

# Almacenamiento y Gestión de la Información en el CESCA



# Almacenamiento y Gestión de la Información en el CESCA

El **Centre de Supercomputació de Catalunya (CESCA)** es un organismo cuyo objetivo fundamental es gestionar un gran complejo de sistemas de cálculo y de comunicaciones para dar servicio a la universidad y a la investigación, basándose en cinco áreas de actividad: las redes de comunicaciones, los portales y repositorios para información universitaria, los sistemas para cálculo científico y almacenamiento de datos, la promoción del uso y los beneficios de estas tecnologías y la operación y el mantenimiento de toda la infraestructura del centro.

El Área de Portales y Repositorios proporciona un servicio de acceso a distintos repositorios de e-información: tesis doctorales, documentos de investigación, revistas científicas y tecnológicas, patrimonio digital, memoria digital de Cataluña y un largo etcétera.

El Área de Cálculo y Archivo provee a la comunidad científica e industrial de Cálculo de Altas Prestaciones (CAP) que permite avanzar en disciplinas tan diversas como el diseño de fármacos mediante el modelaje de moléculas, la búsqueda de nuevos materiales, reducir el impacto del ser humano en el medio ambiente y profundizar en el estudio de nuestras galaxias, generando resultados que, posteriormente, son archivados.

Estas dos áreas de actividad, requieren ambas, aunque en forma distinta, capacidad de almacenamiento. La primera, necesita almacenamiento de gran capacidad por el gran volumen de información que se maneja con una capa de presentación de alto nivel, los sistemas de ficheros, que harán accesible la información al usuario. La segunda, busca disponer de almacenamiento de gran rendimiento por la exigencia del cálculo de altas prestaciones, con una capa de presentación de bajo nivel, los sistemas de bloques, que permiten optimizar la comunicación y transferencia de datos con los supercomputadores del centro.

A nivel tecnológico, la solución consiste en la implantación de un nuevo sistema de almacenamiento, que constituye el soporte para estas dos grandes áreas. No sólo aumentando en prestaciones y capacidad de almacenamiento los sistemas anteriores sino integrando en una única cabina de discos, la base para el área de cálculo y el área de portales y repositorios.

La tecnología propuesta para este proyecto está basada en el fabricante NetApp. Esta elección cumple con los requisitos del proyecto demandados por el CESCA:

- Renovación tecnológica y ampliación de la capacidad de almacenamiento.
- Unificación del almacenamiento para entornos Storage Area Network (SAN) y Network Attached Storage (NAS).
- Dar respuesta a dos entornos muy diferenciados, el entorno de cálculo de altas prestaciones y el de depósitos y portales, con diferentes requisitos de ancho de banda, capacidad y tiempo de residencia de los datos.
- Los datos del área de supercomputación se caracterizan principalmente por necesitar un entorno de alto rendimiento mientras que el área de depósitos y portales, requieren un entorno de gran capacidad.
- Homogeneidad en la gestión de las diferentes "islas" de información.

Atendiendo a estos objetivos, la propuesta de SATEC está basada en un sistema de almacenamiento unificado que permite aglutinar en un mismo sistema entornos heterogéneos en cuanto a la naturaleza de los datos, el tipo de acceso y las prestaciones requeridas por cada uno.

Por tales motivos, tanto la tecnología como el servicio se ajustan perfectamente al cumplimiento de estos objetivos facilitando una infraestructura unificada cuya gestión se simplifica al máximo exponente.

# Objetivo

Fruto de las necesidades de crecimiento, el objetivo inicial del CESCA es la renovación y modernización de sus sistemas de almacenamiento. La propuesta de SATEC, más allá de cumplir con estos requisitos, realiza una propuesta que mejora la solución en otros aspectos como son la consolidación de sistemas, la seguridad y protección de datos, la escalabilidad de la solución y su potencial de cara a futuras integraciones con arquitecturas más complejas.

