

Arquitectura de Computadores

Computación Nube

Pablo Ignacio Rodríguez Fernández



INDICE

INTRODUCCIÓN	3
CARACTERÍSTICAS.....	5
VENTAJAS E INCONVENIENTES.....	6
Las ventajas	6
Las desventajas	7
EJEMPLOS Y APLICACIONES.....	9
CLOUDSIM	9
GOOGLE APPS	10
Amazon EC2	11
SUN GRID ENGINE	11
MEEBO	13
EYEOS	13
AVIARY	14
BIBLIOGRAFÍA.....	15

INTRODUCCIÓN

Para entender la computación nube primero trataremos qué es la computación distribuida, puesto que la computación nube es un tipo de esta.

Se habla de computación distribuida cuando el sistema incluye varios subsistemas separados físicamente. A cada uno de los sistemas con capacidad computacional se les suele llamar *nodos*. La computación distribuida trata de diseñar y programar sistemas distribuidos físicamente, abiertos, escalables, transparentes y tolerantes a fallos. La computación distribuida se diferencia de la paralela principalmente en el nivel de transparencia para el usuario: en la computación paralela hay que considerar explícitamente la existencia de diferentes sistemas, y transferir información entre ellos. En realidad, se diferencian más en el grado de transparencia; los sistemas paralelos son opacos, mientras que los distribuidos tienen diferente grado de transparencia para el usuario.

La tendencia actual es a la creación y diseño de sistemas distribuidos, más que paralelos. Y de estos hay también diferentes tipos (que no son excluyentes, sino que muchas veces se intersecan o se incluyen):

- Sistemas *cluster*: Son ordenadores de bajo coste, homogéneos en la configuración (o muy parecidos), y unidos por una red de alta velocidad. Hay muchas soluciones, incluso basadas en software libre, que permiten construir un clúster de altas prestaciones.
- Sistemas *grid*: se trata de integrar y usar colectivamente recursos procedentes de instalaciones separadas físicamente.
- Las *redes overlay* incluyen diferentes recursos unidos por una aplicación que mantiene simultáneamente la conectividad, y proporciona todos los demás servicios relacionados con el mantenimiento de la red. Se denominan *overlay* por estar contruidos *encima* de otras redes, y, especialmente, tener una arquitectura que es independiente de las mismas.
- Sistemas *entre pares* (P2P): son generalmente redes overlay, es decir, basadas en una aplicación, y se usan principalmente para compartir ficheros. Sin embargo, son un sistema distribuido que, en principio, puede usarse para compartir cualquier tipo de recursos.
- Computación *de ciclos redundantes*: se trata de aprovechar la capacidad de un ordenador que está encendido, pero sin realizar ninguna tarea por parte del usuario. El sistema más popular es SETI@Home, pero hay otros muchos sistemas que funcionan de forma similar. Si este uso de ciclos más o menos redundantes se hace sin el conocimiento del usuario, se trataría de *computación parásita* o *sigilosa*.

Algunos de los sistemas distribuidos mencionados anteriormente pueden hacer uso de servicios web para la comunicación entre los diferentes nodos, como es nuestro caso, de ahí lo de nube, haciendo metáfora con la nube que es internet.

Dentro de la computación grid ha surgido a partir del año 2007 la denominada *computación nube* o cloud computing. La computación nube es un subconjunto de la computación en rejilla o grid en la cual hay una cierta cantidad de recursos a disposición

del usuario, y se puede acceder a ellos en función de las necesidades. La computación grid sería un tipo de tecnología que serviría para manejar la *nube* de recursos.

Lo interesante de la computación nube es la *virtualización* de recursos, a diferencia del grid, que restringe mediante parámetros de calidad de servicio y la cantidad de recursos usables, en la computación nube los recursos se *virtualizan* convirtiéndose no en *servicios* virtuales, sino en ordenadores o periféricos virtuales. Mediante servicios tales como el Elastic Compute Cloud de Amazon o el Google Apps se abstraen diferentes niveles de computación, desde el más bajo hasta el más alto, es decir, se trata de presentar una plataforma, programa, infraestructura o almacenamiento de datos como un servicio.

