

DESCRIPCIÓN PROYECTO CHEAP TB

Información del documento

Proyecto presentado:	CheapTB
Descripción breve:	Diseño y fabricación de un sistema jerárquico de almacenamiento de alta capacidad, de bajo coste de adquisición y funcionamiento
Autor:	S. A. XESTIÓN CESGA
Versión:	Final

Santiago de Compostela, febrero - 2005

Descripción del proyecto¹ y objetivos del mismo

La necesidad de un sistema de almacenamiento masivo de datos de bajo coste para pequeñas y medianas empresas, así como para los laboratorios de investigación, está muy presente en el mercado tecnológico actual. Estas necesidades de almacenamiento son cada vez mayores, del orden de TeraBytes (TB) y las soluciones tecnológicas existentes en la actualidad para resolver este problema tienen un coste elevado. Por ello, se ha diseñado e implementado un sistema de **almacenamiento distribuido basado en tecnologías de bajo coste**. El sistema consta de tres niveles: Un primer nivel de almacenamiento, un segundo nivel de metadatos y un tercer nivel con la interfaz con el usuario.

En los últimos años han aparecido diferentes soluciones para resolver las necesidades de almacenamiento. La mayoría de estas soluciones son cerradas y se venden conjuntamente con el hardware. Sin embargo, se puede aprovechar hardware de bajo coste para crear un almacenamiento de gran capacidad y, además, basado en **código abierto**. La solución adoptada es la de agregar el almacenamiento de varias máquinas conectadas. El almacenamiento se presenta a los clientes como un único espacio. Para facilitar la compatibilidad del sistema con diferentes clientes, la interfaz con el exterior soporta múltiples protocolos de acceso. El hecho de usar hardware de bajo coste podría implicar una pérdida de fiabilidad respecto a otras soluciones. Para evitarlo se permite la posibilidad de duplicación de los posibles puntos de fallo. De esta forma el fallo de un elemento no implica la caída del sistema.

Objetivos

Los objetivos que se han alcanzado en este proyecto se resumen a continuación:

- **Almacenamiento Masivo:** CheapTB soporta el almacenamiento de grandes cantidades de información. Estamos hablando de TB, en un principio, escalable a decenas de TB.
- **Ahorro de Energía:** El sistema consume el mínimo de energía posible. En caso de que los discos de un nodo de almacenamiento no se estén usando durante un cierto tiempo, el sistema apaga dicho nodo, con el consiguiente ahorro energético.
- **Interfaz Conocida:** Para permitir el uso de CheapTB con el mayor número de sistemas posibles sin necesidad de instalar ningún software adicional, el sistema soporta varios sistemas de ficheros en red que se usan habitualmente (NFS, SMB...). Además, al ser fácilmente escalable, la inclusión de nuevos adaptadores es sencilla.
- **Redundancia:** Los datos y metadatos se almacenan en más de una parte del sistema para mejorar la tolerancia a fallos.
- **Fallo parcial:** En caso de que algún disco o nodo falle, el fallo sólo afecta a los contenidos almacenados en ese nodo, garantizando el funcionamiento del resto del sistema.

¹Este proyecto ha sido parcialmente financiado por la Xunta de Galicia, a través del Programa de Tecnoloxías da Información e das Comunicaci3ns do Plan Galego de Investigación, Desenvolvemento e Innovaci3n Tecnol3xica.

Grado de realización en la actualidad

La solución, que se encuentra en explotación, está enfocada al sector empresarial y de investigación y puede crecer en capacidad fácilmente. Puede ser utilizada en actividades profesionales con altas demandas de almacenamiento o en proyectos de renderización de imágenes.

Actualmente se está utilizando para dar servicio a investigadores de tres universidades: Santiago, Coruña y Vigo. Asimismo, está siendo utilizado por el propio CESGA y la UDC como sistema de almacenamiento, además de en el Instituto de Física de Cantabria y en el CSIC (Centro Superior de Investigaciones Científicas).

Existen contactos con diferentes empresas del sector para dar servicios de externalización de copias de seguridad, almacenamiento de películas y datos.

Por otro lado, se trabaja en la mejora del producto, planteando posibles vías, como la implementación de un sistema de ficheros para entornos ofimáticos de alta disponibilidad o la mejora de la fiabilidad del sistema CheapTB mediante mecanismos de tolerancia a fallos.

Resultados conseguidos y valor añadido aportado por el proyecto al cliente

En el mercado existen muchas soluciones para el almacenamiento de grandes cantidades de datos. Normalmente este tipo de soluciones están diseñadas como sistemas de propósito general (para almacenar cualquier tipo de información) lo que hace que tengan un coste muy elevado. Los dispositivos utilizados para almacenar la información son discos magnéticos, cintas magnéticas o una combinación de ambos. Los sistemas basados en cintas tienen el inconveniente del retraso en la localización de la información y en los sistemas de discos el coste de mantenimiento es elevado debido al consumo eléctrico y la disipación térmica. Sin embargo, muchas veces las necesidades existentes en las empresas y en los centros de investigación, se pueden abordar de un modo más simple, lo que, a un tiempo, permite abaratar los costes y aumentar las prestaciones de los sistemas.

