

SISTEMA DE ALMACENAMIENTO BASADO EN ARQUITECTURA DE LA COMPUTACIÓN EN LA NUBE DESDE UNA PERSPECTIVA DOCENTE

ALMACENAMIENTO BASADO EN ARQUITECTURA EN LA NUBE

AUTORES: Rodolfo Matius Mendoza Poma¹Patricio Navas Moya²Maira Natalia Martínez Freire³DIRECCIÓN PARA CORRESPONDENCIA: mathius2875@yahoo.es

Fecha de recepción: 20 - 03 - 2017

Fecha de aceptación: 25 - 04 - 2017

RESUMEN

El presente artículo trata del diseño y la arquitectura de sistema de almacenamiento en la nube donde juegan un papel vital en la infraestructura de computación en la nube con el fin de mejorar la capacidad de almacenamiento, así como la rentabilidad. Por lo general, el sistema de almacenamiento en la nube proporciona a los estudiantes un espacio de almacenamiento eficiente con la característica de elasticidad. Uno de los retos del sistema de almacenamiento en la nube es difícil de equilibrar la capacidad elástica que proporciona gran capacidad de almacenamiento y de inversión de alto costo por ello con el fin de resolver este problema en la infraestructura de almacenamiento en la nube, el servidor de almacenamiento de clúster basado en PC de bajo coste está configurado para ser activado para la gran cantidad de datos para proporcionar a los usuarios en la nube. Por otra parte, una de las aportaciones de este sistema se propone como un modelo de análisis utilizando el modelo de red de encolado, que se basa en una arquitectura destinada a proporcionar un mejor tiempo de respuesta, la utilización del almacenamiento, así como el tiempo en espera cuando el sistema está en funcionamiento. De acuerdo con el resultado del análisis en los ensayos experimentales, el almacenamiento se puede utilizar más de 90% de espacio de almacenamiento. En esta investigación, dos partes se describen como el diseño y la arquitectura de sistema de almacenamiento en la nube clúster basado en PC.

PALABRAS CLAVE: Almacenamiento; Computación; Nube.

¹ Ingeniero en Sistemas, Magister en Sistemas Informáticos Educativos, Docente de la Universidad Técnica de Cotopaxi, Facultad de Ciencias Administrativas, Carrera de Secretariado Ejecutivo Gerencial, Administrador de la Plataforma MOODLE de la Facultad, Latacunga – Ecuador.

² Ingeniero en Sistemas, Magister en Gestión de Bases de Datos, Docente de la Universidad de las Fuerzas Armadas ESPE, Departamento de Eléctrica y Electrónica, Carrera de Ingeniería de Software, Jefe de Laboratorio de Producción de Software e Investigación, Planificador de las áreas de computación.

³ Licenciada en Ciencias de la Educación, Profesora de Enseñanza Media en la Especialización de Informática, Magister en Docencia Universitaria y Administración Educativa, Docente Titular de la Universidad Técnica de Cotopaxi, Facultad de Ciencias Agropecuarias y recursos Naturales, Latacunga – Ecuador. E-mail: maira.martinez@utc.edu.ec

BASED STORAGE SYSTEM ARCHITECTURE OF CLOUD COMPUTING

ABSTRACT

Design and architecture of cloud storage system plays a vital role in cloud computing infrastructure in order to improve the storage capacity as well as cost effectiveness. Usually cloud storage system provides users to efficient storage space with elasticity feature. One of the challenges of cloud storage system is difficult to balance the providing huge elastic capacity of storage and investment of expensive cost for it. In order to solve this issue in the cloud storage infrastructure, low cost PC cluster based storage server is configured to be activated for large amount of data to provide cloud users. Moreover, one of the contributions of this system is proposed an analytical model using queuing network model, which is modelled on intended architecture to provide better response time, utilization of storage as well as pending time when the system is running. According to the analytical result on experimental testing, the storage can be utilized more than 90% of storage space. In this paper, two parts have been described such as design and architecture of PC cluster based cloud storage system.

