



Universitat Oberta
de Catalunya

MIGRACIÓN A EUCALYPTUS

INFRAESTRUCTURA COMO SERVICIO (IAAS)

MEMORIA PROYECTO FINAL MASTER OFICIAL DE SOFTWARE LIBRE
(UNIVERSITAT OBERTA DE CATALUNYA)

ADMINISTRACIÓN DE REDES Y DE SISTEMAS OPERATIVOS
Curso 2010-2011

ALUMNO: Alberto Castillo Mancera (acastilm@gmail.com)

CONSULTOR UOC: Joaquín López Sánchez-Montañés

CONSULTOR EXTERNO: Jordi Massaguer Pla (openTrends Soluciones y Sistemas)



LICENCIA

Este trabajo se publica bajo la licencia CC-BY-SA, disponible de forma completa en la dirección web: <http://creativecommons.org/licenses/by-sa/2.0/legalcode>

En esta otra dirección (<http://creativecommons.org/licenses/by-sa/3.0/deed.es>) se muestra un resumen muy comprensible de la licencia, que reproduzco aquí:

- ***Usted es libre de:***
 - Compartir - copiar, distribuir, ejecutar y comunicar públicamente la obra
 - Hacer obras derivadas

- ***Bajo las condiciones siguientes:***
 - Atribución — Debe reconocer los créditos de la obra de la manera especificada por el autor o el licenciante (pero no de una manera que sugiera que tiene su apoyo o que apoyan el uso que hace de su obra).
 - Compartir bajo la Misma Licencia — Si altera o transforma esta obra, o genera una obra derivada, sólo puede distribuir la obra generada bajo una licencia idéntica a ésta.

- ***Entendiendo que:***
 - Renuncia — Alguna de estas condiciones puede no aplicarse si se obtiene el permiso del titular de los derechos de autor
 - Dominio Público — Cuando la obra o alguno de sus elementos se halle en el dominio público según la ley vigente aplicable, esta situación no quedará afectada por la licencia.
 - Otros derechos — Los derechos siguientes no quedan afectados por la licencia de ninguna manera:
 - Los derechos derivados de usos legítimos u otras limitaciones reconocidas por ley no se ven afectados por lo anterior.
 - Los derechos morales del autor;
 - Derechos que pueden ostentar otras personas sobre la propia obra o su uso, como por ejemplo derechos de imagen o de privacidad.
 - **Aviso** — Al reutilizar o distribuir la obra, tiene que dejar muy en claro los términos de la licencia de esta obra. La mejor forma de hacerlo es enlazar a esta página.

Presentación del Proyecto Final de Master

Este trabajo es la Memoria del Proyecto Final del Master Oficial de Software Libre, impartido por la UOC (Universitat Oberta de Catalunya).

El Proyecto lo comprenden 2 asignaturas:

- **TFM - Adm. de redes y de sistemas operativos:** Trabajo Fin de Master, en el que se aplican conocimientos en gestión de proyectos, adquiridos a lo largo del Master Oficial en Software Libre.
- **Pr. ext. - Adm. de redes y de sistemas operativos:** Trabajo práctico desarrollado bajo la ayuda y supervisión de una empresa externa, en el que se aplican conocimientos técnicos adquiridos en el resto de asignaturas cursadas.

TFM: Proyecto de migración a Cloud Computing (Eucalyptus)

Partiendo de un proyecto ficticio, que nos ha sido encargado por un cliente, se pretende realizar un documento en el que se recoja, con cierto detalle, las fases para llevar a cabo el proyecto, incluyendo todo el ciclo de vida hasta su implantación.

La tecnología central utilizada en este proyecto es lo que se conoce como Infraestructura como Servicio (o por las siglas IAAS en inglés), que se engloba dentro de un término más genérico como es Cloud Computing (o Computación en la Nube).

Hoy en día existen diversas tecnologías de virtualización que nos permiten ejecutar varias máquinas virtuales (cada una con su propio sistema operativo) compartiendo un mismo hardware físico (memoria, CPU, dispositivos E/S,...), y además, de forma simultánea. Gracias a ello, pueden aprovecharse mejor los recursos hardware de que dispone una organización o persona particular.

Las IAAS aprovechan la capacidad de virtualización para crear infraestructuras flexibles y escalables en la “nube”. Recibe el nombre de “nube” porque para el usuario debe ser transparente dónde se está ejecutando una determinada máquina virtual.

En este Trabajo Fin de Master se aborda la migración de la infraestructura de los sistemas informáticos de una empresa hacia un modelo IAAS.

Mi tutor en esta asignatura ha sido Joaquín López Sánchez-Montañés, consultor de la UOC.

Pr. ext.: Práctica externa: Demo Live CD de Eucalyptus.

Esta práctica externa ha sido desarrollada con la supervisión y apoyo de la empresa colaboradora openTrends Soluciones y Sistemas, teniendo asignado como tutor externo a Jordi Massaguer Pla.

En línea con el Trabajo de Fin de Master (TFM), el objetivo de esta práctica es construir un Live CD de demostración del software de gestión de nubes privadas Eucalyptus **[R1]**.

El Live CD debe mostrar en ejecución las principales funcionalidades de Eucalyptus. Para ello, al arrancar un ordenador personal desde el lector de CD se inicia de forma automática una aplicación web, donde el usuario puede ver información de la nube configurada en esta demo (clusters, imágenes de máquinas virtuales y el estado de las instancias) y probar a iniciar o parar instancias de la imagen de máquina virtual que previamente ha sido añadida a la nube.

El Live CD será desarrollado utilizando la herramienta web SuseStudio **[R2]**, y una vez completado se publicará en Suse Gallery **[R3]**, galería de versiones de OpenSuse que usuarios registrados publican para compartirlas. Cualquier usuario puede descargar aquellas que le interesen, incluso probarlas online con la herramienta TestDrive.

Se aporta, como documento anexo [A1], la guía de uso del Live CD.

Índice de contenido

MIGRACIÓN A EUCALYPTUS.....	1
LICENCIA.....	2
Presentación del Proyecto Final de Master.....	3
TFM: Proyecto de migración a Cloud Computing (Eucalyptus)	3
Pr. ext.: Práctica externa: Demo Live CD de Eucalyptus.....	4
1. Introducción al Cloud Computing.....	6
2. Presentación del cliente.....	8
3. Estudio de viabilidad.....	9
3.1. Necesidades y requisitos del cliente.....	9
3.2. Análisis de la situación actual.....	10
3.3. Definición de requisitos del sistema.....	11
3.4. Estudio de alternativas de solución.....	12
3.5. Valoración y elección de las posibles soluciones.....	13
4. Análisis del sistema.....	15
4.1. Definición del sistema.....	15
4.2. Requisitos del proyecto.....	16
4.3. Definición de interfaces de usuario.....	18
4.4. Especificación del plan de pruebas	18
4.4.1. Pruebas unitarias.....	18
4.4.2. Pruebas de integración.....	19
5. Diseño del sistema.....	20
5.1. Arquitectura.....	20
5.1.1. Modelo Eucalyptus.....	20
5.1.2. Opciones de virtualización.....	24
5.2. Especificación de estándares, normas de diseño y construcción.....	25
5.3. Especificaciones de desarrollo y pruebas.....	26
5.4. Requisitos de implantación.....	27
6. Desarrollo e Implantación del sistema.....	28
6.1. Primera fase.....	28
6.2. Segunda fase.....	29
6.3. Tercera fase.....	30
6.4. Cuarta fase.....	31
6.5. Formación.....	32
7. Planificación del Proyecto.....	34
7.1. Fases.....	34
7.2. Equipo humano.....	35
7.3. Costes del proyecto.....	35
CONCLUSIONES.....	37
REFERENCIAS.....	39
ANEXOS.....	39

1. Introducción al Cloud Computing

En este primer capítulo vamos a explicar de forma breve qué es el Cloud Computing, para que nos permita entender mejor el proyecto de IT que se desarrollará en esta Memoria.

Según la definición de Wikipedia **[R4]**, el término Cloud Computing hace referencia a la computación independiente de su localización física, donde servidores compartidos proporcionan, bajo demanda, recursos, software y datos a computadores y otros dispositivos.

Eucalyptus define el Cloud Computing de forma similar: es la exposición de recursos de computación a través de una interfaz de servicio web según son necesitados.

El término “cloud” o “nube” se refiere al hecho de que la infraestructura física subyacente permanece opaca para los usuarios finales. Es también por ello que habitualmente se representa Internet como una nube.

Dentro del Cloud Computing hay que distinguir 3 conceptos:

- **Software como Servicio:** se trata de software accesible por Internet (disponible en la “nube”), que podemos utilizar sin necesidad de ser instalado en nuestro computador. Un ejemplo claro es Google Docs, donde podemos crear documentos en la nube con sólo usar un navegador web.
- **Plataforma como Servicio:** plataformas configuradas con todo el software necesario para un determinado uso. Es habitual que su destino sea el utilizarlas como entorno de desarrollo o ejecución de aplicaciones.
- **Infraestructura como Servicio:** utiliza la virtualización de recursos hardware para proporcionar infraestructuras flexibles y escalables en la “nube”. El modo de facturación de las empresas que proporcionan estos servicios suele estar basado en los recursos utilizados y el tiempo de uso. Uno de los casos de mayor éxito es Amazon EC2.

Sobre el Cloud Computing puede hacerse una clasificación en 3 categorías, según el tipo de nube:

- **Nube Pública:** en lugar de encontrarse en las instalaciones de la organización que los utiliza, los recursos de computación se hallan en instalaciones de terceros, en Internet. En este caso el cliente paga a la empresa proveedora por el uso que haga de esos recursos. Puede entenderse como un alquiler de recursos hardware. Su ventaja radica en que se dispone de una alta escalabilidad, pudiéndose añadir a la infraestructura tantas máquinas virtuales como se necesiten, dando la sensación de poseer recursos infinitos. Los mayores inconvenientes de este tipo de nubes son su alto coste si su uso es continuado en el tiempo, y la privacidad de los datos, que no puede garantizarse.

- **Nube Privada:** en este caso, los recursos de computación son propiedad de la organización que los utiliza, y se encuentran dentro de su Intranet. Consigue que la infraestructura IT sea más eficiente y flexible, a la vez que, al encontrarse en un entorno controlado, puede proporcionar un alto nivel de seguridad sobre los datos sensibles. Tiene el inconveniente de que su escalabilidad queda limitada por los recursos que posee la organización.
- **Nube Híbrida:** este tipo de nube trata de dar respuesta a la limitación de escalabilidad de las nubes privadas. Son nubes privadas que, cuando es necesario, añaden nuevos recursos que se encuentran fuera de su ámbito, alquilándolos a terceros. Es una buena opción cuando se requiere un alto número de recursos por un periodo de tiempo limitado, por ejemplo ante picos de demanda de un determinado servicio.

