

# Creación de Aulas Docentes mediante la Virtualización de Sistemas Informáticos

Javier S. Zurdo  
Universidad Rey Juan Carlos  
C/Tulipán s/n. 28933  
Móstoles - Madrid, España  
javier.zurdo@urjc.es

Alberto Sánchez  
Universidad Rey Juan Carlos  
C/Tulipán s/n. 28933  
Móstoles - Madrid, España  
alberto.sanchez@urjc.es

Miguel S. Zurdo  
Universidad Rey Juan Carlos  
C/Tulipán s/n. 28933  
Móstoles - Madrid, España  
mas.zurdo@gmail.com

## RESUMEN

Este artículo pretende mostrar cómo la implantación de la virtualización a través de un nuevo framework de trabajo propio, ha permitido mejorar la calidad de las prácticas en un aula docente de la Escuela Técnica Superior de Ingeniería Informática (ETSII). Las ventajas de esta manera de trabajar se resumen en la disminución de la burocracia y los procedimientos de administración de sistemas, la reducción de los tiempos de despliegue, los escasos costes económicos y la percepción del incremento de la calidad del servicio.

## Palabras Clave

Virtualización, P2P, clonación, QoS

## 1. INTRODUCCIÓN

La excelencia universitaria suele conseguirse como fruto de la modernización y de la especialización de aquellos aspectos que demanda la sociedad. Para ello, se requiere un fuerte compromiso y esfuerzo de las cuatro partes implicadas en el proceso: Estudiantes, Personal de Administración y Servicios, Personal Docente e Investigador y las entidades externas a la Universidad. El compromiso suele ser lo más fácil de obtener cuando se fijan unos objetivos de mejora determinados. Sin embargo, el esfuerzo es más difícil dado que habitualmente implica un mayor tiempo de dedicación o ayuda económica.

La ETSII tiene como su principal activo la cantidad de jóvenes docentes que capitalizan el esfuerzo humano y el tiempo de dedicación necesario para incrementar la calidad docente con investigaciones punteras. Pero poco se ha hablado de los esfuerzos económicos que son necesarios para la constante renovación de los equipamientos informáticos. Por ello, criterios como la eficacia y eficiencia en la administración de recursos informáticos puede suplir la carencia económica actual. Asumamos las siguientes necesidades que tienen cada uno de los actores de una asignatura, en una carrera universitaria cualquiera:

1. El administrador de recursos asigna aulas a profesores y asignaturas según sus necesidades de software. Evita colisiones y solapamientos. Su objetivo es instalar todo el software posible en todas las aulas docentes.
2. El profesor desea tener preparadas las aulas de prácticas con el software requerido en tiempo y forma.

3. El administrador de sistemas desea tener bajo control todos los ordenadores del parque informático, gestionando las aplicaciones y los permisos de usuario. Evita sobrecargar los ordenadores con software superfluo y respeta el número de licencias contratadas. La estabilidad del aula es su principal objetivo.
4. El alumno desea poder realizar las prácticas en unos ordenadores rápidos y no sobrecargados. En determinados momentos desearía que la aplicación que utiliza pueda estar disponible sin restricción de espacio, licencias o tiempo de acceso.

Manejar este cúmulo de restricciones es tremendamente complicado por lo que se decidió atacar aquel punto que es el cuello de botella y que consume la mayor parte de los recursos económicos y humanos (la gestión y administración de las aulas docentes).

## 2. ANTECEDENTES

La virtualización de sistemas no es un concepto que se haya inventado en estos últimos años. A modo de resumen, la virtualización es una especie de particionado de los recursos para aumentar los ratios de utilización de los mismos. Esta idea ya se utilizaba en 1970 para que los mainframes de gran tamaño fueran utilizados al máximo de sus capacidades. Fue IBM quien empezó a dividir un gran mainframe en un conjunto de ordenadores independientes, buscando conseguir multitarea de un único mainframe.

En la década de los 80-90, con la explosión de los ordenadores personales y el modelo de informática distribuida, el modelo de virtualización comenzó a desaparecer. Era más barato tener un ordenador por persona que comprar un mainframe para dar servicio a todo el personal de una empresa. Ésto, que inicialmente suponía una ventaja económica sustancial, con el paso del tiempo derivó en una redefinición de los problemas de las tecnologías de la información asociadas al coste por computador.

Ejemplos de estos problemas son la baja utilización de los recursos computacionales personales, incremento de los costes de mantenimiento por ordenador, incremento de los costes de la infraestructura de red y servicios, incremento de los costes de formación y de personal, incremento en los costes de riesgos ante desastres o el incremento de tiempo en la resolución de problemas en los escritorios personales.

