

Propuesta de arquitectura para el soporte de una plataforma de Teleformación Moodle

Proposed architecture to support a Moodle Learning Manager System

Juenlis Enrique Coss Piña¹, Adrian Misael Peña Montero¹, Cesar Miguel Calvo de la Paz¹

¹Universidad de las Ciencias Informáticas (UCI)

juenlis@uci.cu, {ampena, [cmcalvo](mailto:cmcalvo@estudiantes.uci.cu)}@estudiantes.uci.cu

Resumen

En los últimos años la enseñanza virtual ha ocupado un lugar significativo en el entorno educativo a escala mundial, lo que ha propiciado la aparición de novedosas tecnologías y numerosas plataformas de Teleformación para el soporte de la misma. En la Universidad de las Ciencias Informáticas (UCI) el desarrollo y perfeccionamiento del proceso de enseñanza aprendizaje ocupa un lugar prioritario, donde el modelo de formación prevé el uso intensivo de las tecnologías de la información y las comunicaciones, para lo que tiene implementada una plataforma de Teleformación donde se gestionan los cursos para la formación de pregrado, postgrado y de segundo perfil para más de 13 mil usuarios entre estudiantes y profesores. Debido al incremento del número de actividades y recursos disponibles en ese entorno virtual y la tendencia al incremento de los mismos, el acceso de usuarios concurrentes se eleva paulatinamente, lo que provoca deficiencias en la estabilidad del servicio y en el rendimiento del mismo. Con este trabajo se pretende proponer una arquitectura que garantice alto rendimiento y escalabilidad para el soporte de una plataforma de Teleformación Moodle de grandes dimensiones utilizando ordenadores de pocas prestaciones.

Palabras clave: clúster, moodle, teleformación, , software libre.

Abstract

In recent years the virtual learning has occupied a significant place in the educational environment on a global scale, which has led to the emergence of many new technologies and e-learning platforms to support it. At the University of the Computer Science development and refinement of the teaching-learning process is a priority, where the model of training provides intensive information technology and communications, which has implemented a learning manager system of which are run training courses for undergraduate, graduate and second-profile for more than 13 thousand users among students and teachers. Due to the increasing number of activities and resources available in this virtual environment and the trend of increasing them access concurrent users rises gradually, causing deficiencies in the stability of service and performance. With this work is to propose an architecture that ensures high performance and scalability to support a large Moodle e-learning platform using computers of few benefits.

Keywords: cluster, e-learning, free software, moodle.

Introducción

Actualmente en todo el mundo está cada vez mas presente el uso intensivo de las Tecnologías de la Informática y las Comunicaciones (TICs) en el proceso de enseñanza-aprendizaje. La Enseñanza Virtual o *e-learning* ha cobrado fuerzas en los últimos años; su aplicación e integración de las nuevas tecnologías en los procesos formativos, se presenta como un gran reto tanto para las instituciones educativas, como para las organizaciones empresariales.

Uno de los factores que se tienen en cuenta para la asimilación de las TICs en la formación a distancia es la selección de una Plataforma de Teleformación, la cual aporte las herramientas y recursos necesarios para gestionar, organizar, coordinar, diseñar e impartir programas de formación a través de la red, y sea lo suficientemente flexible como para poder crear entornos de formación y aprendizaje adaptados a las necesidades de cualquier tipo de organización.(Centro de Educación y Nuevas Tecnologías de la UJI, 2004)

La Universidad de las Ciencias Informáticas tiene implementada para la gestión de cursos a distancia la plataforma de código abierto Moodle, una de las más difundidas en todo el mundo, la cual está diseñada para facilitar la comunicación pedagógica entre los participantes en un proceso educativo, sea este completamente a distancia, presencial, o de naturaleza mixta que combine ambas modalidades en diversas proporciones.

