



Comunicación

038

UN SISTEMA DE INFORMACIÓN EN UNA BOTELLA (O CASI): CONSOLIDACIÓN Y VIRTUALIZACIÓN DE SERVIDORES EN EL MEC

José María Becerra González

Jefe de Servicio de Sistemas Corporativos
Ministerio de Educación y Ciencia

Javier García del Río

Jefe de Servicio de Sistemas Informáticos
Ministerio de Educación y Ciencia

Palabras clave

Virtualización, Consolidación, Máquinas virtuales.

Resumen de su Comunicación

El Ministerio de Educación y Ciencia ha llevado a cabo un proyecto de virtualización y consolidación empleando Vmware ESX 2.5.1 Se discute la arquitectura empleada, tanto física como lógica. Se detalla la experiencia obtenida en el proyecto de virtualización, con especial énfasis en aquellos aspectos más interesantes para aquellos organismos que estén planeando llevar a cabo un proyecto similar.

UN SISTEMA DE INFORMACIÓN EN UNA BOTELLA (O CASI): CONSOLIDACIÓN Y VIRTUALIZACIÓN DE SERVIDORES EN EL MEC

1. ¿Qué es la virtualización?

Un concepto realmente interesante, que ha experimentado un notable desarrollo en los últimos años, y que parece que por fin está encontrando sus caminos productivos y no sólo experimentales. Si leemos algunos análisis de carga de servidores, millones en todo el mundo, podemos encontrarnos con estadísticas interesantes de como, y de forma general, tan solo aprovechamos entre un 20-30% de la capacidad de proceso de estos servidores. Es decir, a cualquier servidor a nivel global le sobran el 70% de sus recursos. Si a esto unimos la proliferación de servidores, dada la reducción de precio que han sufrido, nos encontramos con un parque infrautilizado y con dificultades de mantenimiento.

En estos momentos, sin ser exhaustivos, hay varias tecnologías de virtualización en el mercado: Vmware (de EMC, www.vmware.com) y Virtual Server (de Microsoft). Además, podemos encontrar otras tecnologías open source, como es el caso de Xen, que se basa en un modelo algo diferente al empleado por los 2 anteriores: el modelo de hypervision, en el que el núcleo del sistema operativo a virtualizar debe ser adaptado para la virtualización, frente al modelo de las 2 primeras, en las que se sitúa un sistema operativo intermedio que aísla la capa de hardware de los sistemas operativos a virtualizar (que no necesitan ser modificados), ofreciéndole a estos sistemas operativos unos dispositivos virtuales.

2. El proyecto de virtualización del Ministerio de Educación y Ciencia

Dado que en el Ministerio de Educación y Ciencia (MEC) se daba la situación expresada en la introducción, se decidió abordar un proyecto de virtualización y consolidación, que sirviera de piloto y permitiera evaluar si era posible la virtualización de todo (o casi todo) el sistema de información del MEC. Para ello, se decidió escoger como piloto la implantación del sistema de servidores de aplicaciones OAS (Oracle Application Server) necesarios para integrar las aplicaciones que provenían del extinto Ministerio de Ciencia y Tecnología. Inicialmente se adquirieron 4 servidores, certificados por Vmware, con las características de la Tabla 1 (exceptuando que la memoria RAM era de 4 GB). Las máquinas virtuales que se crearon eran Windows 2003 Server Estándar Edition, con OAS 10g desplegado sobre ellas. Estas máquinas necesitaban la suficiente potencia y recursos como para que el piloto fuera significativo. De la implantación realizada, se pudo ver que el punto débil de la instalación era la memoria RAM. Por lo demás, el sistema funcionó perfectamente, y de hecho, se movieron máquinas virtuales desde una máquina física a otra en un par de ocasiones por problemas puntuales en una de las máquinas físicas mediante el uso de Vmotion. Ante estos buenos resultados, se abordó la virtualización de todos los elementos posibles del sistema de información del MEC, para lo cual se amplió la infraestructura disponible.