En los años 90, la empresa VMware [1] empezó a observar que los ordenadores personales incrementaban la potencia de cálculo de manera exponencial, por lo que en el futuro podrían darse problemas de infrautilización de los recursos

(tal y como sucedió en los mainframe). Fue en 1999 cuando mostró su sistema para la virtualización en sistemas x86, como un medio para solucionar gran parte de los problemas anteriormente indicados. Ofrecían un aislamiento completo de las máquinas, la movilidad entre plataformas y la elección de múltiples sistemas operativos para dar soporte a distintas aplicaciones.

En la actualidad, otras empresas como Microsoft, Sun Microsystems/Oracle, Parallels, etc., han desarrollado sus propias soluciones de virtualización. AMD e Intel, viendo el nicho de negocio que suponía, han implementado continuas modificaciones de los juegos de instrucciones para acelerar la ejecución de máquinas virtuales. Resulta curioso que ambos diseñadores de procesadores hayan llegado a definir las mismas funcionalidades con distintas aproximaciones. Esto derivó a que a finales del 2009 se pudiera considerar como un estándar de facto dichas funcionalidades.

El problema fundamental que tienen todas las soluciones de virtualización comerciales es el coste elevado de entrada a esta tecnología. A modo de ejemplo, se solicitó una serie de presupuestos a partners oficiales de VMware para invertir en la virtualización de aulas docentes. El escalón de entrada supone unos €150.000 para los primeros 100 PCs virtualizados y €25.000 por cada 100 equipos adicionales.

El dimensionamiento de la infraestructura se antoja imprescindible para que los servidores puedan asumir un adecuado crecimiento en años sucesivos [2]. Definamos p. ej., un campus tecnológico con alrededor de 1.000 equipos informáticos en los que se quiere virtualizar en torno a 600 PCs. Asumiendo los parámetros de los presupuestos solicitados, el coste rondaría los €275.000. La diferencia entre los equipos reales del campus y los que se realmente se virtualizan se debe a que no todas las aulas son susceptibles de ser migradas a entornos virtuales. Son claros los ejemplos de asignaturas que hacen uso de tarjetas gráficas, FPGAs, dispositivos hápticos, webcams, programación de dispositivos móviles, etc. y que son inviables su correcto desarrollo en entornos virtuales. Si se quisiera virtualizar 1.000 puestos de trabajo, el coste ascendería a unos €400.000.

Dado que la virtualización sí resuelve problemas muy específicos, se pensó en recurrir a soluciones basadas en Open Source que minimizaran drásticamente los costes económicos. Dividir el problema de la implantación de la virtualización dependiendo de si la docencia es on-line o presencial [4], es un ejemplo de ello. Optimizar las aulas docentes, eliminando recursos inútiles y completando la funcionalidad de un aula ya implantada o haciendo que dichas funcionalidades fuesen independientes de la imagen desplegada previamente [3], también disminuye de manera ostentosa los costes económicos.

Por ello, montamos una pequeña aula docente que implementarse las publicaciones mencionadas anteriormente y reutilizase todos aquellos recursos que la Universidad no saca partido (de las renovaciones o de las desinventarizaciones). Además se creó un framework de trabajo que fuese fácil de utilizar por el administrador y que su funcionamiento resultase transparente a ojos del docente y del alumno.

## 3. EL AULA VIRTUALIZADA

### 3.1 Hardware

Actualmente se dispone de un aula docente con unos 20 ordenadores de características heterogéneas. La totalidad

de ellos tienen una antigüedad superior a 6 años, procesador compatible con Pentium 4 a 3GHz, 512 MBytes de RAM, alrededor de 80 GBytes de disco duro, tarjeta gráfica integradas en placa y tarjeta de red 10/100Mbps. Cabe destacar que el rendimiento medio de escritura de los discos duros ronda los 35MBytes/s.

El servidor virtualización es un Dell PowerEdge T310, con extensiones de virtualización compatibles con un modelo baremetal. Es un equipo QuadCore a 2,53GHz, 8 GBytes de RAM y 450 GBytes de disco duro en RAID 5.

Las comunicaciones entre los distintos elementos del aula se realizan a través de un switch Fast Ethernet de 100 Mbps. Teóricamente la velocidad máxima que se puede alcanzar son 12MBytes/s por nodo (obviamente no hemos tenido en cuenta los payload que los protocolos añaden por sus cabeceras y por la propia congestión de la red universitaria).