KEYWORDS: Storage; Computing; Cloud.

INTRODUCCIÓN

Cloud Computing es una forma de computación con recursos que son escalables y muy flexibles que proveen una diversa gama de servicios utilizando tecnología web [1]. Además de ser una plataforma de servicio de modo emergente, que se encarga de la organización y servicio de planificación basado en la Internet. En general, el concepto de computación en la nube incorpora infraestructura Web, Web 2.0, las tecnologías de virtualización y otras tecnologías emergentes. Con la tecnología de la computación en nube, los usuarios utilizan una variedad de dispositivos, incluyendo PCs, portátiles, móviles y PDAs a los programas de acceso, almacenamiento y aplicaciones de servicios ofrecidos por las nubes en las universidades. Las ventajas de la tecnología de la computación en nube incluyen el ahorro de costos, alta disponibilidad, y fácil estabilidad. El objetivo principal de la computación en nube es proporcionar servicios de las TIC a cada uno de los estudiantes sobre la nube. El servicio modelos se dividen en la nube como Software-as-a-Service (SaaS), que permite a los estudiantes ejecutar aplicaciones de forma remota desde la nube. Infraestructura-as-a-service (IaaS) se refiere a la informática recursos como un servicio. Esto incluye ordenadores virtualizados con potencia de procesamiento garantizada y ancho de banda reservado para el almacenamiento y el acceso a Internet. Platform-as-a-Service (PaaS) es similar a IaaS, pero también incluye los sistemas operativos y servicios necesarios para una aplicación particular. Es decir, PaaS es IaaS con una pila de software a medida para la aplicación dada. Los Datos almacenamiento-as-a-Service (DaaS) proporciona almacenamiento que el consumidor utiliza incluyendo ancho de banda requisitos para el almacenamiento.

Desde el punto de vista de los estudiantes de la carrera de informática, estudian el servicios de cloud computing proporcionan el software de aplicación y sistema operativo de las computadoras de escritorio al lado de la nube, lo que hace que los usuarios puedan usar en cualquier momento desde cualquier lugar y utilizar los recursos de almacenamiento de grandes escala y de computación. En la computación en nube, almacenamiento en disco es el de los mayores gastos. Nube proveedores de servicios de oferta enormes reducciones en los costos de la capacidad, la eliminación del trabajo necesarios para la gestión de almacenamiento y mantenimiento y aprovisionamiento inmediato de capacidad a un costo muy bajo por terabyte. Por lo tanto, el almacenamiento y la computación en datos masivos son el desafío clave para la infraestructura de cloud computing. En esta investigación, nos centramos en el servidor de almacenamiento de clúster basado en PC que proporciona los dispositivos de hardware y software para la gran cantidad de almacenamiento de datos.

DESARROLLO

1. Nube de Almacenamiento

La nube se ha convertido en un nuevo vehículo para la entrega de los recursos, tales como la computación y almacenamiento de clientes en la demanda. Almacenamiento en la nube es uno de los servicios que proporcionan los recursos de almacenamiento y servicio basado en los servidores remotos basados en cloud computing. Nube de almacenamiento será capaz de proporcionar un servicio de almacenamiento a un menor costo y una mayor fiabilidad y seguridad. La ventaja de la nube de almacenamiento es que permite a los usuarios en los datos de acceso de tiempo. Hoy en día, una gran variedad de sistemas de almacenamiento en la nube están disponible, todos con diferente funcionalidad, optimizaciones y garantías. Es decir, las soluciones se centran en un escenario particular y la condición de servicios se adapta. Como resultado, los sistemas de almacenamiento en la nube varían en el formato de datos (por ejemplo, Tecla / Valor vs. tienda fila), optimización de vías de acceso (por ejemplo, leer vs. escribir, unidimensional vs. el acceso multidimensional), distribución (por ejemplo, solo frente a la distribución multi centro de datos), consulta lenguaje, soporte de transacciones, disponibilidad. Arquitectura del sistema de almacenamiento de nube típica incluye un servidor de control maestro y varios servidores de almacenamiento.