En resumen, mediante **cloud computing**, puedes usar programas y servicios que no están instalados en tu computadora (y por lo tanto no tendrás que ocuparte de mantenerlos actualizados) sino en otra computadora y tu puedes acceder a ellos mediante una conexión a Internet. Hay servicios gratuitos y otros pagos. Tus datos se guardan encriptados y puedes acceder a ellos y seguir trabajando desde cualquier computadora (y en algunos casos hasta desde el teléfono móvil) en la que te encuentres, desde cualquier lugar del mundo, con el único requisito de tener conexión a Internet.

El [video sobre la computación nube](#) que aquí pongo es bastante esclarecedor.

CARACTERÍSTICAS

Como ya uno puede imaginarse, una de las principales diferencias de la computación nube es que no hay necesidad de conocer la infraestructura que hay detrás de esta, pasa a ser “una nube” donde las aplicaciones y servicios pueden fácilmente crecer (escalar), funcionar rápido y fallar poco sin conocer los detalles del funcionamiento de esta “nube”.

Las características más destacables podríamos decir que son las siguientes:

- *Virtualización*: las aplicaciones son independientes del hardware en el que corran, incluso varias aplicaciones pueden correr en la misma máquina, o una aplicación puede usar varias máquinas.
- *Multipropósito*: El sistema esta creado de tal forma que permite a diferentes clientes compartir la infraestructura sin preocuparse de ello y sin comprometer su seguridad y privacidad.
- *Transparencia*: el acceso a un recurso debe ser independiente de su situación física; el sistema debe proveer una capa que identifique uniformemente recursos locales y remotos.
- *Apertura*: los sistemas distribuidos deben de estar definidos por protocolos que permitan añadir fácilmente nuevos sistemas.
- *Escalabilidad*: Deben de permitir añadir un número indeterminado de nodos a la conexión, y las prestaciones que se obtengan deben de escalar de un modo razonable con el número de nodos, siendo ideal el escalado lineal.
- *Tolerancia a fallos*: la desaparición de uno o varios nodos no debe acabar con la estructura de la red, y la degradación de prestaciones debe ser razonable.

Hay que tener en cuenta que la regulación de la pertenencia a la red debe ser tal que las demás características sean posibles (vamos, que no haya un usuario autorizado o no que se cargue la tolerancia a fallos o la escalabilidad a base de la apropiación de recursos), y que se dificulte la apropiación de recursos no autorizados por parte de un usuario (autorizado o no).

VENTAJAS E INCONVENIENTES

Las ventajas

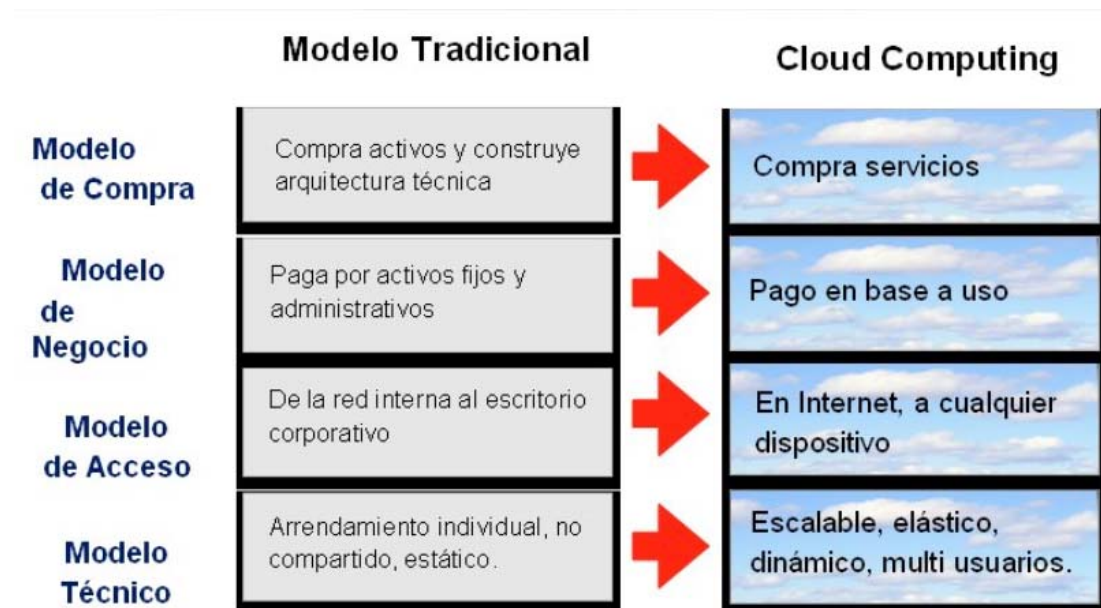
Usar esta tecnología trae muchas ventajas.

