

## Technological design of an e-Learning platform

◆ Marcos Montero Torres y Andrés Navarro-Soto Ballester

### Resumen

Proyecto de diseño e implantación de una arquitectura Moodle con capacidad para dar servicio a un colectivo de 30.000 usuarios bajo criterios de escalabilidad, rendimiento y alta disponibilidad. Al margen del diseño de la arquitectura, se definen y desarrollan unas pruebas de rendimiento que nos permitan mejorar la eficiencia del sistema y determinar de manera fiable la validez de la plataforma respecto a la dimensión del entorno con antelación suficiente a su puesta en explotación. Asimismo, se pretende poder repetir este tipo de estudios de rendimiento para realizarlos.

**Palabras clave:** Moodle, rendimiento, Jmeter, Campus Digital, UPC, UPCnet, Atenea.

### Summary

Design and implementation project for a Moodle architecture capable of serving a group of 30,000 users under scalability, performance and high availability criteria. Aside from the architecture design, performance tests are defined and developed to improve system efficiency and determine, reliably, the validity of the platform as regards the size of the environment well in advance of its launch. It also seeks to ensure that this type of performance study can be carried out regularly prior to any future modifications to the Moodle platform.

**Keywords:** Moodle, performance, Jmeter, Digital Campus, UPC, UPCnet, Atenea.

## 1. Introducción

Durante el curso 2005/2006 la UPC realizó un proyecto piloto de implantación de una nueva plataforma de campus digital basada en Moodle. Una vez decidida su validez como herramienta de e-learning, se procedió a diseñar un proyecto para acometer la extensión del nuevo campus digital (conocido como Atenea) a toda la comunidad UPC.

En su nueva fase, el campus digital debería dar servicio a un colectivo de 30.000 estudiantes y 3.000 profesores, ofreciendo unos 4.000 cursos Moodle. La plataforma definitiva para ofrecer este servicio debía de cumplir una serie de requisitos en cuanto a rendimiento, escalabilidad y disponibilidad que garantizasen el correcto funcionamiento de Moodle bajo elevadas condiciones de carga y permitieran dar respuesta con facilidad a futuras ampliaciones del servicio, ya fuera en número de usuarios, asignaturas o capacidad de carga.

La entrada en funcionamiento del sistema se produjo en septiembre de 2006 y durante los meses anteriores se llevó a cabo el proyecto de diseño de la plataforma y un estudio de rendimiento para garantizar su correcto dimensionado.

## 2. Arquitectura Hardware

### 2.1. Arquitectura del sistema:

La arquitectura definitiva consta de 4 capas diferenciadas:

- **Balaceo de carga:**

Un clúster formado por 2 servidores Linux con LVS + HA que se encargan de balancear el tráfico entre los diferentes Front-ends HTTP. Permiten distribuir la carga de manera "inteligente", así como mejorar la disponibilidad eliminando de forma dinámica los Front-ends que no se encuentren en funcionamiento.

◆  
Durante el curso 2005/2006 la UPC realizó un proyecto piloto de implantación de una nueva plataforma de campus digital basada en Moodle

◆  
La entrada en funcionamiento del sistema se produjo en septiembre de 2006 y durante los meses anteriores se llevó a cabo el proyecto de diseño de la plataforma y un estudio de rendimiento para garantizar su correcto dimensionado



Apache Jmeter ha sido la herramienta seleccionada para realizar las pruebas de rendimiento

Como paso previo para la realización de las pruebas de carga, se elaboró un perfil de uso del campus digital basado en los datos reales de utilización del sistema durante el proyecto piloto

- **Front-ends HTTP**

Está formada por 3 servidores Linux con Debian + Apache.

- **Back-end**

Formada por 2 nodos, uno activo y uno pasivo. Ambos con Linux RedHat y PostgreSQL 8.1.