Las ventajas del almacenamiento en la nube no solo son de orden económico, ya que como servicio público ofrecen una vía rápida y sencilla de ganar agilidad en el negocio y de dotarse de una forma rentable y flexible de las últimas tecnologías, no siempre al alcance de todas las organizaciones. [2]

El almacenamiento en la nube se puede utilizar para copiar imágenes de máquinas virtuales desde la nube a ubicaciones en las instalaciones o importar una imagen de máquina virtual a partir de un lugar en las instalaciones de la biblioteca de imágenes de nubes. Además, almacenamiento en la nube puede

ser utilizado para mover imágenes de máquinas virtuales entre las cuentas de usuario o entre los centros de datos [3].

La información que se maneja en la nube debe ser manipulada cuidadosamente con la finalidad que no caigan en manos inadecuadas que se encargan de robar, o en manos de hackers, así como pueden ser extraviados en agujeros de seguridad.

Para la privacidad, se debe tomar en cuenta prevenciones con la finalidad que la información que se esté utilizando sean confidencial y no caigan en manos de terceras personas.

La conectividad es en donde la velocidad de la conexión a internet es muy importante para el acceso a la información y la disponibilidad de las aplicaciones. La Internet es la llave de acceso a la nube, si no existe conexión a la internet no se puede trabajar con Cloud Computing.

Una vez que somos conscientes de las ventajas del “Cloud Computing” podemos tomar la decisión pero debemos antes analizar el mercado de aquellas empresas que ofrecen este servicio y evaluar las características que nos ofrecen para luego compararlas y escoger la que más beneficios y garantías nos ofrezcan, además debemos fijarnos bien en la solidez que dicha empresa tienen en el mercado. Esto con el fin de estar seguros que la empresa que escogemos nos va a resguardar los datos, y que nuestros clientes y empleados puedan acceder a los servicios de forma efectiva [4].

Se debe tomar en cuenta que antes de comprar algún servicios se debe primero analizar todas las garantías y beneficios que nos ofrecen las empresas con dicho servicio con la finalidad de evitar futuros problemas como robo y pérdida de la información.

2. Modelos de Servicio en la Nube

Cloud Computing es un concepto amplio y utilizado de muchas formas, sin embargo, existe un consenso general respecto a los tres modelos principales que describen la prestación de servicios en la nube [5].

SaaS (Software as a Service).

Es un modelo donde una aplicación es proporcionada por un proveedor de servicios por medio de la Internet en la cual pueden acceder diferentes usuarios. En esta aplicación el cliente la puede usar, pero no puede gestionar ni controlar la infraestructura sobre la que ésta se ejecuta, como hardware, almacenamiento o red.

La empresa que presta el servicio del software es la que se encarga de realizar el soporte y mantenimiento técnico de la aplicación que utiliza el cliente, de acuerdo al contrato establecido entre la empresa y el cliente.

PaaS (Platform as a Service)

Este es un modelo que permite dar soporte al ciclo de vida de cada aplicación, como es en la etapa de construcción, así cuando está en marcha. No es

necesario ningún software para la utilización de este servicio, ya que la solución es por medio de diferentes herramientas vía web. El cliente solo tiene control de las aplicaciones desplegadas.

IaaS (Infrastructure as a Service)

Este modelo proporciona el acceso a recursos mediante virtualización en la nube, mediante la internet, una de las características más importantes es la virtualización del hardware mediante el procesamiento. Para lograr el servicio se debe tener un adecuado espacio en los virtualizadores, conexión de red, un buen ancho de red estable y confiable, direccionamiento IP y balanceadores de carga.

3. Modelos de Despliegue de Cloud

Los Modelos de Despliegue se refieren como se organizan y pone en funcionamiento un sistema. Para Cloud Computing, el NIST define cuatro Modelos de Despliegue, también conocidos como “Tipos de Cloud”, que ofrecen la infraestructura necesaria para soportar los Modelos de Servicio (SaaS, PaaS, IaaS).

