UNIVERSITAT POLITÈCNICA DE VALÈNCIA

etsinf
Escola Tècnica Superior d'Enginyeria Informàtica

Escola Tècnica Superior d'Enginyeria Informàtica
Universitat Politècnica de València

Estudio del impacto de la virtualización en una empresa de monitorización informática

TRABAJO FIN DE GRADO
Grado en Ingeniería Informática

Autor: Guillermo Ibáñez Galiana
Tutor: Xavier Molero Prieto
Tutor empresa: Miguel Ángel Sáez Soro

28 de agosto de 2015

Resumen

Con el fin de dar una visión práctica y aplicable a este estudio, se va a presentar una infraestructura real que hace uso de la virtualización para ofrecer servicios. Esta infraestructura es propiedad de una empresa valenciana que dedica sus esfuerzos a la seguridad de infraestructuras informáticas mediante la monitorización constante de equipos y redes. Aunque no es el objetivo de este trabajo se explicará brevemente en qué consiste esta monitorización para poder relacionarla y usarla como uno de los argumentos que llevó a esta empresa a virtualizar su infraestructura.

Para conseguir resultados óptimos es necesario realizar un estudio que sopese qué tipo de servicios va a soportar el sistema, su disponibilidad, su eficiencia energética y el presupuesto, es así como se decidió migrar la infraestructura. Durante el proyecto se intentarán argumentar todos los aspectos que se tuvieron en cuenta para tomar esta decisión.

Otra parte importante del estudio será el cálculo del PUE para el CPD propiedad de la empresa. El PUE da una idea de la cantidad de energía que se puede ahorrar en un centro de datos bien planificado y basado en la virtualización. Este ahorro es muy relevante para una PYME para la cual el consumo energético supone un gasto importante. El cálculo del PUE es bastante complejo desde el punto de vista técnico por motivos que se explicarán en detalle y ha sido posible gracias a que uno de los principales objetivos de la empresa es la monitorización exhaustiva de todos los elementos de la infraestructura. Por último se presenta un experimento a pequeña escala diseñado para dar una idea de la mejora que supone virtualizar una serie de servicios sobre una infraestructura adecuada sobre un modelo más tradicional.

Palabras clave: monitorización, virtualización, consumo, centro de datos.

Resum

Amb la finalitat de donar una visió pràctica i aplicable a aquest estudi, es va a presentar una infraestructura real que fa us de la virtualització per oferir servicis. Aquesta infraestructura es propietat d'una empresa valenciana que dedica els seus esforços a la seguretat d?infraestructures informàtiques mitjançant la monitorització constant d'equips i xarxes. Encara que no es l'objectiu d'aquest treball, s'explicarà breument en què consisteix aquesta monitorització per a poder relacionar-la i utilitzar-la com un dels arguments que va donar lloc a aquesta empresa a virtualitzar la seua infraestructura.

Per aconseguir resultats òptims es necessari realitzar un estudi que sospesi quin tipus de servicis va a suportar el sistema, la seua disponibilitat, la seua eficiència energètica i el pressupost, es així com es va decidir migrar la infraestructura. Durant el projecte s'intentaran argumentar tots els aspectes que es van tenir en compte per prendre aquesta decisió.

Un altra part important de l'estudi serà el càlcul del PUE per al CPD propietat de l'empresa. El PUE dona una idea de la quantitat d?energia que es por estalviar en un centre de dades ben planificat i basat en la virtualització. Aquest estalvi es molt rellevant per a una PYME per a la que el consum energètic suposa una despesa important. El càlcul del PUE es bastant complex des del punt de vista tècnic per motius que s?explicaran en detall i ha sigut possible gràcies a que un del principals objectius de l?empresa es la monitorització exhaustiva de tots els elements de la infraestructura. Per últim es presenta un experiment a xicoteta escala dissenyat per donar una idea de millora que suposa virtualitzar una sèrie de servicis sobre una infraestructura adequada sobre un model més tradicional.

Paraules clau: monitorització, virtualització, consum, centre de dades.

Abstract

This work presents a real infrastructure witch uses virtualization to offer services from a practical point of view. This infrastructure is property of a Valencian company focused on cyber-security through constant monitorization of hosts and networks. Although it is not its main objective, this work explains how this monitorization is done in order to relate it and use it as one of the arguments that lead the company to virtualize all of its infrastructure.

To achive optimal results it is important to plan a study than takes into account what services the system is going to run, its availability, its energetic efficiency and the budget, and with this it was decided to migrate the infrastructure. Throught this project it is intended to explain every aspect taken into account to take this decision.

Another important part of this study will be the calculus of the PUE ratio for the data center property of the company. This ratio is a good indicative on how much a well planned virtualized infrastructure is able to save in electrical power. This save is quite important for a small to medium company where the expenses related to energy consumption are very relevant. This calculus is quite complex because of reasons that will be explained later on and has been possible thanks to the constant monitorization of the infrastructure. Finally a little experiment has been designed to give an idea on the advantages of virtualization over more traditional solutions.

Keywords: monitorization, virtualization, consumption, data center.

Índice general

Índice de figuras

Índice de tablas

CAPÍTULO 1

Motivación y objetivos

Este primer capítulo consta de las motivaciones que llevaron al desarrollo de este trabajo y los objetivos que se pretenden conseguir.

1.1 Motivación

Desde que el término virtualización apareciese alrededor de 1960 ha ido evolucionando y ampliándose dando cabida a multitud de tecnologías y metodologías. El objetivo de este trabajo no es tanto el explicar en qué consiste la virtualización o sus tipos, como centrarse en un modelo práctico y aplicable a una infraestructura diseñada para ser versátil, eficiente y con las miras puestas en la alta disponibilidad (HA, *High Availability*).

En los últimos años la virtualización se ha convertido en una tendencia al alza a la hora de diseñar un sistema informático medio o grande. Las herramientas o infraestructuras que la soportan son muy variadas y su elección es un punto clave para aprovechar al máximo la potencia del hardware sobre el que se soporta este sistema. Además de mencionar y explicar estas mejoras se ha preparado un experimento para poder realizar una cuantificación de lo que suponen frente a una infraestructura que no utilice virtualización.

Esta búsqueda de la eficiencia tiene dos objetivos principales y de igual importancia; uno de ellos, quizá el más obvio, es el de abaratar costes ,y el segundo reducir las emisiones de gases contaminantes derivados de la producción de electricidad. Esto último se está convirtiendo cada vez más en una preocupación que influye, no solo a la hora de diseñar un sistema, sino incluso a la hora de diseñar la electrónica que gobierna cualquier computadora. Como resultado de esta preocupación se acuñó el término de *Green Computing* o Computación Ecológica, centrada en el desarrollo de la potencia de calculo con el menor consumo posi-

ble. En noviembre de 2007 se anunció que se crearía una lista[1] de los 500 mejores centros de datos desde el punto de vista energético.

Pero estas dos ventajas no son las únicas bazas a favor de la virtualización. A lo largo del proyecto se enumerarán y explicarán los argumentos que llevaron a la empresa propietaria del CPD del estudio a migrar su infraestructura a una virtualizada.

1.2 Objetivos

Este trabajo tiene como objetivo principal analizar el impacto de basar una infraestructura en una solución de virtualización. Para ello se pretende:

1. Explicar en qué consiste el sistema que se va a basar en la virtualización.
2. Presentar las ventajas que se considera se obtienen de la virtualización incluyendo dos ejemplos prácticos.
3. Plantear las dificultades encontradas en el sistema que forzaron el cambio a la virtualización.
4. Exponer las soluciones que se llevaron a cabo.
5. Proponer alguna mejora para el sistema

1.3 Estructura de la memoria

Este trabajo se estructura en seis capítulos que abarcan desde lo más teórico de la virtualización, pasando por la infraestructura en la que se basa el estudio hasta un caso práctico de cálculo y un experimento a pequeña escala. Los capítulos son los siguientes:

1. Capítulo 1: Consta de una introducción al trabajo y expone las motivaciones y los objetivos del mismo.
2. Capítulo 2: En este capítulo se realiza una descripción del sistema informático que se va a usar para el estudio a fin de poder entender mejor qué es lo que se pretendía obtener al basar el sistema en la virtualización.
3. Capítulo 3: Para poder comprender por qué ha influido tan profundamente el uso de la virtualización en una empresa de monitorización informática

[1] `http://www.green500.org/about`

se va a explicar en qué consiste esta monitorización y por qué ha sido relevante, además para poder realizar el trabajo sobre todo en los capítulos de contenido más práctico.

4. Capítulo 4: Como aproximación más práctica en este capítulo se realiza el cálculo del PUE en un experimento que ha durado un año, durante el cual se han recogido muestras de consumo diarias y se han corregido y utilizado para este cálculo. Con el valor del PUE para el CPD del estudio se pretende realizar una comparación con los valores más comunes en otros centros de datos.

5. Capítulo 5: Además del uso del PUE para realizar una comparación orientativa en este capítulo se presenta un pequeño experimento que comparará los consumos de dos máquinas que proveen los mismos servicios, una de ellas usando virtualización y la otra no.

6. Capítulo 6: Se presentan las conclusiones extraídas del trabajo y analiza si se han cumplido los objetivos propuestos. Se plantean mejoras a futuro para la infraestructura.

1.4 Manejo de la bibliografía

Para la realización del trabajo se han consultado ls siguientes referencias bibliográficas:

- Para profundizar en los requisitos del servicio de DNS se ha usado la referencia bibliográfica [1].
- Para introducir la importancia del ahorro energético y para el cálculo del PUE se ha usado la referencia bibliográfica [2]. De ella se ha extraído información al respecto de qué significado tiene la medida y de como conseguir realizar el cálculo lo más preciso posible.
- En el trabajo se explica brevemente el servidor de correo postfix, para lo cual se ha consultado la referencia bibliográfica [3].
- Para que el trabajo fuese posible se desarrollaron algunos agentes para Nagios en SNMP. Para ello y para comprender mejor el funcionamiento del protocolo se ha usado la referencia bibliográfica [4].
- Como se menciona en el trabajo uno de los objetivos que se consiguen mediante la virtualización es la reducción de la huella ecológica. En la referencia bibliográfica [5] se expone como puede realizarse el cálculo y lo importante que es para el medio ambiente mantener a raya el impacto sobre el medio ambiente.

- Teniendo en cuenta el valor del PUE calculado para el centro de datos del estudio y para poder compararlo con los valores para otros CPD se ha utilizado la referencia bibliográfica [6].
- El trabajo en general gira al rededor de la virtualización y una parte imprescindible de esta es el hipervisor. De la [7] se extraen conocimientos de su funcionamiento general y del papel que juega en el conjunto de la virtualización.
- Durante el trabajo se nombre el *multipath* como una técnica para mejorar el rendimiento y la tolerancia a fallos del sistema. La información necesaria se consultó en parte en la referencia bibliográfica [8].
- Toda la información de Nagios que se utiliza en el trabajo se ha obtenido de la experiencia de su uso y de la referencia bibliográfica [9], que se constituye como una de las mejores fuentes en el mercado para aprender sobre la plataforma.
- La distribución de Linux CentOS se nombra como la elección corporativa para el sistema operativo, la información necesaria se obtuvo de la experiencia trabajando con el sistema operativo y de la referencia bibliográfica [10].
- Para realizar alguna de las propuestas de mejora sobre el CPD de la empresa se hace referencia a la referencia bibliográfica [11] que trata de algunas buenas prácticas para el diseño y construcción de centros de datos.
- Sobre las características de la virtualización a nivel corporativo usando vmWare se ha utilizado, además de la experiencia adquirida de su uso, la referencia bibliográfica [12].
- Al trabajar sobre el sistema operativo Linux se han mencionado ventajas derivadas de su uso, alguna de las cuales se han completado usando la referencia bibliográfica [13].
- Para explicar que es la técnica usada en RAID se ha referido a la referencia bibliográfica [14].
- Una de las ventajas derivadas del uso de la virtualización junto con Linux que se han nombrado en el trabajo es la alta disponibilidad, para conocer mejor qué es y como aplicarla se ha referido a la referencia bibliográfica [15].

1.5 Acrónimos y términos utilizados

A continuación se incluyen una serie de acrónimos y términos utilizados a lo largo del trabajo.

- *Blade*: Los servidores de tipo *blade* son un formato de servidor específico que se utiliza a lo largo del trabajo y que se explicará más adelante. Para abreviar, a este tipo de servidor se le llamará simplemente blade de ahora en adelante.
- *Host*: la palabra host puede traducirse como anfitrión y se utiliza en situaciones muy variadas. Durante el trabajo se utilizará para aquellos casos en los que una máquina sirva de soporte a un producto, dispositivo o a otra máquina; como podría ser una máquina virtual.
- *Rack*: aunque se explica más adelante, y en algunos casos se usa la traducción como armario de IT, durante el trabajo se nombrará este elemento como simplemente rack.
- DMZ: Zona desmilitarizada (*Demilitarized Zone*), es la zona de la red que queda fuera de la protección del cortafuegos y expuesta a Internet.
- *Proxy*: De forma general un proxy es un dispositivo que se utiliza como intermediario para conexiones. Durante el trabajo se menciona como un dispositivo que filtra todas las conexiones entrantes y salientes hacia unas determinadas máquinas que se sitúan en la red interna. De esta forma esas máquinas quedan protegidas detrás del cortafuegos pero siguen siendo accesibles.
- *Hotplug* y *hotadd*: Son los términos que usa mvWare para referirse al añadir CPU o memoria RAM en caliente a una máquina virtual. Aunque podrían tener traducción se a preferido dejarlo así para ser más fiel a la tecnología que se usa en el trabajo.
- ESX: para poder asignar máquinas virtuales a un servidor físico se instala en él un sistema operativo propio de vmWare llamado esxi. Más adelante se explica más en detalle este sistema. A lo largo del trabajo se menciona en un par de ocasiones el término esx. Este termino se refiere simplemente a un servidor físico del tipo que sea sobre el que se ha instalado el sistema operativo esxi y al que por lo tanto se le pueden asignar máquinas virtuales.
- iLO: Hace referencia a una tecnología de HP llamada *Integrated Lights Out* por la cual se gestionan las máquinas físicas a través de la red sin necesidad de que haya ningún sistema operativo instalado en el servidor. Esta tipo de solución existe en todos los fabricantes modernos de hardware, en este trabajo se nombra la solución de HP porque es la marca de hardware más utilizada en la empresa.
- RAID [14]: *Redundant Array of Independent Disks* es una técnica utilizada para mejorar el rendimiento o crear discos tolerantes a fallos basada en utilizar varios discos duros independientes como si fuesen uno solo usando diferentes técnicas. Durante el trabajo se mencionan los niveles de RAID 0, 1 y 5, que se explican a continuación:

- RAID 0: Consiste en crear un solo disco usando varios discos para que el resultante tenga la capacidad sumada de todos los discos que lo forman. Mejora el rendimiento pero no es tolerante a fallos. El número mínimo de discos es de dos.
- RAID 1: o RAID en espejo. Consiste en usar un número par de discos para crear copias exactas de uno de ellos. En su forma más básica usa dos discos iguales que contendrán la misma información. Usando el mínimo de dos discos es tolerante a la pérdida de uno de ellos.
- RAID 5: o distribuido con paridad. Utiliza un mínimo de tres discos de los cuales usa la capacidad de uno de ellos para información de paridad de los datos contenidos en los otros dos. Esta información además es repartida entre el conjunto total de discos. La pérdida de discos en un RAID 5 siempre es equivalente a la capacidad de uno de ellos. Es tolerante a la pérdida de uno de los discos, al igual que el RAID 1 con la ventaja de no estar limitado a solo dos discos y tener una perdida menor. Además mejora ligeramente el rendimiento puesto que los datos están repartidos entre el total de los discos.

CAPÍTULO 2

Descripción del sistema

En este capítulo vamos a tratar de describir el conjunto del sistema informático partiendo de los servicios hasta llegar al hardware. Seguiremos un esquema descendente puesto que la base para las diferentes elecciones de diseño es tener claro qué servicios va a ofrecer teniendo en cuenta las limitaciones de presupuesto. Básicamente se va a dividir el sistema en cuatro partes principales: servicios, sistema operativo, plataforma de virtualización y hardware, y se explicarán las relaciones entre estos elementos como argumento para decidir sobre ellos.

