

Simulación de una red Grid con máquinas virtuales para crear un entorno de aprendizaje de la computación de alto desempeño

Simulation of a Grid network with virtual machines to create a learning environment of high performance computing

Simulação de uma rede Grid com máquinas virtuais para criar um ambiente de aprendizagem da computação de alto desempenho

Fecha de recepción: 10 de septiembre de 2015
Fecha de aprobación: 18 de diciembre de 2015

Edwin Eduardo Millán-Rojas*
Adriana Patricia Gallego-Torres**
Diana Carolina Chico-Vargas***

Resumen

Presenta una metodología de construcción de una plataforma para simular una red Grid con máquinas virtuales, con el objetivo de crear entornos virtuales que contribuyan a los procesos de aprendizaje. El diseño utilizado para el desarrollo está basado en un prototipo que contribuye a generar un entorno de simulación de bajo costo a través de la computación de alto desempeño. Como resultado de la investigación se obtuvo un conjunto de máquinas virtuales que pueden ser desplegadas en cualquier equipo de cómputo para simular un entorno Grid; además, se presenta la forma de la configuración de esta red para ser parte integral de la construcción y ejecución de la plataforma como instrumento válido para la enseñanza de la computación de alto desempeño a un bajo coste de hardware y la simulación de los servicios Grid sobre la plataforma.

Palabras clave: Simulación, Computación Grid, Entornos virtuales de aprendizaje.

* M.Sc. Universidad de la Amazonía (Florencia-Caquetá, Colombia). e.millan@udla.edu.co.

** Ph.D. Universidad Distrital Francisco José de Caldas (Bogotá-Distrito Capital, Colombia). adpgallegot@udistrital.edu.co.

*** Universidad de la Amazonía (Florencia-Caquetá, Colombia). d.chico@udla.edu.co.

Abstract

This research presents a methodology for building a Grid platform to simulate a network with virtual machines in order to create virtual environments that contribute to the learning process. The design used for the development is based on a prototype that helps to generate a low cost simulation environment through the high-performance computing. As a result of the investigation a set of virtual machines that can be deployed on any computer equipment to simulate a Grid environment, was obtained. Besides the form of the configuration of this network is to be an integral part of the construction and design of the platform which is presented as valid instrument for the teaching of high performance computing at a low cost hardware, and simulation of Grid services on the instrument platform.

Keywords: Simulation, Grid computing, Virtual learning environments.

Resumo

Apresenta-se uma metodologia de construção de uma plataforma para simular uma rede Grid com máquinas virtuais, com o objetivo de criar ambientes virtuais que contribuam aos processos de aprendizagem. O desenho utilizado para o desenvolvimento está baseado em um protótipo que contribui a gerar um ambiente de simulação de baixo custo através da computação de alto desempenho. Como resultado da pesquisa obteve-se um conjunto de máquinas virtuais que podem ser implementadas em qualquer equipamento de cômputo para simular um ambiente Grid; além disso, apresenta-se a forma da configuração desta rede para ser parte integral da construção e execução da plataforma como instrumento válido para o ensino da computação de alto desempenho a um baixo custo de hardware e a simulação dos serviços Grid sobre a plataforma.

Palavras chave: Simulação, Computação Grid, Ambientes virtuais de aprendizagem.

Cómo citar este artículo:

[1] E. E. Millán-Rojas, A. P. Gallego-Torres & D. C. Chico-Vargas, “Simulación de una red Grid con máquinas virtuales para crear un entorno de aprendizaje de la computación de alto desempeño”, Fac. Ing., vol. 25 (41), pp. 85-92, ene.-abr. 2016.