Para ello, Moodle cuenta con distintos tipos de herramientas de comunicación sincrónica y asincrónica, herramientas para la gestión de los materiales de aprendizaje y herramientas para la gestión de usuarios que incluye sistemas de seguimiento y evaluación del progreso de los estudiantes, así como gran cantidad de módulos y filtros que le aportan funcionalidad a la plataforma y posibilidad de integración con otras aplicaciones. (Dirección de Teleformación, 2007)

Actualmente, en el Entorno Virtual de Aprendizaje de la UCI (EVA UCI), se gestionan todos los cursos del plan de estudios de la formación de pregrado de la carrera de Ingeniería Informática, el plan de estudio de la formación de postgrado para los profesores de la Universidad y los cursos generados por las asignaturas optativas de los segundos perfiles y de la producción principalmente, soportando en total más de 300 cursos, a los cuales acceden en cualquier momento y desde cualquier lugar más de 13 mil usuarios entre estudiantes y profesores.

Dada la variedad de herramientas en el EVA UCI para la comprobación de los conocimientos del estudiante, es posible realizar diferentes tipos de evaluaciones, siendo la más utilizada la realización de exámenes en línea, los cuales en ocasiones son aplicados en una asignatura a todos los estudiantes de un año de la carrera, lo que significa aproximadamente 2 mil usuarios navegando concurrentemente dentro de la aplicación, lo que genera una carga que el sistema no es capaz de soportar.

Otro factor que está elevando paulatinamente el acceso de usuarios concurrentes, es el aumento del número de actividades y recursos disponibles en ese entorno virtual y la tendencia al incremento del uso de los mismos, lo que provoca deficiencias en la estabilidad del servicio y en el rendimiento de este.

Para el soporte de la plataforma, la Dirección de Teleformación de la UCI cuenta con 10 ordenadores y un servidor profesional donde están montadas réplicas del sistema, las cuales son complejas de administrar y brindan un servicio deficiente, incapaz de soportar con estabilidad la concurrencia generada.

Tomando en cuenta la situación explicada anteriormente, se ha planteado el problema fundamental ¿Cómo implementar un sistema que soporte los altos niveles de concurrencias generados en la plataforma de Teleformación de la UCI?

Para solucionar el problema planteado el presente trabajo tiene como objetivo proponer una arquitectura que, aprovechando los recursos disponibles, garantice alto rendimiento y escalabilidad para el soporte de la plataforma de Teleformación de la UCI.

Metodología Computacional

Para lograr implementar un sistema para dar solución al problema planteado se realizó un estudio sobre las tecnologías de *clustering*, caché y almacenamiento compartido, lo cual derivó la decisión de elaborar una propuesta que involucrara el diseño de un clúster heterogéneo¹ que permita incrementar la capacidad de procesamiento usando tecnología estándar, tanto en componentes de hardware como de software, el cual presente combinaciones de servicios con alto rendimiento, equilibrio de carga y escalabilidad.(Sabbut, 2004).

Los motivos para utilizar un clúster de alto rendimiento están dado por el tamaño del problema por resolver y el precio de la máquina necesaria para resolverlo, mediante el cual se puedan conseguir capacidades de cálculo superiores a las de un ordenador más caro que el costo conjunto de los ordenadores del clúster.

Para garantizar esta capacidad de cálculo, los problemas necesitan ser paralelizables, ya que el método con el que los clúster agilizan el procesamiento es dividir el problema en problemas más pequeños y calcularlos en los nodos². Estos últimos son los que contendrían los servidores de aplicación que comparten la carga de trabajo y de tráfico de los clientes, por lo que la solución más razonable suele ser la de redundar los servidores que ofrecen dichos servicios, para que trabajando como un único servidor (servidor virtual), se consiga un sistema escalable³ y de mayor capacidad que mejore el rendimiento del sistema. (Manuel, 2004).

¹ Conjunto de ordenadores con diferentes hardware y sistema operativo unidos mediante una red de alta velocidad que actúan en conjunto usando el poder de cómputo de varios CPU en combinación para resolver ciertos problemas dados.

² Cada una de las máquinas que conforman el clúster.

³ Sistema cuyo rendimiento es mejorado después de haberle añadido más capacidad hardware.

