

2

ALMACENAMIENTO JERÁRQUICO (HSM), HIERARCHICAL STORAGE MANAGEMENT

Manuel Jiménez Iturbide
Responsable de Almacenamiento y Consolidación
Getronics Iberia

La empresa consultora Research ha indicado que el éxito futuro de las entidades públicas y privadas, dependerá de las decisiones que se tomen hoy de acuerdo al conocimiento de la información con que cuentan. La información se ha transformado en la ventaja más crítica que las entidades poseen.

El valor de la información va más allá de simplemente proporcionar acceso rápido a ésta. Las organizaciones también deben evaluar maneras eficientes de unir la información diaria con los procesos y establecer métodos para proveer un seguro acceso a la información requerida en el tiempo preciso y desde cualquier estación de trabajo.

En los últimos tiempos el crecimiento de la información no estructurada por su volumen, complejidad y diversidad ha hecho necesaria una evolución en los sistemas de gestión de su almacenamiento.

Esta evolución en la Administración Pública se ha hecho imprescindible para evitar el caos en los sistemas de información tanto internos como externos, teniendo en cuenta tanto el incremento constante del volumen de información como la velocidad con que esta es necesitada.

Aunque las necesidades internas y externas sean diferentes, parece lógico disponer de un repositorio común donde se almacene y mantenga actualizada toda la información. Posteriormente toda esta información puede ser publicada con los controles de seguridad oportunos a través de los canales (internos / externos) más adecuados en cada caso.

La armonización europea ha llevado consigo oportunidades en cuanto a intercambio y compartición de información particular respecto al cumplimiento de normativas comunitarias.

El aumento del servicio al ciudadano es una constante en el entorno de la Administración Pública en nuestro país en paralelo a un aumento de la transparencia y auditabilidad aspectos claves en la gestión de las administraciones públicas de cara al ciudadano

Hoy las iniciativas más activas son aquellas en las que se mejora la eficiencia de la Administración y se mejora la calidad de servicio al ciudadano.

Impulsar la productividad y modernizar los servicios públicos en una economía impulsada por el conocimiento haciendo “accesible en línea” a las Administraciones Públicas e insistiendo en la necesidad de garantizar la seguridad en el mundo digital son elementos claves a la hora de diseñar los planes de acción necesarios para lograr las características anteriores de eficiencia y calidad.

Todo lo anterior está dentro de la directriz europea para 2005 que dice que Europa deberá contar con:

- Unos servicios públicos en línea modernos.
- Una administración electrónica.
- Unos servicios electrónicos de aprendizaje.
- Unos servicios electrónicos de salud.
- Un entorno dinámico de negocios electrónicos,

Y para hacer todo ello posible ,

- Un acceso de banda ancha ampliamente disponible y a precios competitivos.
- Una infraestructura de información segura.

Un sistema de almacenamiento de la información, totalmente integrado en el funcionamiento de la Administración Pública , podrá gestionar de forma eficiente y eficaz el ciclo com-

pleto de la vida de la información; almacenar de forma eficaz esta y compartirla es la base imprescindible para la gestión de la información en tiempo real, configurando servicios públicos interactivos, accesibles a todos y ofrecidos a través de múltiples plataformas.

A principios de la década de los 90 se estimaba que las grandes compañías perdían un documento cada 5 segundos. Hoy en día nos enfrentamos a una economía dominada por la información, en la que es imprescindible tomar decisiones en tiempo real, para lo que el acceso a toda la información almacenada es esencial para abrir la Administración a los ciudadanos.

Varios conflictos surgen cuando una organización utiliza diversos cauces para manejar la información y los procesos de negocio:

- Dificultad de acceso a la información.
- Existencia de distintas interfase de usuario.
- Incremento de los costes operacionales.

Si afirmamos que las personas son el corazón de los negocios, podemos decir que la información es el cerebro de los mismos. Inventarios, Expedientes, Comunicados oficiales, Publicaciones, Bases de datos, etc... deben tener una disponibilidad 7x24 para asegurar la continuidad del negocio. Sin esta información una empresa no puede sobrevivir, por lo que el almacenamiento y la protección de la información deben ser una de las principales preocupaciones en materia de TI de cualquier compañía.

