Seguridad en Sistemas Distribuidos Federados

Javier Echaiz Pablo M. Davicino* Jorge R. Ardenghi

Laboratorio de Investigación de Sistemas Distribuidos (LISiDi)
LISIDI es un miembro del IICyTI (Instituto de Investigación en Ciencia y Tecnología Informática).
Departamento de Ciencias e Ingeniería de la Computación
Universidad Nacional del Sur
Av. Alem 1253 - (8000) Bahía Blanca - Argentina
Tel/Fax: (+54) 291-4595135/6
{je,pmd,jra}@cs.uns.edu.ar

Resumen

Los sistemas distribuidos federados constituyen actualmente uno de los escenarios de cómputo y/o de almacenamiento de información más usuales, especialmente los sistemas basados en computación grid, cloud y peer-to-peer.

En esta línea de investigación se explorará la problemática de seguridad (todavía abierta) en este tipo de sistemas distribuidos de gran escala.

Palabras clave: seguridad, sistemas distribuidos, sistemas distribuidos federados, grid, cloud, peer-to-peer.

Contexto

El trabajo objeto del presente artículo se desarrolla en el Laboratorio de Investigación en Sistemas Distribuidos (LISiDi) del Departamento de Ciencias e Ingeniería de la Computación de la Universidad Nacional del Sur.

La línea de investigación presentada es parte de los Proyectos "Automatización de la Detección de Intrusos a partir de Políticas de Seguridad" (24/ZN14) dirigido por el Mag. Lic. Javier Echaiz, y "Computación Distribuida de Alto Rendimiento y Disponibilidad" (24/N024) dirigido por el Mag. Ing. Jorge Ardenghi. Ambos proyectos son financiados por la Secretaria General de Ciencia y Tecnología de la Universidad Nacional del Sur, y se encuentran acreditados por la Universidad Nacional del Sur, Bahía Blanca.

*Becario de la Comisión de Investigaciones Científicas (CIC) de la Provincia de Buenos Aires.

Introducción

Un *Sistema Distribuido Federado* (SDF) constituye una gran estructura de nodos, donde los administratores de sistemas (distribuidos geográficamente) están motivados por distintos intereses. Los modelos actuales de arquitecturas y de seguridad para este tipo de sistemas no explicitan quién es confiable para hacer qué y por ello estas relaciones de confianza implícitas son generalmente insuficientes.

Los *resource brokers* son entidades que actúan de parte del usuario, empleando para ello credenciales delegadas por éstos para descubrir y utilizar servicios afines a los requerimientos del usuario. Si bien dichos brokers necesariamente avanzan el estado del arte en cuanto a usabilidad y potencial en los SDF, también ocupan una posición estratégica importante para comprometer la seguridad de los usuarios que intentan proteger.

Los *resource brokers* utilizan servicios de registro/directorios, donde se suscriben tanto brokers como proveedores de servicios. Los brokers confían en que los servicios de registro se mantienen actualizados y que contienen información precisa, y a su vez los servicios de registro confían en que los proveedores de servicio describen en forma precisa los detalles de los servicios, como por ejemplo su política de seguridad. Es por ello que creemos que los SDF deben evolucionar al igual que lo hizo la web, donde en sus orígenes se trataba simplemente de un medio de intercambio de información para volverse un lugar donde los fraudes son moneda corriente.

En estos últimos años el foco de atención de la seguridad en este sentido fue para la autenticación de usuarios, especialmente en los últimos

tiempos, para los equemas de autorización distribuida y combinación de políticas para soportar los diferentes tipos de credenciales [25, 30]. Típicamente las credenciales están basadas en certificados X.509 para autenticarse principalmente frente a usuarios. Para mejorar esta arquitectura, en [4] se propone una etapa de negociación de la confianza como parte del SLA, donde ambas partes presentan un conjunto de CAs en las que confían y la autenticación basada en X.509 es exitosa siempre y cuando exista al menos una CA en la intersección de estos dos conjuntos.

Los protocolos involucrados en la autenticación basada en X.509 implican una relación de confianza simétrica, esto es, si el usuario o el sistema de SDF no está satisfecho, entonces la autenticación no prospera. Sin embargo, existe una relación de confianza asimétrica detrás de los sistemas de autenticación y del software que corre sobre ellos y su estudio formará parte de los objetivos de esta línea de investigación.

