

Despliegue de un Sistema de Almacenamiento Distribuido (Ceph) en el Centro de Datos Universitario. El caso de la Universidad Nacional de Cuyo.

Guillermo Calleja¹ Roberto Cutuli²

Coordinación de Tecnologías de la Información y las Comunicaciones,
¹ Departamento de Sistemas de Información, ² Departamento de Redes y
Telecomunicaciones, Rectorado
Universidad Nacional de Cuyo, Centro Universitario, 5500 Mendoza, Argentina
{gcalleja, rcutuli@uncu.edu.ar}

Resumen. Los desarrollos relacionados con las Tecnologías de la Información y las Comunicaciones, han dado lugar a una paulatina convergencia de redes de infraestructura y servicios de valor agregado. Esto plantea crecientes demandas a los administradores de las redes y los sistemas en general y por supuesto a los responsables de las áreas de Tecnologías de la Información y las Comunicaciones universitarias. Sin embargo, los recursos muchas veces son escasos en términos económicos, de infraestructura, de recursos humanos para su despliegue, operación y administración. Por otro lado resulta necesario implementar de manera adecuada servicios y equipos para Educación a distancia; Repositorios Digitales o e-ciencia en general. En este contexto la infraestructura de almacenamiento de los Centros de Datos junto con las redes de comunicaciones constituyen un elemento central para la interacción de una universidad con otros centros de I+D, la sociedad y los integrantes de la comunidad. En particular, la infraestructura, los protocolos de red y los sistemas de almacenamiento de datos sobre la cual se basa todo el sistema de Tecnologías de la Información y las Comunicaciones juegan un rol central en las organizaciones del siglo XXI y ha sido objeto de atención durante las ediciones anteriores de TICAL. El trabajo presenta los antecedentes del Centro de Datos de la Universidad Nacional de Cuyo, describe la infraestructura actual y plantea el despliegue de sistemas de almacenamiento sobre un sistema distribuido libre conviviendo con otras tecnologías de sistemas de almacenamiento en producción. Se enfatiza la necesidad de contar con recursos dentro y fuera de la Universidad que en muchos de los casos no dependen directamente de los administradores de sistemas de tecnologías de la información y las comunicaciones universitarias.

Palabras Clave: Sistemas de archivos distribuidos; Sistemas de almacenamiento, Infraestructura de centros de datos; Redes de transmisión de datos; Despliegue de nuevas infraestructuras de código abierto; Seguridad y resguardo de la información. Nube. Cloud computing; Big data.

Eje temático: Infraestructura y Desarrollo de Software.

1 Introducción

La Universidad Nacional de Cuyo (UNCuyo) [1] es la más grande del centro oeste de Argentina con más de 5000 puestos de trabajos y múltiples enlaces hacia otras redes institucionales y la propia Internet.

En un trabajo anterior presentado en TICAL 2011 [2] los autores describieron las características de la red de la Universidad, se hizo énfasis en la seguridad de la misma y se analizaron las diferentes herramientas de seguridad implementadas.

Además del tráfico de datos, típico de una red universitaria existen diferentes demandas como telefonía o almacenamiento de contenidos para servicios de valor agregado como Sistemas de Información de Gestión Universitaria, Repositorios Digitales: Biblioteca Digital, Centro de Documentación Histórica, producción de contenidos audiovisuales, Educación a Distancia; Sistemas de Videoconferencia, IPTV y también servicios relacionados con la e-ciencia que plantean nuevas necesidades y requisitos a los administradores de la infraestructura y servicios.

En este trabajo se presenta la infraestructura actual de Almacenamiento del Centro de Datos, sus elementos y el plan de despliegue de un sistema de almacenamiento distribuido.

El trabajo está organizado de la siguiente manera: en la sección 2 se presenta una síntesis de la Infraestructura del Centro de Datos de la Universidad y la demanda de almacenamiento de datos en la Universidad. En la sección 3 se presentan algunos antecedentes sobre los sistemas de almacenamiento en la UNCuyo, en la sección 4 se muestran algunas experiencias similares, en la sección 5 se discute el proyecto para la implementación del sistema de almacenamiento Ceph. En la sección 6 muestra el despliegue en curso. Finalmente, en la sección 7 se presentan las conclusiones de este trabajo.

2 Infraestructura del Centro de datos, consideraciones de procesamiento y almacenamiento de datos. La demanda de almacenamiento de datos.

La red de la Universidad Nacional de Cuyo posee más de 5000 puestos de trabajo conectados con sus usuarios en línea. Estos recursos están distribuidos a lo largo de la red que interconecta a las facultades e institutos dentro o fuera del campus Universitario. Estos puestos de trabajo generan y demandan información que es procesada y almacenada en el centro de datos universitario, el cual cuenta con más del 50% de los servidores que funcionan dentro de toda la Universidad Nacional de Cuyo (UNCuyo).

Actualmente en este centro de datos existen al menos 32 servidores físicos de distintas características, i.e. servidores independientes con hardware de última generación, otros con arquitectura del tipo BLADE[3], y también, se cuenta con un pequeño porcentaje de servidores autónomos con una antigüedad promedio de 10 años sobre los cuales corren sistemas de baja criticidad.

Sobre el 60% de esta infraestructura de servidores físicos están montados los motores de virtualización (Hypervisor), sobre los cuales actualmente están corriendo alrededor de 155 servidores virtuales, el resto de los servidores físicos corren sistemas operativos independientes, por fuera de algún ambiente de virtualización.

Todos estos servidores están almacenados, o guardan sus datos, en infraestructuras de almacenamiento de datos, preparados para tal fin del tipo SAN -Arreglo de discos interconectados con una red de Fibra Óptica- FIBRE CHANNEL STORAGE AREA NETWORK[4]. También son utilizados los sistemas de almacenamiento locales propios de los servidores, los que ofrecen una capacidad total de unos 40TeraBytes.

El Centro de Datos de la Universidad, se muestra un esquema en la Figura 1, depende del área de Redes y Servidores, y ésta, de la Coordinación de Tecnologías de la Información y las Comunicaciones de la Universidad Nacional de Cuyo. Aunque no cuenta con un presupuesto fijo anual para invertir en infraestructura de servidores o redes, normalmente las inversiones que se efectúan para el área se realizan cada un cierto tiempo, lo cual ronda en promedio en unos 5 años. Se planifica la compra de la infraestructura contemplando las necesidades actuales y proyectando un crecimiento de unos 5 años hacia adelante. La última compra planificada de un sistema de almacenamiento fue en el año 2010, esa infraestructura se puso en producción durante el 2011, y tuvo en cuenta una capacidad de crecimiento hasta el año 2015. Pensando en una nueva compra ese año, para ponerla en producción en el 2016 y con una previsión hasta el 2020.