Otro objetivo alcanzado con el despliegue de este proyecto consiste en la flexibilidad de la arquitectura propuesta para su adecuación a nuevos proyectos punteros, tecnológicamente hablando. Un ejemplo de ello, es la necesidad reciente del CESCA de la implantación de nuevos servicios en el mundo de la virtualización, una infraestructura que también se desplegará en el nuevo sistema de almacenamiento con todas las ventajas que ello supone: optimización del almacenamiento de datos para entornos virtuales (tecnologías de deduplicación, "thin provisioning" y backup instantáneos sin pérdida de rendimiento).

Esta propuesta en cuanto a equipamiento y arquitectura de la solución, se acompañó de un diseño de plan de migración que garantizase en todo momento la compatibilidad con los sistemas anteriores y una interrupción de los servicios mínima para evitar, en la medida de lo posible, la afectación a los usuarios finales.

Desde el punto de vista de la explotación de los recursos desplegados, se ha permitido al CESCA el cumplimiento de los siguientes objetivos:

- La simplificación de la gestión de los datos almacenados, la centralización y automatización de los flujos de trabajo en el día a día del almacenamiento.

- Identificación de posibles problemas y sus acciones correctoras asociadas para evitar periodos de inactividad de los servicios.
- Reducción del coste total de propiedad con el aprovechamiento de recursos y aumento en la eficacia de su gestión.
- Una plataforma que se adapte al cumplimiento de las normativas de estandarización aplicadas a organismos de esta índole.

Desde el punto de vista de SATEC, como compañía integradora de soluciones tecnológicas avanzadas, el desarrollo de este proyecto ha supuesto:

- La adecuación tecnológica para entornos muy exigentes en cuanto al acceso y procesamiento de la información.
- Un diseño y gestión de estandarización de una metodología para la migración de información entre diversas infraestructuras cumpliendo con los SLA's necesarios para evitar al máximo el impacto en el usuario final.
- Comprender la idiosincrasia de entidades en las cuales, el acceso a la información en todas sus variantes, es crucial para el desarrollo de su actividad.
- La extrapolación de este tipo de soluciones a grandes compañías y corporaciones.
- Demostrar que nuestra compañía posee una dilatada experiencia en la implantación de soluciones y servicios innovadores que se adaptan a cualquier necesidad tecnológica tanto a nivel nacional como internacional.

# Descripción de la solución

En esta solución están presentes servicios de ingeniería de SATEC centrados en los procesos de migración, integración y control de la solución ofrecida y servicios de mantenimiento, suministro e instalación de la infraestructura tecnológica necesaria.

A nivel de tecnología, la solución se basa en una cabina NetApp, concretamente de la serie FAS3170, con dos controladoras equipadas con 32 GB de RAM cada una (64 en total) una capacidad de más de 10 TB en disco de alto rendimiento y más de 50 TB en disco de alta capacidad y con todos los elementos (controladoras, fuentes, ventiladores) redundados, acompañado de dos conmutadores de Fiber Channel Cisco, cada uno con 64 puertos de Fibra activos.

## Consolidación

La cabina de almacenamiento propuesta dispone de una operativa dual, esto quiere decir que es capaz de operar en un entorno NAS sirviendo directamente ficheros y, de forma simultánea, operar en un entorno SAN presentando LUNs por sistema de bloques. Asimismo, también es capaz de operar con discos de alto rendimiento (FC) y con discos de alta capacidad (SATA).

Todos los servicios de ficheros que se presentaban a través de NFS se migraron de los servidores directamente a la cabina con lo que, aparte de la consolidación de almacenamiento, se ha conseguido la eliminación de los antiguos servidores de ficheros, unificando y simplificando el manejo y la operativa diaria.