La infraestructura que soporta los SDF (usualmente un *middleware*) es necesariamente "pesada' ' y por ello inevitablemente contendrá múltiples vulnerabilidades. Dada su naturaleza distribuida propone un escenario atractivo para un atacante, independientemente de su motivación. Por ejemplo, un atacante podría explotar una vulnerabilidad en la plataforma BOINC [2] y ejecutar código arbitrario en la computadora de un participante.

El problema del middleware complejo se agrava aún más si consideramos la integración de tecnologías. Por ejemplo, actualmente MPI [20, 32, 10] se encuentra integrado con globus [11], permitiendo que jobs fuertemente acoplados se puedan comunicar entre si [11, 10] pero incrementando también el riesgo de ejecución de jobs maliciosos que envíen datos confidenciales a terceras partes. Más aún, la tendencia de los SDF actuales es el empleo de Web Services, tecnología que suele promocionarse como "amigable' ' a los firewalls, pero que al mismo tiempo claramente contribuye con el problema.

Más aún, las arquitecturas de los SDF hacen transparente la elección de los sistemas subyacentes: los recursos son *commodities*; y si por ejemplo pudieran diferenciarse los sistemas componentes con los diversos grados de confianza de sus administradores entonces esta abstracción falllaría.

Los nodos de cómputo de un sistema con estas características suelen obedecer un contrato mutuo: existe la expectativa de que los nodos lleven a cabo correctamente las tareas solicitadas (o devuelvan el correspondiente error), y a cambio los nodos esperan que las tareas solicitadas no les ocasionen daños o comprometan la seguridad de su información. Esta es una dualidad clásica en seguridad distribuida: código confiable en un nodo no confiable (en un sentido) y código no confiable en un nodo confiable (en el otro sentido).

Creemos entonces que para muchas aplicaciones existen problemas de seguridad que hacen que el *deployment* no sea actualmente posible en ambients distribuidos de gran escala y por ello surge la pregunta natural de dónde deberíamos invertir esfuerzos de investigación y desarrollo para obtener los mayores beneficios.

Línea de Investigación y Desarrollo

Esta línea de investigación y desarrollo explora uno de los requerimientos fundamentales relacionadas con la seguridad en sistemas distribuidos federados, la necesidad de establecer las bases de confianza necesaria para los usuarios del sistema.

Es necesario entonces que el subsistema de seguridad incluya las siguientes características:

- Durante su inicialización poder verificar su propia integridad (mediante algoritmos de hash) y en consecuencia poder garantizar la integridad de cualquier (combinación) componente de software.
- Poseer una o más identidades criptográficamente fuerte, capaz de probar dicha identidad de software a un tercero.
- Soportar el almacenamiento seguro de la jerarquía completa de claves.

La *confianza* en el sentido planteado por esta línea significa simplemente la aserción de que una plataforma se encuentra en una dada configuración específica. Si esta configuración constituye o no una configuración *confiable* para el sistema también conforma un tópico de interés para nuestra investigación. Para lograr esta aseverción es necesario *sellar* (mediante criptografía asimétrica) los datos, esto significa encriptarlos de forma tal de que sólo puedan ser accedidos por una entidad en un estado particular (bajo el software apropiado).

A su vez, esta línea de investigación contempla no como alternativa sino como complemento, el empleo de plataformas confiables implementadas mediante virtualización. La posibilidad de correr múltiples máquinas virtuales sobre una plataforma física manejada por un *hypervisor* o un VMM (*Virtual Machine Monitor*), indudablemente suma beneficios de seguridad en sistemas federados al incluir aislamiento de procesos (e.g. [12, 31]), pero por otro lado también plantea nuevos desafíos interesantes.

Figura 1: Virtualización.

Resultados y Objetivos

Esta línea de I+D espera conseguir resultados que tiendan a solucionar los problemas de seguridad actuales en los sistemas distribuidos federados. Muchos de estos sistemas, por ejemplo los sistemas grid, proponen la abstracción de transparencia (los usuarios no saben dónde se almacenan sus datos ni dónde se efectuaron sus cómputos) pero plantean nuevos desafíos de seguridad.