El elemento clave para implementar un servidor virtual de estas características, es el balanceador de carga, cuya función es actuar como *frontend* del clúster, repartiendo las peticiones de servicio que recibe éste a los nodos del mismo; dividiendo el trabajo, gracias a un algoritmo, de la manera mas equitativa posible, para evitar así los denominados cuellos de botella, solucionando uno de los principales problemas de los sitios Web de grandes prestaciones, que es como gestionar las solicitudes de un gran número de usuarios.(LVS Project, 1998).

Las aplicaciones, para ser montadas sobre un clúster, deben atender a cierta flexibilidad, donde la flexibilidad está dada por la capacidad del recurso de ser estático virtualmente y ser dinámico físicamente, lo que se traduce en que la aplicación esté diseñada con una clara separación de las capas de aplicación que permita firmemente configuraciones escalables.

El diseño de la plataforma Moodle permite perfectamente la separación de los elementos de esta, de forma tal, que se puede separar la aplicación del Sistema Gestor de Base de Datos (SGBD), este último almacenará y gestionará la información de los cursos, por lo que debe estar configurado para que acepte peticiones de cada uno de los servidores Web del clúster y optimizado para un alto rendimiento.

Otra de las características a tomar en consideración en la implementación de una plataforma Moodle, es que la misma necesita una carpeta donde almacenar los materiales de los cursos y las sesiones PHP de los usuarios, la cual entonces debe ser común para cada una de las instancias de Moodle instaladas en los nodos del clúster, los cuales deben tener acceso en todo momento a la ultima información sin importar por cual de ellos fue actualizada, solución la cual se resolvería con la implementación de un Servidor de Ficheros que centralice y comparta los datos.

Los sistemas de almacenamiento compartido se caracterizan por ser más lentos que si se tuviera la información local, por lo que para contrarrestar este efecto, es necesario implementar en cada uno de los nodos un sistema de caché en disco que mantenga almacenados localmente una copia de los archivos originales que son accedidos con mayor frecuencia, con el objetivo de agilizar la carga de los mismos una vez que vuelvan a ser solicitados.

El servidor de ficheros debe tener además la capacidad de compartir la carpeta de las sesiones PHP de los usuarios de forma tal que la conexión entre los distintos nodos del clúster y el servidor de fichero sea de una forma sincrónica, posibilitando que cada instancia de Moodle pueda acceder al mismo tiempo a la sesión del usuario sin que esta expire durante el tiempo que esté activo el usuario en el sistema.

Otro de los aspectos a tomar en consideración para alcanzar un mejor rendimiento del sistema es la reducción del ancho de banda consumido, la carga de los servidores de aplicación (nodos) y el retardo en la descarga, lo que se logra poniendo entre los clientes y el balanceador de carga un servicio de proxy-caché inverso (o acelerador Web), el cual almacena copias de los datos que pasan por él, de forma que subsiguientes peticiones pueden ser respondidas

por el propio caché, haciendo que el tiempo de acceso medio al dato sea menor, con la consiguiente mejora del rendimiento de los servidores Web, al evitarles procesar peticiones repetidas que ya han sido incorporadas a la caché.

Diseño gráfico de la solución

La siguiente Figura muestra el diseño gráfico de la solución que se propuso anteriormente.