Por distintos factores como la situación económica de la Argentina, lo cual conllevó a restricciones en la importación de hardware de determinadas características, y la falta de presupuesto para los organismos del sector público, entre estos las Universidades Nacionales, se llegó al 2016, sin la posibilidad de poder efectuar una nueva compra de hardware destinada al almacenamiento de información.

También durante esos 5 años hubo un crecimiento exponencial en la demanda de espacio de almacenamiento, provocado por el enorme crecimiento que tuvo la producción de contenidos multimedia, como así también la digitalización de la mayoría de los procesos administrativos dentro del ámbito Universitario de Argentina.

Ya en el 2016, existe una gran demanda de espacio de almacenamiento. La primera proyección que se efectúa es: contar con un espacio de al menos 100TeraBytes, para ser utilizados a partir de 2017, con la posibilidad de tener un crecimiento, sobre el mismo equipamiento a instalar, de hasta los 200TeraBytes en total. Esto para cubrir las demandas futuras hasta el 2020 inclusive.

Uno de los mayores inconvenientes era el presupuestario, ya que necesitaba de una inversión inicial que estaba muy por encima del presupuesto disponible para el área. También, el sector de infraestructura de redes y centro de datos, tiene una gran demanda presupuestaria, aún no satisfecha a la fecha.

El gran problema que se suscita, es que la necesidad de espacio de almacenamiento es algo que no puede ser ignorado. Más aun habiendo tenido algunos incidentes en áreas tales como el de producción de contenido multimedia, los cuales llevaron a la pérdida total de una cantidad enorme de material en crudo y producciones terminadas, a causa de una política de operación inadecuada, además de no haber contado con un sistema de respaldo adaptado a las necesidades de cada sector. La mayor parte de esos contenidos, era guardado sobre sistemas de almacenamiento compartido entre muchos usuarios, los que no contaban con una política de acceso a la información y resguardo

de la misma, como así también de copias de seguridad que sirvieran en el caso de producirse la pérdida de la información almacenada en los sistemas que se utilizan a diario.

La imposibilidad de contar con presupuesto, y la demanda inmediata por almacenamiento, llevó a efectuar un análisis detallado de: ¿cuál era la demanda específica, por sector y por tipo de almacenamiento, requerido por las distintas áreas? Existían problemas como: restricción presupuestaria, demasiada demora para efectuar una compra por medio de una licitación: i.e., los tiempos administrativos internos son muy grandes, como así también los tiempos para la provisión y puesta en funcionamiento del nuevo equipamiento, etc.

De este análisis de las necesidades técnicas, posibilidades presupuestarias y otras restricciones con las que se contaba, estudiando distintas implementaciones y casos reales, se pone foco en una solución que puede servir a las necesidades y demandas de almacenamiento actuales y futuras de la Universidad. Se decide generar un ambiente de prueba, para luego desarrollar e implementar un ambiente de producción con el Sistema de Almacenamiento Escalable y Distribuido con licencia en Software Libre – CEPH FILE SYSTEM [5]. Este almacenamiento definido por software aporta virtualización al almacenamiento de datos de la organización. Se discute en la sección 5. En la figura 1 puede verse un esquema general del centro de datos de la universidad Nacional de Cuyo.

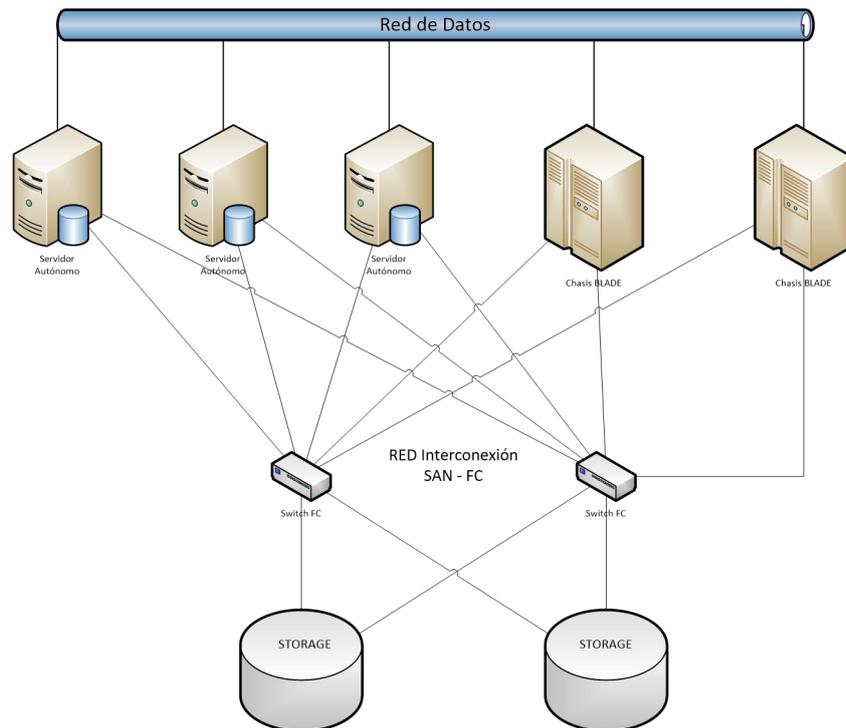


Fig. 1. Esquema general del centro de datos de la Universidad Nacional de Cuyo.

3 Sistemas de Almacenamiento de datos en la UNCuyo

3.1 Antecedentes

Hasta el año 2003 la Universidad Nacional de Cuyo contaba con servidores del tipo DeskTop o Tower con discos interface IDE[6], en los cuales estaban implementados en su mayoría los servidores de red que servían la conectividad, internet y los servicios tales como correo electrónico, proxy y web server. En aquel entonces había algunos servicios de red que se ofrecían solamente por la red local o intranet de la Universidad, como sistema de liquidación de sueldos, sistema de digesto administrativo entre los más importantes. La capacidad de almacenamiento total era de aproximadamente 500GigaBytes.