Cloud Pública

La infraestructura de la nube es establecida para un uso abierto del público en general.

La nube puede pertenecer, ser gestionada y operada por una empresa, una universidad o una organización gubernamental, o cualquier combinación de ellas.

Las infraestructuras siempre se encuentran en las instalaciones del proveedor de cloud [6].

Cloud Privada

La infraestructura de la nube es establecida para uso exclusivo de una única organización que compromete a múltiples usuarios.

Esta infraestructura podrá ser de propiedad, gestión u operación de la organización, de una tercera parte, o de una combinación de ambas.

La infraestructura puede residir fuera o dentro de las instalaciones de la organización [7].

Cloud Híbrida

Las infraestructuras de la nube son una composición de dos o más tipos de las modelos descritas anteriormente (privada, comunidad o pública).

Aun siendo entidades únicas, forman una cierta unidad gracias a la tecnología estandarizada o propietaria que permite la portabilidad de datos y de aplicaciones (p.e., cloud bursting para el balanceo de carga entre nubes) [8].

Cloud Comunitaria

En un modelo de Cloud Comunitaria la infraestructura es compartida y utilizada por un grupo de organizaciones que tienen intereses compartidos, tales como requerimientos específicos de seguridad, políticas o una misión común. Los miembros de la comunidad comparten el acceso a los datos y aplicaciones en la Cloud [8].

4. Arquitectura de un Sistema Académico en una Nube Privada

Arquitectura de red física se basa en la red troncal. Esta arquitectura consiste en la infraestructura de front-end, back-end de la infraestructura y el servidor de almacenamiento de clúster basado en PC. La infraestructura front-end contiene controlador de la nube (CLC), controladores de Cluster (CC), y de la Walrus.

Infraestructura Front-End

El controlador de la nube (CLC) es el punto de entrada en la nube para los usuarios y administradores. Hace interrogaciones gerentes de nodo para obtener información acerca de los recursos, toma decisiones de programación de alto nivel, y las aplica al hacer peticiones a agruparse en un controlador. El Cluster Controller (CC) funciona como el intermediario entre el controlador de nodos y el controlador de la nube. Se reciben solicitudes para asignar imágenes de máquinas virtuales desde el controlador de la nube y, a su vez decide qué controlador de nodos se ejecuta a la instancia de máquina virtual, así como gestiona la red instancia virtual. Todas las imágenes de máquinas virtuales se almacenan en la Walrus que la provisión de un mecanismo para almacenar y acceso a la imagen de la máquina virtual.

Infraestructura Back-End

Infraestructura de back-end consiste en controladores de nodos (NC). Un controlador de nodo se ejecuta en cada nodo que está diseñado para acoger los casos AppScale VM. Estas instancias AppScale VM tiene su propia dirección IP de controlador de nodos. Cada uno del controlador de nodo tiene un dispositivo de almacenamiento local. El almacenamiento local sólo se utiliza para mantener la imagen de la máquina virtual en tiempo de ejecución y para el almacenamiento en caché. Cuando se termina una instancia AppScale VM, Eucalyptus no guarda los datos en el interior la instancia AppScale VM, se necesita un servidor de almacenamiento externo para guardar estos datos de la aplicación en la instancia AppScale VM. Por lo tanto, antes de terminar una instancia, debe cargar los datos al servidor de almacenamiento.

5. Diseño de Servidor Cluster de Almacenamiento

Topología Física de CCPS

En esta sección, se describe cómo construir topología física de CCPS. El marco general de CCPS se compone de tres capas que son capa de servicios de aplicaciones basados, Hadoop Distributed File System (HDFS) capa de banda y la capa de clúster PC. La capa de aplicación basada en web proporciona una interfaz para los usuarios cuyos puede almacenar sus propias aplicaciones