2. Presentación del cliente

ApuestaONline, es una empresa cuyo principal negocio es ofrecer una plataforma web de apuestas sobre todo tipo de deportes.

Se posee una base de datos de usuarios registrados que pueden apostar por un determinado resultado en un evento deportivo, como puede ser un partido de tenis o una carrera de motociclismo, de manera que los resultados menos probables se pagarán más caros si se producen.

Los usuarios pueden acceder al catálogo de apuestas de la semana, y apostar por uno de los resultados posibles. El precio al que se paga un determinado resultado va variando en tiempo real según el número de personas que se decantan por él.

De cada apuesta, la empresa se queda con un porcentaje del importe.

El equipo humano que trabaja actualmente en la empresa lo componen 5 personas con diferentes roles. Sólo una de ellas posee un perfil técnico, y tiene el rol de Administrador de Sistemas y Red.

ApuestaONline cuenta con su propio sistema informático compuesto por las estaciones de trabajo de sus empleados y los servidores donde se alojan el portal web y las aplicaciones internas de la empresa. El hardware que posee es 'en propiedad', y todo el sistema es administrado por un empleado.

3. Estudio de viabilidad

3.1. Necesidades y requisitos del cliente

El negocio de ApuestaONline está en auge. La base de clientes, se está incrementando de manera notable, y cada vez éstos son más activos en la plataforma. Por ello surge la necesidad de escalar la infraestructura para hacer frente a la demanda actual y futura.

En algunas ocasiones se sufre caída de los servicios en momentos de picos de actividad, coincidentes con jornadas en las que se disputan importantes eventos deportivos. La nueva infraestructura debe ser altamente escalable, y capaz de soportar picos sin dejar de dar servicio, ni de que el rendimiento de éste decaiga.

El objetivo del proyecto es realizar el análisis, diseño e implantación de una nueva arquitectura IT altamente escalable y flexible sin que suponga un alto coste económico de mantenimiento. Se persigue que esta infraestructura pueda adaptarse a las necesidades de la empresa en cada momento.

Los servicios informáticos que se tienen en la actualidad deben mantenerse. Así mismo, no se contempla a corto plazo la migración de las estaciones de trabajo (que utilizan en su mayoría software privativo) a soluciones libres, pero la nueva infraestructura debe permitir este cambio en un futuro. Igualmente debe permitir el crecimiento del número de estaciones de trabajo, ya que se contempla la contratación de más personal.

Se valorará positivamente que la solución propuesta aproveche al máximo el hardware existente, teniendo muy en cuenta el coste, que debe minimizarse tanto como sea posible.

En la medida de lo posible debe usarse software libre.

Dado que los ingresos de ApuestaONline provienen únicamente de la plataforma web, ésta debe tener disponibilidad 24x7 y alta disponibilidad. Sólo se contempla la parada del servicio para la realización de tareas de mantenimiento muy concretas, cuando sea indispensable suspenderlo.

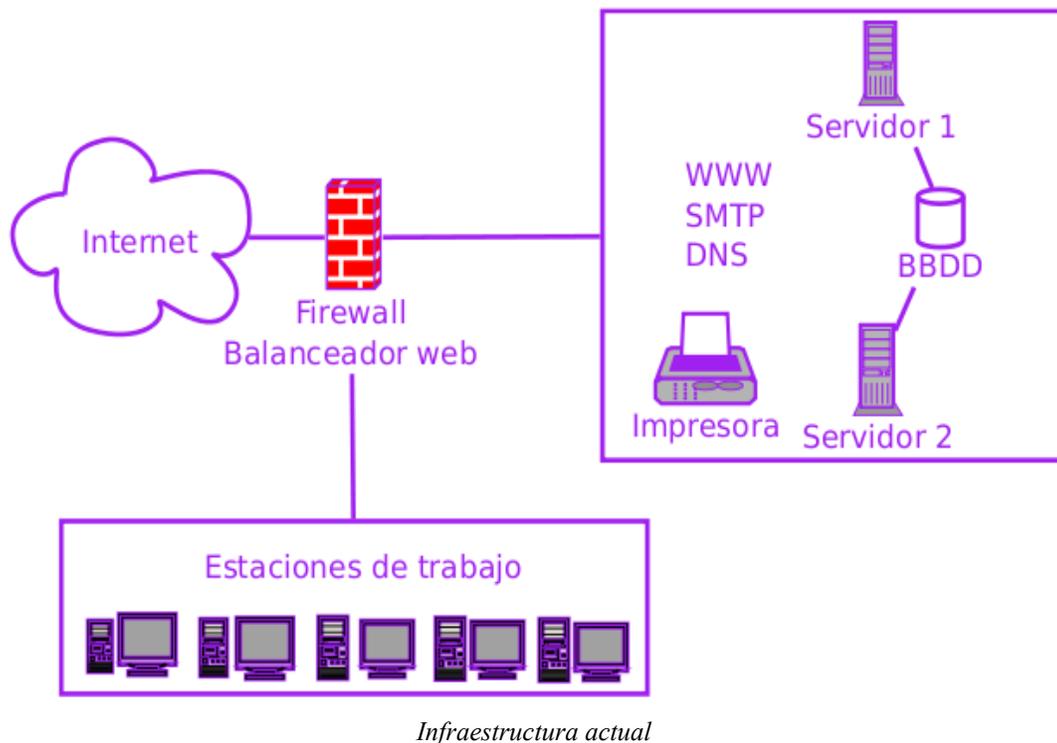
La seguridad de la información y de los sistemas es otro de los puntos claves. Entre otras cosas, las páginas privadas que visiten los usuarios (clientes) deben ser seguras y los servidores deben estar protegidos frente a posibles ataques.

Hay que tener presente que los sistemas guardan información de los clientes, y debe cumplirse en todo momento la LOPD (Ley Orgánica de Protección de Datos).

3.2. Análisis de la situación actual

La infraestructura informática con la que cuenta ApuestaONline para ofrecer sus servicios es muy básica y claramente deficitaria para atender adecuadamente a la demanda actual.

ApuestaONline dispone de la siguiente infraestructura hardware:



Utiliza un firewall (cortafuegos) de “tres patas” para establecer una zona desmilitarizada (DMZ) donde se ubican los servidores web.

Las estaciones de trabajo se encuentran en una LAN protegida por el firewall.

El hecho de que la aplicación web de gestión interna y su base de datos se encuentren en la zona desmilitarizada constituye un riesgo de seguridad si un atacante consigue adentrarse en esta zona.

A continuación se muestra el inventario del hardware que posee ApuestaONline:

- 2 Servidores con 4GB de memoria RAM, 2,40GHz y 400 GB de espacio de almacenamiento (disco duro).
- 1 Servidor SGBD de 4GB de memoria RAM, 2,40 GHz y 1TB de espacio de almacenamiento (disco duro).
- 1 equipo que actúa de firewall entre Internet y la red local, y además balancea las peticiones HTTP dirigidas al portal web hacia los 2 servidores con el objetivo de conseguir Alta Disponibilidad.
- 5 PCs, conectados a la red local, utilizados por los usuarios para realizar su labor diaria, cuyas características principales son:
 - 2GB de memoria RAM.
 - 1,85 GHz de velocidad de procesador.
 - 120 GB de espacio total en disco.
 - Sistema Operativo Microsoft Windows XP.
- Impresora láser

La máquina que actúa como firewall funciona con el sistema operativo Debian Lenny. Las máquinas donde se encuentran los servidores de aplicaciones y el sistema gestor de base de datos funcionan bajo el sistema operativo Ubuntu Server.

Los empleados disfrutan de los siguientes servicios:

- Acceso a internet.
- Servicio de impresión en red.
- Unidad compartida para almacenamiento de ficheros.
- Acceso a aplicación web que les permite gestionar el portal web de apuestas (aplicación backoffice).
- Servicio de correo electrónico, accesible también desde Internet.
- Procesos de backup del sistema de ficheros compartido, así como de las bases de datos de clientes y de la gestión interna de la empresa.

Los clientes disponen de las siguientes funcionalidades:

- Registro online en la plataforma.
- Acceso a área privada (SSL).
- Acceso a Foros.
- Sistema de apuestas, con pasarela de pago.

3.3. Definición de requisitos del sistema

De la información proporcionada por el cliente, podemos enumerar los requisitos que debe cumplir la solución a adoptar:

1. Mantenimiento de servicios actuales: todos los servicios informáticos que se prestan en la actualidad, tanto a clientes como a empleados, deben mantenerse. Entre ellos:
 - servidor de correo electrónico
 - servidor DNS
 - servidor de impresión
 - portal web y aplicación interna (http)
 - unidad compartida de ficheros
 - backup de ficheros y base de datos.
2. Escalable: en caso de ser necesario deben poder añadirse más máquinas a la infraestructura, para lograr una mayor capacidad de procesamiento y así poder atender más peticiones a los servidores web.
3. Alta Disponibilidad: la nueva arquitectura debe tener Alta Disponibilidad (del orden de 9,99% como mínimo), ya que una caída del servicio a los clientes provoca pérdidas económicas directas.
4. Seguridad: dado que el sistema informático guarda información privada de los clientes, se debe asegurar la confidencialidad de ésta.
5. Económico: la solución debe reducir al máximo el coste de implantación y mantenimiento de la infraestructura informática.

3.4. Estudio de alternativas de solución

El principal requisito y que a la vez constituye el objeto de la migración es implantar una infraestructura escalable que se adapte a las necesidades de ApuestaONline.

Se han contemplado 3 posibles soluciones para conseguir éste y el resto de objetivos marcados por el cliente:

Alternativa 1: incrementar notablemente los recursos hardware propios, adquiriendo nuevas y potentes máquinas suficientes para soportar la demanda actual y también los picos que se producen en determinados momentos.

Análisis de costes:

- **Hardware:** se necesita la adquisición de 2 nuevos servidores de altas prestaciones, con un coste orientativo de $3.500 \times 2 = 7.000$ euros.
- **Software:** todo el software utilizado en servidores es copyleft sin costes de licencia.
- **Implantación:** realizada por personal propio y empresa externa.
- **Mantenimiento:** realizado por personal propio.

Alternativa 2: crear una infraestructura de Public Cloud, implantada íntegramente en máquinas de una empresa proveedora de IAAS (Infraestructura Como Servicio).

Análisis de costes:

- **Hardware:** no es necesaria la adquisición de nuevos servidores. Los que se disponen en la actualidad serán utilizados para alojar la aplicación de gestión interna de la empresa, y tareas secundarias.
- **Software:** todo el software utilizado en servidores es copyleft sin costes de licencia.
- **Implantación:** despliegue sobre Amazon EC2 realizado por empresa externa y personal propio.
- **Mantenimiento:** la administración de la infraestructura es realizada por personal propio. El coste viene determinado por el alquiler de recursos a empresa proveedora (Amazon) **[R5]**.
 - 910\$ por instancia reservada, con contrato de 1 año. A este precio hay que sumarle 0,12\$ por hora utilizada.
 - 0,34\$ por hora para instancias On-Demand.
 - A estos costes hay que sumarle el tráfico (0,15 \$ por GB los primeros 10 TB).

