

Laboratorios Remotos sobre Espacios Virtuales

Grosclaude E., Sznek J., Bertogna L., Lopez Luro F., Zanellato C., Sanchez L.
Departamento de Informática y Estadística, Universidad Nacional del Comahue,
Buenos Aires 1400, Neuquén, Argentina
{oso, jsznek, mlbertog, flopez, czanella, lsanchez}@uncoma.edu.ar

Resumen: *Las nuevas tecnologías de Internet permiten el uso de sistemas de software distribuido que proporcionan a los usuarios el acceso en forma remota a laboratorios físicos y virtuales, para llevar a cabo actividades que normalmente son realizadas localmente. Para que el acceso remoto a estos laboratorios sea posible, se debe diseñar y construir una arquitectura de software y hardware que provea las interfaces adecuadas a los usuarios que quieran acceder. En este trabajo se describen las actividades de investigación llevadas adelante en el departamento de Cs. de la Computación dentro del marco de laboratorios remotos, y puntualmente en lo que se refiere a los espacios virtuales generados para la implementación y el uso de los mismos.*

Palabras Claves: Laboratorio Remoto, Virtualización, Alta Disponibilidad, Grid, Seguridad

1. Introducción

Las nuevas tecnologías de Internet permiten el uso de sistemas de software distribuido que proporcionan a los usuarios el acceso en forma remota a laboratorios físicos o virtuales, para llevar a cabo actividades que normalmente son realizadas localmente. Para que el acceso remoto a estos laboratorios sea posible, se debe diseñar y construir una arquitectura de software y hardware en éstos, que provea las interfaces adecuadas a los usuarios que quieran acceder.

En [5] se define un Laboratorio Virtual (LV) como "un espacio de trabajo electrónico para la colaboración y experimentación en investigación u otras actividades creativas, para la generación y distribución de los resultados de investigación utilizando tecnologías de información distribuidas". Un LV no debe ser visto como un reemplazo de un laboratorio real, sino como una posible extensión o complemento a estos últimos, abriendo nuevas oportunidades que probablemente no sean viables de implementar económica y socialmente a través de un laboratorio real.

Se pueden encontrar términos alternativos a Laboratorio Virtual, como Colaboratorio, Laboratorio Remoto y Grupo de Trabajo Virtual entre otros. En [6], se combinan los términos colaborativo y laboratorio para definir lo que se llama un "colaboratorio", como "un centro sin barreras, donde los usuarios pueden realizar sus investigaciones sin importar su ubicación geográfica, interactuando con

colegas, accediendo a instrumental costoso y compartiendo recursos computacionales".

Motivaciones y características generales de LV: En términos generales, se destacan tres fuerzas que motivan la generación y uso de los Laboratorios Remotos Físicos y Virtuales. En primer lugar, un proyecto en particular puede requerir de la estructura de los LV para su implementación, puesto que éstos permitirán un uso eficiente de los recursos, ambientes controlados en cuanto a la seguridad de acceso a los mismos, y mayor flexibilidad a la hora de configurar un laboratorio para realizar las actividades del proyecto. Por otro lado, surge la necesidad del acceso a grandes instalaciones y laboratorios por parte de especialistas, con acceso restringido, desde lugares distantes y fuera de las planificaciones normales. Microscopios electrónicos, Telescopios y Clusters de PCs son ejemplos de recursos que pueden ser accedidos para tareas específicas orientadas a objetivos particulares del especialista que realiza dicho acceso. Por último, la necesidad de interacción y colaboración entre los integrantes de una comunidad de investigación, distantes geográficamente, motiva el desarrollo y uso de los LV, que enriquezcan las formas de comunicación, faciliten los mecanismos para compartir los resultados individuales y permitan construir en forma colectiva y concurrente los resultados del conjunto. Más específicamente, en [5] se mencionan algunos de los factores que motivan la construcción y uso de los LV:

- Ciertos desafíos tecnológicos en las comunidades científicas requieren un esfuerzo que supera la capacidad de un laboratorio tradicional e incluso de una sola nación.
- Los recursos humanos y la expertitud requerida para alcanzar las metas de un proyecto, a menudo se encuentran distribuidas geográficamente en más de una institución.
- El tema en estudio, puede requerir la participación de especialistas de distintas regiones distantes debido a la necesidad de diversidad de los datos o muestras recolectadas, o debido a la existencia específica de dichos datos en determinadas ubicaciones del planeta.
- Para poder llevar a cabo ciertas investigaciones, puede ser necesario o efectivo en cuanto a costos, compartir el acceso a determinados recursos, como instrumentos científicos que son únicos, caros y de difícil acceso, como por ejemplo microscopios electrónicos, aceleradores de partículas, etc.