tales máquinas virtuales (VM), conjuntos de datos y multimedia de datos, entre otros. La capa HDFS soporta el sistema de archivos para la capa de clúster PC. HDFS es como un sistema de archivos a nivel de usuario en el grupo que explota el sistema de archivos nativo en cada nodo para almacenar datos. Los datos de entrada se dividen en bloques, por lo general 64 megabytes y cada bloque se almacena como un archivo separado en el sistema de archivos local. HDFS es ejecutado por dos servicios: el NameNode y DataNode. La NameNode es responsable de mantener el árbol de directorios HDFS, y es un servicio centralizado en el clúster que opera en un solo nodo. Los clientes en contacto con el NameNode el fin de realizar las operaciones del sistema de archivos comunes, como abrir, cerrar, renombrar y eliminar. El NameNode no almacena datos en sí HDFS, sino que mantiene una correspondencia entre el nombre del archivo HDFS, una lista de los bloques en el archivo, y el DataNode en el que se almacenan los bloques. La capa de clúster PC proporciona para almacenar gran cantidad de datos.

Cluster

Hoy en día, muchas organizaciones necesitan sistemas de almacenamiento de terabytes, pero son muy caros y requieren mayor grado de habilidades para su operación y mantenimiento. Por lo general, una PC de escritorio contener más de 100 GB de disco duro (HDD), al menos 256 MB o más de RAM y 2 GHz o superior. Una típica instalación de un sistema operativo y otro software instalación no utilice más de 20 GB de almacenamiento en disco duro. Esto deja en promedio aproximadamente el 80% del espacio de almacenamiento para ser utilizado. Por lo tanto, se propuso un servidor de almacenamiento de clúster basado en PC que es barato y fácil de mantener. Clúster PC a ser popular en el 1990 para el procesamiento por lotes de aplicación de alto rendimiento.

Un cluster de PC es una colección de nodos informáticos, que está interconectado por una alta velocidad de conmutación de red, todos los nodos pueden ser utilizados individualmente o colectivamente como un clúster.

Vista del Sistema de Servidor de Almacenamiento Basado en Pc Cluster

Servidor de almacenamiento de clúster en PC intenta transferir de la computación en cluster para servidor de almacenamiento. El servidor de almacenamiento utiliza componentes de PC de bajo costo. Los archivos de gran tamaño pueden almacenar los datos a través de múltiples nodos. El presente sistema se basa en una arquitectura cliente-servidor. Cada máquina individual de un grupo se conoce como un nodo. En PC clúster se compone de un nombre de nodo como servidor y muchos datos Nodos como clientes. CCPS (Cloud based Cluster PC System) utiliza HDFS (Hadoop Distributed File System) para almacenar datos de la colección de los nodos. Cada nodo tiene su memoria, dispositivos de E/S y sistema operativo propio. Los nodos están físicamente separados y conectados a través de una LAN [9].

En CCPS consiste en un NameNode y DataNodes. NameNode gestiona todo el cluster de PC y mantiene los metadatos de HDFS que contiene la información de los bloques, las ubicaciones actuales de los bloques, y el seguimiento de los estados de todos los nodos del clúster. NameNode no registra ningún cambio en el espacio de nombres del sistema de archivos como abrir, cerrar, cambiar el nombre de archivos y directorios. También determina la asignación de bloques para DataNodes. Los DataNodes realizan el almacenamiento físico de los archivos. DataNodes también realizan la creación de bloques, la supresión y la replicación según instrucciones de la NameNode. En CCPS, los archivos se dividen en bloques que se almacenan como unidades independientes. Cada bloque se replica en un pequeño número de máquinas separadas (típicamente 3) para tolerancia a fallos. En el grupo PC, cada máquina individual de un clúster se conoce como un nodo. Nuestro sistema se basa en una arquitectura cliente-servidor y consiste en un nodo servidor (NameNode) y muchos clientes nodos (DataNodes) con el fin de hacer que funcionen como una sola máquina. NameNode tiene dos tarjetas de redes y uno es conectarse a la infraestructura de aplicaciones para usuario con el interruptor gigabyte.