2.1 Servicios

El primer paso para poder diseñar correctamente la infraestructura que soporte el sistema es considerar los servicios que va a alojar. Si bien es posible que a lo largo del tiempo surjan nuevos servicios que no se tuvieron en cuenta inicialmente, es importante tener claros cuáles serán los principales o los que más consuman para no evaluar a la baja el aprovisionamiento general del sistema. En este caso de estudio la empresa se dedica a la seguridad informática mediante la monitorización activa y pasiva de la red, de los hosts y los servicios. Más adelante se explicará cómo funciona esta infraestructura de monitorización a fin de entender mejor el conjunto.

Inicialmente los servicios que se plantearon fueron el conjunto básico necesario, podríamos decir que genérico, para que la empresa funcione:

- Servidores de DNS internos y externos.
- Servidor de correo interno, en este caso Lotus Notes.
- Servidor para compartir ficheros pesados.
- Servidor de documentación.

- Servidores para alojar los propios productos de la empresa implantados en la infraestructura.
- Servidores de desarrollo.
- Cortafuegos

2.1.1. DNS *(Domain Name Server)*

Para la resolución de nombres la empresa consta de 4 servidores. En realidad son 2 principales, uno externo y otro interno, los otros dos duplican la configuración automáticamente de sus superiores. Aquí se establece el primer mínimo en cuanto a hosts físicos que debe haber. Los 4 servidores van a estar separados en 4 máquinas virtuales distintas en 4 hosts físicos distintos. El motivo es que los DNS externos no pueden compartir host con los internos puesto que los externos estarán en DMZ y los internos no. Además por otra parte no tiene sentido duplicar un servicio para hacerlo tolerante a fallos pero hacer que dependan del mismo host físico, por lo tanto tenemos la necesidad de tener un mínimo de 4 máquinas solo para los DNS.

Aunque hoy en día la resolución DNS es muy rápida y prácticamente inapreciable a medida que la carga aumentaba estos servicios empezaron a consumir bastánte capacidad de computo por lo que se tuvo que asignar a los 2 principales una buena capacidad puesto que sin ellos la empresa se queda sin dar servicio y sin navegación.

2.1.2. Correo electrónico

Para el correo corporativo se optó por la solución de IBM Lotus Notes. Esta solución consta de un servidor principal que corre la aplicación, una máquina para el MTA (*Mail Transport Agent*) y de los clientes que se conectarán. La máquina que alberga la aplicación en sí es considerada crítica y está duplicada. Al igual que en los DNS, estas dos máquinas no pueden compartir host y también se intentó que no compartiesen host con otros servicios críticos como los DNS. Además estas máquinas solo podían alojarse en hosts que no estuviesen expuestos al exterior, de ahí la existencia de otra máquina para el MTA.

La máquina que opera el agente de transporte de email o MTA es una máquina situada en DMZ y también tiene un duplicado, generalmente apagado, que solo se usa si la máquina o el host tuviesen algún problema. Esta máquina hace el papel de *relay* (traducido como mediador o transportista) y solo se encarga de reenviar todo el correo que entra y sale de la empresa. De esta forma se evita exponer al exterior la máquina con la aplicación de correo.

2.1.3. OwnCloud

Ante la necesidad de compartir ficheros con clientes o con empleados fuera de la oficina se implantó un servicio de compartición de ficheros en la nube conocido como OwnCloud. Esta máquina aunque no es considerada crítica, tiene dos requisitos importantes a tener en cuenta: debe estar expuesta al exterior o detrás de un proxy y debe estar aprovisionada con suficiente almacenamiento. Actualmente la máquina está alojada en DMZ pero hay planes de llevarla a servidores internos y crear una máquina que haga de proxy en DMZ.

Poco a poco, con el uso del servicio apareció un problema con el que no se contó inicialmente. Todo el contenido que se almacena en la máquina va cifrado y puesto que los archivos que se suelen subir son bastante pesados el proceso de cifrado es costoso computacionalmente. Las características con las que se aprovisionó la máquina inicialmente pronto se quedaron cortas y hubo que asignarle más procesadores, cosa que se pudo hacer en caliente sin pérdida de servicio gracias a la virtualización.

2.1.4. Servicio documental

Uno de los activos más importantes para una empresa es la inteligencia que se va acumulando a medida que el negocio crece. Esta inteligencia consta de un producto desarrollado, documentación técnica o empresarial, documentos para la resolución de problemas, datos fiscales, proyectos, documentación derivada de I+D+I, conocimiento adquirido sobre clientes, etc. No hace falta decir que esta información es increíblemente valiosa para la empresa y sin ella no podría funcionar. Además, a lo largo de los años el volumen de datos es bastante importante y asegurar un almacenamiento suficiente a la vez de un acceso seguro es un reto para los ingenieros de sistemas.

Este servicio consta de dos partes: un servidor de almacenamiento fuertemente protegido, con unos recursos bastante altos y considerada la máquina más crítica de la empresa, y un servicio de gestión documental.

2.1.5. Producto empresarial

Además de la consultoría y servicios de seguridad la empresa posee un paquete de productos de desarrollo propio que se integran entre sí para ofrecer una gestión más cómoda de la infraestructura. Estas máquinas son consideradas críticas puesto que toda la empresa usa sus servicios y alguna de ellas requiere de un aprovisionamiento de recursos bastante alto. Como se comentará más adelante cuando se explique en qué consiste la monitorización, dentro de esta infraestructura alguno de estos servicios requiere de recursos algo más especiales que un

simple aprovisionamiento de CPU, RAM o disco. Estos factores se tuvieron que tener en cuenta para decidir sobre el hardware que se pretendía utilizar.

2.1.6. Desarrollo del producto

Para el desarrollo del producto empresarial se definieron 3 fases: desarrollo, preproducción y producción. Esto afecta directamente a cómo se distribuyen las máquinas virtuales sobre los hosts físicos puesto que los entornos de desarrollo y preproducción deben estar separados del de producción. Hay que tener en cuenta que en algunas ocasiones la validación del producto exige que la máquina de desarrollo o preproducción disponga de los mismos recursos que la de producción, por ejemplo, para pruebas de estrés, y en ocasiones esto puede ser difícil de conseguir si el producto es muy demandante y es otro factor a tener en cuenta a la hora de distribuir las máquinas y de elegir el hardware.

2.1.7. Cortafuegos

Aunque este punto es obvio puesto que toda empresa debería disponer de un cortafuegos, se ha considerado importante su mención puesto que es una de las pocas máquinas que no son virtuales debido a limitaciones físicas de la virtualización.

2.2 Sistema operativo

Un punto muy importante a la hora de elegir el sistema operativo es considerar los servicios que va a ofrecer y asegurarse de que existe una solución compatible. En el caso que estamos tratando se optó por la utilización de CentOS [10] como sistema operativo. CentOS es un derivado de la distribución Linux [13] Red Hat Enterprise Linux (RHEL[1]) compilado por la comunidad a partir del código liberado por Red Hat. Así pues es un sistema de código abierto que pretende ofrecer un producto empresarial gratuito. Puede sonar extraño nombrar un solo sistema operativo y es porque se intenta que todas las máquinas partan de una plantilla muy básica hecha a medida que se personalizará para cada uno de los servicios. Esto significa ,como ya hemos visto antes, que prácticamente todos los servicios se ofrecerán desde una máquina virtual distinta. Más adelante se explicará el porqué de esta unificación a un único sistema operativo y de separar todos (o casi todos) los servicios.

La elección de CentOS como sistema operativo para casi la totalidad de las máquinas fue debida al mantenimiento. Desde que se migraron todos los siste-

[1] `https://www.centos.org/about/`

mas a máquinas virtuales se intentó ahorrar en mantenimiento, de forma que un pequeño equipo de 5 personas pueda mantener, actualizar o ampliar la infraestructura que consta aproximadamente de 24 máquinas físicas y más de 150 máquinas virtuales. Esta unificación a un solo sistema operativo facilita determinadas tareas como las actualizaciones de seguridad de los SO o de las aplicaciones que corren en ellos. Por otro lado el equipo humano que se dedica a este mantenimiento encuentra más fácil y cómodo centrarse en un solo SO lo que les permite perfeccionar sus habilidades y conseguir que su tiempo de respuesta frente a cualquier emergencia disminuya considerablemente al conocer con exactitud que versiones del SO corren las máquinas. Actualmente se ha creado un proyecto interno con la intención de mantener un repositorio propio para controlar el versionado de todo lo que se instale en los operativos y poder realizar actualizaciones en masa utilizando alguna aplicación del tipo del Red Hat Satellite[2].

Por otra parte la adopción de CentOS facilitó las tareas de investigación que se llevan a cabo cuando se pretende instalar un servicio nuevo o cuando surge alguna emergencia gracias a la enorme comunidad que soporta el desarrollo de este sistema operativo, que recordemos es una versión que parte del código de RHEL. Además gracias a que deriva de RHEL se podría incluso optar por contratar el soporte directo de Red Hat para los problemas más complejos. Al equipo que realiza todas estas tareas se le ofrece por parte de la empresa la posibilidad de certificarse con Red Hat de forma gratuita, superando los cursos y exámenes necesarios para obtener estas importante certificaciones[3]. En el aspecto técnico CentOS ofrece una serie de capacidades que se consideran muy importantes a la hora de instalarlo sobre una máquina virtual:

- Soporte nativo para LVM[4].
- Soporte para *hotplug/hotadd* de núcleos y RAM.
- Soporte para añadir hardware en caliente.

2.2.1. LVM *(Logical Volume Manager)*

LVM es una implementación de una tecnología para administrar volúmenes lógicos para el kernel de Linux. LVM incluye características como las siguientes:

- Redimensionado de grupos lógicos.
- Redimensionado de volúmenes lógicos.
- Instantáneas de solo lectura (LVM2 ofrece lectura y escritura)

[2] `https://access.redhat.com/products/red-hat-satellite`
[3] `http://www.redhat.com/es/services/certification`
[4] `https://www.centos.org/docs/5/html/Cluster_Logical_Volume_Manager/`

- RAID 0 de volúmenes lógicos.

Hay que tener en cuenta que LVM no implementa ni RAID 1 ni RAID 5 por lo que es necesario utilizar software adicional o crear los RAID a nivel físico. En un sistema grande la administración de todos los discos duros necesarios es un trabajo costoso en tiempo y dinero y se hace particularmente complejo si el sistema contiene discos de distintos tamaños o prestaciones. Aunque mantener grupos de discos de las mismas características es una buena práctica hay veces en las que simplemente no es posible. LVM permite agrupar estos discos o fragmentos de ellos de distintas formas en un mismo grupo lo que facilita enormemente la gestión. Como ya hemos mencionado antes estos grupos puedes redimensionarse a voluntad añadiendo más partes (*extents*) de otros discos o reduciendo/eliminando existentes. LVM se compone de tres partes como se aprecia en la figura 2.1:

- Volúmenes físicos (PV *Physical Volume*): son las particiones del disco duro con sistema de archivos LVM.

- Volúmenes lógicos (LV *Logical Volume*): es el equivalente a una partición en un sistema convencional. Un LV es visible como un dispositivo estándar de bloques por lo que puede contener cualquier sistema de archivos.

- Grupos de volúmenes (VG *Volume Group*): es el nivel organizativo superior. Contiene los LV que le asignemos bajo un nombre personalizado, esto es, no tendrá nombre de partición como */dev/sdX* sino que se le puede dar un nombre organizativo del tipo *home*, datos, etc.

2.2.2. *Hotplug y hotadd*

Estas tecnologías permiten añadir núcleos de CPU y memoria RAM en caliente sin necesidad de reinicios y el núcleo de Linux comenzará a usarlos directamente sin necesidad de ninguna configuración adicional. En el caso del estudio esta característica se gestiona desde la plataforma de virtualización ya que no se añaden físicamente más modelos de RAM ni procesadores sino que se asignan más recursos a esa máquina virtual.

2.2.3. Hardware en caliente

Además de CPU y RAM CentOS permite añadir cualquier otro hardware en caliente como adaptadores de red, usb, distintos tipos de lectores, etc.

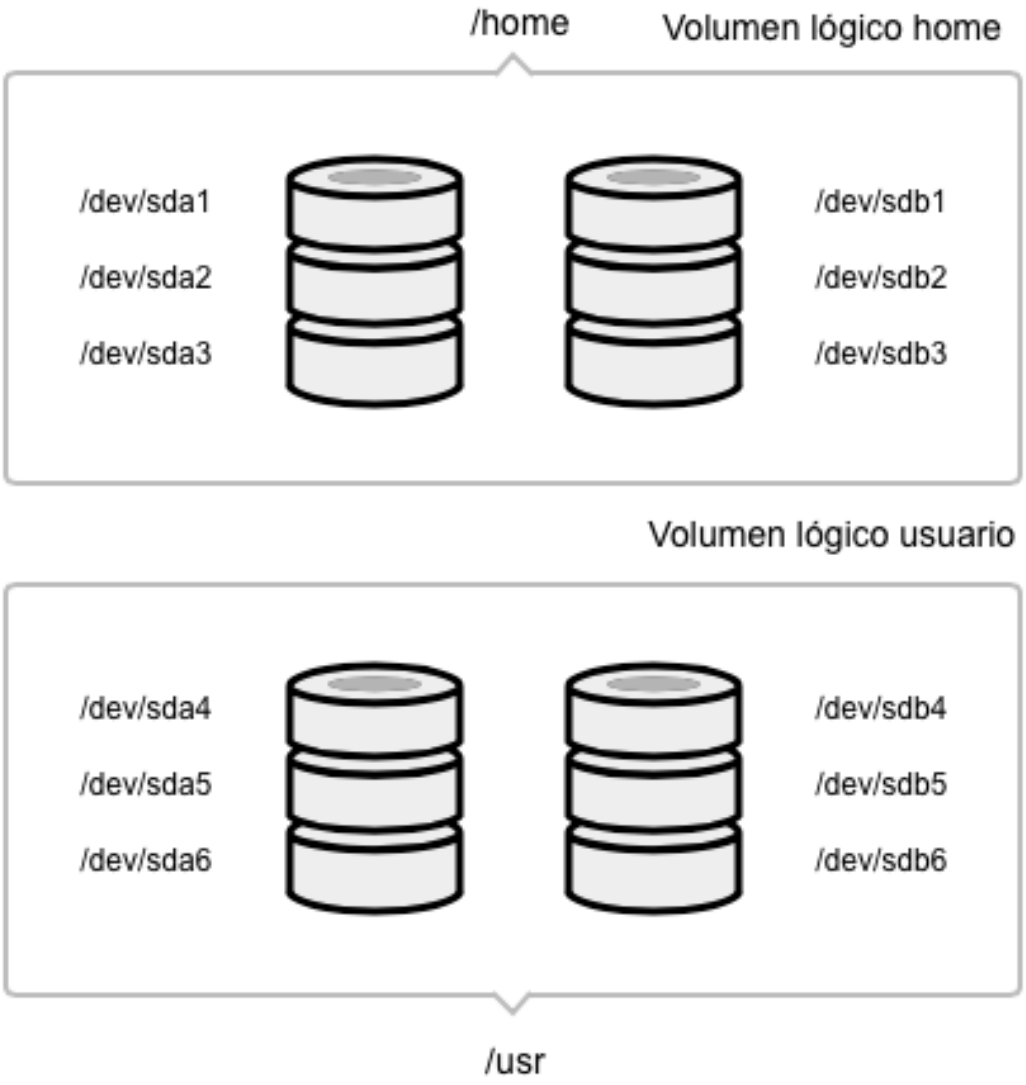

Figura 2.1: Esquema LVM donde se aprecian diversas particiones físicas agrupadas en un solo volumen lógico aumentando su capacidad.

2.3 Plataforma de virtualización

La plataforma de virtualización que se usa en esta infraestructura está basada en la tecnología de vmWare [12]. Esta plataforma de virtualización está constituida por 3 componentes. Por una parte tenemos el sistema operativo que se instala sobre las máquinas físicas, denominado ESXi, concretamente se utiliza la versión 5.5.0. Este sistema operativo se instala directamente sobre las máquinas físicas como cualquier otro sistema pero no ofrece nada por sí mismo. Lo único que permite es, mediante el acceso directo a la máquina una pequeña configuración de seguridad y red. Es lo que se denomina un *Hypervisor* o hipervisor [7].