El volumen de información que poseen los organismos públicos está aumentando de manera exponencial, hasta el punto de que los analistas prevén un crecimiento continuado de la necesidad de capacidad de almacenamiento por encima del 180% anual para, al menos, los próximos 3 años.

Pero el almacenamiento de toda esa información en discos de alto rendimiento tiene un coste elevado, y su uso no está justificado cuando esa información no va a ser utilizada y simplemente se conserva siguiendo un principio de “prudencia”. Es en este punto donde aparece el concepto del ILM (*Information Lifecycle Management*) o ciclo de vida de la información.

ILM (Ciclo de vida de la información)

La información que cualquier organismo genera tiene un ciclo de vida que viene marcado por el uso que podamos hacer de ella en un momento determinado. En la mayoría de los casos el valor de la información va disminuyendo progresivamente a medida que pasa el tiempo, hasta el punto de que llegado un momento su valor puede ser caso residual. El ritmo de maduración de la información varía según su naturaleza y aplicación, lo que añade una dosis de complejidad a su gestión en el tiempo.

Por ejemplo, un estudio de mercado sobre determinados hábitos de los ciudadanos puede ser un documento importante hoy, pero su valor pasados unos años será insignificante, pues los ciudadanos habrán evolucionado y sus hábitos no serán los mismos.

A medida que la información de una entidad va perdiendo valor, los costes derivados de su almacenamiento van perdiendo su justificación, por lo que el reto principal de las Direcciones de TIC es ir trasladando la información a dispositivos más baratos. Pero ¿cómo podemos asegurarnos de que la información necesaria va a estar siempre disponible?, ¿qué metodología debemos seguir para hacerlo?.

ALMACENAMIENTO JERÁRQUICO HSM (HIERARCHICAL STORAGE MANAGEMENT)

Un sistema transparente para las aplicaciones y los usuarios, que mantiene de forma automática la información en el tipo de soporte adecuado dependiendo del número de accesos, antigüedad y propietario de la información, permitiéndonos ahorrar costes

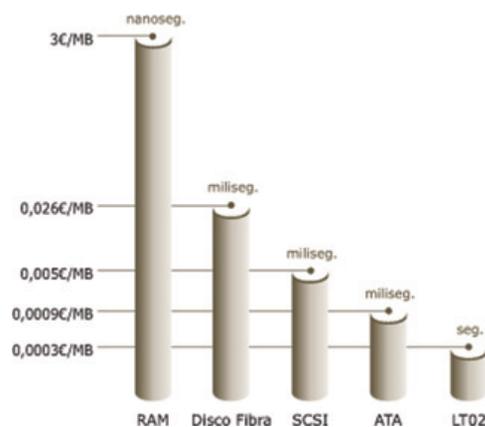
La información que vamos acumulando la tratamos de diferente forma dependiendo del valor que pueda tener para nosotros, no tratamos de igual manera la fecha de nuestro aniversario que el precio de un artículo del supermercado. La fecha de nuestro aniversario la tenemos siempre memorizada (incluso la apuntamos en nuestro calendario) por tener un gran valor para nosotros.

Si jerarquizamos nuestra información personal dándole más o menos importancia a determinados aspectos, ¿no debiéramos hacer lo mismo con la información corporativa?, ¿por qué tratar toda la información de la misma manera?, ¿tiene toda la información el mismo valor?. La respuesta es claramente NO.

Existen en el mercado diferentes *software* gratuitos que nos permiten revisar qué información de toda la que tenemos almacenada utilizamos frecuentemente y cual no se usa nunca una vez que se ha almacenado.

Seguramente no nos hemos parado a pensar en la cantidad de información a la que no accedemos pero que tenemos en nuestros discos de almacenamiento (de gran rendimiento y alto valor económico), pues nos da reparo eliminarla por si la necesitamos en un futuro. ¿No tendría sentido pasar esos datos a un sistema de almacenamiento más barato y liberar espacio de nuestros sistemas caros y de alta velocidad?

HSM, tecnología de gestión derivada de los entornos Mainframes que manejan gran cantidad de datos, es una estrategia de gestión del almacenamiento a través de la cual los datos son automáticamente "migrados", partiendo de unas políticas predefinidas, desde un soporte físico a otro. La información migrada aparece en todo momento como información local para los sistemas y cuando se accede a esta información, ésta es llamada y presentada al sistema independientemente del soporte físico donde se encuentre.