Arquitectura del sistema

1 - Arquitectura física.

La arquitectura física del sistema se muestra en la Figura 1. El elemento base del sistema de virtualización son ordenadores Dell PowerEdge 1850, con las características indicadas en la Tabla 1.

Tabla 1	
Procesador	2 x Pentium IV Xeon con HT a 3.0 GHz, 1 MB de cache L2
Memoria RAM	8 GB, DDR2, dual rank
Discos duros	2 x 73 GB U320 SCSI
Controladora RAID	Si, 256 MB RAM
Tarjeta HBA	Emulex a 2 Gbps
FSB	800 MHz
Red	2 x 1 Gbps en placa base. En algunos casos, se ha incluido una tarjeta extra con 2 bocas x 1 Gbps
Fuentes de alimentación	2, redundantes
Altura	1 U, enracable
Garantía	NBD

En un principio, se planteó el uso de blades para la instalación, pero dado el alcance inicial del proyecto (sólo para migrar los servicios OAS asignados a la parte de Investigación del Ministerio) se descartó por los aspectos económicos (al ser un número reducido de máquinas físicas las necesarias, no compensaba la adquisición de toda la infraestructura asociada a un sistema blade).

Las características, en cuanto a la conexión a la LAN de los servidores, varían según la zona de seguridad en la que se encuentran:

- en la red interna, las máquinas disponen de sólo 2 interfaces de red, 1 para gestión de Vmware, uso de Vmotion y la realización de las copias de seguridad y otro interfaz para uso de las máquinas virtuales. La comunicación se realiza a 1 Gbps. Esta distribución, como se puede ver en la Figura 2 es suficiente, ya que en ningún momento las máquinas virtuales llegan a saturar el canal de comunicaciones.
- en la intranet administrativa, la única máquina disponible cuenta con 4 interfaces de red, 1 para gestión de Vmware, uso de Vmotion y la realización de las copias de seguridad, y otros 3 para uso de las máquinas virtuales, dado que en esta zona las comunicaciones son a 100 Mbps.
- en la DMZ, las máquinas disponen de 4 interfaces de red, 1 para gestión de Vmware, uso de Vmotion y la realización de las copias de seguridad, y otros 3 para uso de las máquinas virtuales, con el objeto de crear distintas zonas de seguridad. También las comunicaciones en esta zona se realizan a 1 Gbps.

La conexión con la SAN (Storage Area Network) es a través de una tarjeta HBA (Host Bus Adapter) Emulex, a 2 Gbps. Almacenando las máquinas virtuales en la SAN permite emplear Vmotion entre las distintas máquinas físicas, y especialmente, con la máquina de reserva, y tener posibilidades de crecimiento sin los problemas que plantearía el uso de discos internos. Los discos internos de las máquinas físicas se emplean para la instalación de Vmware ESX, y para guardar los ficheros de configuración de las máquinas virtuales. Las máquinas virtuales se almacenan en LUN de discos de fibra, que son los que tienen el mejor rendimiento (hemos considerado, a efectos de rendimiento, las máquinas virtuales como bases de datos). Además, al igual que en la conexión a la LAN, se ha separado el acceso de las máquinas virtuales a las LUN en función de la zona a la que pertenecen. En estos momentos, tenemos asignadas LUNs específicas a cada zona: intranet, intranet administrativa y DMZ. Dentro de cada zona, los servidores físicos tienen acceso a todas las LUNs de su zona.

Además, disponemos de un servidor dedicado en exclusiva a la gestión de Virtual Center (sistema de administración y gestión gráfica y centralizado de máquinas virtuales de Vmware), con las características

indicadas en la Tabla 2.

Procesador	1 x Pentium IV 2.8 GHz
Memoria RAM	2 GB, DDR, ECC
Discos duros	2 x 160 GB SATA, 7200 rpm
Red	2 x 1 Gbps en placa base.
Fuentes de alimentación	2, redundantes
Altura	1 U, enracable

La información de Virtual Center se almacena en un esquema en la base de datos de producción del MEC (Oracle), integrándose en el sistema general.