I. INTRODUCCIÓN

Este trabajo pretende construir una propuesta conceptual y metodológica que contribuya a la formación de ingenieros acorde con los desarrollos epistemológicos actuales y con la sociedad de la información y el conocimiento, de tal forma que nos permita elaborar una resignificación de los procesos de enseñanza-aprendizaje y de su relación con los espacios físicos y temporales, propios de la antigua Grecia, donde se dio origen a las aulas de clase tal y como las conocemos hoy en día; aulas que fueron inspiradas en los teatros de la Roma imperial, cuya visión del aprendizaje daba su máxima importancia al contexto físico y a la influencia mediadora de la retórica; en esta época es cuando se establece el papel de maestro-alumno, con Marco Fabio Quintiliano como el principal pedagogo romano de entonces [1]. La irrupción de las Tecnologías de la Información y la Comunicación (TIC) como herramienta educativa supone un punto de inflexión en la forma como hoy se conciben el aprendizaje, las instituciones educativas, los procesos educativos y la gestión del aprendizaje, que evolucionan de forma paralela a las TIC, rompiendo con los esquemas rígidos y la enseñanza tradicional, para adoptar otras formas de construir, gestionar y transmitir los conocimientos. Al respecto, podríamos señalar que el lugar tradicional de las escuelas para aprender y enseñar pasa “de ser un espacio-lugar a ser un espacio-nodo de la red de interacción global. La noción educativa de lugar de aprendizaje no desaparece, sino que evoluciona junto a la sociedad del conocimiento.

Como consecuencia de la evolución acelerada de las tecnologías y de la búsqueda constante por resolver los problemas de aprendizaje y ofrecer distintas alternativas, nuestro trabajo pretende contribuir a los procesos de enseñanza mediados por las TIC, por medio de simulación de una red Grid con máquinas virtuales para crear un entorno de aprendizaje de la computación de alto desempeño. De esta forma, nuestra intencionalidad se basa en el desarrollo de una plataforma para simular una red Grid con máquinas virtuales que contribuya a los procesos de formación [2].

II. ASPECTOS TEÓRICOS DE LA COMPUTACIÓN GRID Y LA SIMULACIÓN

El principio de la computación Grid es la agregación de recursos computacionales heterogéneos distribuidos entre distintas organizaciones para formar un meta ordenador [3-5]. De esta forma hace fácil la utilización de infraestructuras distribuidas. Las mallas computacionales son un modelo que plantea la unión de los recursos computacionales de diferentes organizaciones para resolver problemas fuera del alcance del más potente superordenador individual [3]. La principal diferencia de la Grid con un sistema distribuido está basada en ir más allá de compartir información entre máquinas; en la Grid se realiza una distribución de todos los recursos, todo establecido y controlado de forma estratégica para obtener el mejor provecho del recurso computacional [4].

A. Servicio Grid

La computación Grid se abre paso como una herramienta necesaria para el almacenamiento y análisis de volúmenes elevados de información que otras disciplinas generan [5]; de allí que el conocimiento de la Arquitectura Grid se hace necesario en los entornos académicos, con el propósito de dar a conocer, en un ambiente controlado, cuáles son las experiencias que los estudiantes pueden adquirir en el entorno Grid; pero los costos del Hardware necesario para implementar una arquitectura de este tipo desbordan los presupuestos de las entidades públicas, que, además, carecen de la prioridad que el docente requiere en el proceso de enseñanza.

Un servicio Grid es un *web service* que se ajusta a una serie de convenciones (interfaces y comportamientos) que definen cómo los clientes interactúan con el sistema distribuido [6]; se define además como “mecanismos para crear, gestionar e intercambiar información entre entidades a través de llamadas” [7].

B. Arquitectura Grid

Es una arquitectura abierta basada en un estándar que facilita la extensibilidad, la interoperabilidad, la portabilidad y la comparación de código. También se dice que es una arquitectura de protocolos que definen los mecanismos básicos, los cuales permiten a los usuarios y a los recursos negociar, establecer, gestionar y explotar la compartición de recursos [6, 8]. La arquitectura formada por una Grid permite utilizar los recursos de procesamiento propios, lo cual ayuda a evitar mayor inversión de recursos en infraestructura tecnológica; sin embargo, es costosa, en comparación con las soluciones tradicionales [5, 9].

C. Recursos para la plataforma

El software utilizado en el proyecto es el *Globus Toolkit*, que es un producto desarrollado para funcionar como administrador general de los recursos; se puede descargar de Internet [10]. *Globus* es el middleware de la Grid, es decir, una interface que se ubica entre los recursos finales y la aplicación de usuario; esta herramienta ha surgido como el estándar de facto para aplicaciones Grid.