A partir de 2004 se comenzó a tomar conciencia sobre la importancia de contar con un Centro de Datos y servidores dedicados para tal fin, en ese entonces se efectúa la compra de aproximadamente 10 servidores dedicados a tal fin con las características de ser apropiados para el montaje en rack (RackMount). Estos servidores contaban con sus propios sistemas de almacenamiento local, los que en su mayoría estaban compuestos por 2 discos SCSI[7] de 146GBytes y para los sistemas más críticos, se utilizaron los mismos servidores que contaban con tarjeta controladora RAID[8] en Hardware y 5 o 6 discos de 146GBytes. Entre 2005 y 2009, se incorporaron algunos servidores más al Centro de Datos, los cuales ya en su mayoría contaban con tarjetas controladoras de discos RAID y conjuntos de hasta 6 discos SCSI[7]. La capacidad de almacenamiento total era de aproximadamente 5TeraBytes.

Ya en 2010, se efectúa la nueva planificación de servidores para los próximos 5 años y con ella confecciona y se adquiere la nueva infraestructura de servidores para el centro de datos, la cual contaba con servidores de arquitectura blade: IBM BLADE CENTER y sistemas de almacenamiento interconectados con una red de Fibra óptica (SAN), como se muestra en la figura 1. La capacidad de almacenamiento total era de aproximadamente 30TeraBytes.

Entre 2013 y 2016 se comienza a implementar una solución de almacenamiento con GlusterFS[9], la cual fue utilizada solamente para guardar información de respaldos. La capacidad de almacenamiento total estaba en aproximadamente 40TeraBytes.

En la segunda mitad de 2016, dada la creciente demanda de almacenamiento para objetos multimedia, y dado que la experiencia con GlusterFS no había sido del todo satisfactoria, por los problemas que se tuvieron respecto del hardware sobre el cual se implementó, se comienza a analizar una solución alternativa por intermedio del sistema de archivos distribuido Ceph, el cual reduce el costo de almacenar datos de la organización, aumenta su capacidad para gestionar el crecimiento exponencial de datos de manera eficiente, flexible y automática, y no posee puntos débiles.

3.2 Implementación inicial.

En base a las necesidades cada día mayor de las distintas áreas, por espacio de almacenamiento, luego de analizar los recursos de hardware disponibles en la organización, y pensando en la que la posibilidad de facilitar la provisión futura de hardware en el contexto en la cual nos encontrábamos, se utilizó equipamiento de PC de escritorio que estaba en desuso para las necesidades actuales según los requerimientos de un equipo de escritorio para un usuario final. Se contaban con PC de escritorio con procesador de 32 bits, 1 GigaByte de memoria RAM, un disco duro de 160GigaBytes, tarjeta de red de 1GigaBit/seg.

Las áreas que demandan urgente un sistema de almacenamiento son: Biblioteca Digital – Centro de Documentación Audiovisual – SID UNCuyo[10] (Repositorio de Archivos de Video Históricos), Televisión Digital Abierta (TDA): SeñalU[11] – Canal Televisión Digital de la Universidad Nacional de Cuyo (Archivos de video en crudo y Producciones finales), Dirección de Sistemas de Información y Dirección de Redes y Servidores (Respaldo de Información).

El esquema inicial del Sistema de Archivos distribuido Ceph, estaba formado por un sistema que mantiene el mapa del estado del almacenamiento y el software de Metadatos (Metadata Software), llamado MONITOR-MDS[5], con las características del equipo PC descritas en el inicio de la sección 3.2. 6(seis) PC: Dispositivos de Almacenamiento de Objetos (Object Storage Device): OSD[5], de las mismas características del PC MONITOR, con el agregado de un Disco Duro de 2 TeraBytes de capacidad. Con esta configuración se logró obtener 9 TerBytes de Capacidad.

3.3 Implementación Actual.

Inicialmente se utilizó una configuración básica. Analizando el funcionamiento del sistema, se adoptaron algunas recomendaciones de infraestructura, indicadas en la documentación. Por ejemplo: para lograr la eliminación de puntos débiles, la recomendación es contar con al menos tres equipos con la función de MONITOR, dado que si existe solo uno y falla, el sistema queda fuera de servicio. Entonces, para mejorar el rendimiento del sistema, balanceando las cargas de procesamiento, darle mayor confiabilidad, redundancia y alta disponibilidad, se agregaron 3(tres) MONITORES –MDS, uno en un servidor físico y los otros dos en servidores virtualizados, con diferentes motores de virtualización (Hypervisor) y en distintas infraestructuras. Además, para aumentar la capacidad de almacenamiento se agregaron 2 nodos OSD, a los 6 ya existentes, logrando así una capacidad de almacenamiento de alrededor de 14TeraBytes. Estas funcionalidades agregadas, contribuyen al crecimiento exponencial de la capacidad de almacenamiento de datos de manera eficiente, flexible y automática. La configuración se muestra en la figura 3.