2 - Arquitectura lógica.

La Figura 3 muestra la arquitectura lógica del sistema. Sobre el núcleo de virtualización que presenta Vmware ESX, podemos desplegar múltiples máquinas virtuales

Zona	Número de máquinas físicas	Número de máquinas virtuales	Espacio en MV	Espacio para snapshots	Funciones
Intranet	6	44	1.500 GB	900 GB	DNS, Servidores de Aplicaciones OAS, Servidores de Aplicaciones Java, Servidores de Bases de datos Oracle, DC, Servidores Web, Servidores de Bases de datos SQL Server, Pasarela de Correo Móvil
Intranet Administrativa	1	4	135 GB	390 GB	Servidores de Aplicaciones, Servidores de Bases de datos SQL
DMZ	2	16	190 GB	176 GB	Servidores de aplicaciones Java, Relays de SMTP, DNS, Servidor Web corporativo
Total	9	64	1.825 GB	1.466 GB	

En la versión actual de Vmware ESX, sólo se pueden crear máquinas virtuales con sistemas operativos Windows y Linux, lo cual encaja perfectamente con la actual filosofía del MEC.

3. Aspectos puntuales de la implementación del sistema

A continuación se comentan algunos aspectos importantes que han ido surgiendo en la instalación del

MEC y qué solución se le ha dado.

1 - Seguridad.

Una funcionalidad muy interesante de Virtual Center es la posibilidad de integrar la seguridad de acceso y gestión a las máquinas virtuales en un dominio de Directorio Activo, cosa que se hizo con el dominio corporativo de Windows 2003. Se han creado grupos de usuarios en el dominio, en los que se incluyen los usuarios que van a gestionar sus máquinas, y estos pueden acceder a través del cliente de Virtual Center, para operaciones físicas sobre las máquinas virtuales que tienen asignadas, o por los medios remotos habituales: telnet ssh, terminal server, etc.

En cada una de las máquinas físicas, el acceso se ha establecido en alto, lo que impide el acceso por medios que no empleen SSH. Esto es especialmente importante para las máquinas situadas en la DMZ.

2 - Copias de seguridad de las máquinas virtuales.

Este es un punto tremendamente importante, dado que la facilidad que aporta Vmware para provisionar nuevas máquinas puede llevarnos a un punto en el que la infraestructura sea vulnerable.

Inicialmente, se realizaron pruebas con la utilidad incluida en Vmware: vmsnap.pl, para la realización de imágenes, y vmres.pl, para su restauración, pero sus funcionalidades eran muy limitadas. Tras realizar una búsqueda en los foros de Vmware (muy recomendables tanto para estar al día de nuevas funcionalidades como para resolver problemas), encontramos referencias a 2 herramientas, principalmente: ESX Ranger (<http://www.vizioncore.com/>) y vmbk (<http://www.vmts.net/vmbk.htm>, actualmente en la versión 2.22.0 rev. 5), la primera de pago y la segunda, gratuita.

Se decidió realizar pruebas con la segunda, porque reunía suficientes funcionalidades para la infraestructura del MEC.

Se trata de un pequeño programa en Perl, que aprovecha la API incluida con Vmware, y que permite realizar imágenes (snapshots) en caliente de las máquinas virtuales, almacenándolas en distintos tipos de almacenamiento (recursos CIFS, NFS o FTP), con uso de sesiones (para diferenciar copias de las máquinas virtuales según necesidades) y que se puede incluir en el gestor de tareas programadas (cron) de la máquina física.

La actual implementación consta de un espacio en un NS-702G (NAS de EMC2), que se comparte como un servidor CIFS, para la zona interna, y con un espacio en ese mismo NS-702-G, compartido como recurso NFS para las zonas de la intranet administrativa y la DMZ.