Por otra parte tenemos el Vcenter. Esta aplicación se puede instalar sobre un Windows o un SUSE y controla la comunicación entre las diferentes máquinas con ESXi instalado y permite al operador acceder a los servicios de la plataforma de vmWare mediante el tercer componente: el cliente de escritorio VSphere.

Este cliente de escritorio se conecta al Vcenter y desde él se controlan todas las máquinas virtuales y los ESXi. Permite desde la creación de nuevas máquinas virtuales a la gestión de la red virtual, gestión de aprovisionamiento, terminal para las máquinas virtuales etc. Esta arquitectura permite que un pequeño equipo gestione una cantidad bastante grande de máquinas virtuales y sus respectivos hosts en remoto y de una forma rápida, clara y ordenada. En este punto vemos la primera gran diferencia respecto de un sistema más tradicional en el que se debe gestionar por regla general una máquina física por servicio o agrupar varios servicios en una sola máquina creando así un único punto de fallo en la máquina física.

En la figura 2.2 se puede ver un esquema general de como se estructura la plataforma de virtualización en una vista por niveles. En el nivel inferior se puede ver el hardware sobre el que se soporta la plataforma. Sobre él se instala el ESXi como ya hemos comentado. Sobre este sistema operativo corren las diferentes máquinas virtuales gestionadas por el vCenter desde el nivel superior, al que se conecta el administrador desde el cliente de escritorio.

Algunas de las características que hacen de la plataforma de virtualización de vmWare tan versátil y que se explicarán a continuación son:

- Ahorra energía.
- Reduce la huella ecológica del CPD.
- Facilita la creación de entornos de prueba.
- Reduce la dependencia de un solo fabricante.

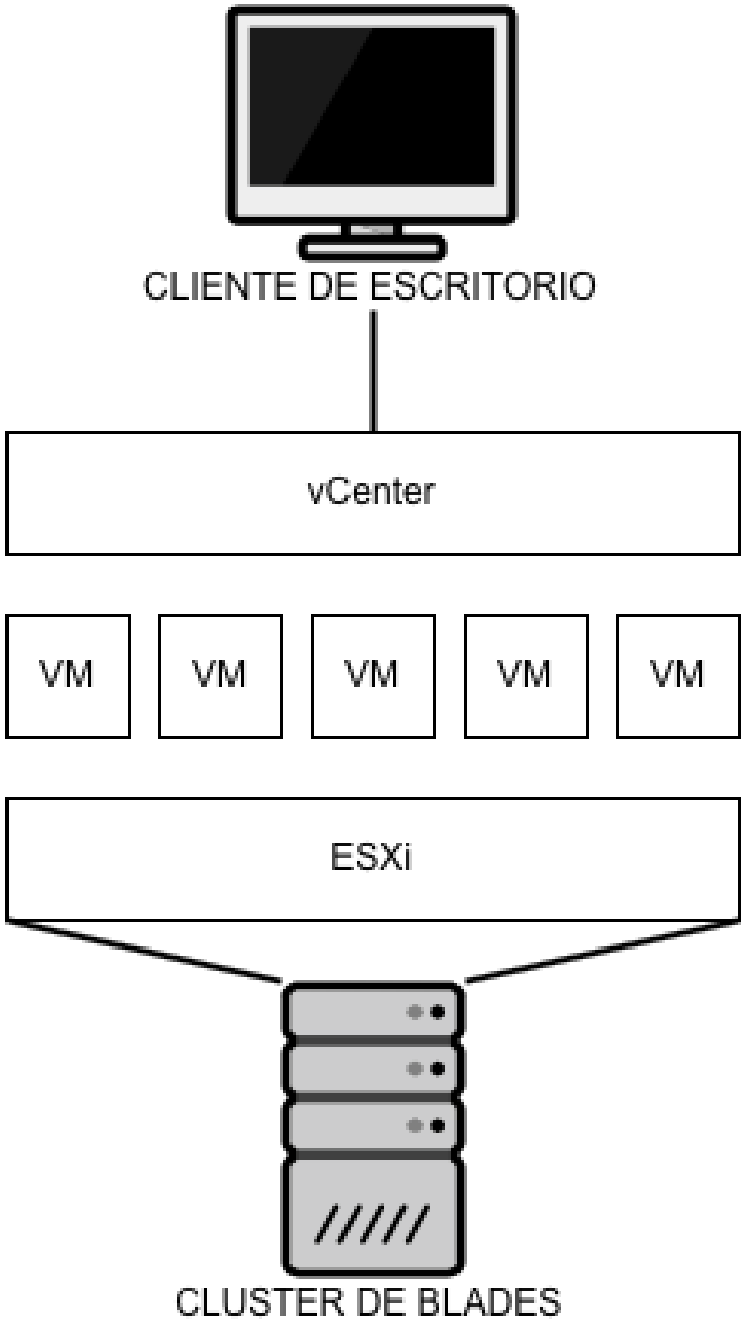

Figura 2.2: Esquema de la plataforma

- Aumenta el tiempo de disponibilidad.
- Facilita la gestión de las máquinas.
- Facilita la contingencia del sistema.
- Facilita realizar las copias de seguridad.
- Facilita la instalación de nuevas máquinas.
- Simplifica la instalación de red.
- Aísla servicios.
- Extiende la vida de aplicaciones antiguas.

2.3.1. Ahorro de energía

Como hemos podido ver hasta ahora un punto muy importante a favor de la virtualización es el ahorro energético. Migrar servidores físicos a máquinas virtuales y consolidarlas en un número más reducido de servidores físicos significa reducir el consumo energético directo de las máquinas y por lo tanto también el consumo del equipamiento de refrigeración.

Las últimas versiones de la plataforma de virtualización de vmWare incluyen una característica llamada DRS del inglés *Distributed Resource Scheduler*. Esta tecnología permite distribuir la carga entre las máquinas físicas para adaptarse a los cambios de carga. Es necesario que el vCenter tenga acceso a la iLO[5] de los hosts para poder realizar acciones sobre ellos a la hora de apagarlos o encenderlos. El principio básico de funcionamiento consiste en utilizar la iLO de los hosts para poder apagar o encender los servidores según la carga total del sistema. El administrador debe establecer unas políticas de apagado y consolidación de forma que el servicio sepa qué máquinas virtuales puede mover y que hosts puede apagar entre otras cosas. Gracias a esta tecnología el sistema es capaz de realizar ciclos de encendido y apagado según varía la carga total del sistema, por ejemplo, en el ciclo de día y noche. Mediante la utilización de vMotion reubicará las máquinas virtuales en un menor número de hosts y apagará los que dejen de tener máquinas encendidas ahorrando así en consumo eléctrico.

2.3.2. Huella ecológica

La huella ecológica o *ecological footprint* [5] en inglés es una medida estándar de la demanda de recursos naturales en contraste con la capacidad del planeta

[5] `http://www8.hp.com/us/en/products/servers/ilo/`

para regenerarlos [4]. Representa la extensión de tierra o agua biológicamente productiva para proveer los recursos consumidos o asimilar los desechos.

Aunque el cálculo de la huella ecológica es muy complejo y en ocasiones imposible podemos aventurar que en el caso de un centro de datos este cálculo se basa en estimar la energía consumida y la extensión de bosque/mar que se necesita para absorber el CO_2 que se deriva de la producción de esa energía. Aun así este calculo es muy impreciso porque no se están teniendo en cuenta los recursos necesarios para la producción de los elementos físicos que conforman el centro de datos o los desechos que se generan de reparaciones o actualizaciones. Además, algunos elementos de los centros de datos son extremadamente contaminantes como las baterías de los SAI o determinados compuestos utilizados en la manufactura de chips y placas electrónicas. Estos deshechos pueden acabar no siendo eliminados correctamente como advierten algunas organizaciones[6]. Hoy en día se realiza un gran, aunque insuficiente, esfuerzo en conseguir producir una electrónica energéticamente más eficiente y libre de compuesto tóxicos no reciclables.

De esto podemos deducir que el uso de menos servidores y electrónica de red tiene como consecuencia directa la reducción de la huella ecológica del CPD de diversas maneras puesto que además de generar menos desechos, el mejor aprovechamiento de la energía conlleva una reducción en la producción de CO_2.

2.3.3. Entornos de prueba

Una vez que el ciclo de vida de un hardware llega a su fin y debe ser actualizado, lo más habitual es desecharlo o acumularlo, lo cual como hemos visto empeora la huella ecológica del CPD. En vez de esto, mediante la virtualización podemos crear fácilmente entornos de prueba, laboratorios o entornos de desarrollo separados del entorno de producción en los que realizar pruebas o experimentos que podrían poner en riesgo a máquinas físicas produciendo incluso pérdidas de servicio. De esta manera, mediante el uso de máquinas virtuales podemos someter al sistema operativo a pruebas que de otra manera serían mucho más costosas y peligrosas. Además, gracias a determinadas técnicas que proveen los entornos de virtualización como los *snapshots*[7] podemos volver a un estado anterior de la máquina en cuestión de minutos haciendo posible repetir un mismo experimento sobre un determinado estado de una máquina en un tiempo antes imposible. De ésta y otras técnicas se hablará más adelante en el capítulo.

[6] `http://www.greenpeace.org/international/en/news/features/poisoning-the-poor-electroni/`

[7] `http://kb.vmware.com/selfservice/search.do?cmd=displayKC&docType=kc&docTypeID=DT_KB_1_1&externalId=1009402`

Otro posible entorno es el de desarrollo. Si la empresa desarrolla su propio software podría crease un entorno de desarrollo y/o preproducción en el que desarrollar las aplicaciones y testarlas en un entorno lo más parecido posible al de producción sin el riesgo inherente de usar este último.

2.3.4. *Vendor lock-in*

Aunque no siempre es malo centrarse en un solo fabricante de hardware o incluso a un modelo concreto de servidor, en un sistema no virtualizado puede resultar un problema. En muchos casos por particularidades de la infraestructura, por compatibilidad con aplicaciones o simplemente por costumbre, se tiende a utilizar una sola marca o modelo. La virtualización añade una capa abstracta sobre el hardware que permite mayor flexibilidad a la hora de elegir. Esto además puede ser una buena herramienta a la hora de negociar cuando llega el momento de ampliar o renovar el hardware.

2.3.5. Alta disponibilidad

Hoy en día casi todas las plataformas de virtualización ofrecen un buen número de características avanzadas que no se encuentran en servidores físicos que ayudan a las empresas a ofrecer alta disponibilidad [15] en sus servicios. El concepto de disponibilidad no debe ser confundido con el tiempo que un servidor lleve encendido, del inglés *uptime* que también es una orden de consola que devuelve el tiempo desde el último reinicio o apagado del servidor. Estos dos términos no significan lo mismo puesto que un servidor puede estar encendido (dentro del *uptime*) pero no disponible, por ejemplo, porque la red esté caída.

Por otra parte la disponibilidad puede ser un cálculo general para la infraestructura o para un servicio concreto. En el sistema del estudio hay definidos para el responsable del CPD unos porcentajes de tiempo que debe cumplir anualmente, en concreto el porcentaje general para el CPD es del 99,99999 % esto es 3,15 segundos como máximo no disponible al año. Para algunas partes del sistema los porcentajes pueden variar tanto a la alza como a la baja dependiendo del activo. Por poner un ejemplo: una máquina virtual que contiene una página web que no es gestionada por la empresa propietaria del centro de datos tiene una disponibilidad más baja que una que esté directamente al cargo de la empresa puesto que el contenido no es controlable y podría causar la pérdida del sistema operativo. De hecho, es bastante común que, debido a una mala programación, determinadas máquinas tengan disponibilidades muy bajas del orden del 99,5 % o 1,83 días sin servicio al año.

Aún así estos porcentajes son alcanzados gracias a técnicas que proporciona la virtualización la cual permite mantener una máquina funcionando bajo cargas

más altas de lo que soporta normalmente o incluso recuperarlas en emergencias. La capacidad de mover una máquina virtual de un servidor a otro es probablemente una de las ventajas más llamativas de la virtualización. Esta técnica permite incluso migraciones en grandes distancias. Por poner un ejemplo más practico de estas dos técnicas: tenemos una máquina que aloja una web sobre un servidor de aplicaciones TOMCAT. Esta máquina tiene asignados 2 núcleos, 4 GB de RAM y funciona entre el 50 y el 60 % de su capacidad siempre. Gracias a que está virtualizada no se están desaprovechando recursos puesto que la capacidad de procesamiento de esos dos núcleos virtuales se comparte con otras máquinas que puedan necesitarlo y solo se refiere al máximo asignado por el administrador. Debido a un pico en las visitas a esa página la máquina puede alcanzar el 100 % de uso de procesador y/o memoria lo cual le impide dar servicio a algunos clientes. Gracias a la virtualización el administrador es capaz de asignarle temporalmente más recursos, ya sean núcleos o memoria RAM y todo esto sin reiniciar ni dejar de dar servicio. Así, la máquina se recuperaría, conseguiría superar el pico sin dejar de dar servicio y cumpliría con sus metas marcadas de alta disponibilidad.

Este último ejemplo ilustra cómo una característica de la virtualización es capaz de ayudar frente a un imprevisto a nivel de software pero ¿qué pasaría si el servidor físico que aloja la máquina cae? ¿Y si ese servidor contiene no una, sino 5 máquinas virtuales? Al utilizar la virtualización, junto con un conjunto bien gestionado de servidores físicos, el hipervisor es capaz de mover en caliente las máquinas virtuales alojadas en uno de los servidores que sufre problemas y reubicarlas equilibrando la carga entre los otros servidores disponibles sin perder servicio, o en un caso más extremo, con una pérdida mínima.

Este tipo de ventajas son increíblemente valiosas para una empresa que se dedica a ofrecer servicios, sobre todo si se dedica a la monitorización, pues no tendría sentido que la máquina que monitoriza el estado de centenares de dispositivos no estuviese disponible. Para esta máquina en concreto el porcentaje fijado de disponibilidad es del 99,99999 % o más. En la práctica cualquier porcentaje de disponibilidad superior significa que el tiempo que puede pasar no disponible tiende a 0, es una máquina crítica y está preparada y aprovisionada para que solo un desastre en el centro de datos pueda detenerla.

2.3.6. Gestión

La virtualización sobre vmWare proporciona una herramienta centralizada para la gestión de todas las máquinas, tanto virtuales como físicas. Mediante esta herramienta se pueden gestionar todos los aspectos de estas máquinas. Por parte de las máquinas físicas y aunque normalmente los fabricantes proporcionan una herramienta web de gestión más completa, para un uso normal permite vigilar los sensores y tareas básicas como reiniciar o gestionar el direccionamiento, etc. Por parte de las máquinas virtuales permite gestionar absolutamente todo em-

pezando por el hardware y proporciona una consola gráfica para acceder a las máquinas incluso aunque no sean accesibles a través de la red. Además facilita mucho las tareas de inventariado puesto que te permite exportar a una amplia variedad de formatos el inventario de máquinas virtuales y hosts personalizando los datos que se mostrarán.

2.3.7. Contingencia

La virtualización ofrece dos componentes importantes a la hora de crear una solución para recuperarse de un desastre. La primera es la capacidad de abstraer el hardware. Eliminando la dependencia de un vendedor en particular o incluso modelo de servidor, la recuperación de una o varias máquinas ya no pasa por tener hardware idéntico, solo suficiente capacidad de cómputo y puede adquirirse hardware más barato puesto que rara vez se usará. En segundo lugar la mayoría de plataformas de virtualización proveen de software que facilita la automatización del proceso de recuperación ante un desastre e incluso realizar pruebas de estrés para ver como se comportaría la infraestructura.