6. Modelo Funcional de Almacenamiento Cluster Pc Basado Servidor

El servidor de almacenamiento suelen procesar muchos puestos de trabajo simultáneos (es decir, de almacenamiento de archivos) cada una de las cuales utiliza para varios recursos compartidos: acceso a los archivos, el tiempo de procesador y ancho de banda de la red [9]. Dado que un trabajo puede utilizar un recurso en cualquier momento, el resto de trabajos deben esperar en una cola para su turno en el recurso. Como trabajos reciben servicio en el recurso, se eliminan de la cola y nuevos puestos de trabajo llegan y se unen a la cola. La teoría de colas es una herramienta que ayuda a calcular el tamaño de las colas y el tiempo que los trabajos pasan en ellos. Nuestro modelo utiliza modelo de red M/M/1 de colas donde M representa la naturaleza de Markovianas o las tasas de llegada y de servicio y 1 indica el número de servidores conectados a la cola, para proporcionar número medio de archivos en el sistema, el número promedio de tiempo de servicio por archivo y la utilización total del sistema. El servidor de almacenamiento de clúster basado en PC se compone de muchos PC. Los datos almacenados en el servidor de almacenamiento de clúster basado en PC pueden acceder desde la red de la nube privada.

Servidor de almacenamiento de clúster basado en PC se compone de N máquinas heterogéneas independientes. Donde 1 es Name Node y los restantes N-1 son máquinas Data Nodes que almacenan un total de B diferentes bloques b_1, b_2, \dots, b_m donde $B > N$. NameNode mantiene diferentes bloques de metadatos $M = \{M_1, M_2, \dots, M_n\}$, donde n es el número total de metadatos. Cada DataNode D_i , hay bloques B_i guardadas en él, y $B_i = \{b_{i1}, b_{i2}, \dots, b_{imi}\}$ es el conjunto de bloques que pertenecen a D_i . Construimos modelo analítico para el conglomerado PC servidor de almacenamiento basado con llegadas de Poisson en el que las peticiones de E / S se emiten al azar en las solicitudes de tiempo han completado.

En el modelo m , la aplicación envía una solicitud de almacenamiento al servidor de clúster basado en PC a través de la red IP. Tasas de llegada de Poisson de m petición arroyos son la tasa promedio de llegada de Poisson donde llegan nuevos puestos de trabajo en la cola y la cantidad promedio de tiempo que se necesita un servidor para procesar este tipo de puestos de trabajo es el tiempo de servicio del servidor. Sea i el número aleatorio de archivos en el sistema.

El sistema se puede modelar, cada estado representa el número de archivos en el sistema. Como el sistema tiene una cola infinita y la población es ilimitado, el número de estados del sistema que puede ocupar es infinito: el estado 0 (no hay archivos en el sistema), el estado 1 (1 archivo), el estado 2 (dos archivos), entre otros. Como la cola nunca será completa y el tamaño de la población es infinita, tasa de llegada, tasa promedio, es constante para cada tipo de estado y el servicio también es constante para todos los estados. De hecho, independientemente del estado, podemos tener sólo dos eventos un nuevo archivo de llegada. Así que si el sistema está en el estado i , va al estado $i + 1$. Un archivo sale del sistema. Así que si el sistema está en el estado i , se pasa al estado $i - 1$.

Evaluación y análisis del modelo

Este modelo tiene dos parámetros. En primer lugar, se especifica la intensidad de la carga de trabajo que en este caso es la velocidad a la que llegan los archivos (por ejemplo, un archivo cada dos segundos). En segundo lugar, hay que especificar el servicio. Se especifica también la demanda que es el requisito de servicio promedio de un archivo. En esta validación se asumen evaluar en diferente tasa de llegada de 5, 10, 15, 20, 25, 30 archivos por segundo, respectivamente, y servicio de tarifa es 32 segundos por archivo. El tiempo medio de respuesta, utilización y espera promedio de tiempo del sistema con diferente tasa de llegada.