Como ejemplo de tarea de snapshot, tenemos la siguiente:

```
/usr/local/bin/vmbk.pl -d /vmbackup_impares/ -s smb=//wuria/vmbackup_impares,user=<dominio>\\<usuario>,password=<clave> -l /usr/local/bin/vmbackup_impares_linux.log -c -4 -b linux -e smtp=<servidor smtp>.mec.es,from=<cuenta>@mec.es,to=<cuenta>@mec.es -6 120000
```

Una ventaja añadida de este sistema de backup es que no utiliza más que el espacio realmente ocupado por la máquina virtual, independientemente del tamaño de los discos asignados a la máquina virtual. Así, por ejemplo, en la intranet, el backup de las máquinas virtuales se realiza sobre un espacio de 780 GB, en el que se albergan 2 copias alternas de las máquinas virtuales.

La restauración de las máquinas virtuales con esta herramienta es muy sencilla, ya que al generarse la

imagen genera también un script de recuperación que va guiando en el proceso de recuperación, pudiéndose usar incluso para generar una nueva máquina a partir de una ya existente.

También se han creado réplicas de las máquinas virtuales en la máquina de reserva, apuntando a los discos de las máquinas virtuales operativas, con el objeto de que si cualquiera de las máquinas físicas que albergan máquinas virtuales dejara de funcionar sin dar tiempo a realizar un movimiento mediante Vmotion, podamos arrancar las máquinas virtuales en la máquina de reserva gracias a que los datos están albergados en el almacenamiento centralizado.

3 - Plantillas y discos de instalación de sistemas operativos.

Una funcionalidad muy interesante de Virtual Center es la generación de plantillas a partir de máquinas virtuales ya existentes, lo que permite disminuir aún más el tiempo de despliegue de nuevas máquinas virtuales. En el caso del MEC se dispone de las siguientes plantillas:

- para Windows 2003 server, con un disco de 8 GB para el sistema operativo. Esta plantilla genera una máquina que no está unida al dominio, pero que si tiene todos los parches importantes, evitando así desplegar máquinas que puedan suponer peligros de seguridad.

- para Red Hat Enterprise Linux, de las que se dispone de varias. Una para bases de datos Oracle, que además del sistema operativo con todos los paquetes y parámetros del sistema ya ajustados, incluye el software de gestión de base de datos Oracle 9.2.0.6, con lo que sólo sería necesario añadir discos para las instancias en el momento necesario.

- también se dispone de plantillas para el servicio de comunicaciones e internet, que incluye los servidores de aplicaciones estándar del MEC: Tomcat y JBOSS, para las aplicaciones Java, y Oracle Application Server, para las aplicaciones de gestión internas.

Todas estas plantillas se guardan en el servidor de Virtual Center, permitiendo que estén accesibles desde todas las máquinas físicas.

También se guardan imágenes ISO de los principales sistemas operativos con los que se trabaja en cada uno de los servidores físicos, con el objeto de tener disponibles paquetes o programas en caso de necesitar actualizar cualquier máquina virtual.

4 - Bases de datos virtuales

La experiencia con bases de datos ha sido algo limitada, debido a que la versión de Vmware actualmente disponible en el mercado (2.5.2) tiene algunas restricciones que dificultan su uso en este campo: no soporta todavía sistemas operativos de 64 bits para virtualizar; no se pueden asignar más de 2 procesadores a una máquina virtual y tiene limitaciones en cuanto al total de memoria RAM que se puede asignar a una máquina virtual dada (3600 MB). Ambas limitaciones son importantes en el campo de las bases de datos de producción corporativas.

Teniendo estas restricciones presentes en la versión actual, se inició la virtualización de 2 instancias de uso interno, bajo Linux (Red Hat) y con la versión 9.2.0.6 de Oracle. Aprovechando la facilidad de creación de plantillas de Vmware, tras instalar y configurar el sistema operativo y el software de gestión de base de datos, se creó una plantilla, que permitirá a posteriori generar una máquina virtual, y sólo tener que añadir discos virtuales. A este respecto, Vmware no establece diferencias entre el uso de discos virtuales o discos raw (discos físicos asignados directamente a una máquina virtual) en cuanto a rendimiento.