2.3.8. Copias de seguridad

Gracias a la capa de abstracción del hardware que proporciona la virtualización las copias de seguridad se pueden realizar con métodos alternativos que facilitan la gestión y permiten realizar las copias en menos tiempo. Hoy en día la mayoría de software de copia de seguridad a nivel empresarial es capaz de integrarse con la plataforma de virtualización para realizar copias completas o parciales de las máquinas virtuales accediendo a los discos virtuales de estas. De este modo una copia completa realizada a nivel de máquina virtual consiste en copiar un solo archivo por disco duro virtual asignado a la máquina lo que mejora los tiempos de copia puesto que reduce el número de archivos a copiar de miles de archivos relativamente pequeños a un par de archivos de gran tamaño. Aquí entra también la habilidad del administrador para aprovisionar correctamente las máquinas virtuales. Algunas recomendaciones a seguir a la hora de aprovisionar de espacio las máquinas virtuales son:

- Intentar asignar el menor número de discos virtuales a una máquina teniendo en cuenta siempre los servicios que va a correr, por ejemplo, si una misma máquina tiene un frontal web y una base de datos se le podrían asignar dos discos virtuales para poder optimizar por separado los servicios.

- Para poder dimensionar una máquina correctamente se debe usar LVM para gestionar el tamaño variable de los discos e intentar siempre no dejar excesivo espacio libre en las máquinas.

- A ser posible gestionar el nivel de RAID sobre cabina para adaptar cada máquina o grupo de máquinas a las prestaciones del RAID configurado. Por ejemplo, una máquina crítica suele correr sobre un RAID 1 o incluso RAID 0+1, el resto de máquinas suelen correr sobre RAID 5.

2.3.9. Crecimiento

Hoy en día con el rápido aumento en las exigencias a las infraestructuras informáticas es muy importante tener una buena planificación de crecimiento. La virtualización ofrece una capacidad de escalabilidad que una infraestructura basada en servidores físicos no puede alcanzar.

Debido a que un número elevado de máquinas virtuales se consolidan en el menor número posible de máquinas físicas la capacidad económica necesaria para ampliar la infraestructura es mucho menor pudiendo adquirir más servidores para un mismo chasis o adquirir un chasis nuevo que bien aprovechado ocupará mucho menos espacio que un conjunto homólogo de servidores tradicionales.

Todo esto sumado a la simplicidad de su despliegue y gestión hacen que aumentar las capacidades del sistema sea mucho más sencillo y no sea necesario incurrir en ventanas de mantenimiento para parar o reiniciar máquinas físicas.

2.3.10. Organización de red

Aunque no es directamente un factor importante para este estudio se considera una ventaja para los administradores de red ya que la virtualización simplifica enormemente el cableado y la electrónica necesarias. En este caso de estudio usando un par de chasis de blades y un par de cabinas de almacenamiento y la dispersión de los servidores es mínima o visto de otra forma, todos los elementos que deben ser conectados a la red se concentran en menos espacio.

La diferencia básica de esta estructura virtualizada de más de 150 máquinas comparada con una infraestructura similar en la que cada máquina fuese un servidor físico es la del ahorro de espacio y número de máquinas. De esta forma, el número de elementos a conectar a la red es mucho menor, simplificando el cableado y la gestión. Para todas las máquinas virtuales la gestión de la red se realiza mediante conmutadores virtuales, exentos de mantenimiento y sencillos de gestionar desde la interfaz del cliente de escritorio. Una vez todas las máquinas están organizadas solo se necesitan unos pocos elementos de red más para completar toda la infraestructura de red.

2.3.11. Aislamiento de servicios

En el mundo de los servidores físicos muchos centros de datos adoptan una filosofía de *un servicio un servidor* para aislar aplicaciones. Pero esto provoca un crecimiento exponencial en el número de servidores, incrementa los gastos y en muchos casos desaprovecha la potencia de las máquinas. La virtualización proporciona aislamiento de los servicios y elimina los problemas de compatibilidad consolidando muchas de estas máquinas en unos pocos servidores físicos. Todo esto además limita los desechos puesto que cada máquina virtual se aprovisiona con la cantidad exacta de CPU, memoria y almacenamiento que necesita.

Aislamiento y estado separado son dos conceptos muy importantes en el mundo de la virtualización; a continuación se explica qué significan y por qué todas las técnicas relacionadas son tan importantes hoy en día.

Cuando hablamos de aislamiento, significa ocultar todos los recursos de una aplicación de todo los demás en el sistema. Por ejemplo, si una aplicación tiene un fichero en disco, o un valor de registro o una conexión, solo los procesos que corren paralelos a la aplicación dentro de su entorno pueden ver estos recursos. Por lo general se tiende a relajar este aislamiento de forma que la aplicación pueda compartir ciertos recursos pero siempre de forma muy controlada.

Usando una máquina virtual para cada servicio tenemos la ventaja de encapsular la mayoría de problemas en una determinada máquina virtual en vez de permitir que el problema se extienda y afecte a otros servicios provocando efectos cascada. Un proceso colgado en la cola y consumiendo todo el procesador o la memoria podría acabar con todo el servidor de no ser por este aislamiento. Pero como ya hemos comentado antes, aprovisionando correctamente las máquinas evitamos malgastar recursos y solo esa máquina se volverá incapaz de continuar funcionando, permitiendo reiniciarla sin afectar a los otros servicios.

Otra ventaja es que cualquier cambio realizado sobre una determinada máquina virtual no afectará a ningún otro servicio, permitiendo, por ejemplo, realizar actualizaciones de librerías sin que los otros servicios sufran efectos secundarios. Esto además es una gran ventaja desde el punto de vista de la seguridad en el que la actualización de seguridad, de una versión a otra de un servicio podría inutilizar otro servicio. Además, el procedimiento permite que gracias a la virtualización a penas haya pérdida de servicio. Si una de estas actualizaciones implica reiniciar la máquina o el servicio, o incluso pararlo un tiempo prolongado que ponga en peligro los márgenes fijados para la HA, se lleva a cabo el siguiente procedimiento:

1. Se crea un clon de la máquina a actualizar. Dependiendo de la máquina esto puede tardar entre unos segundos a unos cuantos minutos, pero en cualquier caso es asumible y no incurre en ningún tipo de pérdida de servicio.

2. Una vez creado el clon se desconectan las tarjetas de red de la máquina a actualizar y se conectan las del clon. En este caso la pérdida de servicio puede ser tan mínima que ni se aprecie.

3. Se procede con la actualización de la máquina. Puesto que tenemos un clon que sigue funcionando el procedimiento de contingencia si algo sale mal es tan sencillo como eliminar la máquina actualizada y repetir el proceso con el clon.

4. Una vez la máquina está actualizada se comprueba que todo es correcto y se realiza el cambio como en el paso dos.

5. Se elimina el clon.

Sin embargo determinadas máquinas no son propensas a actualizarse con este método. Este es el caso de máquinas cuyo estado es cambiante constantemente, por ejemplo, servidores de correo. Aun así, este procedimiento es muy cómodo para máquinas con poco uso o cuyo estado no es cambiante. Dentro de este grupo estarían todas las máquinas de preproducción o desarrollo y máquinas que pueden entrar en mantenimiento viendo reducido pero no eliminado su servicio, por ejemplo, servidores de para almacenamiento web. Estos últimos entran en ventana de mantenimiento y quedan en solo lectura; de esta forma los sitios web que alojan pueden seguir siendo consultados aunque no modificados durante el tiempo que dure la intervención.

A pesar de todas estas ventajas se puede argumentar que esto tiene una desventaja muy clara: la duplicación. Si cada servicio se aísla en una máquina virtual distinta llegamos al punto de duplicar (en el mejor de los casos) alrededor del 80 % de los archivos del sistema operativo. Esto efectivamente podría tener un efecto muy negativo desde el punto de vista del almacenamiento donde se estaría almacenando la misma información un gran número de veces, en ocasiones cientos de veces llegando al punto de que ya no sea un simple problema sino completamente inviable. Afortunadamente esta gran desventaja se tuvo en cuenta desde el principio y la solución de almacenamiento por la que se optó posee una característica que convierte este problema en algo insignificante: la deduplicación.

2.3.12. La deduplicación

La deduplicación es en realidad un conjunto de técnicas destinadas a aprovechar mejor la capacidad de almacenamiento cuando hay una cantidad consi-

derable de datos duplicados. Este concepto también se conoce como compresión inteligente y provee al sistema de almacenamiento de dos ventajas muy notables:

- Ahorro de espacio ocupado gracias a que se guarda solo una copia de los datos independientemente del número real necesario.

- Puesto que solo se guarda una copia solo se envía o se recibe, es decir, se lee o se escribe una de estas copias ahorrando en utlización de tiempo de disco y de conexiones.

El uso de esta técnica es uno de los motivos por los que se tiende a usar el mismo sistema operativo en todas las máquinas. En la vida de un sistema operativo hay un gran número de archivos que se crean durante la instalación y nunca son modificados, o que siendo modificados lo son muy pocas veces. En un entorno en el que más del 90 % de los sistemas operativos tienen una configuración similar puesto que usan la misma versión de la mayoría de paquetería básica se calcula que el ahorro de espacio en disco gracias al uso de la deduplicación es de entre el 50 y el 80 %. No hace falta decir lo importante que se vuelve el uso de esta técnica para ahorrar costes y reducir la huella ecológica que produce el consumo masivo de unidades de almacenamiento.

2.3.13. Aplicaciones antiguas

En algunos casos concretos nos encontramos aplicaciones que solo funcionan en sistemas operativos desfasados o sin soporte pero que se siguen utilizando. En algunos casos por falta de presupuesto para sustituirlas, por ejemplo, si son a medida, en otros simplemente no hay una alternativa mejor. Pero estas aplicaciones necesitan de sistemas operativos que no pueden ser instalados en máquinas modernas o que no aprovecharían todo su potencial, en cualquier caso incurriendo en una pérdida económica que podría no ser desdeñable.

Para estos casos la virtualización ofrece la solución perfecta. Puesto que el problema es adaptar el hardware al operativo, la virtualización no presenta ningún problema a la hora de asignarle los recursos justos y necesarios a ese operativo desfasado. Gracias a esto tenemos una aplicación que de otra manera hubiese tenido que correr sobre un hardware desaprovechado o nos hubiese obligado a mantener un software antiguo pero corriendo sobre una máquina virtual pequeña que apenas supone un impacto sobre nuestra infraestructura. Desde el punto de vista económico supone una ganancia directa puesto que mantener hardware antiguo tiene el problema de que los repuestos son caros o inexistentes y muchas veces adquirir hardware moderno simplemente no es una opción.

2.4 Plataforma hardware

La infraestructura que soporta la virtualización está formada por tres elementos principales: un par de chasis de servidores tipo blade, dos cabinas de almacenamiento y los elementos de red para interconectarlos. A continuación se detallan las características de cada uno y cómo funcionan entre sí.

El cerebro de todo el sistema está basado en el uso de un tipo de servidor llamado comúnmente servidor tipo blade (ver figura 2.3) o simplemente blade, del inglés cuchilla, haciendo alusión a su forma. Este tipo de servidor está diseñado para aumentar el aprovechamiento del espacio, reducir el consumo y simplificar su explotación. Están pensados para ser montados en chasis (ver figura 2.4) los cuales dan cabida a entre 8 y 16 servidores en un espacio de entre 4U y 6U estos incluso pueden añadirse en caliente. Además son más baratos de producir ya que las fuentes de alimentación y la electrónica de red se concentran en el chasis y se comparte entre todos los blades. Por supuesto como cualquier servidor convencional estos chasis incluyen fuentes de alimentación *hot-plug* redundantes, varias tarjetas de red e incluso tarjetas de fibra óptica para almacenamiento. Otra ventaja es que la ventilación es modular. Estos ventiladores se sitúan en la parte trasera del chasis, son *hot-plug* y se pueden adecuar a la carga del mismo dependiendo del número de blades en funcionamiento.

Al hablar de los tipos de servidor se menciona una medida usada para cuantificar el espacio que ocupa un elemento dentro de un rack. Un rack (ver figura 2.5) no es más que un armario diseñado para contener elementos electrónicos varios, como: servidores, PSU, conmutadores, etc. La medida usada es la U o simplemente unidad. Así pues un servidor que ocupe 2U ocupa dos espacios dentro del armario que generalmente son de 42U pero pueden llegar a 46U o 47U no pudiendo exceder nunca los 2000mm de altura externa.

Al respecto de la ocupación del chasis hay que tener algo muy en cuenta. Debido al formato del chasis, incluso aunque no contenga ningún blade sigue ocupando entre 4U y 6U. Si tomamos como ejemplo un chasis de 6U y servidores convencionales de 1U sería necesario ocupar el chasis con, como mínimo, 6 blades, equivalentes a los servidores convencionales para igualar la potencia y el espacio ocupado. Dicho de otra forma, si se opta por el uso de servidores tipo blade hay que planificar de antemano la situación que van a ocupar en el CPD y comparar con la utilización de servidores convencionales en relación al precio, el consumo y la potencia requeridas. Por lo general, si se tienen previsiones de crecimiento y de que la potencia necesaria vaya a ser mayor, es mejor optar por servidores en blade que más adelante permitirían, en el mismo espacio, llegar a completar el chasis con 16 unidades.

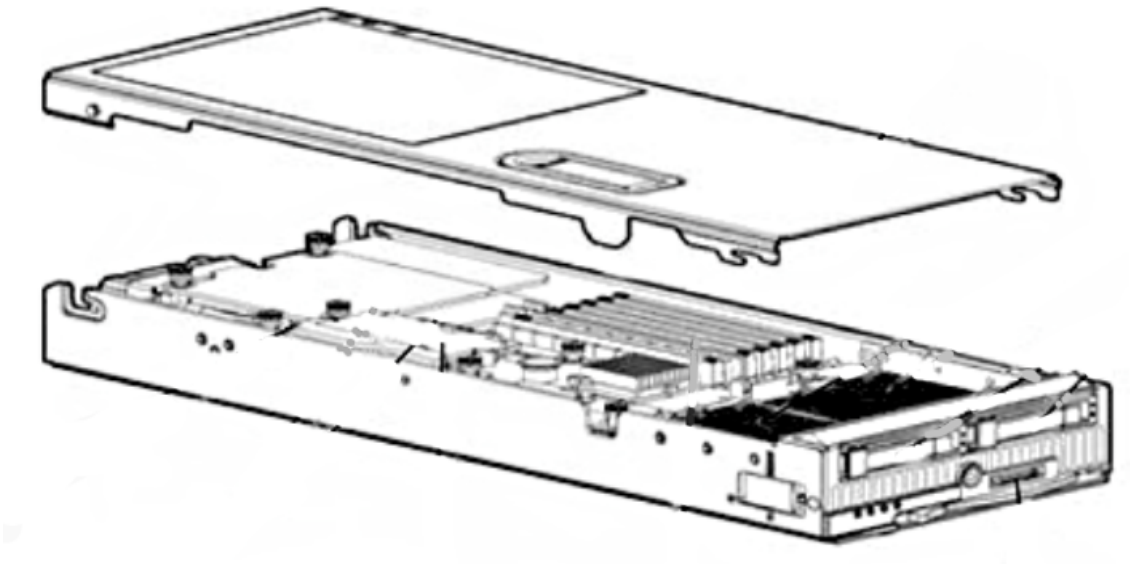

Figura 2.3: Esquema simplificado de un blade HP

Otro punto importante a tener en cuenta y que se deduce fácilmente del formato de los blades es que tienen un punto de fallo común en el chasis. Si el chasis dejase de funcionar se pierden todos los blades de ese chasis. Aunque pueda sonar catastrófico hoy en día los chasis poseen todas sus características redundadas, ya sea ventilación, tarjetas de red, de fibra o fuentes de alimentación.

A nivel de conectividad facilitan enormemente la gestión de la red y las conexiones físicas puesto que toda la electrónica queda concentrada en el chasis. Las tarjetas de red que proporciona el chasis tienen una gestión vía web que permite cualquier configuración que se necesite, por ejemplo, asignar una boca en concreto a un único blade, configurar una de ellas en modo promiscuo, *teaming y bonding*, etc. Para la empresa del estudio la capacidad de la electrónica de red de cambiar a modo promiscuo es muy importante. El modo promiscuo permite a una interfaz de red capturar todo el tráfico que pase por la red a la que pertenezca, incluso aunque no vaya dirigido a ella. En la práctica es utilizada por los productos de seguridad de la empresa para asegurarse que el tráfico de la red es correcto y no contiene conexiones maliciosas. Las otras dos técnicas mencionadas son el *teaming y bonding*. Estas técnicas tratan de solucionar el mismo problema: la perdida de conectividad por el fallo de un elemento de conexión como pueda ser un cable de red. Básicamente consisten en agrupar varias interfaces de red para hacerlas tolerantes a fallos. También pueden ser usadas para equilibrar o repartir la carga entre las interfaces. En el caso de estudio se usan dos chasis, un Dell PowerEdge M1000e con 6 blades Dell M610 y un HP BladeSystem c7000 con 16 blades HP ProLiant BL480c G1.