Es claro que adoptar nuevas tenologías en esta materia presenta el potencial de mover el problema, pues por ejemplo sigue siendo necesario el uso de claves para firmar digitalmente: cambia el foco del atacante, pues pasa de tratar de extraer claves a tratar de lograr que la entidad firmante convalide algo cuando debería no hacerlo.

De todas formas parece prometedor el uso de nuevas tecnologías que extiendan y expandan la seguridad bajo diferentes escenarios, aportando mejoras no sólo en lo relacionado con seguridad sino como un resultado, facilitando arquitecturas y modos de trabajo que previamente no fueron posibles.

Los objetivos de esta línea incluyen entonces el diseño de susbsistemas de seguridad que cumplan con las necesidades planteadas en la sección anterior y sean potencialmente adaptables a distintas necesidades SDF. Inicialmente se espera una implementación exitosa en globus.

Formación de Recursos Humanos

Este trabajo corresponde a una de las principales líneas de investigación del Laboratorio de Investigación en Sistemas Distribuidos (LISiDi) de la Universidad Nacional del Sur. La misma será objeto de estudio de diversas tesinas de grado de las carreras de Ingeniería en Sistemas de Computación y Licenciatura en Ciencias de la Computación, ambas dictadas por el Departamento de Ciencias e Ingeniería de la Computación de la Universidad Nacional del Sur y posiblemente se extienda a varios trabajos de postgrado (actualmente forma parte del trabajo de una tesis de Magister).

A su vez, el LISiDi cuenta con recursos propios, sobre los cuales se despliegan los distintos componentes que permiten trabajar con los sistemas objetos de este estudio. Si bien es posible trabajar en cuestiones de seguridad distribuida sin la necesidad de contar con sistemas distribuidos de gran escala (que se encuentren dispersos geográficamente) es cierto que para validar los modelos propuestos es interesante que los entornos sean efectivamente distribuidos, permitiendo probar los sistemas bajo diferentes dominios organizativos, diferentes CAs (con cadenas de certificación no triviales), atravesando firewalls, etc. Para ello se trabajará en forma colaborativa con otras Universidades, especialmente con el LIDI de la Facultad de Informática de la Universidad Nacional de La Plata, junto con quienes se ha trabajado exitosamente en cuestiones de computación grid.

A continuación se mencionan los cursos de pregrado relacionados con la línea de investigación presentada, dictados por los integrantes del grupo de investigación:

- Sistemas Operativos. Materia obligatoria para los estudiantes de la carrera Ingeniería en Sistemas de Computación, Universidad Nacional del Sur.
- Sistemas Distribuidos. Materia obligatoria para los estudiantes de la carrera Ingeniería en Sistemas de Computación, Universidad Nacional del Sur.
- Sistemas Operativos y Distribuidos. Materia obligatoria para los estudiantes de la carrera

Licenciatura en Ciencias de la Computación, Universidad Nacional del Sur.

- Sistemas Distribuidos I. Materia obligatoria para los estudiantes de la carrera Ingeniería en Sistemas, Facultad de Ingeniería, Universidad Nacional de la Pampa. Esta materia actualmente se dicta por docentes de la propia Facultad de la UNLPam.

- Sistemas Distribuidos II. Materia obligatoria para los estudiantes de la carrera Ingeniería en Sistemas, Facultad de Ingeniería, Universidad Nacional de la Pampa. Esta materia actualmente se dicta por docentes de la propia Facultad de la UNLPam.

- Seguridad en Sistemas. Materia obligatoria para los estudiantes de la carrera de Ingeniería en Sistemas de Computación, y optativa para los de Licenciatura en Ciencias de la Computación, Universidad Nacional del Sur.

Asimismo se han dictado cursos de posgrado relacionados con Seguridad de la Información y Sistemas Distribuidos (especialmente referidos al paradigma grid y peer-to-peer) en varias Universidades Nacionales del país.

Referencias

[1] Mehran Ahsant, Jim Basney, Olle Mulmo, Adam Lee, and Lennart Johnsson. Toward an on-demand restricted delegation mechanism for grids. In *Proceedings of the 7th IEEE/ACM International Conference on Grid Computing*, GRID '06, pages 152–159, Washington, DC, USA, 2006. IEEE Computer Society.

