of data transferred, the effective time interval and the encrypted values of them enables any user to request the verification without revealing the content of data. The overhead of the proposed scheme is negligible because the hash values encrypted and the data to represent the time interval are very small and the verification time only includes the decryption time, the hash verification time and the communication time to query the initial version of data to each domain. The scalability of the proposed scheme is another merit because the number of provenance records or any change of the number does not affect the performance of the scheme directly.

In the future, the quantitative analysis for the proposed scheme will be added.

Reference

[1] Chapman, A., Blaustein, B., and Elsaesser, C., "Provenance based belief", *In Proceedings of the 2nd Workshop on the Theory and Practice of Provenance*, July 2010.

[2] De Keijzer, A. and Van Keulen, M., "Quality measures in uncertain data management", *Scalable Uncertainty Management*, Vol. 4772, 2007, pp. 104-115.

[3] Lang, B., Foster, I., Siebenlist, F., Ananthakrishnan, R., and Freeman, T., "A Multipolicy Authorization Framework for Grid Security", *In Proceedings of IEEE NCA06 Workshop on Adaptive Grid Computing*, 2006.

[4] Barkstrom, B. R., "A mathematical framework for earth science data provenance tracing", *Earth Science Informatics*, Vol 3, No. 3, 2010, pp. 167-169.

[5] Tilmes, C. and Fleig, A. J., "Provenance Tracking in an Earth Science Data Processing System", *Provenance and Annotation of Data and Processes*, 2008.

[6] Anand, M. K., Bowers, S., Altintas, I., and Ludascher, B., "Approaches for exploring and querying scientific workflow provenance graphs", *Lecture Notes in Computer Science*, Vol. 6378, 2010, pp. 17-26.

[7] Anand, M. K., Bowers, S., and Ludascher, B., "Techniques for efficiently querying scientific workflow provenance graphs", *In Proceedings of the 13th International Conference on Ex-tending Database Technology*, 2010.

[8] Szomszor, M., and Moreau, L., "Recording and reasoning over data provenance in web and grid services", *Lecture Notes in Computer Science*, Vol. 2888, 2003, pp. 603-620.

[9] Han, H., Jung, H., Yeom, H. Y., Kweon, H. S., and Lee, J., "Hvem grid : Experiences in constructing an electron microscopy grid", *Lecture Notes in Computer Science*, Vol. 3841, 2006, pp. 1159-1162.

[10] Jagadish, H., and Olken, F., "Database management for life science research", *ACM SIGMOD Record*, Vol. 33, No. 2, 2004, pp. 15-20.

[11] Foster, I., Kesselman, C., Tsudik, G., and Tuecke, S., "A Security Architecture for Computational Grids", *Proc. 5th ACM Conference on Computer and Communications Security Conference*, 1998, pp. 83-92.

[12] Jung, I. Y., Cho, I. S., Yeom, H. Y., Kweon, H. S., and Lee, J., "Hvem datagrid : Implementation of a biologic data management system for experiments with high voltage electron microscope", *Lecture Notes in Bioinformatics*, Vol. 4360, 2007, pp. 175-190.

[13] Moreau, L., Clifford, B., Freire, J., Futrelle, J., Gil, Y., Groth, P., Kwasnikowska, N., Miles, S., Missier, P., Myers, J., Plale, B., Simmhan, Y., Stephan, E., and Den Bussche, J. V., "The open provenance model core specification (v1.1)", *Future Generation Computer Systems*, 2010.

[14] Prat, N., and Madnick, S., "Measuring data believability : A provenance approach", *In Proceedings of the 41st Annual Hawaii International Conference on System Sciences*

(HICSS 2008), 2008.

[15] Hartig, O. and Zhao, J., "Using web data provenance for quality assessment", *In Proceedings of the 1st Int. Workshop on the Role of Semantic Web in Provenance Management(SWPM)*, 2009.

[16] Buneman, P., Khanna, S., and Chiew Tan, W., "Why and where : A characterization of data provenance", *In Proceedings of the 8th International Conference on Database Theory(ICDT 2001)*, 2001, pp. 316-330.

[17] Hasan, R., Sion, R., and Winslett, M., "Preventing history forgery with secure provenance", *ACM Transactions on Storage*, Vol. 5, No. 4, 2009, pp. 12:1-12:43.