Figura 2.4: Vista frontal de un chasis HP donde se aprecian varios blades sin conectar

Figura 2.5: Vista frontal de un rack de HP de 42U con la puerta frontal abierta

Figura 2.6: Vista frontal y trasera de un chasis M1000e cargado

2.4.1. Dell PowerEdge M1000e

Este chasis (ver figura 2.6) es el más moderno y potente de los dos que posee la infraestructura en este momento. Se ha ido rellenando en el transcurso de un año con 4 blades más, empezó con 2 y ahora tiene 6. Como se ha indicado antes infrautilizar un chasis tiene efectos negativos sobre el consumo por lo que se decidió adquirir más unidades para este chasis en vez de adquirir otro chasis completo.

Los blades que componen este chasis, los Dell M610, constan de 2 procesadores Intel Xeon de 6 núcleos con Hyperthreading para un total de 24 procesadores lógicos, 92 GB de RAM, dos HBA para conectividad de fibra, dos controladoras gigabyte ethernet y dos discos SAS 15k 74 GB para el almacenamiento del sistema operativo. Una característica interesante de estos blades que se está probando actualmente en dos de estas unidades es el uso de tarjetas SD para el almacenamiento del SO.

Este modelo consta de dos sockets internos para este tipo de almacenamiento y es capaz de montar RAID sobre estos. Esta técnica crea un almacenamiento más veloz que el tradicional sobre HDD, consume mucha menos electricidad, disminuye notablemente la huella ecológica de cada servidor, aumenta la velocidad de arranque y uso general del almacenamiento interno, abarata los costes de reparación o sustitución y reduce el calor generado por cada unidad. Además la vida útil de este tipo de tarjetas es mayor que la de los discos físicos por lo que disminuye el tiempo medio de reparación (MTTR, *Mean Time To Repair*) y aumenta el tiempo medio entre fallos (MTBF, *Mean Time Between Failures*).

Los blades de Dell se usan principalmente para contingencia y para las máquinas de desarrollo que no tienen la necesidad de HA. Dependiendo de la cantidad de máquinas que se estén utilizando para desarrollo en cada momento, hay un numero de blades encendidos distinto, pero la mayor parte del tiempo hay encendidos dos.

2.4.2. HP BladeSystem c7000

Estos blades son menos potentes que los de Dell se usan para los servicios principales, puesto que soportan menos máquinas se intenta repartir homogéneamente la carga que suponen las máquinas virtuales más pesadas. Las especificaciones para estos blades son: dos procesadores Intel Xeon de 2 sin Hyperthreading o de 4 núcleos con Hyperthreading, entre 32 y 64Gb de RAM, dos HBAs para conexiones de fibra y dos controladoras gigabyte ethernet.

Los blades por si mismos poseen un almacenamiento muy limitado. Cada uno tiene dos discos de alto rendimiento en RAID 1 de 73Gb aproximadamente y que solo contienen el hypervisor de vmWare que ya hemos mencionado en el apartado sobre la plataforma de virtualización.

Esto nos lleva al siguiente elemento dentro del sistema, las cabinas de almacenamiento.

2.4.3. Sistema de almacenamiento

Todas las máquinas virtuales están almacenadas en dos cabinas (ver figura 2.7) de almacenamiento, cada una con unas características y un propósito diferente. En el sistema del estudio tenemos una IBM Storwize V5000 y una EMC Celerra. Por un lado la cabina IBM contiene todas las máquinas virtuales. La cabina EMC está siendo sustituida y su contenido migrado a la cabina IBM.

- La cabina IBM tiene actualmente 43 discos agrupados en distintos RAID con una capacidad total de 28,40 TB utilizables. Utiliza tres tipos distintos de discos: SSDs de 186GB configurados como caché, discos SAS de 10K rpm de 560 GB y discos SATA de 7200 rpm de 2,8 TB.

- La cabina EMC tiene una capacidad total de 60 discos: 15 SATA de 750 GB, 30 de 1 TB y 15 SAS de 270 GB.

Todo el sistema de almacenamiento está conectado a dos conmutadores de fibra los cuales interconectan los chasis con las cabinas. Todo esto está configurado con tolerancia a fallos utilizando *multipath*. El *multipath* se crea utilizando dos conmutadores y ofreciéndole al chasis varios caminos de fibra hasta la cabina; de esta forma si un conmutador, o una tarjeta de fibra en la cabina o en el chasis fallase, aún se mantendría la conectividad por otro camino. Esta técnica además mejora el rendimiento de la conexión al utilizar todos los caminos simultáneamente repartiendo la carga o sumando sus anchos de banda, técnica se denominada Concurrent Multipath Routing [8].

Figura 2.7: EMC Celerra (izq) y IBM Storwize V5000 (der)

El tercer elemento en el sistema es la electrónica de red. Por un lado ya hemos hablado de como el chasis simplifica la electrónica puesto que todos los servidores, y en consecuencia todas las máquinas virtuales, se conectan a la red a través de las tarjetas que se concentran en el chasis. Obviamente esto plantea que podría darse el caso de que las tarjetas del chasis no tengan potencia suficiente para servir a 16 blades con más de 150 máquinas virtuales, pero como ya hemos comentado anteriormente la configuración de la electrónica de red del chasis permite usar técnicas de *bonding o teaming* para solucionar esto, además de que suelen ser tarjetas a 10 Gb/s. Aun así, es un punto a tener en cuenta a la hora de elegir este tipo de servidores para virtualización.

Pero centrándonos en la estructuración de la red tenemos un esquema como el de la figura 2.8. Fuera del esquema quedan algunos servicios como cámaras de seguridad, cerraduras electrónicas, etc. En la figura podemos ver como los distintos elementos están conectados al cortafuegos a través de distintos conmutadores para que las conexiones sean tolerantes a fallos. Los conmutadores separan físicamente las distintas áreas. Por un lado tenemos conmutadores para servidores y por otro, conmutadores para usuarios. También podemos ver dos puntos de acceso wifi, uno que conecta con la red interna y otro con una salida separada a internet y que se utiliza como wifi de cortesía.

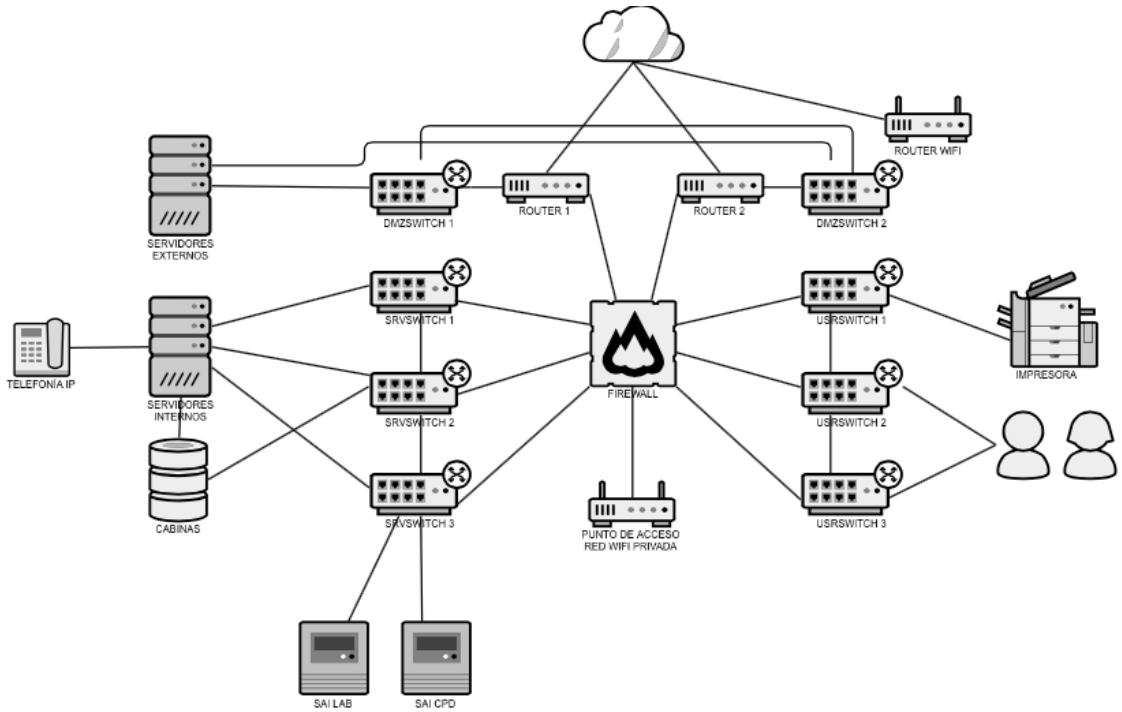

Figura 2.8: Esquema de red

CAPÍTULO 3

La infraestructura de monitorización

3.1 Introducción a la monitorización

La empresa propietaria del CPD del estudio es una empresa que se dedica a la seguridad informática. Dentro de los servicios de seguridad que proporciona integra en uno de sus productos la plataforma de monitorización Nagios [9] con la que se asegura que todos los hosts y servicios necesarios para el correcto funcionamiento de la infraestructura no sufren ningún problema. Además el producto integra otra solución para detectar problemas de seguridad en la red, Snort. El objetivo de este producto es integrar una serie de aplicaciones que cubran todos los aspectos de seguridad de la infraestructura a monitorizar. Además de Nagios y Snort incluye otras aplicaciones como MRTG y GrayLog.

Más adelante se explica qué es Nagios y en qué consiste, pero como acabamos de mencionar hay otras aplicaciones integradas en el producto. Estas aplicaciones tienen propósitos diferentes, como:

- MRTG (Multi Router Traffic Grapher): es una herramienta que se utiliza para supervisar la carga de interfaces, generalmente en routers y conmutadores. Esta herramienta utiliza SNMP para realizar consultas a estos dispositivos y generar con los datos informes en html que pueden ser utilizados en páginas web.

- Snort: es un *sniffer* de red y un detector de intrusos basado en red capaz de monitorizar todo un dominio de colisión. Con la información que recolecta puede generar informes de alertas para que un analista de seguridad pueda determinar si se trata de una amenaza o no. Junto con Nagios es una de las herramientas más importantes del producto, aunque mucho más orientada a la seguridad que Nagios.

- GrayLog: es una solución de almacenamiento y análisis de logs. La idea detrás de GrayLog es conseguir tener unificados en una máquina todos los logs de las diferentes máquinas dentro de la infraestructura y facilitar así su análisis por parte del personal de sistemas o seguridad.

Este producto que se ofrece al cliente está también implantado en la propia infraestructura de la empresa y se usa para monitorizar la disponibilidad de los hosts y los servicios. Pero si se usa correctamente es capaz de ayudar a detectar problemas de seguridad que es el principal propósito de la empresa.

En el siguiente capítulo se explica qué medidas se han utilizado para obtener un cálculo aproximado de la eficiencia energética del CPD (PUE), pero para poder explicarla correctamente es necesario conocer la infraestructura de monitorización sin la cual el cálculo de esta medida hubiese sido muy complejo o imposible. Aunque el producto incluye como ya hemos visto varias aplicaciones vamos a centrarnos en Nagios ya que gracias a ella se han que se hayan podido realizar las mediciones necesarias para el calculo del PUE.

Primero vamos a explicar qué es Nagios y cómo funciona para así poder explicar cómo se monitoriza la infraestructura del estudio y tener una mejor perspectiva del conjunto. Además, se hará hincapié en la importancia del histórico que la aplicación permite almacenar para realizar mejoras y ajustes o detectar problemas.

3.2 ¿Qué es Nagios?

Nagios es un sistema de monitorización de código abierto de hosts y servicios a través de la red. Posee cuatro componentes principales:

- El core, o núcleo de la aplicación.
- Los plugins o checks de servicio y host que se explicarán más adelante.
- El frontal web (ver figura 3.2).
- Los agentes remotos.

Nagios es una parte muy importante del sistema de monitorización puesto que es el primero en avisar de anomalías en las máquinas o servicios. A continuación se va a explicar su funcionamiento e integración con el resto de la plataforma de monitorización.

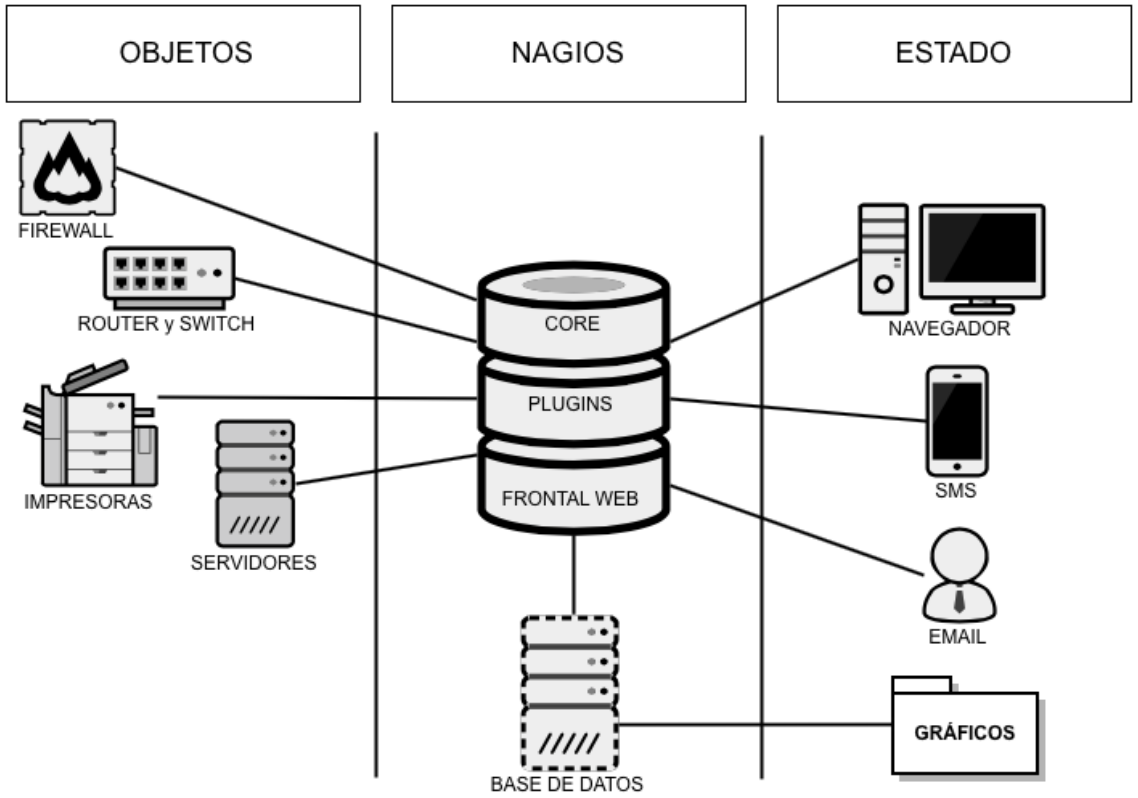

Figura 3.1: Funcionamiento de Nagios

3.3 ¿Cómo funciona Nagios?

Como ya se ha expuesto anteriormente Nagios está compuesto de cuatro componentes principales, el núcleo de la aplicación, los plugins, el frontal web y los agentes remotos. El esquema general de funcionamiento se puede ver en la figura 3.1. El funcionamiento se explicará en detalle más adelante.

3.3.1. Nagios Core

El núcleo de la aplicación es el encargado de ejecutar los archivos binarios necesarios y gestionar todas las actualizaciones de estado que se realicen sobre los hosts y servicios. Además contiene todos los archivos de configuración donde se encuentran definidos todos los hosts, servicios, ordenes, plantillas y relaciones entre estos. Cuando un nuevo host entra en la infraestructura es aquí donde debe darse de alta, asignarle dirección IP, características y relacionarlo con los servicios que se quieren monitorizar.