2. Líneas de Investigación

- Seguridad

Desde el punto de vista de la seguridad se deben considerar los modelos de identificación, autorización, certificación y control de acceso que soportará la arquitectura. El avance masivo de los sistemas de computación distribuidos y la creciente demanda en compartir información online a través de varias aplicaciones por Internet, poniendo énfasis en la interoperabilidad, ocasionaron que se ponga gran atención en los aspectos de seguridad de los servicios web, tecnología empleada para implementar esa interoperabilidad.

Con esta línea de investigación se plantea el desarrollo de un modelo de seguridad para ambientes colaborativos sobre Internet, concentrándonos en particular sobre una parte del proyecto de investigación de Software para procesos colaborativos, que es el Laboratorio Remoto y el acceso a

recursos por Internet mediante servicios web. Ese modelo de Seguridad se basará en el modelo estándar RBAC (Role Based Access Control) [1] usando para su especificación XACML (eXtensible Access Control Markup Language) [2]. La especificación XACML describe cómo se pueden construir bloques para la implementación de los diferentes elementos del modelo RBAC.

Otro de los aspectos de seguridad está relacionado con el análisis de los eventos que suceden en una red, poniendo énfasis en aquellos eventos que significan ataques a la misma. Un método bastante extendido consiste en el análisis de los logs del sistema, que se realiza en forma diferida y posterior a la ocurrencia de un eventual ataque; este método permite reconocer ataques ya efectivizados y sirve para prevenir que no sucedan nuevamente.

Otro método es mediante la implementación de un Sistema de Detección de Intrusos (IDS), que analiza en tiempo real el progreso ciertos eventos de seguridad caracterizados como potenciales intrusiones. Este tipo de herramientas suponen un conocimiento más complejo de las actividades que se realizan sobre el sistema para detectar fallas o eventos anormales realizados sobre el mismo, induciendo de ello la existencia de algún ataque.

Lo que se plantea como trabajo de investigación es la generación de un espacio de prueba y experimentación mediante la implementación de una HoneyNet[3], y el uso de un IDS dentro de ella, de esta manera se pueden estudiar aspectos de seguridad en una red organizacional o corporativa como: análisis de las secuencias que conforman tráfico sospechoso de ataque a los diferentes servicios provistos por la red virtual (ssh, smtp, etc.) y correlación de los paquetes con destino a los distintos nodos de la red con el objeto de poder estudiar técnicas de ataque.

- Virtualización de Recursos

Una de las posibilidades en la implementación de laboratorios virtuales se encuentra en el uso de multiples recursos heterogéneos sobre una infraestructura Grid[9]. Un problema de las plataformas Grid es que no hay una limitación clara de uso de recursos dentro de las organizaciones virtuales, y la ejecución de algunas tareas puede afectar la performance del sistema. Debido a que no hay una manera efectiva de limitar el uso de los recursos no es posible garantizar un entorno adecuado para ejecuciones de altas prestaciones o aplicaciones críticas.

Otro serio problema es que mientras los entornos Grid proveen acceso a distintos recursos con distintos tipos de configuración de software, una aplicación sólo se podrá ejecutar en condiciones específicas y en un entorno adaptado a sus necesidades. La variación de sistema operativo, versiones de middleware, librerías o disposición de sistemas de archivo, ponen barreras a la portabilidad de las aplicaciones.

Como línea de investigación se plantea el desarrollo e implementación de un entorno virtual como recurso de la organización virtual. Este tipo de solución se podría resolver limitando de manera clara el uso de memoria, disco y procesamiento dentro de máquinas virtuales con imágenes de sistemas operativos configurados (kernel, bibliotecas, etc) para el correcto funcionamiento de las aplicaciones

desarrolladas.

- Clustering

Entre otras metas, el proyecto de Laboratorios Virtuales debe proporcionar una infraestructura usable y robusta para apoyo a la enseñanza, meta que no se logra sin un elevado costo en complejidad. Esta complejidad está en oposición con la autonomía que debe tener el sistema para poder actuar en diferentes franjas horarias, con diferentes configuraciones, etc. Vistos los riesgos de disponibilidad que esta complejidad acarrea, dicha infraestructura debe ser protegida mediante un diseño específico de Alta Disponibilidad (AD). Para desarrollar éste, y en consonancia con los objetivos planteados en [8], se profundiza la investigación en principios y técnicas de AD, en particular en lo relacionado con clustering. Ciertos desarrollos implementados en el marco del Software Libre han dado, recientemente, pasos interesantes hacia la solución de determinados problemas. Algunos de ellos (GFS, heartbeat, DRBD) apuntan a necesidades propias de la provisión de infraestructura de AD, mientras que otros (como los proyectos de virtualización o paravirtualización, que han alcanzado un nivel de usabilidad muy alto), indirectamente pueden ofrecer ventajas para el diseño de arquitecturas de AD.