### 3.2 Software

Todos los equipos del aula tienen instalado una distribución GNU/Ubuntu 10.04 Desktop especialmente diseñada para soportar la virtualización de Virtual Box y VMware, tal y como se refleja en las publicaciones antes comentadas [3] [4]. Además tienen activado el demonio Sshd para la gestión remota de los equipos y modificado el gestor de arranque GDM que permite seleccionar un arranque normal o una de las máquinas virtuales previamente enviadas al equipo.

Al contrario que los clientes, el servidor de servicios de virtualización tiene instalado un sistema operativo baremetal VMware ESXI 4, completamente gratuito desde el año 2009. Sobre este "meta" sistema operativo se instalan los demás sistemas operativos que contendrán el resto de los servicios necesarios (servidor web Apache, SFTP, interfaz de administración de clonaciones, etc.)

### 3.3 Procedimiento de uso

El procedimiento para utilizar el aula es el siguiente:

1. El profesor solicita el aula virtualizada y se comprueba si hay colisión con alguna otra reserva previa.
2. Si hay colisión, se debe buscar un nuevo hueco o negociar con el profesor que previamente la tiene reservada.
3. Si no hay colisión, se entregará una copia de la máquina virtual para que el profesor realice la instalación de su software en ella.
4. Una vez instalado todo el software, se despliega la máquina virtual en el aula mediante P2P (bittorrent).
5. Finalmente se le asigna un slot de arranque en el menú GDM de Linux y se prueba el funcionamiento de manera remota en el aula.

Suele ser muy recomendable agrupar software en máquinas virtuales con características comunes, como puede ser programas de simulación de circuitos y VHDL juntos, compiladores, plugins y aplicaciones de áreas docentes similares. Esto genera máquinas virtuales muy especializadas y optimizadas en su rendimiento. La principal ventaja es que podemos tener instalado en la misma máquina física distintas máquinas virtuales simultáneamente, accediendo a una u otra con un simple cierre apertura de la sesión deseada. Además, si una máquina virtual no funciona bien, se puede clonar de nuevo dado que el equipo posee una copia local

limpia de la máquina, siendo un procedimiento ultra rápido con tiempos de restauración menores a 5 minutos.

### 3.4 Coste

Si bien es cierto que el servidor tuvo que ser adquirido nuevo, su coste económico no superaba los €2.000. El resto de equipamiento ha sido reutilizado de equipos de renovación, equipos desinventariados y donaciones del propio personal docente e investigador.

El proyecto se inició en diciembre del 2009, con unos esbozos de las técnicas a utilizar y de las herramientas comerciales y libres que se querían testear. La parte del proyecto dedicada a la programación e implementación del framework no superó los dos meses de trabajo. El coste económico de la implantación y despliegue fue cero principalmente porque se asumió como carga de trabajo en el Proyecto Final de Carrera titulado "Eficiencia en la Administración de Aulas Docentes: Virtualización de un entorno presencial /on-line".

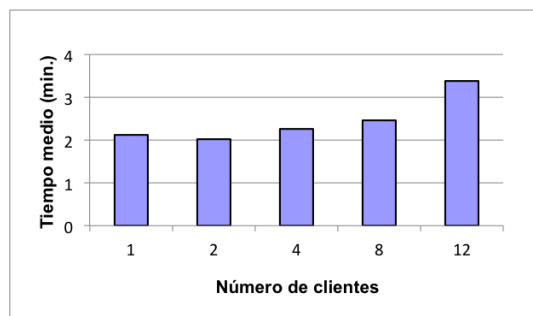
### 3.5 Rendimiento y QoS

Para medir el rendimiento y analizar la calidad de servicio (QoS) de nuestra propuesta, se ha realizado la distribución de las máquinas virtuales en tres escenarios diferentes:

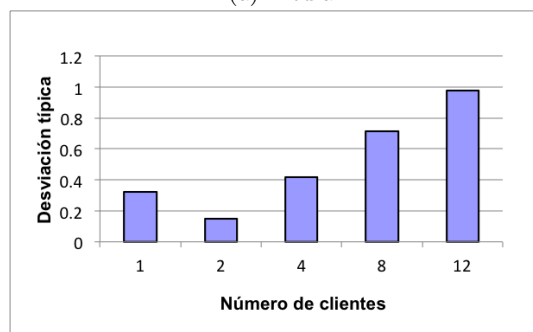
- Fijando el número de semillas a 2 (para provocar que por lo menos una se encuentre dentro del aula donde se quieren realizar las copias) y 512 MB como tamaño de imagen, para medir la influencia del número de dispositivos finales copiando al mismo tiempo.
- Fijando el número de clientes a 8 (debido a los buenos resultados obtenidos) y 512 MB como tamaño de imagen, para medir la influencia del número de semillas iniciales.
- Fijando el número de semillas a 2 y el de clientes a 8, para medir la influencia del tamaño de imagen en el rendimiento.