En el momento de la implantación de Moodle las únicas opciones posibles para usar como SGBD eran MySQL y PostgreSQL. Ninguna de estas 2 SGBDs ofrece todavía un soporte para clustering suficientemente maduro. Por este motivo se optó por la utilización de un sistema sin alta disponibilidad automatizada aunque el nodo de backup, junto con el almacenamiento en SAN, permite una rápida recuperación del servicio en caso de avería hardware.

- **Almacenamiento SAN**

Los datos se encuentran ubicados en una SAN con particiones RAID-6 de 200Gb para la base de datos y 800Gb para el sistema de ficheros "moodledata".

### 3. Herramienta de simulación

La herramienta seleccionada para realizar las pruebas de rendimiento ha sido Apache Jmeter. Los factores fundamentales que se han tenido en cuenta para su elección son los siguientes:

- Posibilidad de establecer y verificar criterios de calidad de uso: por ejemplo, la generación de estadísticas en base al tiempo de respuesta para cada petición HTTP o el número/porcentaje de peticiones erróneas.
- Gestión de concurrencia de usuarios: la herramienta debe ser capaz de simular accesos simultáneos al sistema por parte de usuarios diferentes. Debe poder gestionar su autenticación y las características propias de cada uno de ellos.
- Escalabilidad: es imprescindible que la herramienta sea capaz de generar simulaciones con cargas muy elevadas para responder a las necesidades cada vez mayores del entorno del campus digital.
- Posibilidad de realizar pruebas realistas: no sirve cualquier simulación. Es necesario utilizar patrones de navegación semejantes a los de los usuarios del sistema, tiempos de sesión similares a los suyos e incluso intervalos entre diferentes peticiones lo más próximos posibles a la realidad. En definitiva, es importante poder diseñar perfiles de uso realistas y reproducirlos a gran escala.

### 4. Pruebas de rendimiento

La realización de las pruebas de rendimiento del sistema tenía dos objetivos:

- Certificar la validez de la arquitectura tecnológica escogida para soportar volúmenes de carga adecuados a la magnitud del colectivo UPC.
- Detectar los aspectos mejorables de la plataforma para dotarla de la capacidad de respuesta suficiente y una vez conseguida ésta, tener una idea lo más exacta posible de los elementos susceptibles de mejora en futuras ampliaciones.

#### 4.1. Preparación del escenario de pruebas

Como paso previo para la realización de las pruebas de carga, se elaboró un perfil de uso del campus digital basado en los datos reales de utilización del sistema durante el proyecto piloto. La intención

era conseguir información lo más fiable posible sobre el estilo de navegación de los usuarios en el campus: tiempos de sesión, secciones más accedidas, intervalos entre peticiones, etc.

Una vez determinados los aspectos más relevantes del comportamiento de los usuarios y añadiendo información específica suministrada por usuarios clave del sistema, se elaboraron 2 perfiles de navegación diferenciados para profesores y alumnos.

#### 4.2. Ejecución de las pruebas: Cuellos de botella y Soluciones implementadas

Establecido el escenario de pruebas, comenzó un plan de ejecución en un entorno de test más la realización de 2 pruebas finales en el entorno de explotación ya con su configuración definitiva.

Para las pruebas sobre el entorno de explotación se utilizó una configuración de Jmeter en clúster para poder generar la carga necesaria. El objetivo era verificar si el sistema era capaz de atender 1.000 usuarios trabajando simultáneamente, para lo que se realizaron pruebas de las siguientes características:

TABLA 1. CARACTERÍSTICAS DE LAS PRUEBAS DE CARGA				
	#PC cluster Jmeter	#Hits/seg	#Hits/hora	Usuarios concurrentes
Prueba 1	12	600	2.100.000	1200
Prueba 2	18	800	2.900.000	1500

Se trata de pruebas con 1.200 o 1.500 sesiones de 7 minutos de usuarios diferentes que se ejecutan de manera solapada en el tiempo.

A continuación se resumen las mejoras efectuadas en cada uno de los elementos de la plataforma para conseguir los niveles de rendimiento deseados.

##### Balanceadores de carga:

Las actuaciones sobre ellos consistieron en afinar al máximo los valores de los diferentes time-outs del director y en definir un test adecuado para la comprobación del servicio http.