La innovación de este proyecto se encuentra, por un lado, en la parte económica, no sólo en lo que se refiere a la compra del hardware (discos duros IDE en lugar de cintas), sino también en la relación entre la cantidad de datos que se puede almacenar y el coste global de las tecnologías empleadas. Por otro lado, en la utilización de software de desarrollo propio, ofreciendo una solución muy competitiva frente a otras soluciones comerciales y con una utilidad inmediata para los nuevos sistemas de e-business, así como para las PYMEs que necesiten almacenamiento de ficheros grandes de lectura secuencial. También puede ser utilizado por las distribuidoras de vídeo bajo demanda (radios, televisiones, empresas de teleformación), productoras cinematográficas o empresas que dispongan de oficinas dispersas geográficamente, consolidando su información y posibilitando que esté disponible en todo momento en cualquier parte.

Desarrollo del proyecto, metodología y herramientas aplicadas.

El proyecto, que ha sido desarrollado por el CESGA y el Laboratorio de Fundamentos de Computación e Inteligencia Artificial (LFCIA) de la Universidad de A Coruña, se dividió en las siguientes tareas:

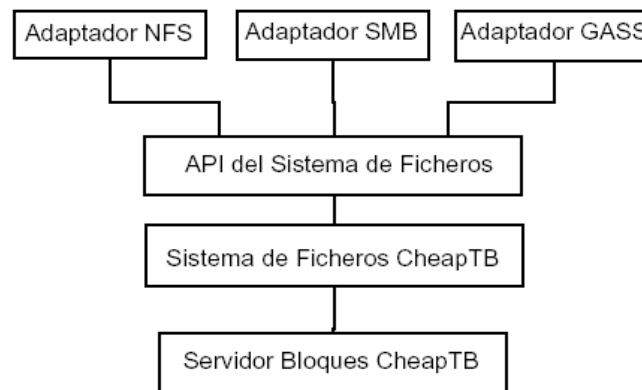
1. Definición de requerimientos del sistema.
2. Diseño del sistema.
3. Diseño y construcción de las aplicaciones de prueba del sistema.
4. Implantación.
5. Pruebas y validación.
6. Estudio económico.
7. Conclusiones.

Se diseñó y desarrolló una plataforma hardware y software, basada en Linux, para la implementación de un sistema de almacenamiento de datos de bajo coste de adquisición (inferior a 6000€ por Terabyte) y mantenimiento reducido.

Así mismo, se definieron los componentes software necesarios para proporcionar al usuario una imagen única del conjunto como servidor de datos NAS (*Network Area Storage*), permitiendo el acceso desde cualquier ordenador conectado a Internet.

El sistema se encuentra estructurado en tres capas:

- **Capa I: Interfaz con el cliente.** El sistema especifica una API común a la que se pueden conectar distintos adaptadores para diferentes protocolos, como NFS, SMB o cualquier otro que se considere necesario.
- **Capa II: Sistema de Ficheros.** El sistema de ficheros gestiona la metainformación sobre los archivos y directorios almacenados. Esta parte del sistema se encarga de asegurar la consistencia de la información almacenada ante las peticiones concurrentes de los clientes. El sistema de ficheros decide la ubicación de los datos, con el objeto de aumentar la localidad de los mismos y conseguir así una mejora en el consumo de energía. Los ficheros tienden a crecer en el mismo disco y nodo donde ya poseen bloques, para que el hecho de utilizar un fichero implique al menor número de sistemas de almacenamiento posibles.
- **Capa III: Dispositivo de Bloques.** El dispositivo de bloques presenta todo el espacio de almacenamiento disponible como un único dispositivo al sistema de ficheros. El sistema de ficheros no necesita conocer la localización de los datos, ni cómo transferirlos y almacenarlos.



Cada capa puede residir en más de un nodo. La capa II garantiza la coherencia de la información almacenada. La metainformación está almacenada en una base de datos distribuida para facilitar el acceso concurrente desde varios nodos. Debido al coste de la comunicación entre los nodos donde reside la base de datos, el sistema de ficheros cuenta con delegaciones. Un nodo puede utilizar copias locales de cierta metainformación para un acceso más rápido mientras no haya más nodos intentando acceder al sistema. Cuando algún otro nodo necesita acceder a esta información, se lo notifica al nodo que posee la delegación para que almacene los cambios en la base de datos y en los nodos de almacenamiento.

El sistema de almacenamiento puede ser accedido a través de adaptadores que implementan protocolos estándar como SMB o NFS. Los adaptadores realizan las peticiones necesarias al sistema a través de una interfaz similar a POSIX, común para todos los adaptadores. El sistema de ficheros realiza las modificaciones necesarias en los metadatos, y obtiene o cambia los datos utilizando los servicios de un sistema de almacenamiento, que presenta una estructura jerárquica de nodos y discos como un único dispositivo de bloques.