[2] David P. Anderson. Boinc: A system for public-resource computing and storage. In *Proceedings of the 5th IEEE/ACM International Workshop on Grid Computing*, GRID '04, pages 4–10, Washington, DC, USA, 2004. IEEE Computer Society.

[3] Michael Bailey, Evan Cooke, Farnam Jahanian, and Jose Nazario. The Internet Motion Sensor - A Distributed Blackhole Monitoring System. In *Proceedings of the 12th Annual Network & Distributed System Security Symposium (NDSS '05)*, pages 167–179, San Diego, California, USA, February 2005.

[4] J. Basney, W.Ñejdl, D. Olmedilla, V. Welch, and M. Winslett. Negotiating trust on the grid. In *In 2nd WWW Workshop on Semantics in P2P and Grid Computing*, 2004.

[5] E. Bertino and J. Crampton. Security for distributed systems: Foundations of access control. In Y. Qian, D. Tipper, P. Krishnamurthy, and J. Joshi, editors, *Information Assurance: Survivability and Security in Networked Systems*, pages 39–80. Morgan Kaufman, 2007.

[6] Andrew Cooper and Andrew Martin. Towards a secure, tamper-proof grid platform. In *Sixth IEEE International Symposium on Cluster Computing and the Grid (CCGrid 2006), 16-19 May 2006, Singapore*, pages 373–380. IEEE Computer Society, 2006.

[7] J. Crampton, H.W. Lim, and K.G. Paterson. What can identity-based cryptography offer to web services? In *Proceedings of 4th ACM Workshop on Secure Web Services*, pages 26–36, 2007.

[8] J. Crampton and G. Loizou. Administrative scope: A foundation for role-based administrative models. *ACM Transactions on Information and System Security*, 6(2):201–231, 2003.

[9] M. Dekker, J. Crampton, and S. Etalle. RBAC administration in distributed systems. In I. Ray and N. Li, editors, *Proceedings of 13th ACM Symposium on Access Control Models and Technologies*, pages 93–102, 2008.

[10] Ian Foster and Nicholas T. Karonis. A grid-enabled mpi: message passing in heterogeneous distributed computing systems. In *Proceedings of the 1998 ACM/IEEE conference on Supercomputing (CDROM)*, Supercomputing '98, pages 1–11, Washington, DC, USA, 1998. IEEE Computer Society.

[11] Ian T. Foster. Globus toolkit version 4: Software for service-oriented systems. *J. Comput. Sci. Technol.*, 21(4):513–520, 2006.

[12] Jason Franklin, Mark Luk, Jonathan M. McCune, Arvind Seshadri, Adrian Perrig, and Leendert Van Doorn. Remote detection of virtual machine monitors with fuzzy benchmarking. *ACM SIGOPS Operating System Review Special Edition on Computer Forensics*, 42(3):83–92, April 2008.

[13] Kevin Fu, M. Frans Kaashoek, and David Mazieres. Fast and secure distributed read-only file system. 20(1):1–24, February 2002.

[14] Kevin Fu, Seny Kamara, and Tadayoshi Kohno. Key regression: Enabling efficient key distribution for secure distributed storage. In *Proceedings of the Symposium on Network and Distributed Systems Security*, February 2006.

[15] Guofei Gu, Bin Zhu, Shipeng Li, and Shiyong Zhang. PLI: A New Framework to Protect Digital Content for P2P Networks. In *Proceedings of the 2003 International Conference on Applied Cryptography and Network Security (ACNS'03)*, October 2003.

[16] Jun Ho Huh and John Lyle. Trustworthy log reconciliation for distributed virtual organisations. In Liqun Chen, Chris J. Mitchell, and Andrew Martin, editors, *Trust '09: Proceedings of the 2nd International Conference on Trusted Computing*, Lecture Notes in Computer Science, pages 169–182, Berlin, Heidelberg, April 2009. Springer-Verlag.

[17] Jun Ho Huh, John Lyle, Cornelius Namiluko, and Andrew Martin. Managing application whitelists in trusted distributed systems. *Future Generation Computer Systems*, In Press, Accepted Manuscript, 2010.

[18] Yanlin Li, Jonathan M. McCune, and Adrian Perrig. SBAP: Software-Based Attestation for Peripherals. In *Proceedings of the 3rd International Conference on Trust and Trustworthy Computing (Trust 2010)*, June 2010.