Alternativa 3: se trata de una alternativa mixta, consistente en crear una infraestructura de Private Cloud, desplegada sobre recursos hardware propios. Cuando sea necesaria mayor capacidad de procesamiento puede escalarse la infraestructura contratando el uso de recursos a una empresa proveedora de IAAS. Esta opción da la sensación de contar con recursos infinitos, que estarán disponibles según necesidades.

Análisis de costes:

- **Hardware:** adquisición de 2 nuevos equipos de gama media, con un coste orientativo de $1.300 \times 2 = 2.600$ euros.
- **Software:** todo el software utilizado en servidores es copyleft sin costes de licencia.
- **Implantación:** realizado por empresa externa y personal propio.
- **Mantenimiento:** la administración de la infraestructura es realizada por personal propio. Instancias adicionales ejecutadas en Amazon, según precios publicados en Amazon **[R5]**.

3.5. Valoración y elección de las posibles soluciones

Alternativa 1

La primera alternativa supone una inversión importante en la adquisición de máquinas, que no es asumible de inmediato por nuestro cliente, obligándole a recurrir a financiación externa.

Tiene la ventaja de que la infraestructura es gestionada íntegramente por la empresa, y por lo tanto se tienen el control completo de los sistemas. También puede verse como un inconveniente, ya que la empresa tiene que preocuparse de su gestión contratando a personal con experiencia.

Esta solución no es muy escalable, pues para incrementar los recursos se necesitará siempre adquirir nuevo hardware (CPUs, memoria, capacidad de disco,...).

Alternativa 2

La segunda alternativa proporciona una solución flexible y muy escalable. Por contra, toda la información del cliente queda en manos de la empresa proveedora de los servicios de Cloud Computing. Al ser una Public Cloud, los datos no están del todo protegidos. Otro inconveniente muy a tener en cuenta es que a largo plazo puede resultar una solución costosa.

Alternativa 3

La tercera alternativa es la implantación de una *nube híbrida*, que pretende beneficiarse de ambos modelos descritos en las anteriores. Por un lado se aprovecha el hardware existente en el cliente, utilizando tecnología de virtualización para dar flexibilidad a la infraestructura. Y por otro lado, permite escalar los recursos de forma rápida, alquilando éstos a empresas proveedoras.

Esta alternativa permite que la empresa pueda ofrecer sus servicios con hardware propio, a la vez que puede alquilar nuevos recursos a otra empresa pagando sólo por el tiempo y recursos utilizados (tarificación por tiempo y tipo de recursos).

Ésta es la solución aprobada por nuestro cliente.

Existen varios paquetes software para gestionar una infraestructura de este tipo, como pueden ser Open Nebula **[R6]**, Nimbus **[R7]** o Eucalyptus **[R1]**. Se ha decidido utilizar este último, ya que es el que mayor compatibilidad ofrece respecto a los servicios de Amazon EC2. De hecho, el proyecto Eucalyptus surgió como una alternativa a Amazon y sus servicios web son compatibles con el API de Amazon.

Eucalyptus es publicado bajo una licencia libre (GPL), por lo que el coste en licencias es nulo. Además recibe apoyo de varias empresas, como HP que ha optado por este software para ofrecer una oferta integral hardware+software+servicios. También Canonical ha integrado Eucalyptus en su popular distribución GNU/Linux Ubuntu.

4. Análisis del sistema

4.1. Definición del sistema

Las estaciones de trabajo se encuentran protegidas detrás un firewall, constituyendo una red de área local (LAN).

Los servidores también se encuentran separados de la red externa (Internet) por medio de un firewall, que ha sido configurado para aceptar peticiones entrantes a determinados servicios.

El sistema actual dispone de 2 aplicaciones web:

1. **Portal web de apuestas:** aplicación J2EE desplegada sobre el servidor de aplicaciones JBoss. A esta aplicación se accede vía web, y tiene una zona pública (accesible por protocolo HTTP) y una zona privada y segura (accesible por protocolo HTTPS) en la que es necesario autenticarse con un usuario y contraseña.
La información de cuentas de usuario y contraseña residen en la misma BBDD que utiliza el resto de la aplicación, no existiendo ningún directorio LDAP externo.
2. **Aplicación interna:** aplicación J2EE desplegada también sobre JBoss, a la que sólo debe accederse desde la red privada de la empresa. Todas las páginas son privadas y seguras (protocolo HTTPS). El sistema de autenticación está basado en usuario/contraseña.
Cada usuario tiene asociado uno o varios roles, y en función de éstos podrá acceder a unos determinados módulos de la aplicación.
Su Base de Datos almacena información de nóminas de empleados, contratos con clientes, facturas, etc.

Además el sistema ofrece otra serie de servicios, como son:

1. Servidor de nombres de dominio (servicio DNS).
2. Servidor de impresión: gestiona la cola de impresión de la impresora existente.
3. Servidor de ficheros: unidad compartida con sistema de ficheros jerárquico donde se establecen permisos de lectura/escritura a nivel de directorio.
4. Servidor de correo electrónico: recibe correo entrante y emite correo de salida de las cuentas de correo electrónico de los empleados.

Para resolver los problemas que presenta el sistema actual nuestra empresa, **Cloud Computing Solutions**, propone la implementación de IAAS (Infraestructura Como Servicio), que permite optimizar el uso del hardware existente, a la vez que proporciona facilidades para escalar los recursos, de forma que en determinados momentos pueda aumentarse la capacidad de cómputo, por ejemplo ante picos de visitas por parte de usuarios que interactúan con la plataforma web.

Dados los insuficientes recursos hardware de los que dispone en la actualidad **ApuestaONline**, recomendamos la adquisición de nuevas máquinas que formarán parte de su 'nube privada'.

Para maximizar la alta disponibilidad de la plataforma, se hace necesario contar con un segundo servidor de base de datos. Este servidor será una réplica del primero, y debe establecerse el método de sincronización más óptimo. Entre las opciones disponibles, se valorará aquella en que toda actualización en el servidor principal se realice inmediatamente en el servidor secundario, o bien que cada cierto intervalo de tiempo el servidor secundario pregunte por los cambios que se han producido en el servidor principal y entonces se actualice. La razón de ser de este servidor es hacer de sustituto del servidor principal cuando éste no esté operativo, bien por fallos o caídas, bien porque se están llevando tareas de actualización o mantenimiento offline. Esta medida elimina la necesidad de parar el servicio cuando se programan actualizaciones o backups offline.

La 'nube privada', será implementada con el software Eucalyptus. Éste tiene un API compatible con Amazon EC2 (Elastic Compute Cloud), lo que permitirá a **ApuestaOnline** contratar a Amazon la instalación y ejecución de máquinas virtuales adicionales en su infraestructura (nube pública), pagando sólo por los recursos consumidos en un determinado periodo de tiempo. Esta opción proporciona un importante ahorro económico para la empresa, ya que para dar respuesta a una importante demanda puntual no necesitará invertir en nuevas máquinas, que la mayor parte del tiempo se encontrarán infrutilizadas. En lugar de ello, 'alquila' recursos a un tercero.

La nube privada implementada con Eucalyptus permitirá añadir nuevos recursos hardware de coste medio, en lugar de invertir en grandes y caros mainframes.

4.2. Requisitos del proyecto

A continuación se detallan los requisitos del proyecto. La mayoría de ellos ya son satisfechos por el sistema actual, y deben ser tenidos en cuenta para que se mantengan en el nuevo sistema:

- **Requisitos legales.** La información manejada por los sistemas informáticos de ApuestaOnline se encuadra en los niveles básico y medio que define la Ley Orgánica de Datos de carácter Personal (LODP), por ello deben cumplirse todas las medidas que establece dicha ley.
- **Requisitos de propiedad intelectual y licencias.** El software utilizado debe tener una licencia libre reconocida por la Open Source Initiative **[R8]**. Si no existe una alternativa libre para un determinado propósito debe justificarse debidamente.
- **Requisitos de acceso único.** Debido a que las estaciones de trabajo de los usuarios utilizan el sistema operativo Microsoft Windows, el acceso al sistema de ficheros compartidos se realiza mediante Samba. Cada empleado posee una cuenta de usuario con sus propios permisos sobre directorios y ficheros.
- **Requisitos de acceso web y correo electrónico.** La plataforma web de apuestas contiene páginas públicas accesibles por cualquier visitante mediante el protocolo http,

y a la vez zonas de acceso restringido y personales, a las que sólo usuarios registrados pueden acceder. Para éstas últimas debe utilizarse el protocolo seguro https (SSL), que garantiza la privacidad de la información que viaja por la red.

La aplicación web de uso interno utiliza protocolo seguro https, y el acceso a cualquier módulo exige la autenticación del usuario. Un empleado sólo tendrá acceso a aquellas funcionalidades que su perfil de usuario tenga habilitadas.

Los servicios de correo electrónico también serán cifrados con SSL.

- **Requisitos sobre la base de datos.** Las bases de datos contienen información sensible de clientes y empleados, por lo que sólo deben ser accedidas mediante jdbc desde las aplicaciones web desplegadas (portal de apuestas y aplicación interna), o bien desde la red interna con un mecanismo de autenticación de usuario/contraseña.
- **Requisitos del sistema de seguridad.** El diseño de la red privada de la empresa debe contar con un firewall y una zona desmilitarizada (DMZ) donde se ubicarán los servidores que deben dar servicios al exterior. Tanto los servidores como las estaciones de trabajo reciben una dirección ip privada mediante DHCP y usan un servicio de nombres interno (DNS).
- **Requisitos de impresión en red.** Cualquier empleado, desde su estación de trabajo, puede enviar documentos a la impresora no existiendo ninguna limitación de cuota. Uno de los servidores hace de servidor de impresión (CUPS) gestionando la cola de peticiones.
- **Requisitos de servicio de acceso para publicar datos en el portal web.** El perfilado de usuarios del portal web está basado en roles. Así, aquellos usuarios que tengan asignado el rol apropiado pueden acceder al módulo de administración del portal. El acceso a este módulo les permite, entre otras opciones, añadir contenidos al portal o publicar nuevas apuestas. Existe implementado un sistema de workflow que establece los estados por los que debe pasar un contenido antes de ser publicado. Estos son: creado, editado, aprobado y publicado. Para cada transición de estado el sistema registra qué usuario y en qué momento la realizó.
- **Requisitos de gestión de las copias de respaldo e información de los sistemas.** Está establecido un mecanismo de backup diario, tanto de la base de datos como del sistema de ficheros compartido.
- **Requisitos tecnológicos, mantenimiento y administración.** El personal técnico de la empresa debe recibir la formación necesaria para administrar la nueva infraestructura. Esta formación debe estar más enfocada en la administración de la herramienta Eucalyptus, ya que el resto de tecnologías utilizadas en el nuevo sistema ya son utilizadas en la actualidad, y por ello conocidas por el técnico.
Dado el crecimiento experimentado por la empresa en cuanto a empleados y clientes se contempla la contratación de más administradores de sistema.