Los diferentes recursos que administra el *Globus* se pueden definir de la siguiente forma: Gestión de recursos (protocolo de gestión de recursos en la grilla o *Grid resource management protocol*), Servicios de Información (servicio de descubrimiento y monitorización o *Monitoring and Discovery Service*), y Gestión y movimiento de datos (acceso global al almacenamiento secundario, *Global access to secondary storage* y FTP en grilla, *GridFTP*).

D. Simulación

La simulación se ha utilizado como herramienta en los procesos de la enseñanza y el aprendizaje, en los cuales juegan un papel muy importante las TIC para llevar al estudiante a un entorno capaz de cambiar sus conocimientos y fortalecer la práctica como elemento de aprendizaje [11]. Diferentes trabajos [12-15] hacen relación a la simulación como elemento fundamental en la consolidación de los procesos de enseñanza y aprendizaje en diferentes entornos.

La simulación consiste en situar a un educando en un contexto que imite algún aspecto de la realidad y en establecer en ese ambiente situaciones

probleáticas, similares a las que él deberá enfrentar en un ambiente real, y someterlo a los posibles casos que puede encontrar [2]. Según Martín y Barneto: “las simulaciones permiten que los alumnos hagan ‘descubrimientos’ de gran interés en ámbitos físicos complejos sin la necesidad de difíciles tratamientos matemáticos, apoyándose únicamente en la comprensión del fenómeno y en la aplicación de un método de trabajo eficaz” [16]. Para lograr estos propósitos se debe contar con herramientas que permitan virtualizar la plataforma de simulación.

E. Máquinas virtuales

Con el fin de implementar la plataforma de pruebas y realizar el presente proyecto se emplearon herramientas para construir una red de máquinas virtuales. La ejecución de la plataforma es soportada por el software VMWARE, que se puede usar a partir del licenciamiento gratuito para realizar pruebas al producto con fines académicos [12]; adicionalmente, se requiere el uso de una herramienta para la configuración y la administración de la Grid. Por lo anterior, se requiere el *Globus Toolkit*; la página de *Globus* ofrece toda la documentación necesaria para su implementación [10]. La utilización de la herramienta *Globus* se determinó debido al uso y a la plataforma que actualmente tiene la Universidad Distrital y el Grupo de Investigación GICOGÉ [17]. Para implementar la Grid, el software tiene un módulo de realización de pruebas, el cual no requiere un recurso adicional para la implementación.

III. ENTORNOS VIRTUALES DE APRENDIZAJE EN COMPUTACIÓN DE ALTO DESEMPEÑO

El desarrollo de la humanidad ha estado marcado por los productos resultantes de la ciencia y la tecnología, dentro de los cuales las TIC ocupan un lugar muy importante, ya que han transformado la forma de comunicarnos, de acceder al conocimiento, a la cultura, etc. De acuerdo a Stavrou (2001), las TIC “son una oportunidad para el desarrollo de los países y una de las condiciones esenciales para participar activamente en el mercado global, donde es necesario cerrar las brechas para que el intercambio de bienes y servicios y de capitales se dé con el menor número de asimetrías posibles”.

Otro de los ejes fundamentales de nuestro trabajo es la construcción de un entorno virtual de aprendizaje (EVA), definido como un espacio o comunidad organizada con el propósito de aprender. Para que sea lugar de aprendizaje han de estar presentes ciertos componentes, que se definen desde una óptica interdisciplinar: a) funciones pedagógicas (actividades de aprendizaje, situaciones de enseñanza y materiales de aprendizaje, apoyo y evaluación, entre otros), b) las tecnologías apropiadas y c) la organización social de la educación (espacio, calendario y comunidad) [12]. El uso de las herramientas de telecomunicación en el proceso de enseñanza y aprendizaje es la diferencia clave [13].