5 - Almacenamiento

El almacenamiento es una de las claves del sistema virtual por lo que se decidió implementarlo en el almacenamiento corporativo del MEC, constituido por un CX-500 de EMC, con DAEs con discos de fibra (para las máquinas virtuales) y discos SATA (para los backups). La implementación del sistema virtual ha supuesto un tremendo esfuerzo en el sistema de almacenamiento, ya que la facilidad de despliegue de nuevas máquinas tiene la contrapartida de que hay que provisionar una gran cantidad de espacio (ver tabla 3), tanto para disco asignado a máquinas virtuales como disco para los snapshots de éstas. Por otro lado, el uso de una SAN (ya que se disponía de una) facilita todos los temas de Disaster Recovery, al permitir compartir LUN entre máquinas físicas para realizar movimientos de máquinas virtuales entre ellas usando Vmotion, capacidad de crecimiento (en teoría) sin límite, uso de herramientas de copia de la misma SAN para poder llevar los snapshots a otras ubicaciones, etc.

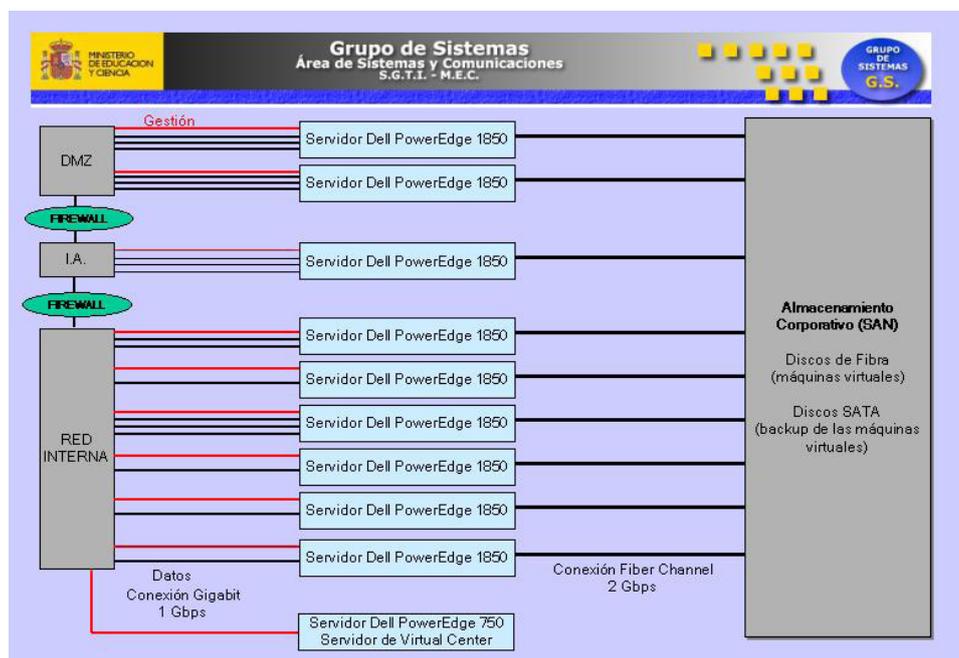


Figura 1



Figura 2

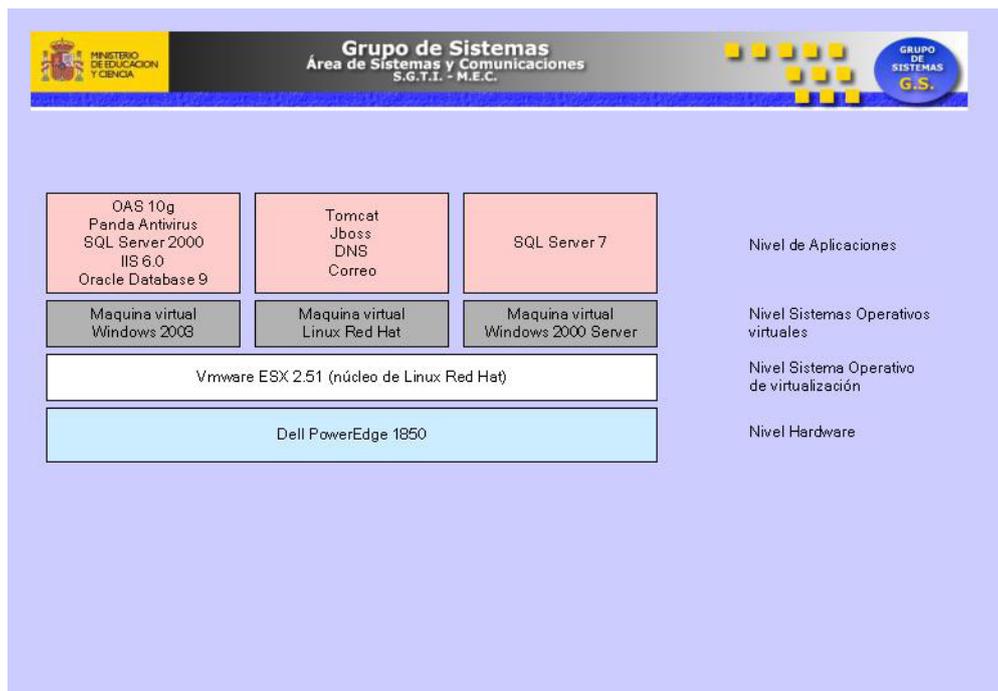


Figura 3

CONCLUSIONES Y FUTURO

La virtualización de los servidores del MEC ha aportado al sistema una elevada flexibilidad (ha permitido ajustar los recursos a las necesidades reales), elevada disponibilidad (se pueden replicar máquinas activas para dar servicios de alta disponibilidad con un tiempo de recuperación prácticamente despreciable ante pérdida o desastre en lugar que ir al procedimiento tradicional de backup/restore mucho más complejo y extenso en el tiempo; además, el uso de Vmotion permite mover máquinas virtuales entre máquinas físicas en caso de cualquier problema con la máquina física que las alberga, en caliente), disminución del parque de máquinas