- **Requisitos de seguridad.** Debe contarse con herramientas de monitoreo de la red y los sistemas, que permitan detectar intrusiones. Debe haber un log de eventos del sistema que permitan reconstruir lo ocurrido recientemente.
- **Requisitos de la infraestructura.**
 - **Autoescalable.** La nueva infraestructura debe ser autoescalable. Esto significa que debe ser posible configurarla de tal manera que, ante aumentos de tráfico hacia el portal web de apuestas, se añadan instancias del servidor web automáticamente (con RightScale **[R15]**), y cuando el tráfico disminuya se detengan estas máquinas virtuales.
Este requisito exige además que el balanceador web sea configurado adecuadamente para repartir el tráfico entre todas las máquinas virtuales disponibles.
 - **Compatibilidad.** Las máquinas virtuales (AMIs) deben ser compatibles con Amazon EC2.

4.3. Definición de interfaces de usuario

No debe existir limitación en cuanto al sistema operativo utilizado por las máquinas clientes (empleados y usuarios del portal). Las aplicaciones web serán accedidas, como hasta ahora, mediante los protocolos HTTP y HTTPS, desde los navegadores web Firefox e Internet Explorer (versión >6), Google Chrome y Opera.

La aplicación web interna tiene definidos diferentes roles. La propia aplicación gestiona qué recursos pueden ser accedidos por un determinado usuario en función de los roles que tiene asociados.

4.4. Especificación del plan de pruebas

La migración que se propone es progresiva, de forma que se van migrando servicios al nuevo sistema, y cuándo estos han sido probados convenientemente se ponen en Producción. Las pruebas unitarias y de integración se realizarán durante todas las fases definidas en la migración.

4.4.1. Pruebas unitarias

Se realizarán pruebas de los servicios instalados en cada servidor de desarrollo, comprobando que su funcionamiento es correcto. Los servicios que deben ser objeto de estas pruebas unitarias son:

- Servicio de nombres de dominio (DNS): debe comprobarse la conectividad de los equipos hacia el servidor DNS, y que éste resuelve las direcciones correctamente.
- Servicio de impresión en red (CUPS): debe comprobarse el correcto funcionamiento de la cola de impresión, y la comunicación del servidor con la impresora.
- Servicio web http (Apache Tomcat): las aplicaciones desplegadas deben estar plenamente operativas y la comunicación con las bases de datos exitosa.
- Servicio DHCP: comprobación que los equipos reciben direcciones ip privadas y quedan conectados a la red.
- Servidor de ficheros: los empleados deben poder acceder en todo momento al sistema de ficheros compartidos.
- Balanceador web: ante cambios en el número de servidores web el balanceador debe repartir la carga entre todos ellos, descartando los servidores que se encuentren detenidos o pausados.
- Componentes Eucalyptus: debe verificarse que existe comunicación entre todos los componentes de Eucalyptus (Cloud Controller, Walrus, Cluster Controller y Node Controller).
- Máquinas virtuales: verificar que pueden instanciarse las imágenes de máquina virtual registradas en la nube.

4.4.2. Pruebas de integración

Antes de que un servidor pase a Producción deben realizarse las pruebas necesarias para asegurar que su interacción con el resto del sistema es correcta. Entre otras, serán necesarias pruebas de conectividad entre nodos, conexión a base de datos, acceso a sistemas de fichero externos, comprobación de reglas de firewall, etc.

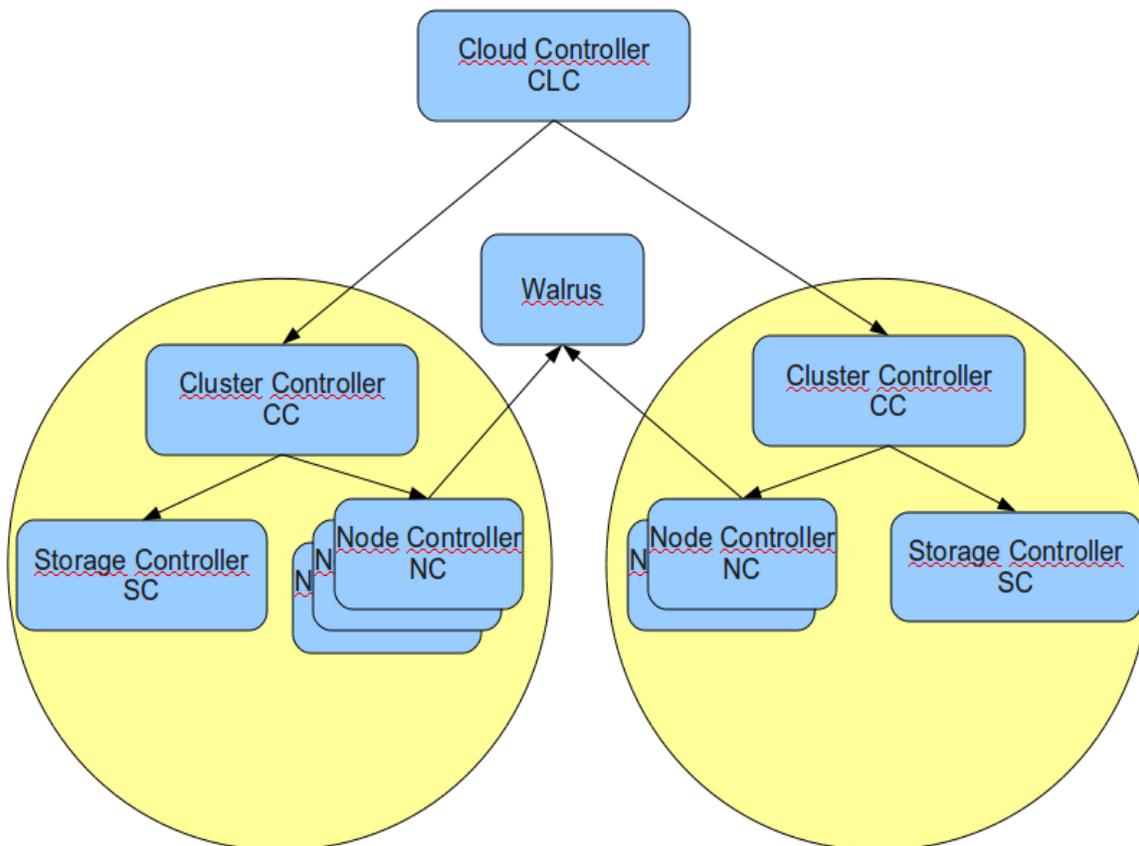
5. Diseño del sistema

5.1. Arquitectura

5.1.1. Modelo Eucalyptus

Eucalyptus es un software para la implementación y administración de nubes privadas, que hace uso de la virtualización de recursos hardware para añadir flexibilidad a la infraestructura.

La arquitectura de Eucalyptus consta de 5 componentes principales, que se describen brevemente a continuación:



Modelo de Arquitectura Eucalyptus

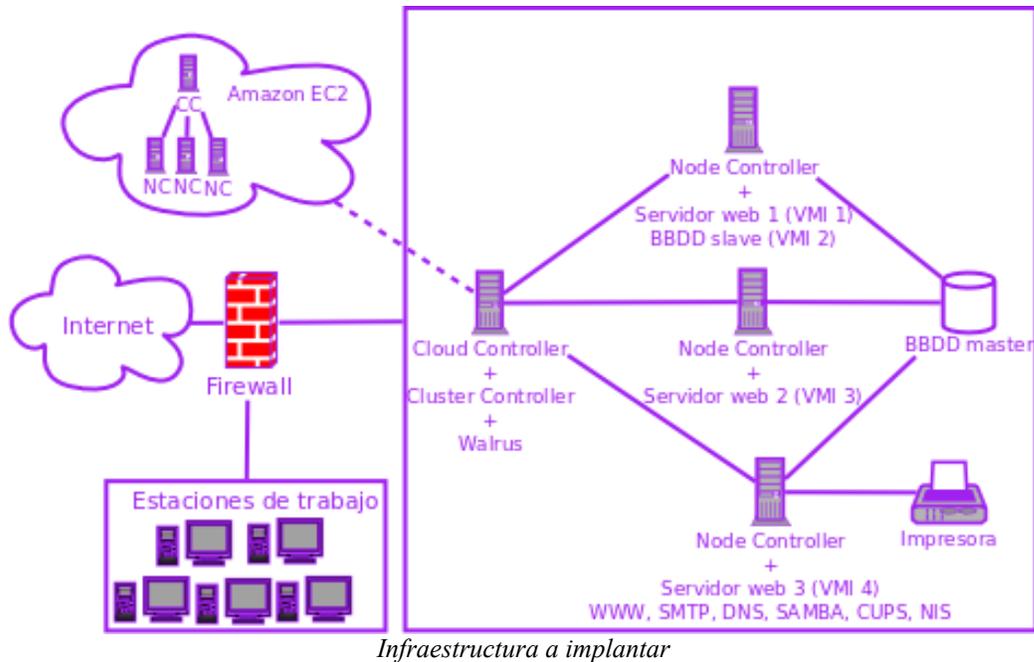
- **Cloud Controller (CLC):** se trata del componente de más alto nivel, el punto de entrada de usuarios y administradores de la nube. Se interactúa con él a través del API Amazon EC2 (que es pública), desde un interfaz web o por línea de comandos.

El CLC se comunica directamente con los Cluster Controller(CC) para obtener información de la nube y gestionarla.

- **Cluster Controller (CC):** recoge información sobre un conjunto de máquinas virtuales, y programa la ejecución de éstas en los Node Controlers (Ncs). Normalmente se instala en una máquina que hace de front-end del cluster.
- **Node Controller (NC):** se ejecuta en cada nodo que aloja máquinas virtuales. Los NCs controlan la ejecución, inspección y parada de las instancias de máquinas virtuales. Además, consulta y controla el software de virtualización en el nodo en el que se ejecuta (host OS y el hypervisor), para dar respuesta a las peticiones del Cluster Controller (CC).
- **Storage Controller (SC):** es un dispositivo de bloque que puede ser añadido a una instancia del sistema de ficheros. No puede ser compartido con otras instancias, pero permite crear un snap-shot (foto) y guardarlo en un sistema de almacenamiento central como Walrus.
- **Walrus:** utilizable por los usuarios como almacén de datos persistentes, organizados como *buckets* (cubos) de objetos. Podemos crear, eliminar, listar buckets, y operaciones de put, get y delete sobre objetos, además de establecer políticas de control de acceso. Posee un interfaz compatible con Amazon's S3. Es útil para guardar imágenes de máquinas virtuales o datos de los usuarios.