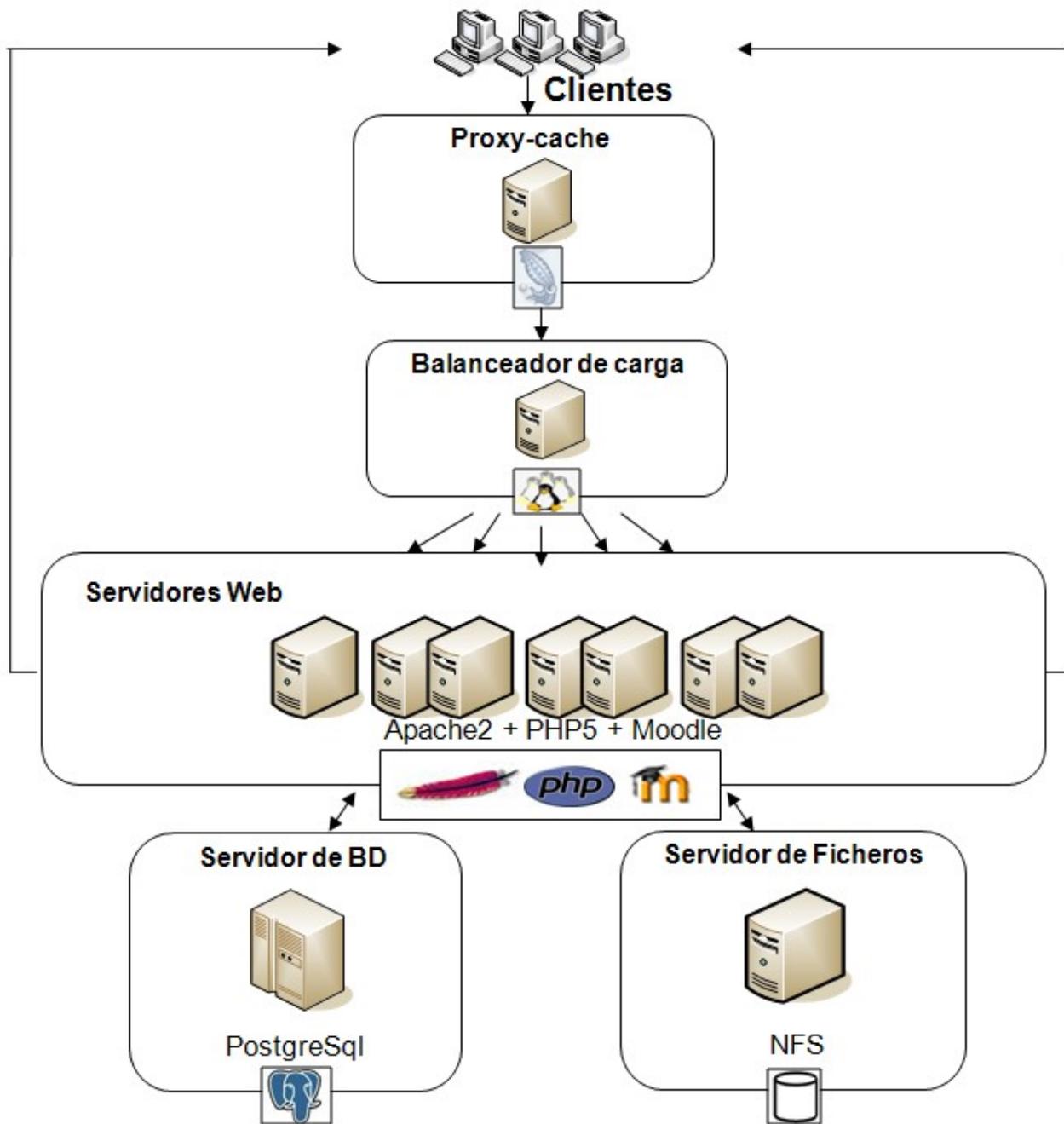


Fig. 1. Diseño gráfico de la solución.

Implementación de la seguridad para la solución

Es de vital importancia tratar el problema de la seguridad informática dentro de la solución. Para ello, resulta importante establecer políticas de seguridad, las cuales van desde el monitoreo de la infraestructura de red, los

enlaces de telecomunicaciones, la realización del respaldo de datos y hasta el reconocimiento de las propias necesidades de seguridad, para establecer los niveles de protección de los recursos. (Ciberhabitat, 2002)

Seguridad ante ataques por la red

La seguridad ante ataques por la red se garantizará individual para cada una de las aplicaciones que prestan servicio en los servidores involucrados en la solución y la implementación de un *firewall* por software en cada uno de estos.

La configuración del servidor Web debe incluir las siguientes reglas:

Permitir conexiones desde cualquier dirección IP de la universidad solo a los servicios que se están prestando en el servidor (HTTP y HTTPS).

Otorgar adecuadamente permisos de lectura y escritura solo donde se requiera.

La configuración del servidor de base de datos debe incluir las siguientes reglas:

Restringir las conexiones desde cualquier computadora.

Brindar acceso a la base de datos solo a los nodos del clúster donde se encuentran los servidores web y a la máquina del Administrador de Red que se encarga de monitorear el estado del sistema, siempre con los permisos únicamente necesarios.