físicas y rapidez en la provisión de servidores a las unidades peticionaria así como una importante reducción del gasto. Ha quedado fuera de este proceso de virtualización toda la plataforma de bases de datos corporativas, por los problemas reseñados para la versión actual de Vmware: falta de soporte de software de 64 bits, y la imposibilidad de asignación de más de 2 procesadores por máquina virtual. El resto de servidores se encuentran migrados o en proceso de virtualizar.

Entre los temas pendientes, y que estamos en estos momentos abordando para mejorar la infraestructura, se encuentran:

- Redundar las bocas de red de las máquinas físicas, con el objeto de aumentar la disponibilidad de los servicios implantados, así como las conexiones de fibra con el CX-500.
- Estudio del uso de GFS (Global File System) para que las máquinas virtuales con sistema operativo Linux puedan compartir discos de la SAN en modo raw.
- Pruebas con Xen, como alternativa para máquinas virtuales con sistema operativo Linux.
- Hacer pruebas con discos SATA para almacenar las máquinas virtuales, en aquellos casos en que estas máquinas no necesitan de un alto rendimiento.

- Integración de los snapshots de las máquinas virtuales en el sistema de copia corporativo, para tener implementado un sistema de Disaster Recovery global.

- En cuanto esté disponible la versión 3.0 de VMware, iniciaremos pruebas con objeto de migrar las bases de datos corporativas a un sistema virtual

ENLACES DE INTERÉS

VMware:

<http://www.vmware.com/>

Virtual Server 2005 de Microsoft: <http://www.microsoft.com/windowsserversystem/virtualserver/default.aspx>

Xen Hypervision System:

<http://www.cl.cam.ac.uk/Research/SRG/netos/xen/>

Vmbk:

<http://www.vmts.net/vmbk.htm>

ESX Ranger:

<http://www.vizioncore.com/>