Como tema de investigación principal en esta línea se encuentra la utilización de máquinas virtuales como recursos redundantes de clusters de AD, con especial atención a sus consecuencias en cuanto a separación de dominios administrativos entre sistema host y sistema de servicio.

3. Resultados Esperados

- Seguridad

Se espera realizar avances en el desarrollo de un Servidor de Seguridad que permita atender los requerimientos de los clientes mediante servicios web utilizando el modelo estándar RBAC basado en esquemas XML (X-RBAC). Se presentará una arquitectura para representar el modelo RBAC extendido para su posterior implementación en el ambiente del Laboratorio Remoto.

Se ha iniciado la implementación de una honeynet que permita estudiar y analizar las secuencias que conforman tráfico sospechoso de ataque a los diferentes servicios provistos por la red virtual (ssh, smtp, etc.). Todo el tráfico será registrado en logs para una posterior implementación de herramientas de análisis de los mismos.

- Procesamiento en entornos grid sobremáquinas virtuales

Se encuentra en desarrollo la creación de redes de recursos virtuales dentro de una organización virtual en entornos Grid[4]. Los usuarios podrán acceder a los recursos en forma interactiva a través de interfaces web. En este desarrollo se busca realizar una configuración en forma segura y con mínima intervención de los administradores locales en cada organización física. Se está aplicando la solución en varios casos de uso como por ejemplo distintos entornos de altas prestaciones y laboratorios virtuales de redes. Actualmente se están desarrollando módulos de administración estándar de dicho entorno virtual e investigando en el desarrollo de planificadores inteligentes basados en políticas.

- Laboratorio de programación con robots y administración de Redes

En [7] se realizó una demostración construida en base a la arquitectura en la que se está trabajando, para la realización de una actividad de teleprogramación de un Robot Lego en lenguaje C, de manera colaborativa y utilizando conexiones de internet de baja velocidad. Se encuentra en adaptación para ser generalizado a escenarios donde se plantea el acceso interactivo a una variedad de recursos remotos de manera colaborativa, usando tecnologías accesibles por nuestros alumnos.

También se llevaron adelante experiencias con generación de entornos para la simulación de distintas alternativas en la conexión de dispositivos de red, con la mezcla de dispositivos físicos y virtuales. El costo y disponibilidad de los componentes utilizados para la construcción y para la interconexión de las redes de computadoras justifican la implementación de estos laboratorios virtuales.

- Alta Disponibilidad

Para maximizar las prestaciones y la utilización de recursos, se busca configurar una metodología de uso general en el ámbito de LVs para dotar de una infraestructura de Alta Disponibilidad a los laboratorios. Se espera obtener mayor velocidad en la instalación, y simplicidad en la administración de esta infraestructura, por medio de la separación de dominios administrativos al utilizar máquinas virtuales como ambiente de servicio.

4. Referencias

- [1] D.F. Ferraiolo and D.R. Kuhn "Role Based Access Control" 15th National Computer Security Conference. 1992
- [2] OASIS XACML Technical Committee ."eXtensible Access Control Markup Language (XACML)". On-Line <http://docs.oasis-open.org/xacml/cd-xacml-rbac-profile-01.pdf>
- [3] The Honymet Project, "Know your Enemy: learning about security threats". Second Edition. Addison-Wesley. Pub. 2004.
- [4] Keahey, K., I. Foster, T. Freeman, and X. Zhang."Virtual Workspaces: Achieving Quality of Service and Quality of Life in the Grid". Scientific Programming Journal, vol 13, No. 4, 2005.
- [5] Report of the expert meeting on Virtual Laboratories. Unesco, 2000, Paris.
- [6] William A. Wulf, James D. Myers. "Collaboratories: Doing science on the Internet". IEEE Computer, 1996.
- [7] Del Castillo Rodolfo, Grosclaude Eduardo, A. López Luro, Francisco Rodríguez, Jorge Sanchez Laura, Zanellato Claudio, Bertogna Leandro "Experiencia con Laboratorio Remoto Colaborativo", TE&ET 2006.
- [8] Eduardo Grosclaude, Soporte de Sistemas para Procesos Colaborativos, WICC2005.
- [9] I. Foster, C. Kesselman, S. Tuecke."The Anatomy of the Grid: Enabling Scalable Virtual Organizations". International J. Supercomputer Applications, 15(3), 2001.