##### Front-ends HTTP:

- Instalación de un acelerador PHP. (PHP eAccelerator)
- Procesos de Apache.

Es el foco prioritario de las mejoras. La capacidad de concurrencia del web server depende, en gran medida, de la RAM disponible. Debido a las limitaciones impuestas por la combinación Moodle-PHP (uso del módulo prefork) la cantidad de peticiones simultáneas que puede atender el sistema tiene como límite la memoria disponible dividida entre la RAM ocupada por un proceso HTTP. Se observó que un proceso Apache ocupaba entre 10 y 20Mb de RAM.

Como medida para aumentar la capacidad del sistema se instaló en los nodos Front-end un segundo servidor Apache compilado sin la mayor parte de módulos que lleva un servidor estándar, de manera que sus procesos sólo ocupan 2Mbytes de RAM.

Este servidor pasa a ser el que atiende en primera instancia. Para cada petición que le llega determina el tipo de contenido, sirve la petición si se trata de contenido estático, y en otro caso, la pasa al servidor apache estándar.

##### Back-end:

En el Back-end se trabajó en buscar la configuración necesaria para garantizar que el SGBD pudiera soportar un número muy elevado de conexiones. Aunque la experiencia ha demostrado que el número de conexiones a la Base de datos suele mantenerse en niveles bastante bajos, interesaba garantizar que en caso necesario, y con recursos de memoria y CPU suficiente, podría atenderse hasta 1.500 conexiones.

Para las pruebas sobre el entorno de explotación se utilizó una configuración de Jmeter en clúster para poder generar la carga necesaria

Como medida para aumentar la capacidad del sistema se instaló en los nodos Front-end un segundo servidor Apache compilado



## 5. Conclusiones

Las cifras extraídas durante el primer cuatrimestre de uso del nuevo Campus Digital son suficientemente clarificadoras y refrendan la validez de los resultados obtenidos con las pruebas de rendimiento y las mejoras en el diseño de la plataforma.

En la actualidad cada día utiliza la plataforma Atenea 1/3 del total de alumnos de la UPC; algunos más de una vez. En una semana 2/3 han accedido al sistema en algún momento, y en un mes la práctica totalidad del alumnado ha accedido al campus al menos en una ocasión.

En los días laborables el número de sesiones Moodle se encuentra en valores cercanos a 1.000 de manera sistemática entre las 10:00h y las 22:00h y en épocas de uso intensivo de la plataforma se han producido picos de 2.000 usuarios simultáneos en el sistema sin impacto en la carga o en la velocidad de respuesta de los servidores.

Al margen de las conclusiones relativas a la validez de la plataforma, hay otras lecciones extraídas de este proyecto quizá no tan evidentes pero no menos importantes:

- Elaborar unos perfiles de uso ajustados a la realidad es un factor crítico para realizar pruebas de rendimiento con garantía de éxito. Aunque se trata de pruebas de laboratorio, es necesario reproducir el escenario de la forma más realista posible.
- Una prueba de rendimiento sólo tiene validez mientras no se produzca ninguna variación en los elementos del sistema. Cualquier cambio requiere de nuevas pruebas para que podamos estar seguros de que la plataforma seguirá comportándose correctamente.
- Es recomendable disponer de un entorno de test estable. Basta con que sea fácilmente comparable al de explotación para que sea posible extrapolar los resultados obtenidos en las pruebas y minimizar la necesidad de recurrir al entorno real, cuya disponibilidad para pruebas suele estar muy restringida.

◆  
En la actualidad  
cada día utiliza la  
plataforma Atenea  
1/3 del total de  
alumnos de la UPC

◆  
Es recomendable  
disponer de un  
entorno de test  
estable

**Marcos Montero Torres**  
(marcos.montero@upcnet.es)

Gestor de Proyectos

**Andrés Navarro-Soto Ballester**  
(andres.navarro-soto@upcnet.es)  
Responsable de Proyectos Tecnológicos

UPCnet (Universidad Politécnica de Cataluña)