Eucalyptus provee un interfaz web en el Cloud Controller (CLC), que permite gestionar algunos aspectos de la nube. Por ejemplo: añadir o eliminar usuarios, definir un cluster o walrus. Hay que decir que las funcionalidades que ofrece son muy limitadas.

En el siguiente esquema se muestra la configuración de los componentes de Eucalyptus aplicados a nuestro caso particular:



Como puede apreciarse, se dispone de una máquina central que hace de varios roles: Cloud Controller, Walrus y Cluster Controller. En la nube privada, inicialmente sólo será necesario un cluster.

En el resto de servidores se instala el Node Controller, y es donde se ejecutarán las máquinas virtuales.

Para montar esta infraestructura serán necesarias las siguientes máquinas:

Nombre máquina	de	Características	Servicios
CloudController		Memoria: 2 GB RAM Procesador: 2,00 GHz Almacenamiento: 1 TB	Eucalyptus Cloud Controller Eucalyptus Cluster Controller Eucalyptus Walrus
Server1		Memoria: 4 GB RAM Procesador: 2,40 GHz Almacenamiento: 400 GB	Eucalyptus Node Controller Servidor web (<i>*máquina virtual</i>) BBDD slave (<i>*máquina virtual</i>)
Server2		Memoria: 4 GB RAM Procesador: 2,40 GHz Almacenamiento: 400 GB	Eucalyptus Node Controller Servidor web (<i>*máquina virtual</i>)

Server3	Memoria: 4 GB RAM Procesador: 2,40 GHz Almacenamiento: 400 GB	Eucalyptus Node Controller Servidor web (* <i>máquina virtual</i>) Apache Tomcat (servidor de aplicaciones) SMTP (servidor de correo) DNS (servidor de nombres de dominio) SAMBA (compartición de ficheros con sistemas Windows) CUPS (servidor de impresión) NIS (Sistema de Información de Red)
BBDD	Memoria: 4 GB RAM Procesador: 2,40 GHz Almacenamiento: 1 TB	Sistema Gestor de Bases de Datos.

En la tabla anterior, se han marcado en color aquellas máquinas que actualmente posee la empresa, por lo tanto serán 2 nuevas máquinas de similares características las que deben adquirirse y sumarse al parque informático.

Detallamos a continuación la función de cada uno de los servidores:

- **CloudController:** se instalan en él los componentes de más alto nivel del software Eucalyptus, que permiten administrar la nube. Estos son:
 - CloudController: gestiona toda la nube.
 - ClusterController: gestiona el único cluster de máquinas que tendremos inicialmente.
 - Walrus: permite almacenar información de usuarios así como máquinas virtuales.
- **Server 1:** se instala Node Controller que maneja 2 máquinas virtuales. Éstas son un servidor web que despliega una instancia del sitio web, y un SGBD utilizado como réplica de la BBDD principal.
- **Server 2:** se instala Node Controller que maneja 1 máquina virtual, que es un segundo servidor web que despliega otra instancia del sitio web.
- **Server 3:** se instala Node Controller que maneja 1 máquina virtual, que es un tercer servidor web que despliega otra instancia del sitio web. Esta máquina virtual será arrancada sólo cuando sea necesario (en picos de actividad o para sustituir a alguno de los servidores web que se encuentre inoperativo). Además, fuera del control de Eucalyptus se instalarán las aplicaciones que usan los propios empleados, como son: servidor de correo, intranet, servidor de impresión, y otros servicios como servidor DNS, DHCP y NIS.
- **BBDD:** se dedica en exclusiva para la base de datos maestra.

Para homogeneizar la infraestructura y sea más sencilla su gestión se recomienda que todos los servidores usen el sistema operativo Ubuntu Server 10.04 LTS (Long Term Support) **[R9]**. Al tratarse de una versión LTS tenemos la seguridad de que Canonical

ofrece soporte durante 5 años para la distribución Server y 3 para Desktop, desde el momento en que esta versión fue lanzada (mediados de 2010).

Eucalyptus permite flexibilizar y optimizar los recursos, por ejemplo arrancando instancias extras de máquinas virtuales cuando se necesiten, y asignarles una cantidad de memoria determinada, en función de las necesidades.

Como ya se ha comentado, si los recursos computacionales propios no fueran suficientes ante algunas situaciones, la empresa podrá ampliarlos haciendo uso de Amazon EC2, quién les facturará por recursos y tiempo consumidos. El soporte a este servicio queda fuera del alcance de este proyecto.

5.1.2. Opciones de virtualización

Eucalyptus construye la nube haciendo uso de la virtualización del hardware, esto es, trabaja con máquinas virtuales que pueden ejecutarse concurrentemente sobre una misma máquina física. Actualmente es compatible con dos sistemas de virtualización: Xen (Paravirtualización) y Kvm (Virtualización completa), ambos software libre.

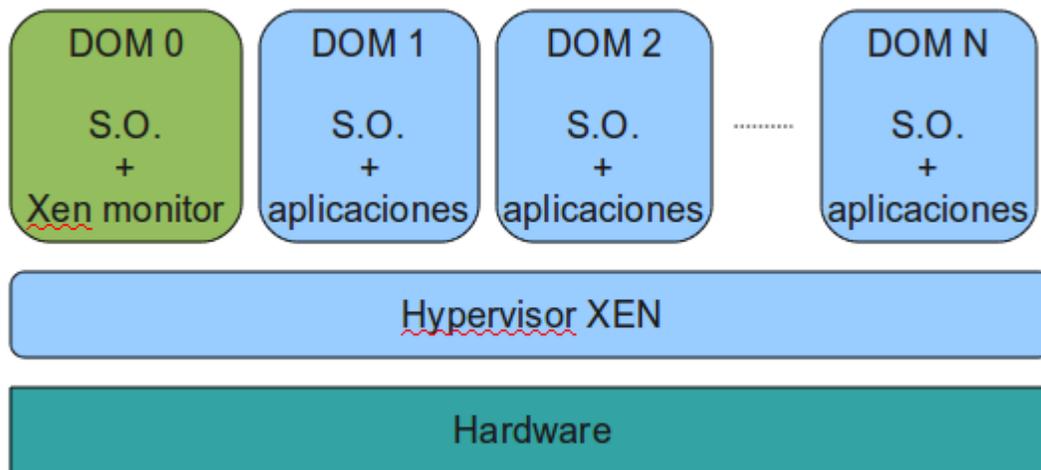
XEN

Xen **[R10]** utiliza un modelo conocido como paravirtualización. Es una capa que se ubica entre el hardware y el sistema operativo, tratando al sistema operativo anterior como una máquina virtual con privilegios (host OS). El resto de sistemas operativos invitados (guest OS) se ejecutan sin privilegios.

Xen utiliza el término *Dominio* para referenciar a las diferentes máquinas virtuales existentes en un mismo equipo hardware. Así, al sistema operativo que se ejecuta tras el arranque (boot) se le denomina Dom0, y al resto DomU (los invitados).

Requiere instalar un nuevo núcleo Linux adaptado para que arranque Xen (el hipervisor).

La siguiente imagen representa un equipo virtualizado utilizando el hipervisor Xen:



Modelo de XEN

KVM

KVM [R11] se encuentra dentro de la categoría de Virtualización completa.

A diferencia de Xen, KVM se integra en el kernel Linux como un módulo más, por lo que no es necesario utilizar un kernel diferente o aplicar parches al existente. Podemos incluirlo en la ruta /etc/modules para que sea cargado al arrancar el sistema, o bien añadirlo manualmente con el comando modprobe <nombre del módulo>.

KVM está basado en las extensiones de virtualización hardware que ofrecen las tecnologías Intel-VT y AMD-V, disponibles en procesadores modernos. En la BIOS de nuestro sistema podemos encontrar la opción de habilitar esta característica.

Con KVM los sistemas operativos invitados aparecen como un proceso más. Se instala sobre sistemas Linux, pero soporta diferentes sistemas operativos invitados (GNU/Linux, Microsoft Windows, Solaris y BSD). Éstos no requieren ser modificados, al contrario que Xen.

En cuanto al software de virtualización, se elige KVM en lugar de XEN por 2 principales motivos. En primer lugar por la facilidad que proporciona el que venga integrado como módulo del kernel Linux, y en segundo lugar por el hecho de poder ejecutar sistemas operativos de diferente naturaleza.

5.2. Especificación de estándares, normas de diseño y construcción

- Gestión de la Documentación. Se creará una wiki con el software Trac [R12], donde se publicará toda la documentación generada en el proyecto de migración. El control de versiones sobre los contenidos de esta wiki serán gestionados por Subversión [R13]. Diariamente se realizará backup de estos contenidos.

Las dudas o problemas que surjan durante el proyecto serán reportadas en la wiki como 'tickets', que serán asignados a la personas que deban resolverlos.

- **Documentos de diseño.** Serán generados en formato estándar ODF, compatible con procesadores de texto como OpenOffice.org.

Todo documento llevará la siguiente información:

- Título
- Responsables
- Personas o colectivo a quien va dirigido

Así mismo deben llevar un control de cambios o revisiones, que incluya los siguientes datos:

Fecha	Versión	Descripción de cambios	Realizada por

Debe utilizarse la herramienta de OpenOffice para control de revisiones, cuando un autor aplica cambios (añade, elimina o modifica) en el contenido.

Las versiones finales de los documentos deben convertirse en formato PDF, de sólo lectura, y su publicación debe ser comunicada a todos los interesados.

- **Diagramas de diseño:** Para la representación de diagramas se utilizará el lenguaje UML **[R14]** (diagramas de componentes, diagrama de despliegue,...), que es un estándar *de facto* ampliamente utilizado.

5.3. Especificaciones de desarrollo y pruebas

- Los componentes de Eucalyptus sólo pueden ser desplegado sobre máquinas con sistema operativo Linux. Se contará con Ubuntu Server tanto para el entorno Desarrollo como para Producción.
- Eucalyptus es compatible con 2 sistemas de virtualización: KVM y XEN. KVM se carga como módulo del kernel Linux, por tanto sólo puede ser instalado en sistemas Linux, pero puede ejecutar instancias de otros sistemas operativos. Funciona con arquitectura x86 y x86_64, siempre que tengan extensiones de virtualización (la mayoría de las nuevas CPUs las tienen). XEN requiere que el sistema operativo invitado sea modificado para ejecutarse en él. El software de virtualización seleccionado es KVM.
- Si en los servidores existe software privativo debe contarse con las licencias necesarias para utilizarlo tanto en Desarrollo como en Producción.
- Debe contarse con todo el software y parches necesarios, entre otros:
 - **Sistema Operativo** para los servidores: última versión de Ubuntu Server con soporte de larga duración (10.04 LTS).