En cuanto al aspecto económico su precio se reduce mucho puesto que su capital de trabajo se convierte en gasto operacional. Además al ser multiusuario los costos y recursos se reparten entre todos.

Si nos ocurre cualquier problema, su redundancia nos permite continuidad y recuperación inmediata.

Son escalables, con lo que tenemos una infraestructura para constantes cambios en demanda, es decir, si la demanda aumenta hasta el punto de no poder dar servicio se puede ampliar sin ningún problema, de forma que se pueda dar el servicio y el resto de usuarios no se vean afectados en ningún momento.

La siguiente gráfica muestra una comparativa con respecto a los modelos tradicionales en la que se ven claramente bastantes ventajas de la computación nube.



Si miramos desde el punto de vista del usuario también tenemos bastantes ventajas. Una de ellas es que puedes acceder a tus datos de cualquier ordenador en el mundo. Con tu nombre de usuario y tu contraseña, puedes leer tus correos, contestarlos, añadir contactos y luego seguir con tu vida.

Otra ventaja es que por lo general no necesitas comprar nada; estos servicios suelen ser gratuitos y son por lo general servicios agregados por los grandes portales o pagados con la publicidad que se allí se muestra.

También estos servicios se actualizan constantemente, añadiendo nuevas funcionalidades sin que tengas que estar pendiente de andar instalando actualizaciones en tu PC.

Otra de las aplicaciones ideales de la computación nube son los backups (copias de seguridad) de computadoras.

Las desventajas

Sin embargo, tener tu información en la nube también tiene sus desventajas, las cuáles es bueno conocerlas.

Dado que la computación en nube no permite a los usuarios poseer físicamente los dispositivos de almacenamiento de sus datos (con la excepción de la posibilidad de copiar los datos a un dispositivo de almacenamiento externo, como una unidad flash USB o un disco duro), deja la responsabilidad del almacenamiento de datos y su control en manos del proveedor.

La computación en nube ha sido criticada por limitar la libertad de los usuarios y hacerlos dependientes del proveedor de servicios. Algunos críticos afirman que sólo es posible usar las aplicaciones y servicios que el proveedor esté dispuesto a ofrecer. Así, el London Times compara la computación en nube con los sistemas centralizados de los años 50 y 60, en los que los usuarios se conectaban a través de terminales "tontos" con ordenadores centrales. El Times argumenta que la computación en nube es un retorno a esa época.

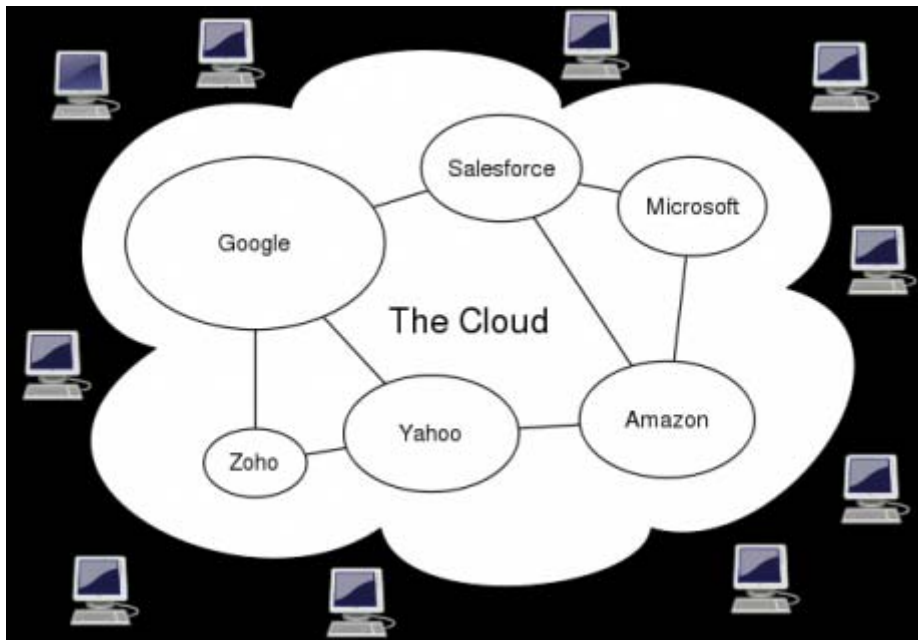
A parte de esto también hay que tener en cuenta una serie de falacias de la computación distribuida, suposiciones que se suelen hacer, y que es necesario superar a la hora de diseñar y programar un sistema distribuido:

- *La red es fiable:* en general, lo es, pero desde el nivel más bajo (un paquete puede rebotar si un router está congestionado, por ejemplo, o un nodo puede desaparecer), al intermedio (ancho de banda insuficiente), hasta el más alto (nivel de presentación), pueden ocurrir múltiples errores.
- *La latencia es nula:* La latencia nunca es nula. Los paquetes están un tiempo no nulo en el medio de transmisión. Esto suele ser importante sobre todo a la hora de paralelizar una aplicación secuencial, pero también si los ordenadores en una aplicación distribuida están en diferentes segmentos de red; la sincronización de los mismos puede ser un problema.
- *El ancho de banda es infinito:* metas lo que metas en la red, cabe. Sin problemas. Lo cierto es que, cuando una red aumenta el número de nodos, llega un momento que alguno de los enlaces puede saturarse.
- *La red es segura:* este es uno de los grandes problemas que nos encontramos, puesto que la red no es, ni mucho menos segura.
- *La topología no cambia:* en ciertas aplicaciones, es más o menos seguro decir que la topología no va a cambiar, y más si el medio no cambia. Pero los nodos se añaden y desaparecen de la red continuamente, igual que los recursos que dependen de ella.

- *Hay un solo administrador:* A veces no hay un solo administrador ni siquiera en cada nodo.
- *El coste de transporte es nulo:* este es un tanto complicado, y no tiene una interpretación única. Para empezar, está claro que mandar un paquete por la red cuesta electricidad; pero también significa que, teniendo en cuenta que ni la latencia es nula ni el ancho de banda es infinito, cada paquete enviado tiene un coste en términos de ocupación del medio; en un medio compartido, habrá que asumir ese coste si abusamos de una conexión. Por otro lado, según varían (y aumentan) los precios de la electricidad hay que plantearse qué tipo de servidor nos da el mejor precio para el transporte de la información.
- *La red es homogénea:* ni lo son las conexiones, ni los nodos conectados a la misma. El problema no es tanto asumir que es así, sino aprovechar las capacidades de todos los nodos y adecuar el paso de mensajes a las características de la red. En algunos casos se puede asumir que es más o menos homogénea, pero en general, la inhomogeneidad hará que trabajar con aplicaciones síncronas (es decir, aplicaciones en que se espera que todos los nodos intercambien información en un punto determinado del cálculo) sea prácticamente imposible (o posible, pero con un desperdicio de CPU considerable).

Todas estas falacias no son inconvenientes como tal, pero pueden llegar a serlo si no se tienen en cuenta a la hora de hacer nuestro servicio.

EJEMPLOS Y APLICACIONES



CLOUDSIM

CloudSim apoya la investigación y desarrollo en el campo emergente de la computación nube, y ofrece una serie de nuevas características:

- Apoyo para la modelización y simulación de gran escala de la infraestructura informática de nubes, incluyendo centros de datos en un único nodo de computación física.
- Autónomo de la plataforma para la modelización de los centros de datos, agentes de servicio, la programación, y la asignación de políticas.
- Disponibilidad de motor de virtualización, que ayuda en la creación y gestión de múltiples, independientes y co-organizado servicios virtualizados en un nodo del centro de datos.
- Flexibilidad para cambiar entre el espacio compartido y de tiempo compartido de asignación de núcleos de procesamiento a los servicios virtualizados.

Estas características de peso de CloudSim podría acelerar el desarrollo de nuevos algoritmos, métodos y protocolos en computación en la nube y, por lo tanto, contribuir a la rápida evolución del paradigma.