Los ambientes virtuales permiten la simulación de modelos (industrias productivas, robots, simuladores de vuelo, escenarios de una sala de guardia de emergencias, simulación de mercados competitivos, etc.) y representan una de las perspectivas más ricas que ofrece la tecnología. En ellos, los actores del proceso educativo pueden ubicarse contextualmente en situaciones dinámicas, propias de la vida profesional, y analizar la evolución de los fenómenos y de las potenciales respuestas técnicas que se puede dar a cada situación [14].

En general, los EVA son dominios en línea que permiten la interacción sincrónica y asincrónica entre el profesorado y el alumnado; además, contienen recursos de aprendizaje que pueden utilizarse por los estudiantes en cualquier momento.

IV. METODOLOGÍA PARA EL DESARROLLO DE LA PLATAFORMA PARA LA SIMULACIÓN DE UNA GRID

La metodología adoptada requiere la aplicación de conocimientos interdisciplinarios, habilidades y técnicas de diseño de ingeniería y de las ciencias experimentales, para lo cual se utilizaron métodos cuantitativos y cualitativos. El enfoque metodológico es de tipo teórico, holístico, descriptivo y explicativo. Concretamente, la metodología utilizada consiste, esencialmente, de tres fases: análisis, desarrollo e implementación; la estrategia de trabajo se basó en efectuar una reformulación constante de los aspectos teóricos, metodológicos y de las pautas de validez.

Fase 1: Análisis

Este proyecto necesitó conceptualizar y analizar los conocimientos que requiere la implementación de cualquier proyecto desarrollado en un entorno Grid, esto es: programación en java, programación de servicios web, conocimientos en XML, UNIX o LINUX, redes, Globus y sistemas operativos. Las máquinas virtuales requieren el sistema operativo Scientific Linux; una vez se configura la máquina base, se puede proceder a replicarla para elaborar toda la red. La guía de instalación del sistema operativo se puede encontrar en scientificlinux [15].

A continuación, se describe el diseño lógico, funcional y físico de la aplicación desarrollada para este trabajo de investigación.

Fase 2: Desarrollo

Una vez instalado el sistema operativo se procede a la configuración de la plataforma Grid. Para el proyecto se toma una máquina base con la instalación elemental del Globus Toolkit; la guía para la configuración de esta se encuentra en la página citada [10]. Una vez configurada la máquina base con el Globus, se puede replicar para la elaboración de la plataforma (MVBase.zip).

Se procede luego a configurar 5 máquinas virtuales que emulan el desarrollo de las supercomputadoras en un entorno real. Las máquinas construidas fueron CA (Autoridad Certificadora), que certificará todos los usuarios de la Grid; la máquina donde se administran los DNS; la WSGRAM; la máquina que servirá de PROXY, y una máquina para el servicio de ftpGRID. Esta infraestructura virtual es la base de la plataforma para simular el comportamiento de una Grid.

Se selecciona la máquina administradora, la CA (Autoridad Certificadora), y se le asigna la IP 10.10.10.10 para la simulación. Se requiere una entidad certificadora de los miembros que conforman o quieren acceder a la Grid. Seguido de esto se define el administrador del DNS, por medio del webmin, el cual deberá adicionar la zona directa del DNS a la CA. Se selecciona el nombre y se asocia a la IP; deben ser distintos a la de la máquina CORE y la CA.

Comandos de configuración de las máquinas virtuales para la Grid

Se procedió, a partir del Shell de comandos de cada máquina virtual, la configuración de ellas, para ser certificadas por la CA (entidad certificadora), con el propósito de colocar en un dominio confiable los procesos que se pueden ejecutar en la plataforma y por fuera de ella.

Arquitectura de la plataforma

El desarrollo del prototipo presenta una arquitectura distribuida (ver Fig. 1) y se puede emplear para lograr aplicaciones didácticas y prácticas de conceptos como lo son servicios web, arquitecturas de hardware y software, programación de sistemas distribuidos y servicios Grid; de esta forma se alcanza la aplicación de la teoría a elementos que permiten visualizar los conceptos de la computación Grid.