Para realizar la monitorización Nagios tiene varias opciones:

- Check SNMP[1]. Utilizando el protocolo SNMP (*Simple Network Management Protocol*) es capaz de realizar consultas básicas a casi cualquier dispositivo conectado a la red incluidos impresoras, conmutadores, cámaras de vigilancia... Este tipo de monitorización no provee de mucha información y puede ser costoso de configurar. Se suele utilizar solo en dispositivos simples que no permiten la instalación de un agente.

- Checks mediante agentes. Estos checks proveen de mucha más información que los sencillos por SNMP pero pueden resultar pesados e incurrir en una sobrecarga que podría no ser asumible por la máquina monitorizada. Para poder realizar checks de este tipo es necesaria la instalación de un agente remoto en la máquina a monitorizar. Este agente se encarga de ejecutar los plugins en la máquina remota y devolver el resultado al núcleo de la aplicación.

La configuración general del núcleo de Nagios, incluyendo los usuarios y los permisos, se puede realizar de una manera más tradicional editando directamente los ficheros de configuración a través de una consola, o utilizar aplicaciones de terceros. En este caso se usa una aplicación que además permite dar de alta hosts y servicios y probar la configuración antes de aplicarla. Esta configuración puede llegar a ser muy compleja dependiendo de la cantidad de hosts y servicios que se den de alta.

Para organizar toda la monitorización se pueden crear grupos de hosts, por ejemplo, electrónica de red, y definir relaciones de parentesco las cuales sirven, por ejemplo, para que si en una relación padre-hijo cae el padre, el hijo no notifique una caída puesto que no necesariamente ha caído él también. Esto evita eventos críticos en cascada.

3.3.2. El protocolo SNMP

Es necesario entender correctamente el protocolo SNMP [4] para poder entender como se realizó la monitorización, para ello vamos a profundizar en el protocolo. SNMP (*Simple Network Management Protocol*) es un protocolo de la capa de aplicación destinado a la gestión y administración de dispositivos conectados a la red. Estos dispositivos pueden ser: routers, conmutadores, impresoras, estaciones de trabajo, etc Utiliza el puerto 161 UDP para las consultas directas y el puerto 162 UDP para las inversas, también llamadas *Traps*.

Los dispositivos administrados contienen una base de información de administración (*Management Information Base, MIB*) organizada jerárquicamente y que contiene los objetos disponibles en el dispositivo. Esencialmente estos objetos son variables que contienen un valor asignado por el propio dispositivo. Algunos de

[1] `https://www.ietf.org/rfc/rfc1157.txt`

ellos pueden ser modificados vía el protocolo. Por ejemplo en la administración de una impresora se podría consultar un objeto, o variable llamada *vacío* con posibles valores 0 o 1 indicando cuando uno de los cartuchos de tinta está vacío.

A nivel técnico SNMP presenta un problema principal importante y es que los nombres de los objetos no son descriptivos, de hecho ni si quiera son texto. Cada objeto pertenece a un nivel en la jerarquía de la MIB y posee un número entero que lo identifica dentro de su nivel. Puesto que es jerárquica llegar a un objeto de nivel tres, implica pasar por los dos niveles superiores y su nombre sería la combinación de los números identificativos de los objetos ((*OID, Object Identifier*)) por los que se va pasando en la jerarquía. Por lo tanto un nombre para un posible objeto de nivel tres podría ser 1.3.2 y devolver un valor menos indicativo aun como por ejemplo, 35.

Para poder recorrer la MIB y todas las variables con sus resultados existe un comando llamado snmpwalk. Este devuelve una salida como la siguiente (los datos se han truncado):

```
# snmpwalk -v 2c -c public localhost system
SNMPv2-MIB::sysDescr.0 = STRING: Linux mago.aut.uah.es 2.6.0-test11 #27 Tue
Dec 16 11:39:03 CET 2003 i686
SNMPv2-MIB::sysObjectID.0 = OID: NET-SNMP-MIB::netSnmpAgentOIDs.10
SNMPv2-MIB::sysUpTime.0 = Timeticks: (120246) 0:20:02.46
SNMPv2-MIB::sysContact.0 = STRING: Root (configure /etc/snmp/snmp.local.conf)
SNMPv2-MIB::sysName.0 = STRING: mago.aut.uah.es
SNMPv2-MIB::sysLocation.0 = STRING: Unknown (edit /etc/snmp/snmpd.conf)
SNMPv2-MIB::sysORLastChange.0 = Timeticks: (4) 0:00:00.04
SNMPv2-MIB::sysORID.1 = OID: IF-MIB::ifMIB
SNMPv2-MIB::sysORID.2 = OID: SNMPv2-MIB::snmpMIB
SNMPv2-MIB::sysORID.3 = OID: TCP-MIB::tcpMIB
SNMPv2-MIB::sysORID.4 = OID: IP-MIB::ip
SNMPv2-MIB::sysORID.5 = OID: UDP-MIB::udpMIB
SNMPv2-MIB::sysORID.6 = OID: SNMP-VIEW-BASED-ACM-MIB::vacmBasicGroup
SNMPv2-MIB::sysORID.7 = OID: SNMP-FRAMEWORK-MIB::snmpFrameworkMIBCompliance
SNMPv2-MIB::sysORID.8 = OID: SNMP-MPD-MIB::snmpMPDCompliance
SNMPv2-MIB::sysORID.9 = OID: SNMP-USER-BASED-SM-MIB::usmMIBCompliance
SNMPv2-MIB::sysORDescr.1 = STRING: The MIB module to describe generic objects
for network interface sub-layers
```

El resultado del comando continua hasta que recorre todo el MIB. Como se puede observar el principio del resultado son datos sobre la configuración del protocolo en la máquina, la fecha y la consulta realizada. En la versión de SNMP 2c además de devolverse el número identificativo se devuelve un nombre.

En el caso del estudio se crearon varios checks *ex profeso*, utilizando SNMP, para recabar la información necesaria de consumos y realizar el cálculo del PUE en el capítulo siguiente. Estos dos checks tienen una sintaxis como la siguiente:

```
check\_snmp -H <IP> -v <version> -c <community> <OID>\\
```

Donde check_snmp es un script de consola que contiene una serie de ordenes y que realiza un snmpwalk contra el dispositivo proporcionado.

- <IP>es la dirección Ip asignada al dispositivo.
- <version>indica la versión del protocolo que puede ser la 1, 2c, 3 o 3c. Actualmente la versión 1 es aceptada por todos los dispositivos que admitan SNMP, la versión 2c no está tan extendida pero es fácil que el dispositivo la acepte. Las versiones 3 y 3c apenas tienen aceptación.
- La <community>es una palabra a modo de contraseña que el dispositivo necesita recibir para poder contestar las peticiones. En algunos casos la community por defecto es *public* pero en otros es una palabra algo más difícil de encontrar. En el caso de estos SAI no se consiguió encontrar la palabra para la community y se utilizó *public*. En estos casos un snmpwalk contra el dispositivo no devuelve todos los valores de la MIB, pero para estos SAI fue suficiente.
- <OID>corresponde con el objeto que se necesita, en este caso, la salida del SAI después de las perdidas por baterías y su propia circuitería.

En la práctica averiguar cual era el OID para la medida que se necesitaba fue costoso puesto que los SAI devolvían una gran cantidad de objetos muy cambiantes y en el caso de uno de los dos no aceptaba la versión 2c del protocolo por lo que la única manera de obtener el OID correcto era comparar con los valores que aparecen en la página web para la gestión del dispositivo con los cientos de valores que devolvía el snmpwalk.

3.3.3. Funcionamiento de Nagios

Para entender mejor la figura 3.1 a continuación se explica el funcionamiento de Nagios con más detalle.

1. Nagios Core consulta mediante SNMP o mediante plugins instalados en el cliente el estado y el valor de aquello que se quiera monitorizar. Hay tres posibles respuestas: OK, CRITICAL y UNKNOWN. Junto con el estado de la consulta se recibe, si el plugin lo requiere, un resultado que puede ser numérico o un texto informativo. Un resultado de un check_cpu para una máquina Linux podría ser: OK, 4.0 2.0 3.0
2. El resultado recibido se actualiza en el frontal web y en caso de estar configurado, se almacena en base de datos, a través de un plugin llamado ndo2db, para crear un histórico que es accesible también desde el frontal web.

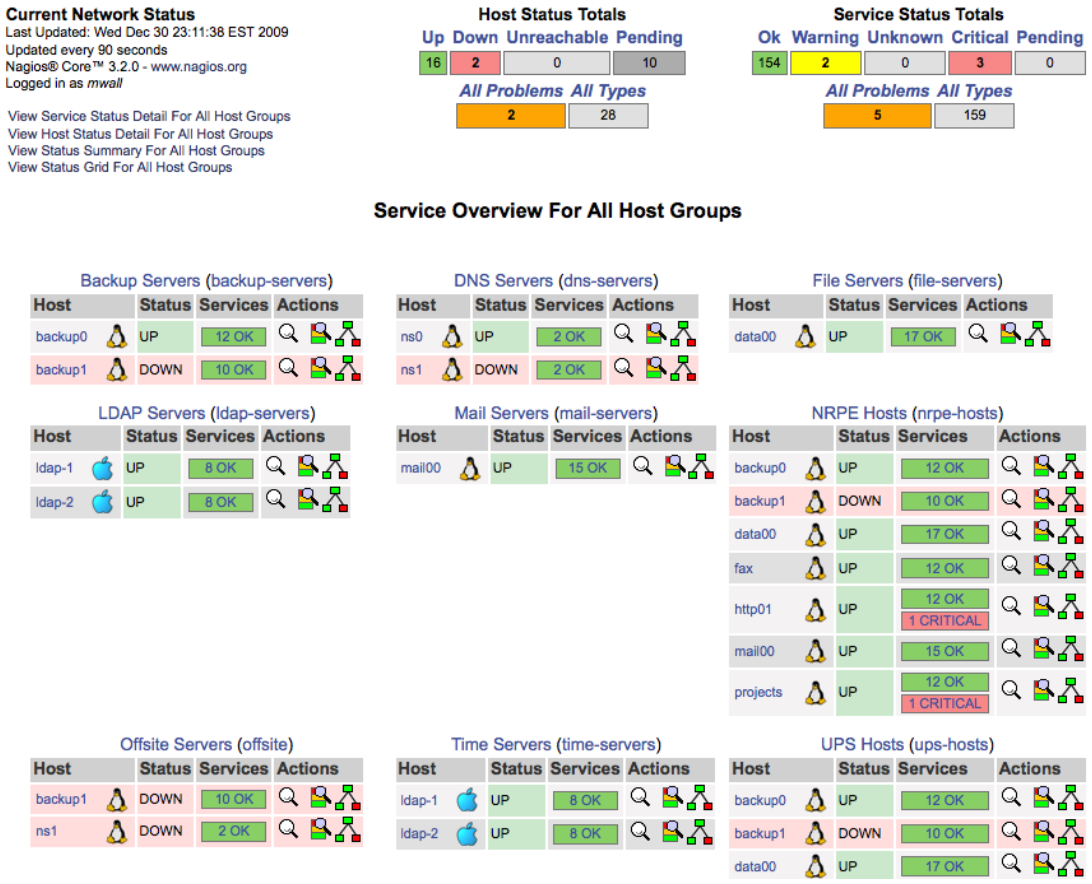

Figura 3.2: Frontal web de Nagios con diversos hosts incluidos en grupos.

3. Si el resultado es un CRITICAL, cada servicio monitorizado tiene una configuración de reintento para evitar falsos positivos. Si el servicio estuviese configurado con tres reintentos y el tiempo entre checks tres minutos: el primer check se ejecutaría a los tres minutos, de ser CRITICAL el tiempo para volver a ejecutarse sería de un minuto en vez de tres y se reintentaría una vez más un minuto después. Si después de los tres reintentos sigue en estado CRITICAL se notificaría en el frontal web. A los estados intermedios antes del último reintento se les llama *SOFT STATE*, al último intento sin variación de estado se le llama *HARD STATE*.

4. Además del estado actualizado en el frontal web, Nagios puede ser configurado para enviar los CRITICAL por correo electrónico o mensaje de texto.

CAPÍTULO 4

El PUE como medida de referencia

En este capítulo se va a explicar qué es el PUE y por qué se ha elegido como indicador para el calculo de la eficiencia energética.

4.1 ¿Qué es el PUE?

El PUE se define como el ratio entre el total de la energía consumida por el CPD y la energía consumida por el equipamiento [2], tal cual se muestra en la siguiente ecuación:

$$PUE = \frac{Consumo\ total}{Consumo\ del\ equipamiento\ de\ IT} \tag{4.1}$$

Para entender esta fórmula hay que definir lo que entendemos por energía total consumida y por energía consumida por el equipamiento informático así como el método por el cual se han obtenido:

- La energía total consumida por el CPD es la suma de todo el equipamiento (servidores, elementos de red...) y todos los demás elementos que permiten su funcionamiento (refrigeración, iluminación, pérdidas por baterías...).
- La energía consumida por el equipamiento es la suma del consumo de todos los servidores y elementos de red.

Los valores que puede tomar el PUE van de 1,0 a infinito. Idealmente el PUE se acercará a 1,0 lo que indicaría una eficiencia energética del 100 % (esto es, toda la energía se usa por el equipamiento de IT). Actualmente es difícil encontrar datos del PUE para distintos centros de datos. Algunos trabajos indican valores para el PUE para un centro de datos común sobre el 3,0 pero con un diseño adecuado se podría acercar al 1,6.

4.2 Calculo en el centro de estudio

Para el CPD de este estudio las medidas se han tomado gracias a los dos SAIs que alimentan el conjunto. La carga está convenientemente distribuida entre los dos. Uno de ellos soporta las dos unidades de aire acondicionado, la iluminación, el sistema de extinción de incendios, las cámaras de seguridad y los cierres de seguridad. Esta unidad se encuentra separada de la sala principal en la que se encuentra el otro SAI y todo el equipamiento en una sala contigua. La otra unidad soporta los servidores y electrónica de red.

Para el calculo del PUE se consideró un nivel 2 de precisión que establece The Green Grid. En este nivel se requiere de dos medidores de consumo diferenciados para el equipamiento y para la refrigeración, afortunadamente como ya hemos comentado esto se consigue gracias a los dos SAIs. Además el nivel dos exige tomar medidas diariamente o cada hora. En este caso todas las medidas se tomaron a las 12:00 todos los días. La elección de la hora depende de la variación de uso del sistema, como este sistema estaba más ocioso por la noche se decidió tomar la medida en hora punta, desfavoreciendo el resultado del PUE aunque se consideró más relevante a la hora de utilizar la medida para mejorar la eficiencia del CPD.

Mediante este nivel 2 se puede generar un valor del PUE adecuado para intentar mejorar la eficiencia del CPD pero este trabajo puede resultar difícil puesto que no especifica qué parte del equipamiento consume más energía. La organización espacial del CPD juega un papel muy importante a la hora de conseguir un mejor rendimiento en los sistemas de refrigeración. Un diseño común en los centros de datos consiste en enfrentar filas de armarios de equipamiento y refrigerar el pasillo que crean, a este pasillo se le llama pasillo frío. Por cada pasillo frío habrán dos pasillos calientes, estos son los pasillos que se forman en la parte trasera de las filas de armarios. Esta organización se basa en que los servidores y eletrónica de red recogen el aire frío por la parte delantera y expulsan el aire caliente por la parte trasera. En el caso del CPD del estudio no se llevó a cabo correctamente este diseño y aunque los armarios están posicionado correctamente la distribución de los elementos de refrigeración parece aleatoria lo que claramente disminuye su eficiencia.

Además respecto a la colocación del equipo de IT no se tuvo en cuenta que máquinas generan más calor y la distribución no sigue ningún patrón aparente. Este equipo se sitúa en una de las filas, en la otra se sitúan todos los componentes de red. La estructuración y organización de los componentes de red tampoco es la más conveniente puesto que los cables se organizan de forma arcaica y aleatoria en las traseras de los armarios formando una capa que acumula el calor que estos dispositivos intentan evacuar. Este es un punto importante a mejorar. El proble-

ma de esto es que una reestructuración masiva del CPD requeriría una parada prolongada que no se puede realizar actualmente.