Todos los componentes desarrollados en el marco de la Guía CloudSim son liberados como "código abierto" bajo la licencia GPL para fomentar la innovación y pasar a los usuarios con plena libertad.

Como una herramienta de simulación para una tecnología emergente, hay varias cuestiones relacionadas con Cloud Computing que aún en desarrollo. El equipo

CloudSim ha hecho todo lo posible para hacer frente a tantos temas como sea posible, teniendo en cuenta no sólo la investigación desarrollada en los laboratorios de redes, sino también las recientes investigaciones relacionadas con las nubes que se han publicado en revistas académicas y conferencias. Sin embargo, puede haber características que requieren investigaciones de nubes que no están presentes en esta versión de CloudSim. Por otro lado, puede haber características que no son requeridas por los investigadores Cloud.

El software puede ser descargado desde <http://www.gridbus.org/cloudsim/>

Para una información más detallada, podemos consultar el pdf de Rajkumar Buyya: [“Modeling and Simulation of Scalable Cloud Computing Environments and the CloudSim Toolkit: Challenges and Opportunities”](#)

También podemos encontrar algunos ejemplos aquí: <http://www.buyya.com/gridbus/cloudsim/examples/index.html>

GOOGLE APPS

Contiene las siguientes aplicaciones a parte del correo gmail:



Google Sites: El servicio Google Sites constituye un modo sencillo de crear y compartir páginas web



Google Docs: crea documentos, hojas de cálculo y presentaciones y compártelos.

Google Docs y Hojas de cálculo (oficialmente **Google Docs & Spreadsheets**) es un programa gratuito basado en Web, para crear documentos en línea. Incluye un Procesador de textos, una Hoja de cálculo y un Programa de presentación básico. Este sitio se originó una vez comprado el editor en línea Writely por parte de Google.

Se puede crear documentos desde la misma aplicación o Importarlos en los formatos soportados, además de poder exportarlos en diversos formatos estándar o enviar por email. Los formatos con los que trabaja son: txt, doc, rtf, odt y sxw en cuanto a documentos de texto se refiere, ppt y pps son los formatos de presentaciones que tolera y csv, xls y ods en cuanto a hojas de cálculo.

También existe la posibilidad de Colaboración de grupos de trabajo, además de poder compartirlos con múltiples usuarios al mismo tiempo.

Existen ciertos límites al manejar documentos por cada cuenta. Los documentos de texto pueden tener hasta 500 KB más 2 MB para imágenes. Cada hoja de cálculo puede tener hasta 10.000 filas, 256 columnas, 100.000 celdas y 40 hojas 80 páginas. Un usuario puede tener un

total de 5,000 documentos y presentaciones, 5,000 imágenes y 1,000 hojas de cálculo. Sólo pueden abrirse hasta 11 hojas al mismo tiempo. Sólo pueden importarse Presentaciones de hasta 10 MB.



Google Calendar: organiza reuniones, establece horarios y publica información sobre eventos.

Google Calendar es una agenda y calendario electrónico desarrollado por Google. Funciona como una agenda electrónica normal aunque incluye algunas peculiaridades, por ejemplo permite sincronizarlo con los contactos de Gmail de manera que podamos invitarlos y compartir eventos, además de poder hacer búsquedas de eventos que podrían interesarnos



Google Talk: envía mensajes instantáneos a tus compañeros de trabajo o realiza llamadas de voz gratuitas de equipo a equipo.



Página de inicio: consulta tu calendario y tus documentos, añade gadgets y realiza búsquedas en Internet desde un mismo lugar.

Amazon EC2

Amazon Elastic Compute Cloud (Amazon EC2) es un servicio web que proporciona la capacidad de cálculo variable en la nube. Está diseñado para hacer la web de computación escalable más fácil para los desarrolladores.