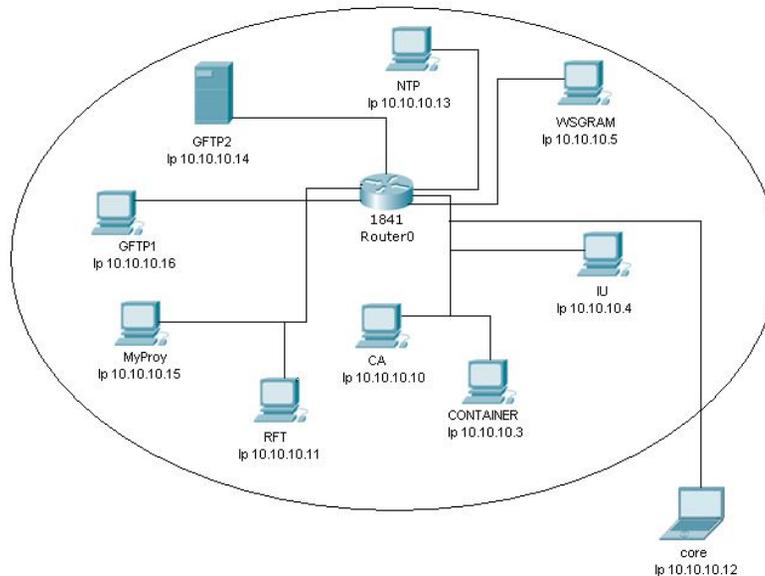


FIG. 1. Plataforma Grid con máquinas virtuales.

Basados en la necesidad de generar un entorno para el aprendizaje relacionado con la computación de alto desempeño (Grid) se logró poner a funcionar cada una de las máquinas virtuales en línea. Se crearon las máquinas virtuales con su respectiva configuración; de esta forma se realiza la transmisión de datos entre ellas, para lo cual fue necesario identificar cada una de las máquinas con una dirección IP. Se configuró un servidor DNS para resolver los nombres de las máquinas, y se encriptaron las transmisiones dentro de la grilla, logrando solo comunicar las máquinas que fuesen certificadas por la autoridad CA. De esta forma se construye una plataforma para simular de una red Grid con máquinas virtuales para crear un entorno de aprendizaje de la computación de alto desempeño.

Fase 3: Implementación

En la implantación y validación se contó con la ayuda de un grupo de docentes del programa de Ingeniería

de Sistemas de la Universidad la Amazonía; sus observaciones y sugerencias fueron tomadas en cuenta para mejorar el sistema y apuntar hacia su óptimo funcionamiento.

V. EVALUACIÓN Y RESULTADOS

Un objetivo de las pruebas de software es descubrir defectos en el software, es decir, que su comportamiento es incorrecto, no deseable o no cumple su especificación. La prueba de defectos está relacionada con la eliminación de todos los tipos de comportamientos no deseables del sistema, tales como caídas del sistema, interacciones no permitidas con otros sistemas, cálculos incorrectos y corrupción de datos.

Se encontraron dificultades en el momento de implementar la plataforma, dado los conocimientos específicos del software para el desarrollo del

ambiente de simulación; sin embargo, el trabajo en sí del montaje de la plataforma es un escenario lleno de recursos para el aprendizaje de la computación de alto desempeño Grid.

VI. CONCLUSIONES Y TRABAJOS FUTUROS

El trabajo que aquí se presenta se propuso establecer un entorno para el aprendizaje de la computación de alto desempeño a través del uso de máquinas virtuales. Lo anterior permite realizar trabajos en un entorno de bajo coste y accesible a todos los centros de educación técnica superior interesados en enseñar sobre computación en sus diferentes niveles.

El uso de la tecnología Grid despierta gran interés en la comunidad académica y científica, por la capacidad de procesamiento y almacenamiento que puede lograrse con los recursos que tienen las organizaciones. Con la simulación y el entorno de aprendizaje desarrollado se pueden abrir caminos para que esta tecnología se masifique y pueda ser puesta en marcha en diferentes tipos de empresas.