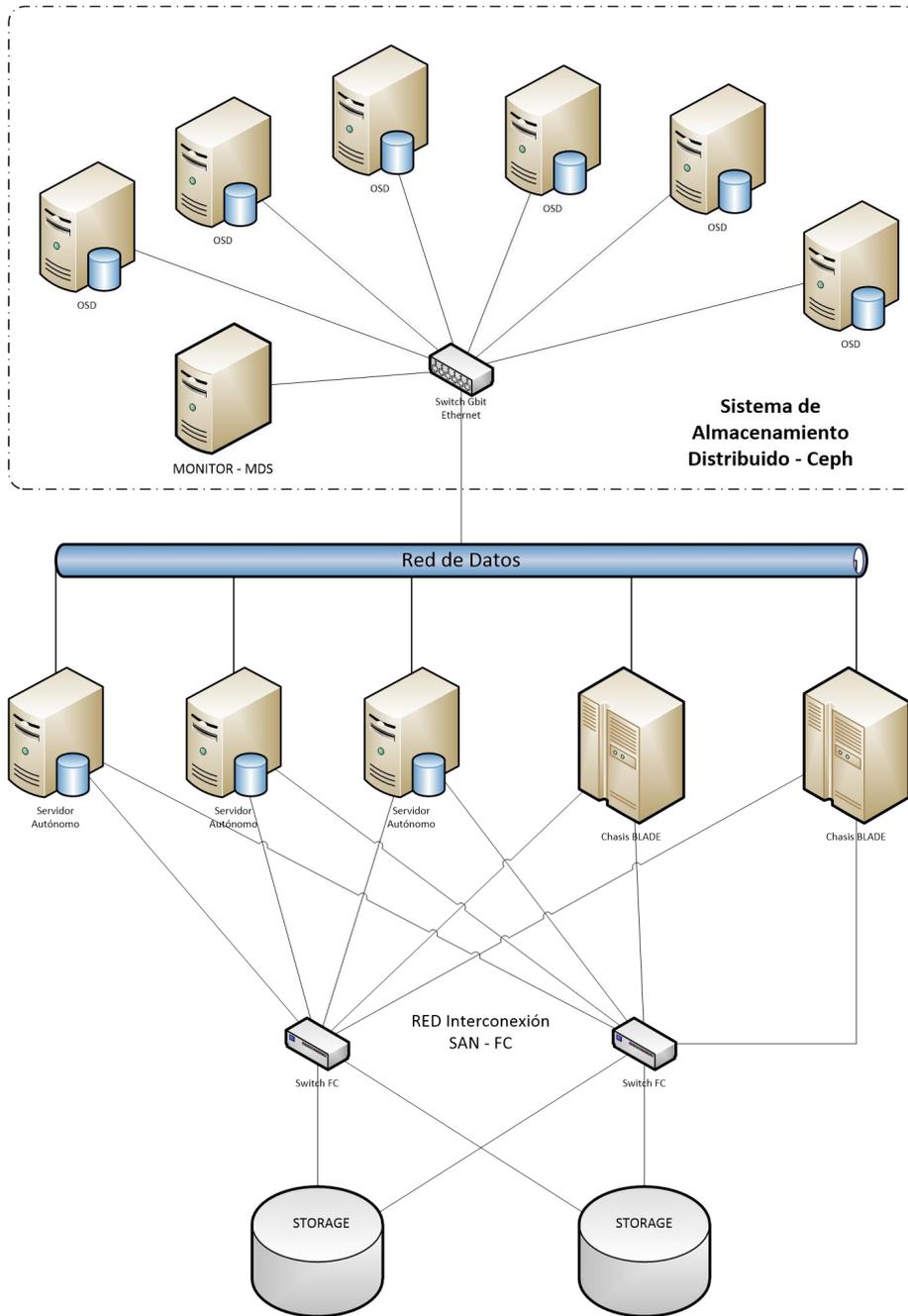


Fig. 2. Esquema Inicial de (Ceph File System) – 1Monitor-MDS+6OSD.

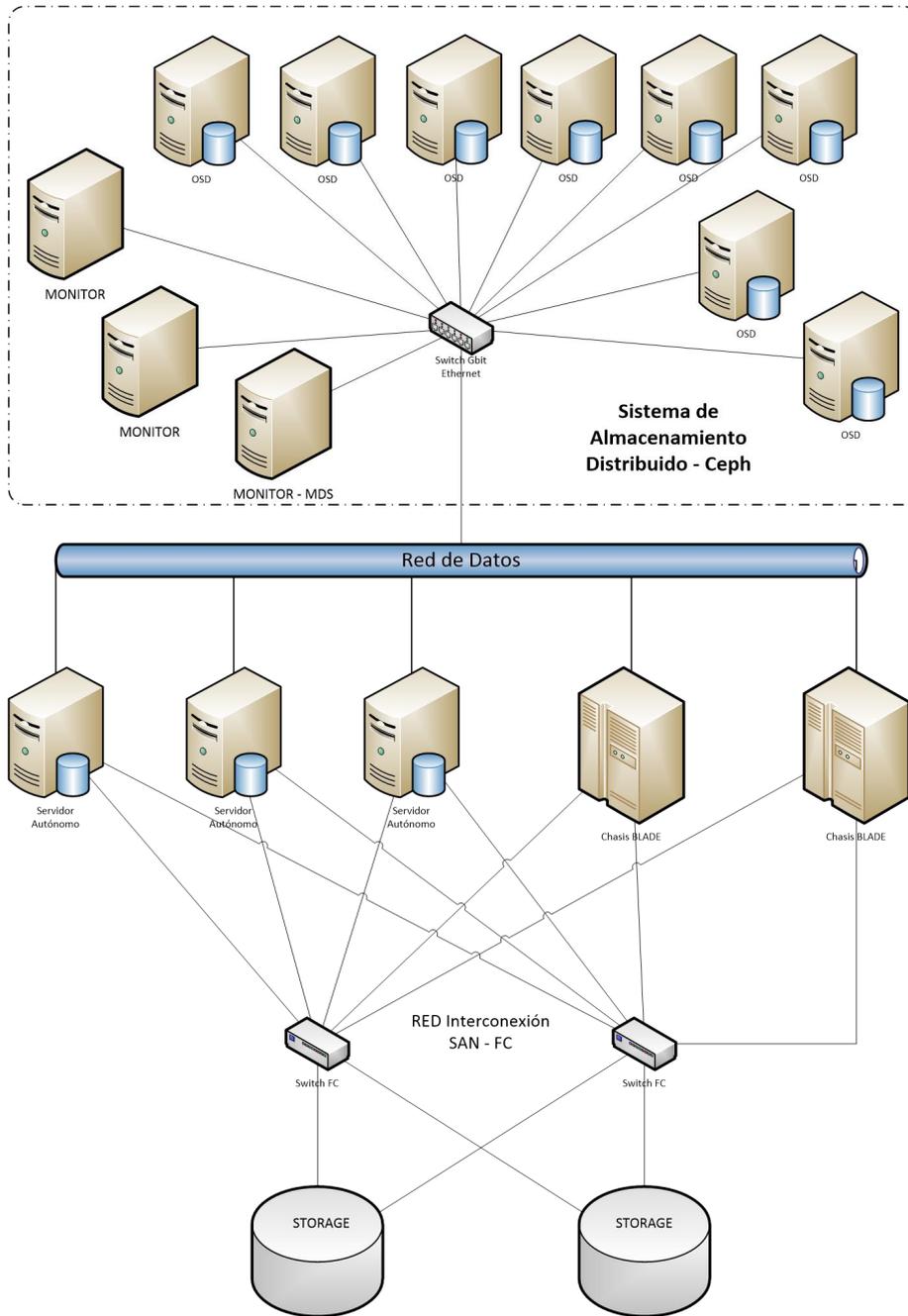


Fig. 3. Esquema Actual de (Ceph File System) – 3 Monitor-(1)MDS+8 OSD.

4 Experiencias Similares

Se han relevado, en ámbitos universitarios y también en empresariales, algunos proyectos e implementaciones de Sistemas de Archivos Distribuidos con licencia de Software libre. Uno de ellos es el propuesto en la Universidad Politécnica de Madrid [12], basada en OpenStack[13] para entornos de Big Data[14], en donde en el escenario presentado, constituye un despliegue inicial a modo de maqueta. Sin embargo, para escenarios de producción de grandes centros de datos, recomiendan una solución basada en la redundancia de recursos y la alta disponibilidad.