4.2.1. Datos de consumo

A continuación se detallan los datos recogidos para realizar el calculo del PUE. Todos los valores corresponden al consumo instantáneo del SAI correspondiente a la refrigeración y del SAI correspondiente al equipamiento de IT a las 12:00. Con estos valores se calculará la media anual de cada uno de los consumos registrados por cada SAI y así calcular el valor final del PUE. Para poder representar más fielmente el valor del PUE se ha decidido sacar la media anual puesto que las estaciones afectan enormemente al consumo del equipo de refrigeración.

Las tablas 4.1 y 4.2 representan el consumo total en kW a las 12:00 para todos los días del año entre el 1 de Agosto del 2014 y el 31 de Julio del 2015 recogidas con las técnicas explicadas en el capítulo anterior. A la hora en la que se recogían las muestras, además, se revisaba la infraestructura al completo para asegurar que la medida no era distorsionada por alguna avería o situación anómala. Algunas de las medidas eran consultadas a posteriori por la imposibilidad de anotar la medida en el momento de la ejecución del check, por ejemplo, por carga de trabajo del técnico o por ser día festivo. Las medias x e y en la última fila de cada tabla muestran la media mensual. Para el cálculo del PUE (ver ecuación 4.4) se calcula por separado el consumo total del equipamiento de IT (ver ecuación 4.2) y el consumo total como la suma entre el consumo total del equipamiento de IT y el consumo total de la refrigeración (ver ecuación 4.3).

$$Consumo\ del\ equipamiento\ de\ IT = y = 11,8417736kW \tag{4.2}$$

$$Consumo\ total = x + y = 9,85669035 + 11,8417736 = 21,698464kW \tag{4.3}$$

$$PUE = \frac{21{,}698464}{11{,}8417736} = 1,8323660571 \tag{4.4}$$

Como podemos ver el PUE para el CPD del estudio es de 1,8, no está mal si tenemos en cuenta que no se tuvieron demasiadas consideraciones al diseñarlo en un inicio aunque es ampliamente mejorable y se propondrán en el siguiente capítulo una serie de mejoras aplicables [6].

Tabla 4.1: Consumos refrigeración

	Ago	Sep	Oct	Nov	Dic	Ene	Feb	Mar	Abr	May	Jun	Jul
1	14,9	13,4	13,5	10,0	6,2	5,3	5,9	5,1	6,0	9,2	14,9	13,4
2	14,6	14,7	13,3	11,2	5,0	5,8	4,6	5,9	7,1	10,1	14,2	14,2
3	14,1	13,3	12,7	9,5	5,6	5,5	4,3	5,3	6,8	11,8	14,0	13,4
4	13,1	13,7	13,2	8,6	6,1	5,7	5,1	5,7	7,0	11,4	14,0	14,9
5	14,5	13,6	12,3	7,2	6,4	5,9	5,4	5,0	7,3	10,7	13,9	14,6
6	14,1	13,1	13,7	5,6	5,3	5,5	5,0	5,0	8,6	11,4	13,2	14,7
7	14,9	13,8	13,5	6,6	5,0	4,1	4,0	5,7	7,8	11,7	13,0	14,3
8	14,7	14,9	13,0	8,4	6,6	5,2	4,7	5,8	8,9	11,8	14,7	14,5
9	13,4	13,5	12,5	9,4	6,6	5,0	4,4	5,6	8,9	13,0	13,1	13,6
10	14,6	14,0	13,2	9,5	5,8	5,1	4,1	5,9	7,7	11,6	13,0	13,8
11	14,4	13,9	13,0	9,2	6,8	4,9	4,5	5,2	8,1	12,8	14,8	13,4
12	13,5	14,6	13,5	9,8	5,2	6,0	5,9	5,3	7,7	11,4	14,7	15,0
13	13,9	14,6	12,5	5,6	5,7	4,6	5,6	5,5	7,2	11,4	14,9	13,7
14	13,7	14,7	12,3	5,4	5,5	4,4	4,8	5,3	8,5	14,6	14,8	14,1
15	14,0	14,8	12,6	5,4	6,0	5,4	4,5	5,1	8,3	13,5	14,2	14,3
16	13,5	13,0	13,6	7,4	5,5	4,2	5,2	5,8	7,9	13,0	13,6	13,8
17	14,5	13,6	12,2	8,0	5,3	5,4	6,0	5,2	7,7	14,3	14,1	13,7
18	14,1	14,2	12,3	8,5	5,3	5,2	5,8	5,2	7,3	13,5	13,9	13,4
19	13,8	14,5	13,0	6,5	5,7	5,0	4,1	5,9	8,7	14,0	14,3	13,0
20	14,0	14,4	13,6	6,0	5,1	5,1	5,8	5,5	9,6	14,0	14,5	15,0
21	13,1	13,2	13,2	6,7	5,6	4,7	4,3	5,6	9,3	14,0	13,2	14,4
22	14,4	14,0	12,1	6,6	5,3	5,0	5,7	5,6	8,7	13,1	14,4	14,3
23	14,1	13,9	12,7	5,1	5,2	4,3	4,7	5,7	8,1	13,1	13,5	14,0
24	14,9	13,0	12,3	5,8	5,6	4,1	4,9	5,6	8,7	13,5	13,9	13,5
25	13,8	13,8	13,6	5,9	5,2	5,5	5,1	5,7	9,0	14,0	14,2	14,7
26	14,6	14,5	12,4	7,0	5,7	5,4	4,2	5,9	8,9	13,4	14,8	14,7
27	13,7	13,1	12,5	7,2	5,8	4,3	4,1	5,6	10,0	14,7	14,8	14,9
28	13,1	14,1	12,4	6,1	5,1	4,9	4,8	5,1	8,4	13,7	13,2	13,6
29	13,9	14,3	12,0	5,4	5,2	4,6		5,2	9,5	13,2	13,0	13,2
30	14,3	14,4	11,2	5,8	5,9	5,2		5,6	10,0	14,5	14,2	13,3
31	13,2		10,1		5,6	5,6		5,8		14,6		14,2
x	14,0	14,0	12,7	7,3	5,6	5,1	5,0	5,5	8,2	12,8	14,0	14,1

Tabla 4.2: Consumos equipamiento IT

	Ago	Sep	Oct	Nov	Dic	Ene	Feb	Mar	Abr	May	Jun	Jul
1	11,1	11,6	11,8	12,6	13,0	11,3	11,7	11,2	11,1	12,3	11,2	11,4
2	11,6	12,0	12,5	11,2	11,8	11,3	11,2	12,3	11,3	12,0	11,0	11,9
3	11,4	11,5	11,5	12,3	12,1	11,0	12,5	11,4	11,7	12,0	11,3	11,1
4	11,8	11,7	11,9	11,9	12,8	11,9	11,0	12,6	12,7	12,2	11,1	11,2
5	11,5	11,2	12,9	11,2	12,5	11,5	12,1	12,5	11,4	11,4	11,0	11,3
6	11,6	11,8	12,9	12,8	11,5	11,4	11,4	11,9	12,3	12,2	11,9	11,2
7	11,3	11,5	11,1	12,4	12,7	12,1	12,1	12,9	12,7	11,5	11,6	11,1
8	11,2	11,8	12,8	13,0	12,0	12,2	11,9	12,7	11,8	11,2	11,7	11,6
9	11,2	11,7	11,7	12,7	11,9	11,4	11,8	12,6	11,8	11,7	11,6	11,6
10	11,6	11,2	11,1	12,9	11,9	11,7	11,1	12,6	11,5	11,4	11,1	11,1
11	11,2	11,4	11,2	12,8	12,6	12,9	12	11,9	11,6	11,4	11,9	12,0
12	11,5	11,8	12,4	11,2	12,3	12,6	12,5	11,2	11,8	12,6	11,4	12,0
13	11,2	12,0	11,8	12,9	11,4	11,4	11,2	12,4	12,5	12,1	11,7	11,4
14	11,1	11,6	12,3	12,8	12,0	12,1	12,2	12,9	11,4	12,5	11,6	12,0
15	11,9	11,2	11,0	11,4	11,7	12,3	12,6	12,6	12,0	12,9	11,3	11,8
16	11,2	11,2	12,7	11,2	11,2	12,6	12,2	13,0	11,3	12,9	11,4	11,3
17	11,5	11,5	11,4	12,9	11,3	11,1	11,4	12,1	12,2	12,0	11,3	11,3
18	11,1	11,8	11,2	11,9	11,6	11,4	12,4	11,6	12,5	12,4	11,3	11,5
19	11,9	12,0	12,5	12,6	11,4	11,2	12,7	12,9	11,6	12,3	11,4	11,2
20	11,3	11,6	11,9	12,1	12,1	12,3	12,7	11,7	11,8	11,2	11,5	11,7
21	11,2	11,0	11,1	12,7	12,3	11,6	12,4	12,1	12,7	12,2	11,5	11,9
22	11,0	11,4	12,3	12,4	12,2	11,8	11,2	12,2	11,6	12,9	11,6	11,5
23	11,4	11,1	12,5	12,3	11,6	12,2	12,2	11,8	12,9	11,7	11,3	11,7
24	11,8	11,2	12,1	12,3	12,2	11,8	12,6	11,6	11,7	12,3	11,5	11,1
25	11,5	11,6	12,1	12,9	12,3	12,4	11,4	11,6	11,8	12,7	11,4	12,0
26	11,1	12,8	11,3	11,8	11,9	12,1	13,0	11,1	13,0	12,5	11,5	11,9
27	11,3	11,9	11,9	11,4	11,0	11,8	12,9	12,8	11,5	11,7	11,2	11,7
28	11,3	13,0	12,8	11,0	11,3	12,1	11,1	12,1	11,5	12,4	11,0	11,4
29	11,7	11,6	12,7	12,0	12,4	11,3		12,0	12,4	11,3	11,8	11,4
30	11,7	12,1	12,1	12,6	12,3	12,0		12,8	11,5	11,1	11,1	11,1
31	11,5		12,7		11,5	11,3		12,7		12,5		11,2
y	11,4	11,6	12,0	12,2	12,0	11,8	12,0	12,0	11,9	12,0	11,4	11,5

Tabla 4.3: Media del consumo de refrigeración

Ago	Sep	Oct	Nov	Dic	Ene
14,0451613	13,9533333	12,7096774	7,31333333	5,64193548	5,06129032
Feb	Mar	Abr	May	Jun	Jul
4,91071429	5,49677419	8,25666667	12,8064516	14,0333333	14,0516129
Media x					
9,85669035					

Tabla 4.4: Media del consumo del equipo de IT

Ago	Sep	Oct	Nov	Dic	Ene
11,4096774	11,66	12,0064516	12,2066667	11,9612903	11,8096774
Feb	Mar	Abr	May	Jun	Jul
11,9821429	12,1870968	11,92	12,0483871	11,4066667	11,5032258
Media y					
11,8417736					

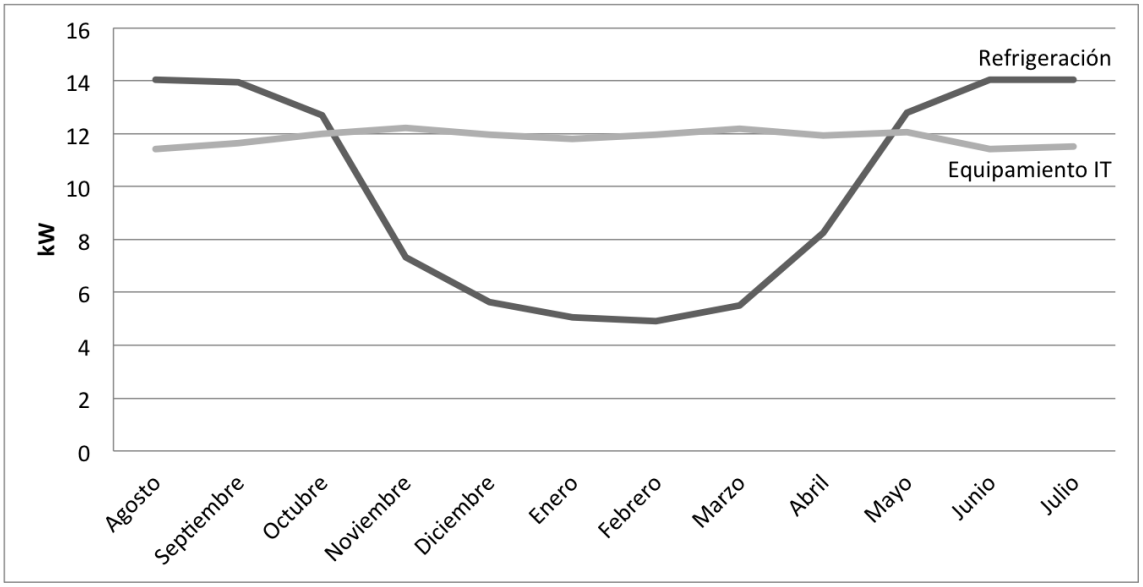

Figura 4.1: Consumos CPD

CAPÍTULO 5

Impacto de la virtualización

En este capítulo se explican los motivos que llevaron a la empresa a realizar un cambio drástico en la infraestructura y se realiza un experimento a pequeña escala para comparar los mismos servicios en una máquina física y en otra virtualizada.

5.1 Necesidad de la virtualización

En el CPD del estudio antes de tener la infraestructura virtualizada se había recurrido a una solución más convencional con servidores físicos. Las causas que forzaron el cambio fueron diversas pero entre ellas habían limitaciones insalvables:

- En el recinto del CPD no cabían más armarios de TI.
- En los armarios existentes no cabían más máquinas.
- Era imposible desde el punto de vista económico ampliar el recinto del CPD.
- En los meses calurosos el sistema de refrigeración a penas era capaz de mantener la temperatura adecuada.
- En muchos casos varios servicios compartían la misma máquina física y como hemos comentado antes esto acarreaba problemas.
- El consumo total del CPD empezaba a ser un problema económico tanto del equipamiento de TI como para el equipamiento de refrigeración.
- Alguna de las máquinas que se utilizaban habían dejado de venderse o de recibir soporte hardware lo que implicaba comprar equipamiento cada pocos meses.

Figura 5.1: Vista frontal HP BL480c g1

- Debido a esto último el mantenimiento del sistema operativo se hacia tedioso puesto que se tenían que mantener diversas versiones de varios sistemas operativos para asegurar la compatibilidad.

5.2 El experimento

Las máquinas más modernas a las que aún se les podía dar algún uso se almacenaron para fines de contingencia o laboratorio. Para poder realizar una comparación cuantitativa entre virtual y real se realizó un experimento a pequeña escala con una máquina que fue cedida por la empresa para el tiempo que durase el experimento. La escala del experimento está marcada básicamente por la limitación en cuanto a máquinas físicas de las que podía disponer. En el momento del experimento solo quedaba un servidor en blade libre y esto limitaba la escala. Además debido al propio formato de estos servidores, para realizar la prueba éste debía ir conectado al chasis que aloja todos los demás servidores de producción por lo que se debían extremar las precauciones. En cuanto al otro servidor y debido al esquema de red que se quería reproducir para el experimento, y por motivos de política de seguridad, éste no podía salir del CPD por lo que se debía llegar a una solución para automatizar el muestreo.

En concreto las máquinas que se compararon fueron:

- Dell PowerEdge 2950 (ver figura 5.3)
- HP BL480c g1 (ver figura 5.1 y 5.2)

La elección de estas máquinas no es aleatoria, ambas máquinas ofrecen sobre el papel las mismas prestaciones y poseen las mismas características. La Dell PowerEdge 2950 es una de las últimas máquinas que se compraron del formato más

Figura 5.2: Servidor HP BL480c g1 sin la tapa

tradicional de 2U y la HP BL480c g1 pertenece al primer chasis de blades que se adquirió. El experimento consiste en someter a ambas máquinas a una carga genérica exactamente igual al mismo tiempo y medir sus consumos cada 5 minutos a lo largo de una prueba de 1 hora.