## Escalabilidad

El sistema actual dispone de una capacidad de crecimiento de hasta 252 discos SATA y 168 discos FC, que proporcionarían 300 TB. La capacidad actual se podría triplicar

problemas. Además, las cabinas de gama superior del fabricante disponen exactamente del mismo sistema operativo, de modo que, ante la posible necesidad de una renovación, toda la infraestructura de datos se podría mantener en los discos actuales pasando éstos a ser operados por controladoras de gama superior.

Estas dos características garantizan una vigencia absoluta para los próximos años.

## Protección de datos

El sistema exclusivo de RAID-DP (Doble Paridad) proporciona una protección de los datos en la que cada conjunto lógico de discos puede tener fallo en dos de sus discos de forma simultánea y seguir manteniendo la integridad de los datos mientras estos se reconstruyen en los discos. Esta configuración, permite una protección infinitamente más elevada que la que proporcionan los RAIDs más tradicionales, como es el RAID 5, que sólo tolera el fallo de un único disco y mucho más eficiente (con una protección similar) que otros mecanismos conocidos como el RAID 1 que mantiene una copia en espejo en un segundo conjunto de discos pero que penaliza mucho en la ocupación ya que cada disco tiene que estar duplicado.

## Proceso de migración

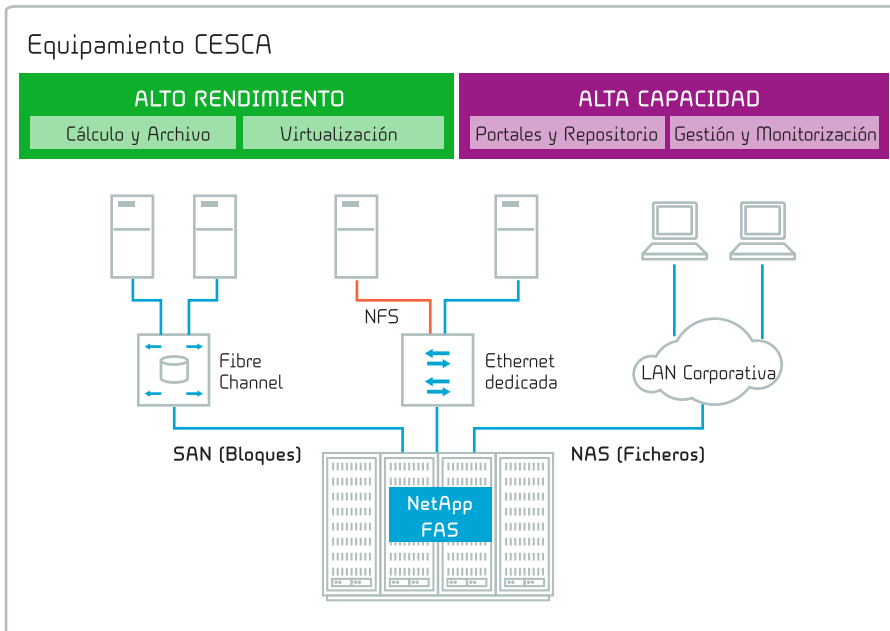
Uno de los puntos claves de la solución ha sido el proceso de migración de datos, así como su integración con los distintos elementos de red y su paso a producción.

Debido al gran volumen de datos que manejan las aplicaciones del CESCO y, más en concreto, de algunas de ellas que utilizan varios TB de información cada una, el proceso de migración no se ha podido plantear como un corte de servicio seguido

equipamiento. Los datos se han migrado durante los períodos de producción conjuntamente con la prestación del servicio de forma que, sólo al final de cada migración, se ha tenido que realizar un corte de servicio mínimo para poner todos los datos ya migrados en producción desde el nuevo sistema de almacenamiento. Para tener actualizados los datos que se modificaban durante el grueso de la migración, se han utilizado técnicas que permiten la sincronización de cambios mediante incrementales.