Según la ecuación $T = \frac{N}{\lambda} = \frac{1}{\mu - \lambda}$, donde T = Retraso medio del archivo, N = Promedio del número de archivos en el sistema, μ = Tarifa media del servicio, λ = Tasa de llegada de Poisson, el tiempo medio de respuesta es 0,03704, 0,04545, 0,05882, 0,08333, 0,14286 y 0,5 respectivamente. De acuerdo con los resultados, el tiempo medio de respuesta depende de la tasa de llegada, cuando la tasa de llegada es aumento, el tiempo medio de respuesta se puede aumentar.

Según la ecuación Utilización = $\sum_{i=0}^N \rho$ donde ρ = La fracción de tiempo en que el servidor está ocupado, el promedio de utilización del servidor es 0.15625, 0.3125, 0.46875, 0.625, 0.78125 y 0,9375 respectivamente. La utilización media depende de la tasa de llegada del sistema, cuando la tasa de llegada es aumento, la utilización media del servidor también se puede aumentar.

De acuerdo con los resultados obtenidos, el tiempo medio de espera en función de la tasa de llegada del sistema. Por lo tanto, el modelo analítico propuesto es

más eficaz en la disminución de la tasa de llegada que el aumento en la tasa de llegada.

CONCLUSIONES

Según este diseño y arquitectura de servidor de almacenamiento de clúster basado en PC se configura en máquinas de computación de bajo costo como servidor de almacenamiento de la nube.

Se puede utilizar para almacenar gran cantidad de capacidad de procesamiento de datos y de alto rendimiento.

Este trabajo se propone un modelo de análisis utilizando modelo de red de colas M/M/1, que se inspira en la arquitectura destinada a proporcionar un mejor tiempo de respuesta, la utilización de almacenamiento, así como la espera del turno cuando el sistema está funcionando.

Como puede verse a partir de los resultados experimentales, el almacenamiento se puede utilizar más de 90% de espacio de almacenamiento. Como trabajos futuros, los mayores desafíos del sistema son de alta escalabilidad y tolerancia a fallos de Name Node.

BIBLIOGRAFÍA

Abadi, D.J. (2009). "Data management in the cloud: limitations and opportunities," *Eng datos IEEE. Toro.*, vol. 32, no. 1, pp. 3 - 12, 2009.

Areitio, J. (2010). "Protección del Cloud Computing en seguridad y privacidad," *Revista española de electrónica*, vol. 2, no. 666, pp. 42-48, 2010.

Barrios, R. and Casadei, L. (2014). "Promoviendo el uso de Google Drive como herramienta de trabajo colaborativo en la nube para estudiantes de ingeniería," *Revista de Tecnología de Información y Comunicación en Educación EDUWEB*, vol. 8, no. 1, pp. 43-56, 2014.

Joyanes, L. (2009). "La Computación en Nube (Cloud Computing):," *Revista cuatrimestral de las Facultades de Derecho y Ciencias Económicas y Empresariales*, vol. 1, no. 76, pp. 95-111, 2009.

Morales, E.R. (2010). "Una experiencia personal: Google Drive para el portafolio virtual docente-estudiante en la docencia Universitaria," *Revista Tecnológica ESPOL*, vol. 28, no. 2, pp. 163 - 176, 2010.

Murazzo, M., Rodríguez, N., Villafañe, D. and Gallardo, D. (2013). "Desarrollo de aplicaciones colaborativas para Cloud Computing," in *XVIII Congreso Argentino de Ciencias de la Computación*, La Plata, 2013.

Santamaria, F.A., Rodríguez Amézquita, J.C. and Molina, F.A. (2014). "Administración y Configuración de una Nube Privada y Segura y Personalizada para Pymes," in *CIIS 2014*, Tunja, Boyacá, 2014.

Sinisterra, M.M. and Henao, T.M. (2012). "Clúster de balanceo de carga y alta disponibilidad para servicios web y mail," *Informador Técnico*, vol. 2, no. 76, p. 93, 2012.

Zhang, Cheng and Boutaba, R. (2010). "Cloud computing: state-of-the-art and research challenges," *Journal of internet services and applications*, vol. 1, no. 1, pp. 7-18, 2010.