Otra Implementación de Ceph, es en la Universidad de Illinois[15]. La aplicación en esta organización, está orientada a entornos de Cloud Computing[16] y Big Data.

Por último, en el ámbito Universitario, la Monash University[17], implementó Ceph en 2012. Esta es una de las Universidades del hemisferio sur que tiene implementado el mayor almacenamiento sobre esta tecnología.

Además se pueden encontrar muchas otras implementaciones empresariales[18], del sistema de Archivos Ceph, entre las que se listan las siguientes:

CERN, CISCO, WESTERN DIGITAL, ZTE, CATALYST, RACKSPACE, CLOUDIN, PROXMOX, TERADATA, SWITCH, DEUTSCHE TELEKOM, BLOOMBERG, UKFAST, DESPEGAR.COM, etc.

Finalmente es de destacar que existen similitudes en los objetivos del despliegue de esta tecnología, más allá de la finalidad de las organizaciones donde se implementa. Estas son: Cloud Computing y Big Data. Es importante tener en cuenta las experiencias anteriores, ya que existen aspectos similares en las organizaciones donde se va a desplegar la tecnología, para lo que podrían tomarse como referencia los trabajos desarrollados en organizaciones semejantes.

5 Proyecto de despliegue del Sistema de Almacenamiento Distribuido en la UNCuyo

5.1 Objetivos del despliegue de Ceph.

- Adecuar el centro de datos y el sistema de red, para poner en producción, en donde sea adecuado y en forma masiva, el Sistema de Almacenamiento Distribuido Ceph.
- Hacer convivir el sistema con otros tipos de almacenamiento según las características y usos de cada uno de ellos.
- Desplegar el sistema, en donde sea posible efectuar la implementación, en todos los sitios fuera del centro de datos principal, de acuerdo a las necesidades, y según los requerimientos, la infraestructura instalada, y los recursos humanos disponibles, y su capacitación.
- En todo lugar dentro de la UNCuyo, donde se requiera y adapte, hacer masiva la utilización de este tipo de almacenamiento.

- A mediano plazo, participar en el desarrollo de proyectos de Almacenamiento Distribuido nacionales e internacionales.
- Fortalecer y difundir los sistemas de Almacenamiento Distribuido de Software Libre y sus aplicaciones.
- Proveer y/o capacitar sobre los servicios de Almacenamiento Distribuido, en la región donde se encuentra la UNCuyo.

5.2 Etapas, alcances y condiciones de la Implementación

- 1ra etapa: Desarrollar una maqueta del sistema de almacenamiento distribuido. Esto es necesario para analizar las necesidades del hardware requerido, las configuraciones utilizadas, el rendimiento según los recursos utilizados, y la demanda efectuada del sistema.
- 2da etapa: De acuerdo a análisis de los datos recabados en la 1er etapa, optimizar la implementación y aplicarla a las demandas existentes que se adapten a este tipo de almacenamiento, por ejemplo: configurar uno o varios sistemas para que comiencen a utilizar esta implementación y observar su comportamiento.
- 3ra etapa: De acuerdo a lo efectuado en las etapas anteriores y al análisis de los resultados obtenidos, optimizar la configuración y el hardware necesario, para que el sistema de almacenamiento distribuido esté montado sobre una infraestructura de servidores adecuada a tal fin. Para esto habrá que adaptar o actualizar los servidores y los sistemas operativos, configurar los servicios, ponerlos en funcionamiento, finalmente, monitorearlos y controlarlos.
- 4ta etapa: Una vez que el sistema está implementado en el centro de datos principal, y que los servidores y sistemas de respaldo de información lo tengan incorporado, desplegar el servicio, para que los usuarios que están por fuera del esquema de servidores y que demanden su utilización, tengan acceso al mismo.
- 5ta etapa: Efectuar el despliegue a nivel de Facultades e Institutos que demanden de este tipo de almacenamiento. Esta se llevará a cabo luego de verificarse la demanda de cada uno de ellos, confirmarse la disponibilidad de hardware en los respectivos centros de datos, y cumplirse la capacitación de los encargados de los servidores y las redes de las Unidades Académicas, Facultades o Institutos.
- 6ta etapa: Por último, ofrecer el servicio a los sectores que lo demanden y no hayan sido alcanzados en las etapas anteriores.

5.3 Consideraciones adicionales:

El Sistema Ceph, de Almacenamiento Distribuido y bajo licencia GPL implementado en la UNCuyo, es un sistema que escala sin límites en recursos físicos y sin puntos de fallo. Está diseñado para utilizarse con gran cantidad de datos, y enfocado en el uso de Big Data y Cloud Computing. Si hacemos foco en el rendimiento del sistema, tendremos que implementarlo sobre hardware de alto desempeño y de gran capacidad, por lo que si efectuamos una comparación de costos, con los sistemas de almacenamiento propietarios, veremos que Ceph, no resulta más económico que una solución propietaria.

Si partimos de la base en la cual disponemos del hardware y queremos reutilizarlo, e ir agrandando nuestra infraestructura de almacenamiento según lo vayamos demandando, Ceph es una muy buena opción, ya que podemos ir adaptando nuestra infraestructura y haciéndola crecer, de acuerdo al uso intensivo que hagamos de este tipo de almacenamiento.

Finalmente, la ventaja que tenemos con Ceph respecto de los sistemas de almacenamiento propietario, es que posee una gran flexibilidad para crecer y redimensionarse sin depender de hardware homogéneo ni propietario, liberándonos así de la dependencia de un único fabricante.

6 El despliegue de Ceph: en curso

6.1 Implementación para el Centro de Documentación Audiovisual - CDA

La implementación de Ceph, en la Universidad se vio impulsada a partir de necesidades específicas, y el planteo de tres puntos fundamentales, que se describen a continuación:

- i) El CDA, dispone de videos históricos institucionales, los que en su gran mayoría están almacenados en cintas VHS[19], y tienen gran probabilidad de sufrir pérdida de información y/o daño físico. Estos videos contienen documentales, programas de televisión, y otros videos institucionales de la vida universitaria. Una buena parte de esta información, estaba almacenada en discos duros (HDD[20]) externos, en formato de archivos multimedia.
- ii) Como se mencionó anteriormente, la implementación de un sistema de almacenamiento de mayor capacidad, se ve dificultado por los altos costos en el equipamiento (el hardware), y la capacitación de las personas a cargo del mismo (el recurso humano). A partir de un análisis detallado de las necesidades actuales de almacenamiento de datos, resulta que no es primordial contar con altas velocidades de transferencia de archivos para ese sistema, sino que prima contar con grandes capacidades de almacenamiento.