En el caso de la infraestructura SAN (Fiber Channel), al ser los mismos servidores los que siguen prestando el servicio desde la nueva cabina, el reto está en migrar una estructura de bloques entre dos cabinas de fabricantes distintos. En este caso, el software de replicación directa entre cabinas no es compatible y, por tanto, hay que recurrir a software de terceros. La solución fue simplificar este proceso a la vez que hacerlo más robusto y seguro: para cada uno de los servidores se montó de forma simultánea la actual LUN de la cabina antigua y la nueva LUN de la cabina actual y utilizando el propio software de mirroring del que disponen los servidores de tipo Linux y UNIX, hacer una sincronización en despejo de las dos LUNes. Esta sincronización ha sido el propio proceso de replicación que se ha llevado a cabo con los servidores en producción y sin detectar pérdidas de servicio durante el mismo. Una vez concluida la migración de datos, la parada momentánea del servicio y la desconfiguración del mirror (parte más crítica del proceso por su complejidad) han culminado con la puesta en marcha del servicio nuevamente ya desde la nueva cabina.

# Beneficios



Podríamos decir que la solución implantada aporta un conjunto de beneficios que se destacan a continuación:

- Aumento de capacidad de procesado y memoria de la actual cabina así como de espacio en disco.
- Eficiencia en el almacenamiento de la información, ofreciendo deduplicación de la información.
- Ahorro energético. Menor número de discos necesarios para almacenar la misma información, por tanto, menor consumo energético.
- Consolidación de la actual solución SAN en un único filer de almacenamiento.
- Consolidación de todos los servidores NAS en el propio filer.
- Gestión única de toda la infraestructura de almacenamiento.
- Escalabilidad de la solución para un crecimiento futuro a medio y largo plazo.
- Mecanismos de backup propios de la cabina e integración con software de terceros.

Para la infraestructura NAS (NFS), se disponía de un conjunto de servidores que a su vez prestaban servicio a otros servidores vía NFS. En este caso, al ser sistemas de ficheros, la migración de datos revestía menos complejidad ya que, aprovechando también la capacidad de los servidores de tipo Linux desde los propios servidores origen, se montaron los volúmenes destino por NFS y, con el software de replicación Rsync, se transfirieron los datos en tasas de replicación bajas para no afectar al servicio. En este caso, el paso clave de la migración fue integrar la propia cabina con el controlador de dominio LDAP (un cluster de RedHat también implantado por SATEC) ya que ahora es la cabina la que sirve directamente los ficheros por NFS. Al hacerse esta integración al final de la migración de datos, también ha garantizado unos cortes de servicio mínimos dentro de las ventanas estipuladas por el CESCA.

## Integración y control

El último paso ha consistido en integrar la cabina en la red de forma completa. Su gestión se realiza desde un único interfaz web permitiendo, de forma sencilla e intuitiva, realizar la mayoría de operaciones como la creación de nuevos agregados, volúmenes, LUNes y su publicación a través de distintos protocolos. Asimismo, se puede realizar una monitorización completa del estado de la cabina y de sus recursos utilizados y disponibles.

Paralelamente, a través del protocolo SNMP, se ha configurado su monitorización que podía integrarse con cualquier software de monitorización estándar, como NAGIOS u otros. La funcionalidad de AutoSupport genera avisos mediante traps SNMP y correos electrónicos (tanto a cuentas de correo del CESCA como de SATEC) ante el fallo o la anomalía de algún elemento de la cabina. Así, por ejemplo, si un disco falla, además de que la reconstrucción del mismo se inicia de forma instantánea en otro disco de Spare, se genera un email avisando de dicho fallo para que se pueda empezar a gestionar su reposición de manera inmediata.

Mediante este despliegue, SATEC aborda un proyecto diferente gestionando una compleja infraestructura como es la del CESCA, muy distinta al resto de compañías y necesidades del entorno Enterprise, demostrando, una vez más, que SATEC está preparada para afrontar proyectos de gran envergadura en entornos de índole similar y a todos los niveles.