La configuración del servidor de salvadas debe incluir las siguientes reglas:

Restringir los accesos desde cualquier computadora.

Brindar solo acceso de conexión a las máquinas que integran la solución.

La configuración del *firewall* en cada uno de los servidores debe incluir:

Política de denegación por defecto.

Permitir conexiones solo a las direcciones de IP que lo requieran según el servicio que se esté prestando en el servidor.

Permitir solo el acceso a los puertos que se requieran según el servicio que se esté prestando en el servidor.

Permitir conexiones solo desde direcciones IP autorizadas al servicio SSH.

Protección y alerta ante los ataques más comunes como son los de negación de servicios (DoS) y la penetración al sistema por intrusos.

Monitoreo de los servicios que prestan los servidores

Mantener un monitoreo sobre los servicios y recursos del sistema nos permitirá detectar problemas y obtener información oportuna para el análisis del comportamiento del mismo, siendo necesario la implementación de un sistema que mantenga un monitoreo continuo para cada uno de los servidores y sus servicios; y en caso de existir

fallos que informe inmediatamente a los administradores. Para esto se puede usar el mecanismo que se prefiera, aunque se sugiere el envío automático de correo y pager.

Es importante también realizar monitoreo al rendimiento de los servidores. El administrador del sistema decide con que frecuencia será realizado. Este monitoreo debe ser, principalmente, a los recursos CPU, memoria, acceso a disco, y tráfico de red.

Conclusiones

En este trabajo se analizaron aspectos relacionados con las tecnologías de *clustering*, caché y almacenamiento compartido, que permitieron elaborar una propuesta de arquitectura de un clúster para el soporte de la plataforma de Teleformación de la UCI, que es capaz de soportar los altos niveles de concurrencias generados esencialmente en los momentos de mayor estrés cuando se realizan exámenes en línea para todos los estudiantes de un año de la Universidad.

Referencias Bibliográficas

- Centro de Educación y Nuevas Tecnologías de la UJI. Selección de un entorno virtual de enseñanza/aprendizaje de código fuente abierto para la Universidad de Jaime I, 2004. [Disponible en: http://cent.uji.es/doc/eveauji_es.pdf]
- Ciberhabitat. Seguridad Informática ¿Qué, por qué y para qué?, 2002. [Disponible en: <http://www.ciberhabitat.gob.mx/museo/cerquita/redes/seguridad/intro.htm>]
- Dirección de Teleformación. Selección de un Entorno Virtual de Aprendizaje para la Universidad de las Ciencias Informáticas, 2007. [Disponible en: <http://teleformacion.uci.cu/mod/resource/view.php?id=8849>]
- Dodo. Cache web, 2007. [Disponible en: http://es.wikipedia.org/wiki/Caché_web]
- Javier. Sistema de archivos distribuidos, 2006. [Disponible en: http://es.wikipedia.org/wiki/Sistema_de_archivos_distribuido]
- Jiménez, M.M. and G.B. Pujol. Arquitecturas de Clustering de Alta Disponibilidad y Escalabilidad, 2003 [Disponible en: <http://idefix.eup.uva.es/Manuales/Clustering/ACADE-LVS-memoria.pdf>]
- LVS Project. Sitio Oficial del proyecto Linux Virtual Server, 1998. [Disponible en: <http://www.linuxvirtualserver.org/>]
- Manuel. Clúster de alto rendimiento, 2004. [Disponible en: http://es.wikipedia.org/wiki/Cluster_de_alto_rendimiento]
- Pedro Paredes, Juan. Alta disponibilidad para Linux, 2001. [Disponible en: <http://es.tldp.org/Presentaciones/200103hispalinux/paredes/pdf/LinuxHA.pdf>]
- Sabbut. Cluster (informática), 2004. [Disponible en: [http://es.wikipedia.org/wiki/Cluster_\(informática\)](http://es.wikipedia.org/wiki/Cluster_(informática))]
- Suárez, José M. Sistema HA bajo Linux, 2002. [Disponible en: http://www.goa.es/docs/linux_ha_apache.pdf]