- iii) Hay disponibilidad de una buena cantidad de Hardware que actualmente no puede ser destinado a estaciones de trabajo de un usuario corriente, i.e. PC obsoletas o en estado de obsolescencia para ciertos usos, las que debido a sus características tales como: capacidad de memoria RAM, de almacenamiento en disco HDD, de velocidad de procesamiento y arquitectura de la CPU, no pueden ser destinadas para el uso habitual de una estación de trabajo. Tampoco este equipamiento se puede actualizar, debido a que la mayor parte del hardware que las compone, está discontinuado, y no pueden ponerse al día con la constante demanda respecto de la actualización tecnológica.

Por lo que, ante la necesidad planteada, y el análisis resultante, se procede a desplegar una solución de acuerdo a los elementos tecnológicos disponibles y realizando una mínima inversión.

Se comienza con la implementación de Ceph, efectuando la instalación en la infraestructura disponible que se describe a continuación: 7 CPU de escritorio, marca HP (modelo Compaq dc7800 Small Form Factor), con 1 GB de memoria RAM, discos duros (HDD) de 160 GigaBytes o 250 GigaBytes, procesadores Dual Core de 1.80GHz.

Se invierte inicialmente en la adquisición de 6 discos duros SATA[21] de 2 TeraBytes cada uno. Se decide adquirir este tipo de dispositivo de almacenamiento, con la finalidad de que sean elementos de hardware común, disponible en el mercado local y a una relación precio/beneficio adecuada, manteniendo el concepto de “bajo costo”. La capacidad de almacenamiento del disco seleccionado es el HDD de mayor capacidad soportado y recomendado por el fabricante de la PC. Además, por una cuestión de disponibilidad y costos en el mercado local, es la capacidad de dispositivo más adecuada a la relación costo/beneficio/disponibilidad en la región.

También se compra un Switch Gigabit Ethernet con 24 puertos de 10/100/1000Mbps. La marca del dispositivo es TPLink y el modelo es de prestaciones media/baja.

La implementación inicial se muestra en la figura 2, está compuesta por 6(seis) OSD (Object Storage Device), y 1(un) monitor con funciones de administrar la MDS (Metadata Software), todos dentro de la infraestructura del sistema de archivos Ceph. Tal como lo indica el sistema de archivos Ceph, se puede definir la cantidad de copias que van a existir, para cada objeto almacenado, estos son los PGS[5] (Placement Groups), se decidió efectuar 4 copias, lo que implica una pérdida significativa del almacenamiento total existente, ya que cada objeto es replicado 4 veces. Esto nos permite contar con una redundancia muy adecuada al soporte a fallas de los OSD.

Con esta configuración quedó un espacio de almacenamiento inicial total de 9,6TBytes. Si bien los 6 OSD con discos de 2TBytes, da un total de espacio de almacenamiento de 12TBytes, la configuración aplicada, deja en promedio una pérdida de espacio de alrededor del 20%.

Se define como única red de interconexión, tanto para la administración como para la transferencia de información, una de clase C (/24). Se configura la seguridad de acceso de cada cliente mediante archivos KEYRING[5], ubicados: en el MONITOR y en c/u de los clientes. La configuración mediante estos archivos, indica que tipo de permisos tiene cada cliente y a que objetos puede acceder.

Se efectúa una implementación sencilla con la finalidad que el sistema pueda ser administrado y mantenido por personal idóneo en el tema, pero no por expertos. Se utiliza la herramienta `ceph-deploy`[5], que permite efectuar distintas tareas de instalación de manera estándar, sin la necesidad de tener que configurar manualmente toda la parametrización de la instalación. Toda la configuración se realiza desde el servidor dedicado para MONITOR, (y en concordancia con el funcionamiento de Ceph). Este mismo componente es el que brinda el servicio de publicar el medio de almacenamiento hacia aquellos servidores (que son clientes del sistema de archivos) y que lo utilizarán para almacenar su información.

Se decide implementar Ceph, a partir de la necesidad almacenar archivos de multimedia, (videos, fotos, audios). En esta, se define un único pool de espacio de almacenamiento (área de almacenamiento), ya que permite crecer sin la necesidad de indicar a que pool se va a asignar cada nuevo OSD que se agrega. Con esto, cada OSD se agrega directamente a un único espacio de almacenamiento. Luego cada servidor cliente de ese sistema de archivos distribuido, utiliza un módulo del núcleo del sistema operativo, en el espacio de usuario, llamado FUSE[22], esto se muestra en la figura 4. El software es `fuse-block`[5], otorga los permisos definidos en el monitor, y monta el sistema de archivos Ceph.

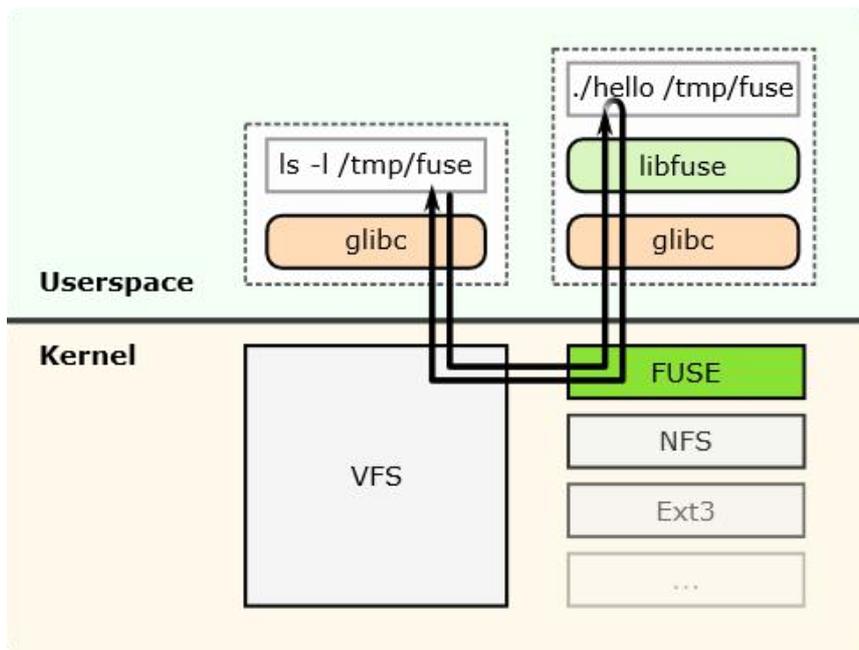


Fig. 4. Arquitectura del Sistema (*FUSE*).