Sus características generales son:

- Procesador: Intel Xeon Processor X5365
- RAM: 64 GB ECC
- Almacenamiento local: RAID 1 SAS 15k HDD
- Almacenamiento externo: RAID 5 SATA 10k HDD

Para tener una mejor referencia se han comparado los resultados publicados por SPEC, concretamente utilizando las medidas SPECint_rate2006 y SPECint_rate_base2006. Hay que tener en cuenta que el hardware de las dos máquinas que se comparan en este estudio es exactamente igual que las máquinas comparadas en el benchmark del SPEC excepto por la cantidad de memoria que en el caso del SPEC es de 16 GB en vez de 64 GB. Los resultados del SPEC, publicados en su página oficial, para estas dos máquinas son:

HP BL480c

SPECint_rate2006 = 116 SPECint_rate_base2006 = 97,5

Dell PowerEdge 2950

SPECint_rate2006 = 112 SPECint_rate_base2006 = 99,1

Para la prueba se instaló en el DELL el sistema operativo elegido corporativamente, CentOS 6.6. Y en el blade de HP el sistema operativo de vmWare, el ESXi sobre el que correrán las máquinas virtuales. Estas máquinas virtuales correrán el mismo operativo que el servidor físico en la misma versión y con toda la paquetería igual.

Sobre el esx correrán 2 máquinas virtuales: una para DNS y correo y otra para Nagios y SNORT.

Sobre el servidor físico se instalarán una serie de servicios que se usan dentro de la empresa, estos son:

- Servidor de DNS (bind)
- Servidor de correo (postfix)
- Nagios
- SNORT

De estos servicios ya se han explicado Nagios y SNORT, a continuación de explica brevemente que son bind y postfix.

- El servidor de DNS bind [1] (*Berkeley Internet Name Domain*) es el servidor DNS estándar en sistemas Linux. Posee características como autosincronización con servidores DNS secundarios, gestión de zonas, etc.
- Postfix [3] es un servidor de correo para el enrutamiento y envío de correo electrónico de código libre. Es el servidor de correo por defecto en la mayoría de distribuciones de Linux.

Para la prueba se creó ex profeso una carga masiva de datos para inyectar y procesar a través de la interfaz promiscua. Además se añadieron una serie de hosts y servicios falsos todos apuntando a la propia máquina con un check especial de Nagios que devuelve la versión de la aplicación llamado check_dummy. Este check se conecta a la intefaz de loopback y siempre genera un OK. La interfaz de loopback es la interfaz de red que conecta con el propio dispositivo. Si bien es cierto que esta carga no representa la realidad, la infraestructura necesaria para realizar una prueba de esfuerzo real no es accesible y generar una carga elevada de checks falsos es la mejor aproximación que se ha podido realizar con los recursos de los que se dispone. Aun así no hay que olvidar que lo que se pretende es realizar una simulación, lo más aproximada posible a la realidad pero no se pretende medir el rendimiento individual sino comparar el de dos máquinas.

A lo largo de la hora que dura la prueba se repite la inyección de datos en la interfaz promiscua y se lanzarán los checks con una frecuencia de 1 minuto. No hay que olvidar que todos los datos recogidos por la interfaz son escritos en una base de datos optimizada para cada máquina al igual que el resultado de los checks a través de ndo2db. Además se realizarán peticiones aleatorias a archivos publicados vía web con una frecuencia de 20 por segundo, intentado simular una carga de un grupo pequeño de administradores usando la plataforma.

El consumo de ambas máquinas se recoge con una resolución de 5 minutos a lo largo de 1 hora. En el caso del blade de HP la recolección de los datos es

Figura 5.3: Servidor Dell PowerEdge 2950

muy sencilla puesto que la gestión del chasis ya provee de esta información en vivo o incluso en gráficas históricas. En el caso del servidor de Dell que cuenta con 2 fuentes de alimentación de 750 W conectadas a una PDU con protección y que incluye un visor en el que ver el consumo actual. Esta PDU además se puede conectar a la red y se monitorizó vía SNMP desde la infraestructura de monitorización de la empresa y se recogieron los datos al finalizar la prueba.

Una PDU (*Power Distribution Unit*) es un dispositivo con diversas salidas diseñado para distribuir la corriente eléctrica entre varias máquinas. En muchos casos estos dispositivos poseen características como protección contra sobretensiones o monitorización de tensión y consumos. En el CPD del estudio cada rack tiene dos circuitos eléctricos redundantes cada uno conectado a una PDU. El modelo de PDU usado solo es capaz de dar información del consumo total, esto fue un problema puesto que todas las PDU de todos los racks tenían más de una máquina enchufada y el tiempo de la prueba tuvo que posponerse hasta que se pudo liberar una PDU.

En la tabla 5.1 pueden verse los resultados de la prueba para las dos máquinas. Las tablas representan el consumo puntual en vatios cada 5 minutos a lo largo de la prueba de 1 hora para las dos máquinas y finalmente la media de los consumos para poder realizar una comparación cuantitativa entre ambas. Recordemos que las dos máquinas poseen características casi idénticas y sus resultados en el benchmark SPEC son prácticamente los mismos, por lo que en teoría deberían dar un resultado muy similar.

Como puede verse en la tabla 5.1 de los dos servidores, el blade de HP consume 38 W menos de media ofreciendo exactamente los mismos servicios y procesando exactamente los mismos datos. Hay que tener en cuenta que determinadas funciones como la conectividad a través del conmutador interno del chasis no cuentan en el computo de energía y aunque ha resultado imposible calcular el impacto que tendrían sobre la prueba podríamos aproximarlo. Teniendo en cuenta que:

Tabla 5.1: Consumos durante el experimento

Servidor	Dell PowerEdge 2950	HP BL480c
5	312	252
10	335	285
15	345	320
20	337	318
25	315	265
30	301	313
35	350	300
40	324	311
45	332	273
50	309	251
55	334	256
60	299	292
Media	324.4166667	286.33333

- El chasis de blades de HP cuenta con 16 blades que comparten los conmutadores de fibra y de ethernet.
- El consumo máximo de energía del chasis es de 7950 W con los 16 blades conectados y las 6 fuentes de alimentación de 2650 W configuradas en pares para tolerancia a fallos.
- Si tenemos en cuenta que los dos servidores comparados tienen el mismo rendimiento se necesitarían 16 Dell PowerEdge para igualar la capacidad de computo del chasis entero, con la diferencia de que cada Dell es capaz de consumir 750 W máximo que multiplicado por las 16 máquinas da un total de 12000 W, casi el doble que el chasis.

Con estas consideraciones en cuenta queda claro que el rendimiento energético del chasis es muy superior a una instalación homologa a base de servidores tradicionales.

Hay que recordar que por la forma de trabajar de esta empresa es necesario desplegar una serie de sondas y de servidores de tratamiento de datos en los clientes. Este despliegue es mucho más sencillo gracias a la virtualización. Inicialmente este despliegue suponía la inclusión de una nueva máquina física en la infraestructura del cliente aplicando obviamente las políticas de seguridad del cliente. En algunos casos esto era un problema puesto que las preferencias del cliente podían no ser fáciles de cumplir en cuanto al tipo de máquina, sistema operativo, medidas de seguridad o incluso fabricante del hardware.

En la mayoría de los casos el cliente ya disponía de algún tipo de plataforma de virtualización sobre el que poder desplegar las máquinas, las cuales pueden

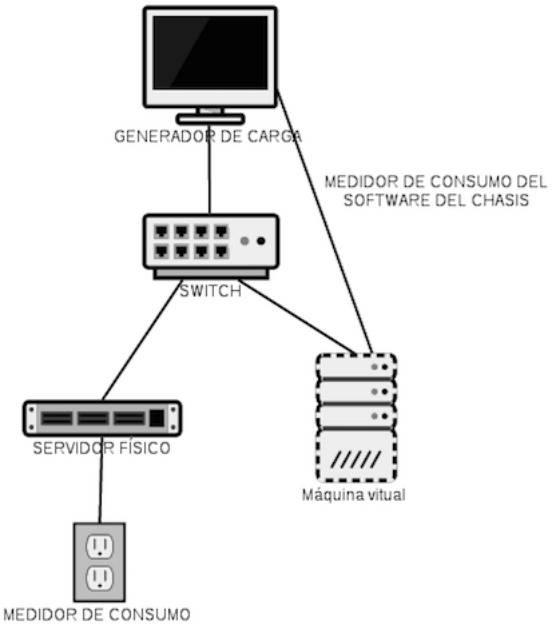

Figura 5.4: Esquema del experimento

ser fácilmente exportadas a casi cualquier formato de máquina virtual. En los casos en los que esto no era posible solía ser más fácil desplegar sobre el hardware que el cliente elija el sistema operativo de vmWare que proporciona una capa de abstracción entre el hardware y la máquina virtual.

5.3 Mejoras propuestas

Después de analizar en profundidad el sistema se han seleccionado una serie de mejoras aplicables al sistema. Algunas de ellas influyen en la organización del propio centro de datos y otras en la organización de la infraestructura de virtualización.

Por una parte las que influyen en la organización del centro de datos se refieren a como están organizadas espacialmente las máquinas y los elementos de refrigeración, para los cuales existen una serie de buenas practicas [11]. El CPD del estudio consta de diez racks, 4 de ellos de comunicaciones, 5 de sistemas y el décimo es el SAI. La organización de los racks responde al diseño de "pasillo frío y pasillo caliente". Esta organización lo que pretende es organizar las máquinas enfrentadas en filas de racks dejando un pasillo en medio donde se crea una corriente de aire frío que las máquinas absorben y expulsan por detrás calentado. A la zona trasera de las filas se le denomina pasillo caliente puesto que el aire calentado que expulsan las máquinas fluye por este pasillo hasta el retorno de las máquinas de refrigeración. Este diseño es muy eficiente, y en un CPD con solo un pasillo frío debería dar muy bien resultado, si se lleva a cabo correctamente. En

el CPD del estudio hay tres problemas que reducen la eficiencia del sistema de refrigeración:

1. Para que esta distribución funcione con máquinas de aire acondicionado de tipo *split* hay que tener en cuenta que los racks tengan las puertas, tanto la trasera como la delantera, fabricadas en malla metálica perforada para que el aire pueda entrar y recorrer el interior fácilmente antes de salir calentado por la parte trasera. En el caso de los racks de comunicaciones que contienen los conmutadores y routers, la puerta frontal es de cristal y la trasera es una plancha de metal no perforado por lo que no dejan que el aire frío entre en el rock y enfríe la electrónica. Por ello las puertas permanecen siempre abiertas. Esta es la primera mejora que se puede realizar: sustituir estos armarios por unos adecuados al diseño del CPD.

2. A raíz del problema anterior se añadió la una máquina de aire acondicionado más. El problema fue que por espacio no pudo situarse siguiendo el diseño inicial del CPD y se colocó detrás de los armarios de comunicaciones rompiendo el flujo de aire del centro del CPD. Este problema tiene una solución compleja puesto que implicaría cambiar maquinaria e incluso obra en algún muro del CPD.

3. El tercer problema se ha ido agravando con el tiempo. A medida que se han ido añadiendo más máquinas el cableado en la parte trasera de los racks ha ido aumentando y no se ha ordenado como es debido resultando en una barrera de cables que frena la salida del aire caliente que desprenden las máquinas. La solución a este problema es sencilla aunque tediosa y debe ser planificada cuidadosamente para no comer errores al desconectar el cableado. La reorganización del cableado debería tener en cuenta las capacidades de los racks de organizarlo y el suelo técnico del CPD por el cual se puede organizar mejor el posible excedente en longitud de los cables.

Por otra parte, y aunque mucho menos grave, hay alguna mejora que se podría realizar en la plataforma de virtualización. Es una mejora sencilla de aplicar aunque requiere de una buena planificación. A medida que la infraestructura ha ido creciendo se han creado decenas de máquinas virtuales nuevas que se han asignado a servidores dentro del chasis con relativa aleatoriedad. Si bien es cierto que si la máquina nueva es una máquina crítica, sí que se han tenido algunas consideraciones a la hora de asignarla a un esx, o incluso en algún caso especial un esx solo contiene una única máquina virtual, en otros casos las máquinas han sido asignadas sin demasiada consideración. Una posible mejora sería realizar un estudio de requisitos de las máquinas virtuales que corren en la infraestructura para intentar repartirlas más adecuadamente entre los diferentes esx.

CAPÍTULO 6

Conclusiones

La conclusión más evidente a la que se llega después de leer este trabajo es que la virtualización supone una gran ventaja a la hora de gestionar los recursos de los que se disponen lo cual es muy importante para una empresa que intenta reducir al máximo su consumo energético. Dada la amplia variedad de soluciones a la hora de diseñar un sistema es necesario realizar un estudio extenso para poder decidir cual usar. Como se ha explicado durante el trabajo las técnicas que proporciona la virtualización, aplicadas correctamente en un entorno adecuado, pueden constituir ventajas decisivas para decantarse por esta solución ya sea propietaria como vmWare o de código abierto como KVM. Este entorno, desde los servicios a la plataforma hardware, debe ser analizada en detalle para que todas las piezas que acaban formando el sistema casen a la perfección. En el caso particular de una empresa dedicada a la monitorización informática como parte de una solución de seguridad, utilizar la virtualización supone facilitar enormemente el despliegue de las sondas necesarias para llevar a cabo la recolección y el tratamiento de la información, tanto en la propia infraestructura como en infraestructuras en clientes.

Volviendo al caso del ahorro energético, se ha demostrado, con el experimento a pequeña escala, que se pueden ofrecer los mismos servicios consumiendo menos energía eléctrica con una buena planificación a la hora de repartir estos servicios en diferentes máquinas y mediante el uso de una plataforma de virtualización bien planteada. Este ahorro energético no es solo beneficioso a nivel económico para la empresa, sino que contribuye a mejorar la utilización de los recursos energéticos del planeta, incluso a reducir la huella ecológica al utilizar menos hardware.

Como reflexión final, después de todo el estudio realizado y después de haber trabajado con la infraestructura durante un año y medio hemos sido capaces de apreciar el enorme beneficio que una infraestructura virtualizada bien diseñada puede suponer para los grandes centros de datos que tanto proliferan hoy en día.

Bibliografía

[1] Albitz, Paul y Liu, Cricket. *DNS and BIND.* O'Reilly, 2001.

[2] Victor Avelar (Schneider Electric), Dan Azevedo (Disney), Alan French (Emerson Network Power). *PUE: A Comprehensive Examination of the Metric*. The Green Grid. October, 2012.

[3] Dent, Kyle. *Postfix : the definitive guide.* O'Reilly, 2004.

[4] Douglas R Mauro, Kevin J Schmidt. *Essential SNMP 2nd Edition.* O'Reilly Media. September, 2005.

[5] Brad Ewing, Anders Reed, Alessandro Galli, Justin Kitzes, Mathis Wackernagel. *Calculation methodology for the national Footprint accounts*. Global Footprint Network October, 2010.

[6] Greenberg, S., Mills, E., Tschudi, W., Rumsey, P., and Myatt, B. *Best Practices for Data Centers: Lessons Learned from Benchmarking 22 Data Centers. Proceedings of the 2006 ACEEE Summer Study on Energy Efficiency in Buildings.*

[7] Tim Jones, Senior Principal Software Engineer, Emulex Corp. *La anatomía de un hipervisor Linux*. Consultado en `http://www.ibm.com/developerworks/ssa/library/l-hypervisor/index.html`. IBM May, 2009.

[8] N. Kammenhuber. *Traffic-Adaptive Routing, Chapter 6.2* . Technische Universität München December, 2007.

[9] Wojciech Kocjan. *Learning Nagios 3.0.* October, 2008.

[10] Peter Membrey, Tim Verhoeven, Ralph Angenendt. *The Definitive Guide to CentOS.* Apress, 2009.

[11] Parise, G. ; Parise, L. *IEEE Transactions on Industry Applications, Volume: 49, issue: 4.* April, 2013.

[12] Ros, Josep. *Virtualización corporativa con VMware.* Torredembarra, 2009.

[13] Ellen Siever, Arnold Robbins, Stephen Figgins, Robert Love. *Linux.* Anaya Multimedia, 2010.

[14] Derek Vadala *Managing RAID on Linux.* O'Reilly, 2002.

[15] Sander van Vugt *Pro Linux High Availability Clustering.* Apress, 2014.