Cuando hablamos de semilla nos referimos a un cliente con la imagen completa lista para compartir. Para cada escenario y prueba determinada, se han realizado 5 tests espaciados en el tiempo, seleccionando la media y la desviación típica como descriptores representativos de cada prueba. La distribución de las máquinas virtuales se realiza a través de la red de interconexión de área local que posee la Universidad, siendo los test afectados por tráfico real de otros usuarios. En función del experimento concreto realizado, algunos de los ordenadores del aula piloto se utilizarán como semillas, mientras otros harán la función de clientes o dispositivos finales a los cuales hay que distribuir la máquina virtual.

La Figura 1 muestra los resultados obtenidos para una imagen de 512 MB con 2 semillas como servidores. El número de dispositivos finales variará en función de los distintos test realizados (1, 2, 4, 8, 12 equipos clientes). Como se puede observar, al aumentar del número de clientes, el tiempo requerido por cada uno de ellos para obtener la máquina virtual también se incrementa, debido a que más elementos comparten la misma red de interconexión y el acceso a la misma semilla. Sin embargo, el rendimiento global, no se ve penalizado, ya que el ancho de banda agregado conseguido al añadir nuevos clientes aumenta de forma lineal hasta llegar al máximo ancho de banda de la red de interconexión



(a) Media



(b) Desviación típica

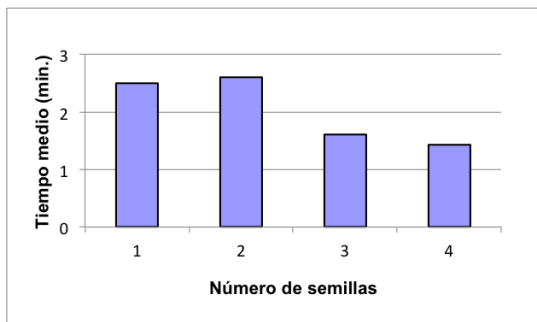
Figura 1: Tiempo de ejecución dependiendo del número de clientes para una imagen de 512 MB con 2 semillas

(en este caso 10 MB/s reales) o por el acceso a disco en las semillas.

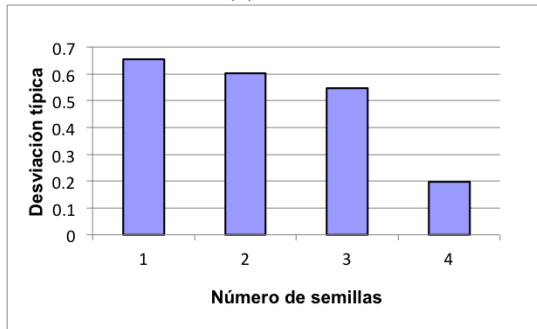
La Figura 2 muestra los resultados obtenidos para una imagen de 512 MB en función del número de semillas iniciales utilizadas, distribuyendolas a 8 dispositivos finales. Como se puede observar, el incremento del número de semillas va acompañado de un descenso en el tiempo necesario para realizar la distribución a cada uno de los clientes, debido a que no se sobrecargan los recursos donde se encuentran almacenadas las semillas.

La Figura 3 muestra los diferentes resultados obtenidos al distribuir imágenes de diferente tamaño en 8 dispositivos finales a partir de 2 semillas iniciales. Como cabría de esperar al aumentar el tamaño se aumenta el tiempo necesario para realizar la distribución en cada cliente. Así el tamaño se convierte en un factor fundamental, ya que cuanto mayor sea, mayor será el tiempo necesario para transmitirlo y más afectará a los accesos a disco de las semillas. Esta mayor saturación provoca que el aumento de tiempo en función del tamaño sea algo mayor que lineal. A pesar de ello, hay que tener en cuenta el potencial y la viabilidad del trabajo desarrollado, ya que posibilita la clonación de una imagen de 4 GB a 8 equipos en menos de 23 minutos, frente a las soluciones tradicionales que requieren muchísimo más tiempo.

Estos resultados muestran como las técnica implementada mejora la distribución de las máquinas virtuales, pero para asegurar la QoS, debemos garantizar una buena adaptación a los cambios en el sistema. En este sentido, es conveniente analizar no sólo el tiempo medio sino también el grado de dispersión de los resultados observados en cada operación de distribución realizadas por cada cliente (básicamente lo cercano que se encuentra respecto al tiempo medio). Las Figuras 1b, 2b y 3b proporcionan un descriptor de esta dispersión.