- **Software de gestión de nube** (IAAS): Eucalyptus 2.0.1 y Euca2ools
- **Sistema Gestor de Base de Datos:** PostgreSQL 9.0.
- **Servidor de aplicaciones J2EE:** Apache Tomcat 6.0.29. Se despliegan tanto la aplicación web interna como el portal web de apuestas.
- **Agente de transporte de correo** (MTA): Exim 4.73
- **Servidor de impresión:** CUPS 1.4
- **Compartición de ficheros** en red Windows: SAMBA 3.5.6

5.4. Requisitos de implantación

El personal técnico del cliente debe involucrarse en todas las fases del proyecto, incluida la propia implantación. Este punto es imprescindible para que, tras la finalización de la migración, el cliente pueda tomar el control del nuevo sistema informático. Recibirá formación así como amplia documentación del nuevo software y servicios implantados.

Las nuevas máquinas deben estar disponibles y revisadas con anterioridad a su uso, así como el software descrito en el punto 5.3.

Debe contarse con un plan de marcha atrás para cada una de las fases de implantación definidas, para que ante cualquier eventualidad sea posible restablecer el sistema a un estado anterior.

Los usuarios del sistema no deben apreciar ningún cambio en la forma en que acceden a los servicios. La migración debe ser totalmente transparente para ellos.

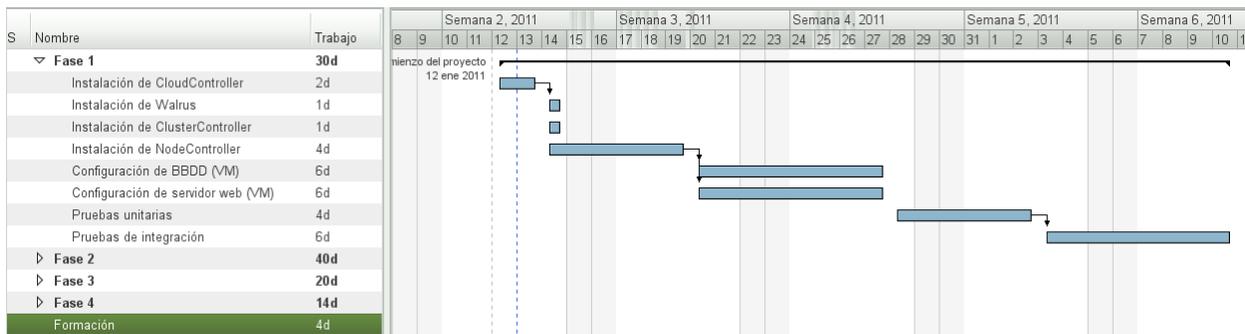
Antes de cada pase a Producción los usuarios deben estar advertidos. Si detectan algún problema deben avisar de inmediato al equipo técnico para que pueda ser atendido.

6. Desarrollo e Implantación del sistema

La migración del sistema actual a la nueva infraestructura se hará de forma paulatina. El procedimiento será implantar un determinado componente o servicios en una máquina destinada como entorno de Desarrollo. Se realizarán las pruebas unitarias sobre los servicios instalados, y cuando hayan sido convenientemente probados esta misma máquina pasa a Producción, ejecutando previamente las pruebas de integración correspondientes. Por ello, las fases de Desarrollo e Implantación las consideraremos de forma conjunta.

La implantación se ha definido en 4 fases que se detallan a continuación:

6.1. Primera fase



Planificación Fase 1

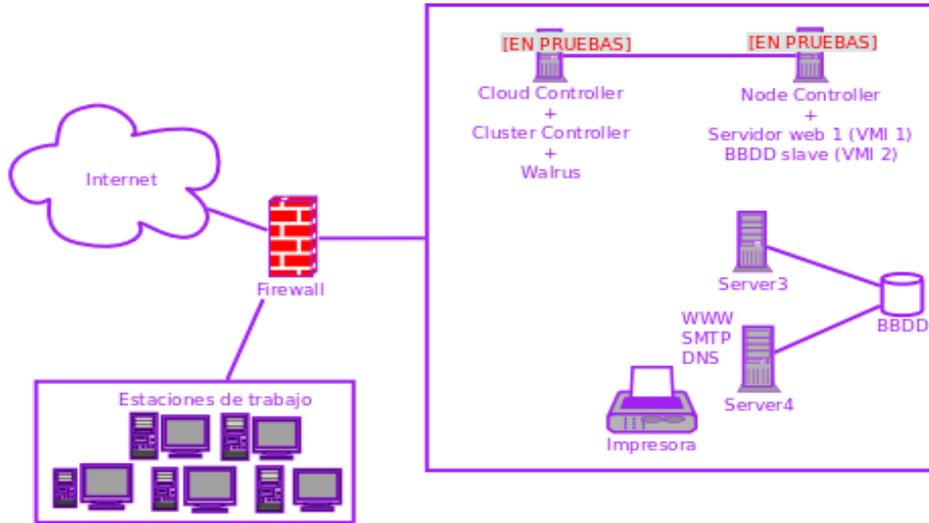
Para esta primera fase el cliente ha adquirido previamente 2 nuevas máquinas, a las que nombraremos CloudController y Server1. En ambas se instala como único sistema operativo la distribución Ubuntu Server 10.04 LTS.

Sobre CloudController se instalan y configuran los componentes: Cloud Controller, Cluster Controller y Walrus.

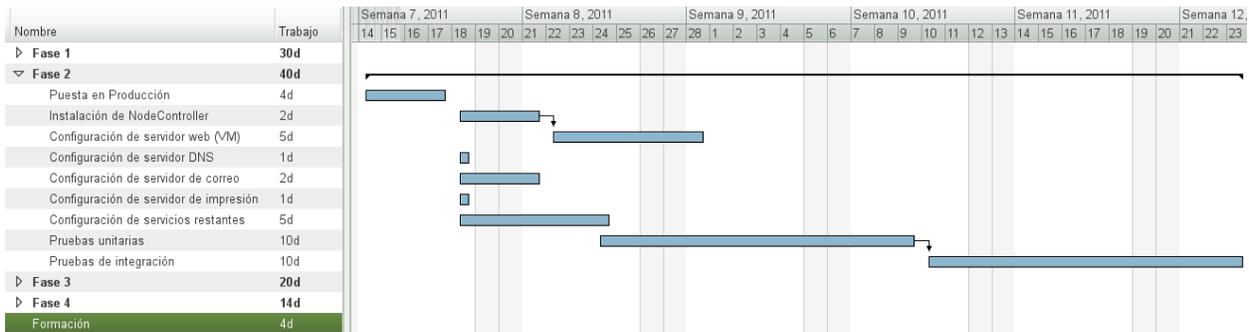
Sobre Server1 se instala y configura el componente Node Controller, y se crean 2 máquinas virtuales. Una de ellas albergará una instancia de la aplicación web, y la otra una instancia de una base de datos de desarrollo, que utilizará el mismo esquema que la base de datos real. Es importante que en el entorno de desarrollo se trabaje con las mismas versiones del software que se utilizarán en producción. En el caso del SGBD debe utilizarse PostgreSQL 9.0. Se debe hacer un volcado (de producción a desarrollo) del contenido de ciertas tablas que son necesarios para que la aplicación funcione correctamente.

La aplicación web que se ejecuta bajo Server1 sólo será accesible desde la red local, en ningún caso desde el exterior. Se configurará para que use la BBDD de desarrollo.

Al finalizar la fase tendremos esta foto de la arquitectura:



6.2. Segunda fase



Planificación Fase 2

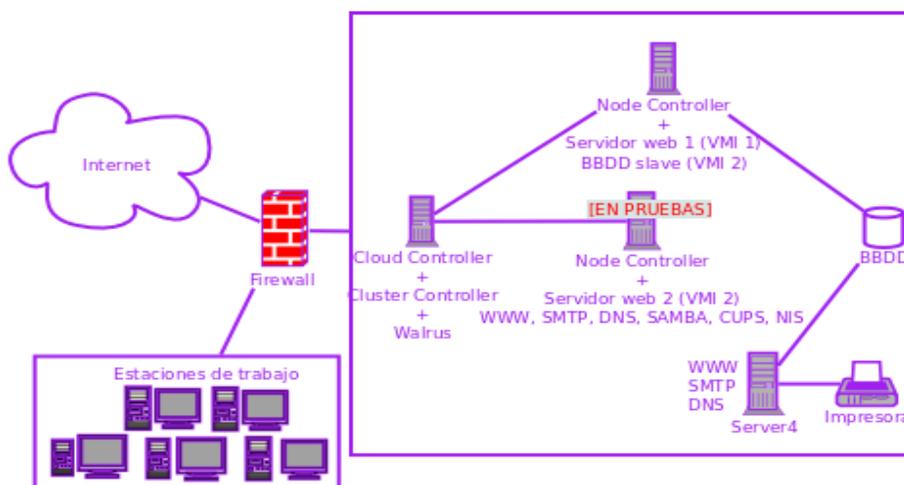
Una vez certificado que la aplicación web funciona correctamente en *Server1*, se realiza el pase a producción. Esto implica configurarla para que utilice la BBDD de Producción y hacerla accesible desde fuera de la red local. Debe configurarse el balanceador web para que añada esta instancia a la lista de servidores web de la aplicación.

En esta segunda fase se trabaja en la adecuación del antiguo servidor *Server3* como otro nodo de la nube.

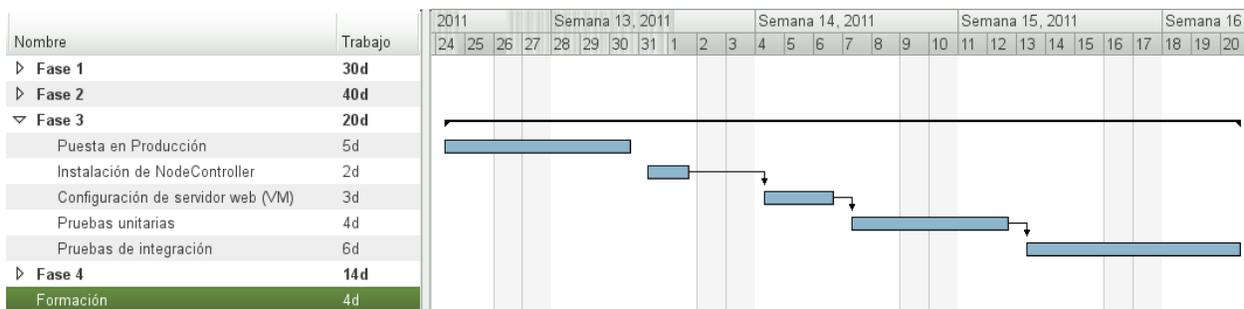
Sobre *Server3* se instala el componente Node Controller de Eucalyptus y se crea una máquina virtual que ejecute una instancia de la aplicación web.