Existen limitantes en los servicios de almacenamiento a gran escala debido a los recursos de hardware que se usaron en la simulación, pero es viable el aprendizaje de estos servicios con volúmenes pequeños de información, dada la escalabilidad que presenta la arquitectura Grid.

En próximos trabajos se pueden desarrollar guías basadas en la plataforma para la ejecución de todos los servicios Grid que trae el software Globus, y realizar nuevas plataformas con middleware diferentes para realizar comparaciones y profundizar en los conceptos teóricos de la computación.

REFERENCIAS

[1] S. Bonner, *La Educación en la Roma antigua*, Barcelona, 1984.

[2] R. Salas Perea y P. Ardanza Zulueta, “La simulación como método de enseñanza y aprendizaje” *Educación Médica Superior*, vol. 9, nº 1, pp. 3-4, 1995.

[3] P. A. Bernstein, “Middleware: a model for distributed system services,” *Communications of the ACM*, vol. 39, nº 2, pp. 86-98, 1996. DOI: <http://dx.doi.org/10.1145/230798.230809>.

[4] M. B. González, J. M. Muñoz y M. Montenegro, “Computación de alto rendimiento para el procesamiento de imágenes digitales [High-performance computing for digital image processing],” *Ventana Informática*, vol. 30, pp. 55-65, 2014.

[5] J. Velandia y J. Castillo, “Plataforma de computación Grid para redes inalámbricas de geosensores” *Gerencia Tecnológica Informática*, vol. 11, nº 31, pp. 35-44, 2012.

[6] G. Becerra y J. Pérez, “Servicio Grid para la clasificación no supervisada de imágenes satelitales utilizando autómatas celulares,” *Ciencia e Ingeniería Neogranadina*, vol. 19, nº 2, pp. 59-76, 2014.

[7] X. U. Shukla, P. S. Parsania y K. C. Kamani, “Applications of Grid Computing in Agriculture: An Indian Scenario,” *Gujarat Journal of Extension Education*, vol. 25, pp. 58-60, 2014.

[8] G. Isaza y N. Méndez, “Arquitecturas y modelos de programación en computación grid,” *Scientia et Technica*, vol. 37, pp. 433-437, 2007.

[9] M. Madoz, “E-Learning: Extensiones de los entornos virtuales para el manejo de actividades experimentales en Informática,” *In XIV Congreso Argentino de Ciencias de la Computación*. 2008.

[10] Globus.or, «Globus Toolkit Downloads,» [En línea]. Available: <http://toolkit.globus.org/toolkit/downloads/>.

[11] J. Cascón y C. Guarch, *Nuevas tecnologías para el aprendizaje*, Ediciones Pirámide, 1997.

[12] D. Baker, “Coronal mass ejections, magnetic clouds, and relativistic magnetospheric electron events: ISTP,” *Journal of Geophysical Research: Space Physics*, vol. 103, nº 8, pp. 17279-17291, 1998. DOI: <http://dx.doi.org/10.1029/97JA03329>.

[13] Y. Montes, “Los Procesos de Enseñanza y Aprendizaje del Idioma Inglés a través de los Entornos Virtuales de Aprendizaje,” *Revista Ciencia Unemi*, vol. 8, nº 13, pp. 47-55, 2015.

[14] L. Bertogna, R. Del Castillo, F. López Luro y C. Zanellato, “Prácticas Remotas sobre

- Laboratorios físicos y virtuales,” 2004. [En línea]. Available: http://sedici.unlp.edu.ar/bitstream/handle/10915/21273/Documento_completo.pdf?sequence=1.
- [15] scientificlinux, “scientificlinux.org,” [En línea]. Available: <https://www.scientificlinux.org/downloads/sl-mirrors/>.
- [16] A. Urquidi Martin y M. Calabor Prieto, “Aprendizaje a través de juegos de simulación: un estudio de los factores que determinan su eficacia pedagógica,” *Revista Electrónica de Tecnología Educativa*, vol. 47, pp. 1-15, 2014.
- [17] J. N. Pérez Castillo, “GICOGE,” [En línea]. Available: <http://gicoge.udistrital.edu.co/>.