Una vez concluida la instalación del sistema de archivos distribuido, se procede a instalar un servidor con el Sistema de catálogo de archivos multimedia. Para ello se utiliza el software `MEDIADROP`[23], que es la herramienta que utiliza el usuario

final para administrar los contenidos multimediales (Archivos de Audio/Video). Este servidor MediaDrop monta parte de su sistema de archivos, enlazando ese espacio de almacenamiento local con el sistema de archivos definido en Ceph. Se define un espacio de almacenamiento inicial de aproximadamente 9,5TBytes. Una vez configurado, se da acceso al personal del CDA para que realice la copia de los archivos multimedia desde sus PC, con sistema operativo Windows, accediendo a través de la red a ese MediaDrop, compartiendo su espacio de almacenamiento por intermedio de un SAMBA[24], el cual está instalado en el mismo servidor.

En el mismo servidor de MediaDrop, se ofrece un servicio de ftp, para que el acceso a la lectura de los archivos multimedia, por parte de los clientes, se efectúe mediante un protocolo adecuado que soporte una tasa de transferencia sostenida superior al de otros protocolos tales como http. Esta configuración mejora la reproducción de videos del catálogo MediaDrop, al reproducir los contenidos de manera estable y fiable, permitiendo mayor cantidad de conexiones hacia el servidor, sin degradar la calidad del servicio. Este esquema se muestra en la figura 5.

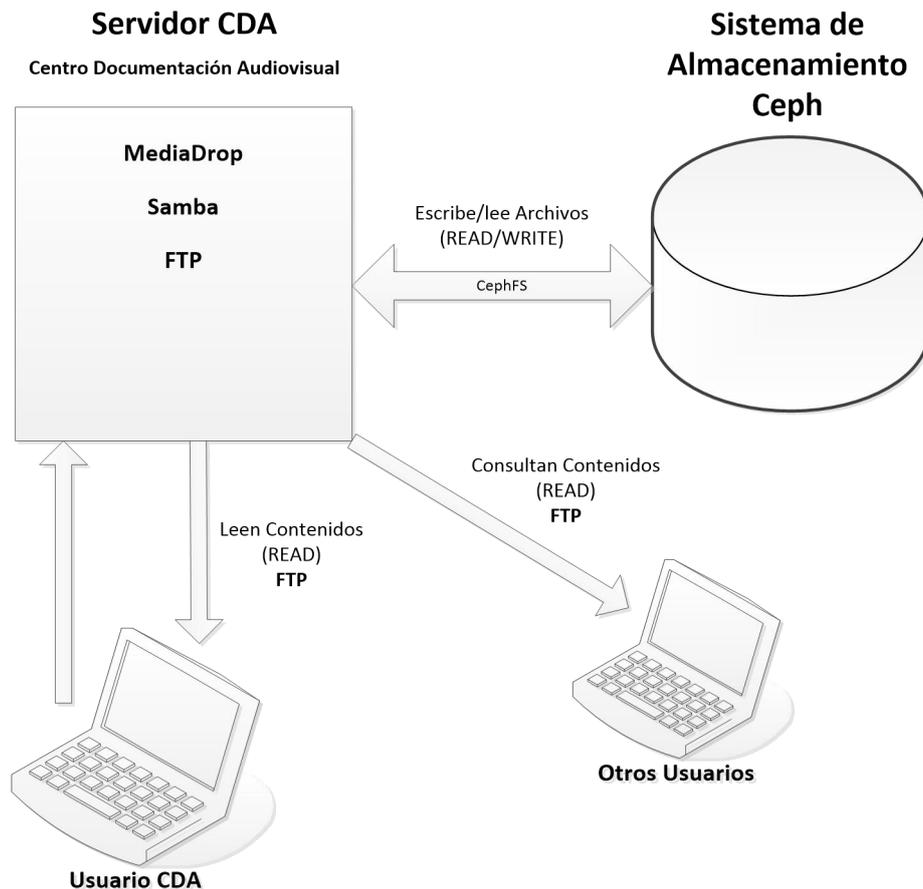


Fig. 5. Implementación en el CDA (Centro Documentación Audiovisual).

La configuración de los monitores quedó de la siguiente manera:

- Monitor 1: Monitor principal, y MDS (Metadata Software).
- Monitor 2: Monitor, y acceso desde el servidor cliente mediante fuse-block.
- Monitor 3: Monitor, y Ceph-Dash[5] (herramienta de monitoreo de estado del cluster)

Luego de efectuar esta modificación, que mejoró el acceso por parte de los clientes y conllevó a mayor utilización del espacio de almacenamiento del sistema de archivos, se detecta que el clúster se está quedando sin espacio de almacenamiento disponible, por lo que se agregan 2(dos) OSD mas, a los 6(seis) existentes hasta ese momento, llevando el tamaño de almacenamiento a 14,5TBytes. Una recomendación muy importante indicada en la documentación de Ceph, es que jamás se debe sobrepasar el 85% de la utilización del espacio total de almacenamiento, ya que el proceso de agregar nuevos nodos OSD se dificultaría, y como consecuencia de esto se puede incurrir en la pérdida de información, motivo por el cual, se debe estar monitoreando continuamente este parámetro y deben tomarse las medidas adecuadas para que esto no suceda.