También, *Server3* se configura con todos los servicios que tenía la empresa para sus empleados (intranet, servidor de correo, servidor de impresión, etc.), para así sustituir a *Server4* en la fase siguiente. Los servicios que proveerá este servidor son cruciales para el buen funcionamiento de toda la empresa, por ello los esfuerzos en la realización de pruebas deben ser mucho mayores.

Al finalizar la segunda fase el estado del sistema será la siguiente:



6.3. Tercera fase



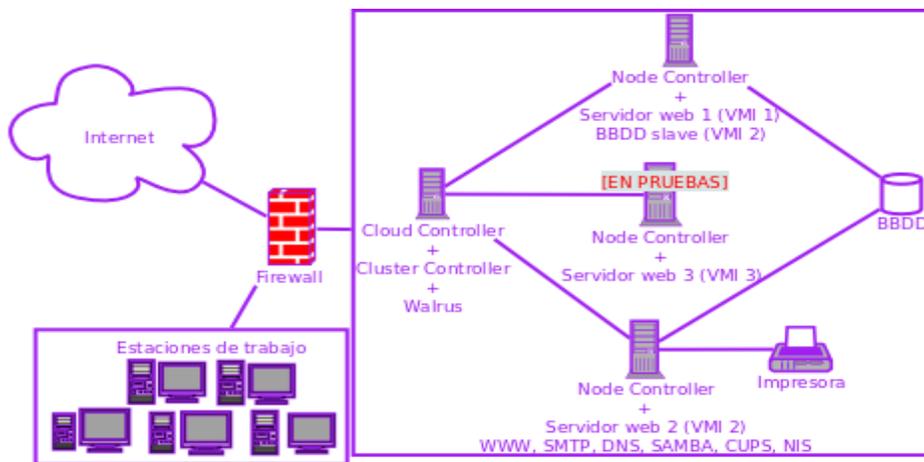
Planificación Fase 3

En esta tercera fase *Server3* pasa a Producción sustituyendo a *Server4*, dándose de alta como un nuevo nodo de la nube privada, y ofreciendo los servicios que ya disfrutaban los empleados.

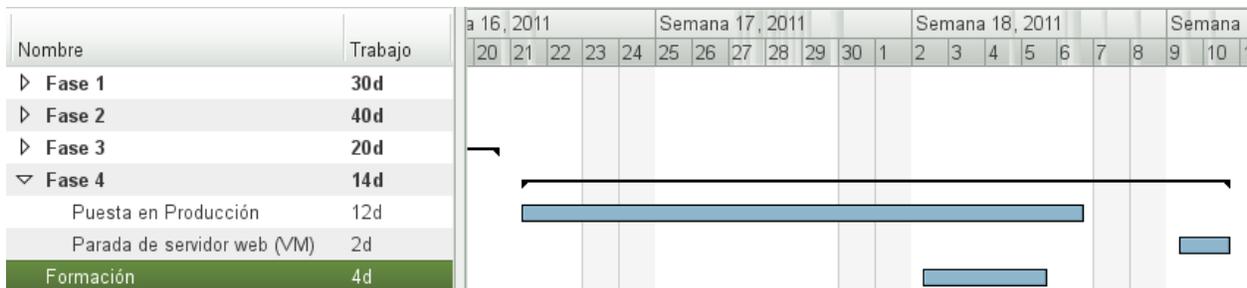
Seguidamente, *Server4* se configura adecuadamente como un nuevo nodo de la nube, pero no activo para Producción.

La sustitución de *Server4* por *Server3* debe realizarse fuera del horario laboral, en fin de semana, ya que será necesaria una ventana horaria en la que los servicios estarán parados hasta completar la migración.

Esta fase es crucial, por ello si surgen problemas que impidan el normal funcionamiento de la empresa se procederá a dar marcha atrás hasta que éstos sean solventados.



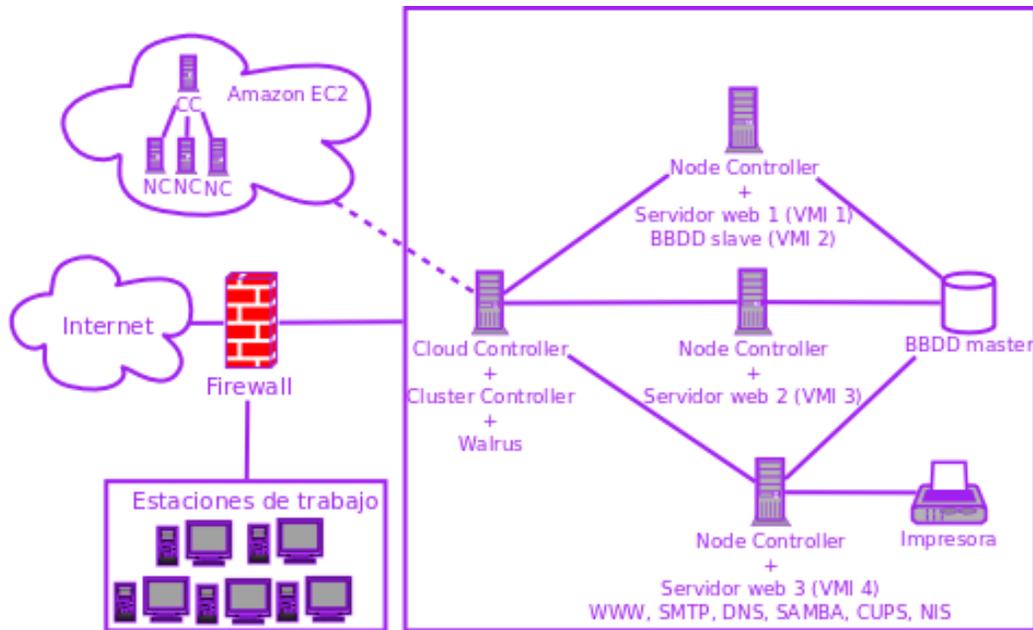
6.4. Cuarta fase



Planificación Fase 4

Esta es la última fase, en la que el *Server4* se pone en Producción, empezando a atender peticiones de los usuarios de la plataforma. Dada la experiencia acumulada en las fases anteriores, este hito será más sencillo de completar con éxito.

Como parte de esta fase, en el *Server3* será parada la máquina virtual para que toda la capacidad del hardware esté dedicada al resto de servicios (servidor de correo, intranet, DNS, CUPS,...). Podrá iniciarse de nuevo en el momento en que se haga necesario.



6.5. Formación

El destinatario de la formación será el técnico con el que cuenta la empresa. Al tratarse de un perfil técnico ya habituado a trabajar sobre sistemas GNU/Linux, la formación impartida se centrará principalmente en la administración de la nube privada y en su integración con los servicios de Amazon.

La formación técnica que recibirá abarca los siguientes temas:

- **Monitorización.** El alumno aprenderá a manejar herramientas que le permitirán monitorizar el uso de CPU, tráfico de red y de Entrada/Salida de las máquinas virtuales.
- **Administración básica.** Se enseñará al alumno la administración y configuración básica del software Eucalyptus: creación de usuarios, acceso SSH a máquinas virtuales, iniciar/detener máquinas virtuales, etc.

- **Autoescalado y balanceo de carga.** Se tratan los temas de balanceo de carga entre las máquinas virtuales en ejecución, y la ampliación o reducción automática de máquinas virtuales en función de las necesidades.

En este proyecto no se contempla la actualización de las aplicaciones que utilizan los usuarios, ni la forma en que interactúan con ellas. Por esta razón, el cambio de infraestructura debe resultar transparente, así que no se contempla en estos momentos ninguna acción formativa.

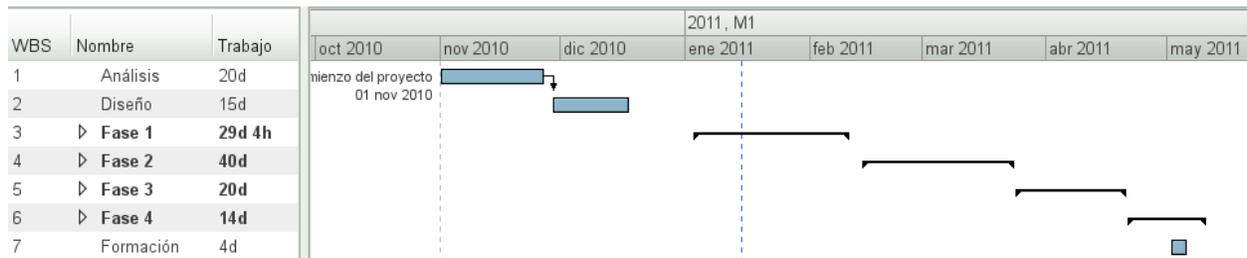
7. Planificación del Proyecto

7.1. Fases

El ciclo de vida de este proyecto comprende 5 grandes bloques de actividades:

1. **Toma de requisitos:** Ésta es una fase preliminar, en la que se llevan a cabo varias entrevistas con los interlocutores del cliente para recoger información de la situación actual y de los requisitos del nuevo sistema. Se ha establecido que los interlocutores deben ser los directores y el técnico.
2. **Análisis:** En esta fase, Cloud Computing Solutions realiza un estudio de la infraestructura del cliente, así como de las aplicaciones informáticas utilizadas por sus usuarios y clientes. Se debe realizar un análisis muy detallado ya que el nuevo sistema debe ofrecer los mismos servicios a usuarios y clientes, y no debe haber pérdida de información.
3. **Diseño:** Basándonos en la información generada de la fase anterior, se diseña la nueva infraestructura. Se establece la arquitectura con los diferentes nodos, definiendo qué servicios ofrece cada máquina.
4. **Desarrollo e Implantación:** En esta fase se realiza la implantación del sistema diseñado, que se hará de forma progresiva estableciendo mecanismos de “marcha atrás” por si se encuentran problemas que impidan el normal funcionamiento de la empresa.
5. **Formación:** El técnico de nuestro cliente se involucrará en el proyecto desde el primer momento. Además, durante la fase anterior recibirá formación y documentación sobre el software Eucalyptus, para que una vez terminada la implantación, tenga los conocimientos necesarios para gestionar la infraestructura.

La planificación prevista para todo el proyecto se resume en el siguiente gráfico:



En esta planificación no se contempla el servicio de soporte y mantenimiento de la infraestructura por un periodo de tiempo. Éste será opcional para el cliente, para el cual se redactará un contrato específico.

7.2. Equipo humano

Dada la experiencia de **Cloud Computing Solutions** sobre el software Eucalyptus, así como en el resto de paquetes a instalar, el equipo técnico es reducido.