(a) Media



(b) Desviación típica

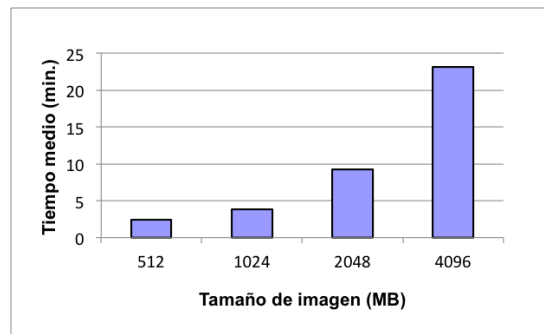
Figura 2: Tiempo de ejecución dependiendo del número de semillas para una imagen de 512 MB con 8 clientes

Como se puede observar, en cualquiera de los casos la dispersión es relativamente pequeña comparado con el valor medio obtenido correspondiente, asegurando la estabilidad de las operaciones realizadas y por tanto incrementando la QoS del servicio de distribución de máquinas virtuales proporcionado. Además, el hecho de que cada cliente se comporte como servidor de sus propias descargas realizadas permite tener un sistema de clonación en alta disponibilidad.

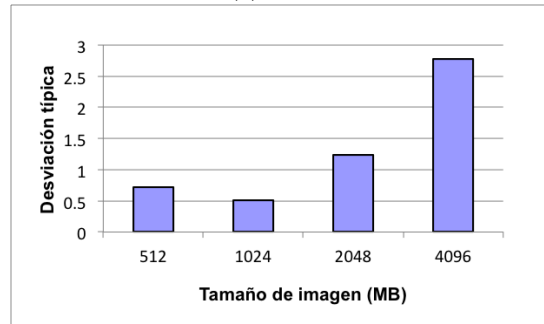
#### 4. CONCLUSIONES

Las principales ventajas de usar un framework de virtualización como éste se resumen en:

- Un interfaz de clonación simple y sencillo.
- Alta disponibilidad tanto en la escalabilidad del sistema en clonaciones masivas como en pérdidas de conectividad de red (no todas las soluciones actuales lo resuelven).
- Dotar al profesor de libertad total para instalar el software necesario en sus clases docentes.
- Dotar al alumno on-line de máquinas virtuales donde realizar las prácticas docentes.
- Dotar de un banco de pruebas sin miedo a que el equipo quede inoperativo.
- Uso de máquinas virtuales especiales para actuar en modo cluster, computación intensiva, grid, ...
- Coste económico irrisorio con tecnología 100% local.
- Compartir con otras Universidades o instituciones docentes los recursos generados entre todas ellas.



(a) Media



(b) Desviación típica

Figura 3: Tiempo de ejecución dependiendo del tamaño de imagen con 2 semillas y 8 clientes

El principal inconveniente de esta solución es el consumo algo más elevado de memoria RAM. Aun así podemos afirmar que en las prácticas del Máster de Informática Gráfica, Juegos y Realidad Virtual necesitamos una máquina virtual con software muy pesado como Autodesk 3DStudio / AutoCAD / Maya y Blender simultáneamente. Dicha máquina funcionó con escasos 512 MBytes de RAM, cumpliendo sobradamente las necesidades solicitadas por los profesores y los alumnos. Aún así, el coste de implantar una virtualización comercial (€200.000) versus nuestro framework con actualizaciones de memoria RAM y disco duro (€18.000), sigue siendo 11 veces más barato y permite mantener parte del cómputo en los nodos locales y no en los servidores.

#### 5. AGRADECIMIENTOS

Queremos agradecer a Enrique Ruiz y a Juan Salvador (PAS informáticos de la URJC), todos aquellos consejos y experiencias que han aportado durante la duración del proyecto.

#### 6. BIBLIOGRAFÍA

- [1] VMware, Inc. <http://www.vmware.com>; accessed January 7, 2011.
- [2] J. Ros. Virtualización Corporativa con VMware. *N cora 2009*. ISBN: 978-84-613-3888-7.
- [3] J. S. Zurdo and M. S. Zurdo. Eficiencia en la administración de aulas docentes. *JITICE'10 I Jornadas en Innovación y TIC Educativas*, 2010.
- [4] M. S. Zurdo and J. S. Zurdo. Virtualización de aulas docentes en un entorno presencial / on-line. *JJI'10 I Jornadas de Jóvenes Investigadores*.