[18] Bose, R., "A conceptual framework for composing and managing scientific data lienage", *In Proceedings of the 14th International Conference on Scientific and Statistical Database Management*, 2002, pp. 15-19.

[19] Cui, S., Widom, J., and Wiener, J. L., "Tracing the lineage of view data in a warehousing environment", *ACM Transactions on Database Systems*, Vol. 25, No. 2, 2000, pp. 179-227.

[20] Gadang, S. S., Panda, B., and Hoag, J. E., "Provenance tracking with bit vectors", *In Proceedings of the Fourth International Conference on Information Assurance and Security(ISIAS 2008)*, 2008, pp. 132-137.

[21] Sahoo, S. S., Sheth, A., and Henson, C., "Semantic provenance for escience : Managing the deluge of scientific data", *IEEE Internet Computing*, Vol. 12, No. 4, 2008, pp. 46-54.

[22] The Globus Security Team, "Globus Toolkit Version 4 Grid Security Infrastructure : A Standards Perspective", *Technical Report*, Globus Alliance, 2005.

[23] Braun, U., Shinnar, A., and Seltzer, M., "Securing provenance", *In Proceedings of the 3rd USENIX Workshop on Hot Topics in Security*, 2008.

[24] Welch, V., Siebenlist, F., Foster, I., Bresnahan, J., Czajkowski, K., Gawor, J., Kesselman, C., Meder, S., Pearlman, L., and Tuecke, S., "Security for Grid Services", *In Proceedings of the 12th International Symposium on High Performance Distributed Computing (HPDC-12)*, 2003.

[25] Gil, Y. and Artz, D., "Towards content trust of web resources", *Journal of Web Semantics*, Vol. 5, No. 4, 2007, pp. 227-239.

[26] Simmhan, Y. L., Plale, B., and Gannon, D., "A survey of data provenance techniques", *Technical Report TR618*, Computer Science Department, Indiana Univ., 2005.

[27] Crypto++ library 5.5.2. http://www.cryptopp.com/.

[28] Open provenance model. http://twiki.ipaw.info/bin/view/OPM.

[29] Quickhash library. http://www.slavasoft.com/quickhash/index.htm.

Author Profile

Im Y. Jung

Im Y. Jung received her MS and Ph.D. degrees in Computer Engineering from Seoul National University in 2001 and in 2010 each. She worked with Electronics and Telecommunications Research Institute as a researcher from 2001 to 2004. She is a Post-Doc Researcher in Seoul National University. She currently teaches and manages projects for storage system and distributed implementation of electronic commerce system. Her research interests are distributed computing systems, cloud computing, data and system security, energy efficient system and storage system.

Hyeonsang Eom

Hyeonsang Eom is an assistant professor at the School of Computer Science and Engineering, Seoul National University. He received his B.S. degree in Computer Science and Engineering from Seoul National University (SNU), Seoul, Korea, in 1992. He received his M.S. and Ph.D. in Computer Science from the University of Maryland at College Park, Maryland, USA, in 1996 and 2003, respectively. He worked as an intern in the Data Engineering Group at Sun Microsystems, USA, in 1997. Before becoming a professor at SNU in 2005, he worked as a senior engineer in the Telecommunication R&D Center at Samsung Electronics. His research interests are in the areas of Cloud Computing, High Performance Storage Systems, Energy Efficient Systems, Fault Tolerant Systems, Digital Rights Management and Information Dynamics.

Heon Y. Yeom

Heon Y. Yeom is a Professor at the School of Computer Science and Engineering, Seoul National University. He received his BS degree in Computer Science from Seoul National University in 1984 and his MS and PhD degrees in Computer Science from Texas A&M University in 1986 and 1992 respectively. From 1986 to 1990, he worked with Texas Transportation Institute as a Systems Analyst, and from 1992 to 1993, he was a Research Scientist at Samsung Data Systems. He joined the Department of Computer Science, Seoul National University in 1993, where he currently teaches and researches on distributed computing systems, energy efficient system, multimedia systems and transaction processing.

◈ 이 논문은 2011년 07월 28일 접수하여 2011년 09월 05일 게재확정되었습니다.