Para el desarrollo exitoso del proyecto, **Cloud Computing Solutions** formará el siguiente equipo de trabajo:

Perfil	Responsabilidades
1 Jefe de proyecto	Máximo responsable del proyecto por parte de Cloud Computing Solutions. Se asegurará del cumplimiento de plazos de entrega y gestionará al equipo de personas a su cargo. Será el interlocutor directo con ApuestaONline .
1 Técnico de Sistemas Senior	Experto en las tecnologías a implantar y en la configuración de equipos y redes. Participará activamente en la instalación y configuración del nuevo sistema.
1 Técnico de Sistemas Junior	Realizará las tareas técnicas que le sean encomendadas, dando apoyo al Técnico de Sistemas Senior.
1 Administrador Senior de BBDD	Su principal labor consistirá en implantar la base de datos de desarrollo, así como configurar una réplica de la base de datos de Producción. Se ocupará también de configurar la sincronización de las instancias de base de datos y el backup del contenido de sus tablas.

Por su parte, **ApuestaONline** se compromete a ceder al proyecto parte de las horas laborales del técnico que tiene en plantilla, que ayudará al equipo de trabajo y será el destinatario de la formación técnica que **Cloud Computing Solutions** imparta. De esta forma, aseguramos el traspaso de conocimiento necesario para el mantenimiento de la infraestructura.

7.3. Costes del proyecto

El coste del proyecto está directamente relacionado con las tarifas de cada uno de los integrantes del grupo de trabajo y el tiempo dedicado al mismo.

El software que instalará **Cloud Computing Solutions** será software libre sin costes de licencia. En este presupuesto no se tiene en cuenta el precio de renovación de licencias del software que el cliente tiene actualmente, que en caso de darse correrá al cargo de este último.

A continuación se detalla el coste del proyecto expresado en euros (sin iva incluido), desglosado en las fases definidas en este documento:

WBS	Nombre	Inicio	Fin	Trabajo	Duración	Desperdicio	Coste	Asignado a	% Completa
1	Análisis	nov 1	nov 26	20d	20d	102d	12000	Jefe de proyecto, Técnico Senior	0
2	Diseño	nov 29	dic 17	15d	15d	102d	9000	Jefe de proyecto, Técnico Senior	0
3	▼ Fase 1	ene 3	feb 10	29d 4h	29d	63d	15994,79		0
3.1	Instalació	ene 3	ene 3	1d 4h	6h 51min	90d 1h	874,79	Administrador BBDD, Jefe de proyecto, Técnico Junior, Técnico Senior	0
3.2	Instalació	ene 3	ene 4	1d	1d	90d 1h	540	Administrador BBDD, Jefe de proyecto, Técnico Junior, Técnico Senior	0
3.3	Instalació	ene 14	ene 14	1d	1d	82d	540	Administrador BBDD, Jefe de proyecto, Técnico Junior, Técnico Senior	0
3.4	Instalació	ene 14	ene 19	4d	4d	73d	2160	Administrador BBDD, Jefe de proyecto, Técnico Junior, Técnico Senior	0
3.5	Configurac	ene 20	ene 27	6d	6d	73d	3240	Administrador BBDD, Jefe de proyecto, Técnico Junior, Técnico Senior	0
3.6	Configurac	ene 20	ene 27	6d	6d	73d	3240	Administrador BBDD, Jefe de proyecto, Técnico Junior, Técnico Senior	0
3.7	Pruebas u	ene 28	feb 2	4d	4d	63d	2160	Administrador BBDD, Jefe de proyecto, Técnico Junior, Técnico Senior	0
3.8	Pruebas d	feb 3	feb 10	6d	6d	63d	3240	Administrador BBDD, Jefe de proyecto, Técnico Junior, Técnico Senior	0
4	▼ Fase 2	feb 14	mar 23	40d	28d	34d	21600		0
4.1	Puesta en	feb 14	feb 17	4d	4d	58d	2160	Administrador BBDD, Jefe de proyecto, Técnico Junior, Técnico Senior	0
4.2	Instalació	feb 18	feb 21	2d	2d	51d	1080	Administrador BBDD, Jefe de proyecto, Técnico Junior, Técnico Senior	0
4.3	Configurac	feb 22	feb 28	5d	5d	51d	2700	Administrador BBDD, Jefe de proyecto, Técnico Junior, Técnico Senior	0
4.4	Configurac	feb 18	feb 18	1d	1d	57d	540	Administrador BBDD, Jefe de proyecto, Técnico Junior, Técnico Senior	0
4.5	Configurac	feb 18	feb 21	2d	2d	56d	1080	Administrador BBDD, Jefe de proyecto, Técnico Junior, Técnico Senior	0
4.6	Configurac	feb 18	feb 18	1d	1d	57d	540	Administrador BBDD, Jefe de proyecto, Técnico Junior, Técnico Senior	0
4.7	Configurac	feb 18	feb 24	5d	5d	53d	2700	Administrador BBDD, Jefe de proyecto, Técnico Junior, Técnico Senior	0
4.8	Pruebas u	feb 24	mar 9	10d	10d	34d	5400	Administrador BBDD, Jefe de proyecto, Técnico Junior, Técnico Senior	0
4.9	Pruebas d	mar 10	mar 23	10d	10d	34d	5400	Administrador BBDD, Jefe de proyecto, Técnico Junior, Técnico Senior	0
5	▼ Fase 3	mar 24	abr 20	20d	20d	14d	10800		0
5.1	Puesta en	mar 24	mar 30	5d	5d	29d	2700	Administrador BBDD, Jefe de proyecto, Técnico Junior, Técnico Senior	0
5.2	Instalació	mar 31	abr 1	2d	2d	14d	1080	Administrador BBDD, Jefe de proyecto, Técnico Junior, Técnico Senior	0
5.3	Configurac	abr 4	abr 6	3d	3d	14d	1620	Administrador BBDD, Jefe de proyecto, Técnico Junior, Técnico Senior	0
5.4	Pruebas u	abr 7	abr 12	4d	4d	14d	2160	Administrador BBDD, Jefe de proyecto, Técnico Junior, Técnico Senior	0
5.5	Pruebas d	abr 13	abr 20	6d	6d	14d	3240	Administrador BBDD, Jefe de proyecto, Técnico Junior, Técnico Senior	0
6	▼ Fase 4	abr 21	may 10	14d	14d		8400		0
6.1	Puesta en	abr 21	may 6	12d	12d	2d	7200	Jefe de proyecto, Técnico Senior	0
6.2	Parada de	may 9	may 10	2d	2d		1200	Jefe de proyecto, Técnico Senior	0
7	Formación	may 2	may 5	4d	4d	3d	2240	Técnico Senior	0

Cloud Computing Solutions establece que el precio establecido en cada fase debe ser entregado por parte de **ApuestaOnline** en los 7 primeros días de cada fase.

CONCLUSIONES

En la fase final de este proyecto en la que me encuentro, estoy convencido de que elegí uno de los proyectos más interesantes ofertados por la UOC este curso 2010.

Este proyecto me ha permitido conocer una tecnología que cada vez está más de actualidad, como es todo lo relacionado con el Cloud Computing, profundizando más concretamente en IAAS, o Infraestructura Como Servicio.

He descubierto que existen diversos tipos de virtualización (Paravirtualización, Virtualización completa asistida por hardware, ...) así como varias soluciones software, entre ellas: XEN y KVM. Previamente sólo conocía la existencia VMWare y Virtual Box (con el que había "jugado" un poco).

En esta Memoria he podido aplicar, de forma 'teórica', los conocimientos que he adquirido de Eucalyptus, a un caso práctico de migración de los sistemas informáticos de una empresa, conociendo las ventajas reales de IAAS.

En el desarrollo de la práctica externa (creación de Live CD demo de Eucalyptus) me he encontrado con numerosos problemas que bloqueaban mi avance, pero que a base de investigar, de preguntar en foros o seguir por caminos alternativos, finalmente he conseguido resolver.

He comprendido la arquitectura que implementa Eucalyptus (y por ende Amazon EC2), y la función de cada uno de sus principales componentes, aunque por falta de tiempo no he podido profundizar en todas sus funcionalidades y su compatibilidad con Amazon. Uno de estos casos es Walrus, un servicio de almacenamiento compatible con Amazon's S3, que puede utilizarse para guardar datos de usuarios o máquinas virtuales. Tampoco he probado todos los modos de red configurables en Eucalyptus para las máquinas virtuales.

Otro aspecto que me parece muy interesante y útil pero en el que no he podido profundizar, es en la posibilidad que brindan algunas herramientas de terceros de configurar la nube como auto-escalable. Esta cualidad consiste en la posibilidad de que, sin intervención humana, la nube aumente su capacidad (añadiendo nuevas máquinas virtuales) o la disminuya (parando máquinas virtuales en ejecución) en función de ciertos parámetros monitorizados, como puede ser el tráfico de red detectado. Esta opción permite que la capacidad de la nube se ajuste a la demanda de recursos, con un servicio 24x7.

Este paradigma de infraestructura, y en concreto Amazon EC2 por hacer públicas algunas de sus especificaciones, abre nuevas posibilidades en la introducción de herramientas

de terceros, que pueden ser explotadas por otras empresas creando un ecosistema de herramientas y servicios alrededor de Amazon EC2 u otras soluciones de Cloud Computing.

A partir de la entrega de esta Memoria y de la Práctica externa, considero que tengo una buena base para seguir investigando acerca de Eucalyptus y Amazon EC2, así como otros paquetes software de gestión de nubes.

REFERENCIAS

- [R1] Eucalyptus, <http://open.eucalyptus.com/>
- [R2] Suse Studio, <http://www.susestudio.com>
- [R3] Suse Gallery, <http://susegallery.com/>
- [R4] Definición de Cloud Computing (wikipedia), http://en.wikipedia.org/wiki/Cloud_computing
- [R5] Costes de Amazon EC2, <http://aws.amazon.com/es/ec2/#pricing>
- [R6] OpenNebula, <http://www.opennebula.org/start>
- [R7] Nimbus, <http://www.nimbusproject.org/>
- [R8] Open Source Initiative, (<http://www.opensource.org/>)
- [R9] Ubuntu Server, <http://www.ubuntu.com/server>
- [R10] XEN, <http://www.xen.org/>
- [R11] KVM, http://www.linux-kvm.org/page/Main_Page
- [R12] Trac, <http://trac.edgewall.org/>
- [R13] Subversion, <http://subversion.tigris.org/>
- [R14] UML, http://en.wikipedia.org/wiki/Unified_Modeling_Language
- [R15] RightScale, <http://www.rightscale.com/>

ANEXOS

- [A1] Guia_DEMO_LIVE_CD_EUCALYPTUS.pdf