Amazon EC2 es una sencilla interfaz de servicio web que le permite obtener y configurar con el mínimo coste. Nos proporciona un control completo de sus recursos de computación y nos permite ejecutar Amazon's proven computing environment. Amazon EC2 reduce el tiempo requerido para la obtención y el arranque de nuevas instancias de servicio a los minutos, lo que le permite ampliar rápidamente la capacidad, tanto hacia arriba como hacia abajo, como tus necesidades de computación cambien. Amazon EC2 cambia la economía de la computación, ya que le permite pagar sólo por la capacidad que realmente utilizamos. Amazon EC2 proporciona a los desarrolladores las herramientas para crear aplicaciones con capacidad de recuperarse del fracaso y aislarse de escenarios que fallan normalmente.

SUN GRID ENGINE

Según la pagina de sun, define Sun Grid Engine como:

“Sun Grid Engine Software es líder en el mundo (y más ampliamente desplegado) administrador de sistemas distribuidos. El software Sun Grid Engine soporta escalabilidad extrema, simplifica la administración y apoya una diversidad de escenarios. Cuenta con la industria líder en fiabilidad y ofrece un rápido retorno de la inversión. Originalmente lanzado por Sun en septiembre de 2000, la última versión es Sun Grid Engine 6.2 Update 4 y está construido sobre el proyecto Grid Engine de código abierto.

El lanzamiento de Sun Grid Engine 6.2 Update 4 es el paso evolutivo de los productos Sun Grid Engine, hacia el cloud computing. Cloud computing se ha hecho popular por su paradigma de pago por uso y porque puede transformar el centro de datos de HPC en un centro de beneficios, aprovechando las inversiones en TI que ya están en su lugar.”

Las características de este producto de Sun son:

Features & Benefits	
Service Domain Manager (SDM) Cloud Adapter and initial Power Saving support. It provides an interface to manage Amazon Elastic Compute Cloud (EC2) Amazon Machine Images (AMI). Supports cloud-only and hybrid (local+cloud) Sun Grid Engine environments.	The ability to provide elastic expansion and contraction based on demand fluctuations.
SGE Inspect provides a graphical monitoring interface for Sun Grid Engine clusters and Service Domain Manager instances, all in a single window. It enables to monitor the service framework and the resources being managed by it.	Provides simpler management of both SGE and SDM. It visualizes key indicators in easy-to-read 2D graphs and charts. It also provides historical documentation of management decisions for future investments when needed.
Exclusive Host Scheduling The cluster can be configured to allow jobs to request exclusive use of various components of a given execution host, An exclusive job will only be scheduled on execution hosts that have no jobs currently running. For a parallel job running on multiple machines, this rule applies for the slave tasks of the job as well.	Offers predictable performance guarantees are required. The tight integration of parallel systems is made much easier and can eliminate the need for additional, complex scripts.
Graphical Installer complements the existing command line based installation routine and simplifies initial cluster setup with easy-to-navigate displays. It also verifies installation options.	Helps new users to more easily install the software. See video with demo or Podcast Daniel Templeton interviews Lubos Petrik .
Job Submission Verifier (JSV) is an automatic filter for the system administrator to control, enforce and adjust jobs submissions. The basic idea is that on both the client side and the server side, the administrator has the ability to	JSVs open up a new world of possibilities that didn't exist before, and it will largely end the need for qsub wrapper scripts. This lowers the barrier for users of High Performance Applications, where small typos or inaccurate resource requests from

configure scripts that can read through job submission options and accept, reject, or modify the job submission accordingly.	users can be automatically corrected.
Scalability	Scales up to 63,000 core CPUs

MEEBO

Meebo es un programa de mensajería instantánea que soportan múltiples servicios IM, incluyendo Yahoo! Messenger, Windows Live Messenger, AIM, ICQ, Jabber y Google Talk. Meebo hace los servicios de mensajería instantánea más accesibles a los usuarios que no pueden o no desean descargar el software necesario para ejecutarlos.

Además Meebo permite crearte una cuenta del propio Meebo, con la que podrás unificar todas tus cuentas de servicios de mensajería instantánea. Meebo es un poco más básico (es menos vistoso y tiene algunas aplicaciones menos) que otros programas como Windows Live Messenger , pero, a pesar de ello, no se olvida de aplicaciones como multi-conferencia o videoconferencia.

EYEOS

Eyeos es un escritorio web cloud computing de código abierto. No obstante, eyeos se puede utilizar de diferentes maneras en función de tus necesidades. Es por ello que creo que es mejor introducir eyeos en sus diferentes variantes, dependiendo del uso primario que se le quiera dar al sistema. Tal y como pone en su página web, a continuación vemos sus variantes: uso personal, desarrolladores, educación y administraciones públicas.