[19] Giorgia Lodi, Leonardo Querzoni, Roberto Baldoni, Mirco Marchetti, Michele Colajanni, Vita Bortnikov, Gregory Chockler, Eliezer Dekel, Gennady Laventman, and Alexey Roytman. Defending financial infrastructures through early warning systems: the intelligence cloud approach. In *CSIIRW '09: Proceedings of the 5th Annual Workshop on Cyber Security and Information Intelligence Research*, pages 1–4, New York, NY, USA, 2009. ACM.

[20] Steven Manos, Marco Mazzeo, Owain Kenway, Peter V. Coveney, Nicholas T. Karonis, and Brian Toonen. Distributed mpi cross-site run performance using mpig. In *Proceedings of the 17th international symposium on High performance distributed computing*, HPDC '08, pages 229–230, New York, NY, USA, 2008. ACM.

[21] Mirco Marchetti. *Distributed architectures and algorithms for network security.* PhD thesis, Doctorate School in Information and Communication Technologies, University of Modena and Reggio Emilia, February 2009.

[22] Mirco Marchetti, Michele Messori, and Michele Colajanni. Peer-to-peer architecture for collaborative intrusion and malware detection at a large scale. In *Proc. of the 12th Information Security Conference (ISC 2009)*, Pisa, Italy, September 2009.

[23] Jonathan M. McCune, Yanlin Li, Ning Qu, Zongwei Zhou, Anupam Datta, Virgil Gligor, and Adrian Perrig. TrustVisor: Efficient TCB reduction and attestation. In *Proceedings of the IEEE Symposium on Security and Privacy*, May 2010.

[24] Bryan Parno, Jonathan M. McCune, and Adrian Perrig. Bootstrapping trust in commodity computers. In *Proceedings of the IEEE Symposium on Security and Privacy*, May 2010.

[25] L. Pearlman, V. Welch, I. Foster, C. Kesselman, and S. Tuecke. A community authorization service for group collaboration. In *Proceedings of the 3rd International Workshop on Policies for Distributed Systems and Networks (POLICY '02)*, POLICY '02, pages 50–60, Washington, DC, USA, 2002. IEEE Computer Society.

[26] D. J. Power, E. A. Politou, M. A. Slaymaker, and A. C. Simpson. Securing web services for deployment in health grids. *Future Generation Computer Systems*, 22(5):547–570, 2006.

[27] H. Rowe and J. Crampton. Avoiding key redistribution in key assignment schemes. In *Proceedings of the Fourth International Workshop on Mathematical Methods, Models, and Architectures for Computer Network Security*, pages 127–140, 2007.

[28] Weidong Shi, Hsien-Hsin Lee, Guofei Gu, Laura Falk, Trevor Mudge, and Mrinmoy Ghosh. InfoShield: A security architecture for protecting information usage in memory. In *Proceedings of the 12th International Symposium on High-Performance Computer Architecture (HPCA'06)*, March 2006.

[29] Guoqiang Shu, Dongluo Chen, Zhijun Liu, Na Li, Lifeng Sang, and David Lee. Vcstc: Virtual cyber security testing capability - an application oriented paradigm for network infrastructure protection. In *20th IFIP Int. Conference on Testing of Communicating Systems, LNCS volume 5047*, Tokyo, Japan, June 2008.

[30] Mary R. Thompson, Abdelilah Essiari, and Srilekha Mudumbai. Certificate-based authorization policy in a pki environment. *ACM Trans. Inf. Syst. Secur.*, 6:566–588, November 2003.

[31] Amit Vasudevan, Jonathan M. McCune, Ning Qu, Leendert van Doorn, and Adrian Perrig. Requirements for an Integrity-Protected Hypervisor on the x86 Hardware Virtualized Architecture. In *Proceedings of the 3rd International Conference on Trust and Trustworthy Computing (Trust 2010)*, June 2010.

[32] Xiaohui Wei, Hongliang Li, and Dexiong Li. Mpich-g-dm: An enhanced mpich-g with supporting dynamic job migration. In *Proceedings of the 2009 Fourth ChinaGrid Annual Conference*, CHINAGRID '09, pages 67–76, Washington, DC, USA, 2009. IEEE Computer Society.