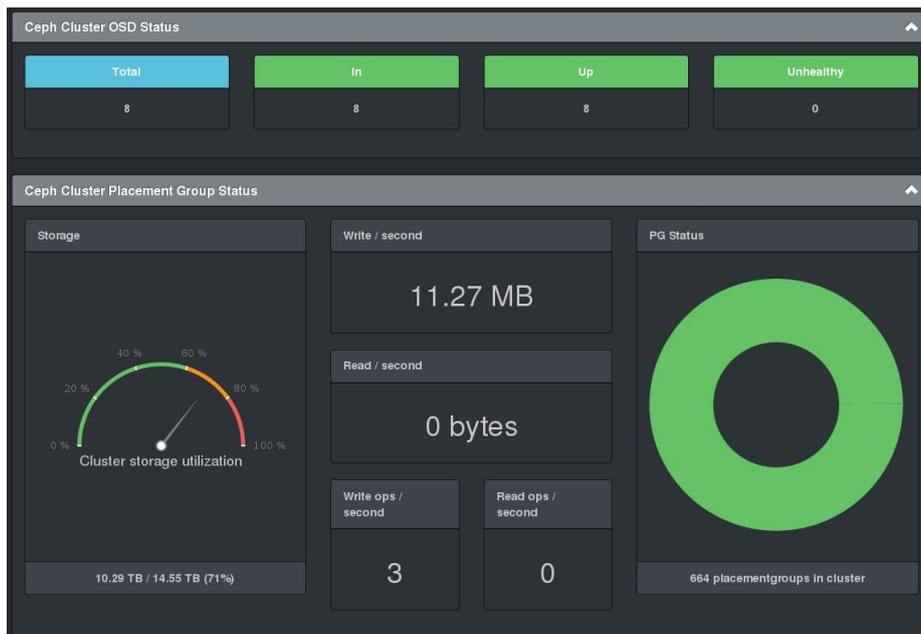


Fig. 8. Estado del Sistema de Almacenamiento - Aplicación (*ceph-dash*).

Luego de efectuar estos cambios se llegó a una situación estable con funcionamiento óptimo. Asimismo se tuvo que realizar una optimización de la configuración de los pesos de cada nodo OSD dentro del cluster(sistema de almacenamiento distribuido), ya que había que balancear la carga de almacenamiento de los nodos existentes con los nuevos nodos incorporados.

Teniendo en cuenta que el cluster le asigna un peso de 1 a cada nodo OSD que se inicia, siendo ese valor el de menor peso, se modifica la configuración para que se redistribuyan las cargas de forma equitativa entre todos los nodos OSD. Este procedimiento se tiene que efectuar manualmente cada vez que se incorpora un nuevo nodo OSD al cluster, el valor aplicado es REWEIGHT, la distribución quedaría entonces como se muestra a continuación en la figura 9:

```
[clus@monitor2 ~]$ ceph osd df
ID WEIGHT REWEIGHT SIZE USE AVAIL %USE VAR PGS
0 1.81850 0.89999 1862G 1295G 566G 69.57 0.98 157
1 1.81850 1.00000 1862G 1253G 608G 67.30 0.95 174
2 1.81850 1.00000 1862G 1274G 587G 68.46 0.97 175
3 1.81850 0.89999 1862G 1361G 500G 73.11 1.03 180
4 1.81850 0.79999 1862G 1335G 526G 71.73 1.01 143
5 1.81850 0.89999 1862G 1318G 543G 70.79 1.00 153
7 1.81850 0.95000 1862G 1365G 496G 73.33 1.04 175
8 1.81850 0.95000 1862G 1329G 532G 71.38 1.01 171
TOTAL 14896G 10533G 4363G 70.71 MIN/MAX VAR: 0.95/1.04 STDDEV: 2.04
```

Fig. 9. Estado de la re-distribución del espacio en el Sistema de Almacenamiento (*Ceph*).

7. Conclusiones

La principal motivación para el despliegue del Sistema de Almacenamiento Distribuido (*Ceph*) es la expansión sin límites del espacio de almacenamiento disponible para todos los usuarios de la UNCuyo, haciendo hincapié en su utilización con gran cantidad de datos, y enfocado en el uso de Big Data y Cloud, ya sea para usuarios finales, sistemas en funcionamiento en el Centro de Datos principal o en los Centros de Datos secundarios de la Universidad Nacional de Cuyo.

Es difícil establecer durante cuánto tiempo habrán limitaciones económicas-financieras por lo que esta solución aprovecha al máximo los recursos existentes actuales y minimiza los costos del hardware y la dependencia que se tienen sobre las infraestructuras de almacenamiento propietarias.

Un punto muy crítico y aplicable a todos los casos, es que requiere capacitación y entrenamiento, por lo tanto uno de los principales costos que tenemos junto con la actualización del hardware, es la capacitación del recurso humano que estará a cargo de administrar esta infraestructura.

Referencias

1. Reseña Histórica de la UNCuyo: <http://www.uncuyo.edu.ar/paginas/index/resena-historica>.
2. R. Cutuli, C. Catania y C. García Garino: Problemas y herramientas en la seguridad de redes de transmisión de datos universitarias. El caso de la Universidad Nacional de Cuyo. Primera Conferencia de Directores de Tecnología de Información. TICAL 2011, Panamá, Junio 2011
3. Servidores BladeCenter IBM:
ftp://ftp.software.ibm.com/la/documents/stg/bladecenter/hs22_es.pdf
ftp://ftp.software.ibm.com/la/documents/gb/mx/BC_H.pdf
4. Fibre Channel Storage Area Networks (SAN): <http://www-03.ibm.com/systems/storage/san/>
5. Proyecto Ceph: <http://ceph.com/>, <http://docs.ceph.com/docs/master/>
6. IDE: https://es.wikipedia.org/wiki/Integrated_Drive_Electronics
http://www.ele.uri.edu/courses/ele408/s2001/projects/roland_ide/IDE.pdf
7. SCSI: https://es.wikipedia.org/wiki/Small_Computer_System_Interface
<http://www.t10.org/scsi-3.htm>
8. RAID: <https://es.wikipedia.org/wiki/RAID>
9. Gluster File System: https://es.wikipedia.org/wiki/Gluster_File_System
10. SID UNCuyo: <http://sid.uncu.edu.ar/sid/cda/>
11. SeñalU: <http://www.senalu.tv/>
12. Universidad Politécnica de Madrid: <http://www.upm.es/>
13. OpenStack: <https://www.openstack.org/>
14. Big Data: https://es.wikipedia.org/wiki/Big_data
15. Illinois University: <http://illinois.edu/>
16. Cloud Computing: https://es.wikipedia.org/wiki/Computaci%C3%B3n_en_la_nube
17. Universidad Monash: <http://www.monash.edu/>
18. Implementaciones empresariales de Ceph: <http://ceph.com/users/>
19. VHS: <https://es.wikipedia.org/wiki/VHS>
20. HDD: https://es.wikipedia.org/wiki/Unidad_de_disco_duro
21. SATA: https://es.wikipedia.org/wiki/Serial_ATA
22. FUSE: https://es.wikipedia.org/wiki/Sistema_de_archivos_en_el_espacio_de_usuario
23. Mediadrop: <http://mediadrop.video/>
24. Samba: <https://www.samba.org/>