Uso personal

“La idea que hay detrás de eyeos es que el conjunto del sistema vive en el navegador web. El cliente sólo debe tener un navegador web para poder trabajar con eyeos y sus aplicaciones, incluyendo Office y PIM. Esto es aplicable tanto por los ordenadores modernos como los obsoletos.”

Desarrolladores

“eyeos proporciona una herramienta para crear fácilmente nuevas aplicaciones. También puedes utilizar eyedesigner para diseñarlas inicialmente. Si estás interesado en programar aplicaciones para eyeos te recomendamos leer la guía de introducción a las aplicaciones y mirar el Centro de Documentación.”

Educación

“eyeos puede proporcionar a escuelas y universidades una completa plataforma web donde estudiantes, profesores y padres tendrán un escritorio personal y colaborativo para trabajar, comunicarse entre ellos y estar organizados dentro y fuera de la escuela.

Los estudiantes y profesores tendrán un intuitivo sistema de escritorio para trabajar con los recursos de la escuela y comunicarse con otros estudiantes y profesores.”

Administración

“Con un servidor central, eyeos puede proveer ayuntamientos, redes de bibliotecas públicas, puntos gratis de Internet y otros entornos públicos con el sistema perfecto para sus usuarios, que tendrán un espacio central para trabajar y comunicarse con los gerentes / responsables centrales, registrándose una vez y utilizándolo posteriormente desde cualquier punto.”

Para más información consultar su web: www.eyeos.org

AVIARY

Aviary es el proyecto de una pequeña empresa llamada Worth1000, LLC, con asiento en Long Island, Nueva York. Su misión, declara, es posibilitar la creación de imágenes, vídeo y audio de toda clase a muy bajo costo, para que se puedan compartir. El sistema es capaz de seguir las producciones mientras son usadas y vueltas a compartir a través de la Web.

Con ese propósito, Worth1000 está construyendo un interesantísimo paquete de aplicaciones que para funcionar solamente requieren tener instalado Adobe Flash 9. Es posible importar imágenes u otro tipo de ficheros y guardarlos o compartirlos en el perfil de usuario.

Ahora mismo podemos contar con seis aplicaciones funcionales:

- Phoenix: un simple pero potente programa de edición de imágenes. Con él se pueden manipular fotografías e imágenes con la precisión de las principales herramientas de Photoshop.
- Peacock: un generador algorítmico de texturas. Experimentar con imágenes basados en píxeles en formas completamente nuevas.
- Toucan: un creador de paletas y pinceles.
- Raven: un editor de vectores. Con Raven podrás crear gráficos vectoriales escalables totalmente apropiados para logotipos o diseños de camiseta.
- Myna: un editor de audio. Puedes remezclar las pistas de música y clips de audio. Aplicación de efectos de sonido y grabar su propia voz o de instrumentos.
- Falcon: un programa para marcar imágenes. Para capturar rápidamente imágenes y páginas web desde tu navegador y luego corta, cambia el tamaño o marca.

Todas estas aplicaciones podrán intercambiar datos entre sí, e interactuar con ellos, de manera que pueden intervenir en un mismo proyecto.

BIBLIOGRAFÍA

<http://www.slideshare.net/gio2345/computacion-en-nube>

<http://geneura.ugr.es/~jmerelo/asignaturas/AAP/AAP-Tema-1.mhtml>

<http://dotsub.com/view/21008c5c-0973-48c2-9e3a-9d16f6539ecb>

http://es.wikipedia.org/wiki/Computaci%C3%B3n_en_nube

<http://jorescobar.com/2008/10/21/como-estar-seguro-en-la-computacion-nube/>

<http://www.mujeresdeempresa.com/blog/?p=430>

<http://aws.amazon.com/ec2/>

<http://www.google.com/intl/es/apps/>

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

<http://enterthegrid.com/primeur/09/articles/monthly/AE-PR-05-09-44.html>

<http://www.buyya.com/papers/CloudSim-HPCS2009.pdf>

<http://es.eyeos.org/>

<http://www.sun.com/software/sge/>

<http://aviary.com/>