Los soportes físicos para el almacenamiento jerárquico se deben elegir dependiendo de:

- Coste por MB
- Tiempo medio de acceso a la información

La información crítica deberá ser alojada en un soporte con tiempo de acceso muy bajo, lo que elevará el coste por MB.

La información con pocos accesos o no tan crítica puede ser alojada en soportes con tiempos de acceso mayores, ya que el coste del MB es menor y nos permitirá reducir el Coste Total de la Propiedad de nuestros sistemas.

SOLUCIONES HSM DE GETRONICS

<p>Reducción TCO (Coste Total de la Propiedad)</p>	<ul style="list-style-type: none"> • Migración automática de los datos con menos uso a dispositivos de menor coste. • Reduce tiempo de backup y recuperación de desastres. • Minimiza las intervenciones manuales del administrador (reduce costes de gestión). • Mayor eficiencia de utilización de los sistemas de almacenamiento
<p>Rápido acceso a datos</p>	<ul style="list-style-type: none"> • Acceso inmediato a los datos almacenados, sin importancia donde residan. • Los datos con alta importancia estarán almacenados en el soporte más rápido. • Transparente a usuarios y aplicaciones.
<p>Retención de datos de acuerdo a las necesidades del negocio</p>	<ul style="list-style-type: none"> • Migración y de-migración de datos flexible, de acuerdo a las políticas definidas por el administrador. • Minimiza el riesgo de pérdida de datos (borrado accidental, sobre escritura de cintas, etc...).

Optimizando el almacenamiento para diferentes entornos y tipos de aplicaciones, como por ejemplo:

- **Entornos Servidor de Correo**
 - Microsoft Exchange Server
 - Lotus Notes
 - Bloomberg Mail
 - Unix Sendmail
 - Instant Messaging
- **Entornos Base de Datos Oracle**
- **Extensión de filesystem hacia almacenamiento infinito**
 - Windows NT/2000
 - Sun Solaris
 - HP-UX
 - Tru64
 - SGI
 - IRIX
 - RedHat Linux

COMPARATIVA, NAS VS HSM

Las soluciones de almacenamiento HSM, como hemos dicho anteriormente, pueden suponer una reducción drástica de nuestros costes de almacenamiento, mediante la jerarquización de la información y su almacenamiento en los dispositivos más apropiados.

Para que podamos apreciar hasta donde puede llegar esta reducción, vamos a establecer dos escenarios diferentes para una misma empresa. Esta empresa tiene una información que ocupa 4Tb y cuenta con unos 400 usuarios.

Solución NAS

Servidor de almacenamiento NAS para 400 usuarios.

El servidor de ficheros consta de dos discos duros internos SCSI para mantener el sistema operativo en mirror y toda la información que es compartida en la red, la salvaguarda en un almacenamiento SCSI con discos SCSI.

Problemas que se presentan

- La infraestructura ya no es capaz de seguir creciendo.
- El coste del mantenimiento de la infraestructura es excesivo.
- El tiempo necesario para la realización del backup del sistema es demasiado grande.



Solución HSM de Getronics

Para una arquitectura con la problemática arriba definida, Getronics podría integrar una solución compuesta por:

- Solución *DiskXtender* de *Legato* Solución HSM que permite tener disponible un disco virtual de capacidad infinita para dar servicio a los clientes, haciendo invisible qué tipo de almacenamiento realmente está guardando dichos datos.
- *Array* de discos SCSI con discos ATA de 1TB de *Microstorage*. Proporciona una solución de gran capacidad a un costo muy bajo. Ideal para soluciones HSM de almacenamiento secundario.
- Autocargador de tecnología LTO2 con capacidad de 3,2TB de *Certance* con tecnología punta ante fallos. Seguridad y capacidad en tecnología LTO al menor precio.



Beneficios obtenidos con HSM

- Sólo en la implementación de la solución conseguimos una reducción de costes de un 52%, y en futuras ampliaciones los precios por MB se reducen en torno a un 82%.
- La ventana de Backup se REDUCE EN UN 75%, cuantificar los costes que tiene el tiempo de levantar un servicio crítico para dar servicio por un desastre en los sistemas sólo puede determinarlo usted.

- **Transparencia absoluta hacia los usuarios.**
- En una infraestructura con almacenamiento con discos de fibra óptica, el coste por MB de la solución es de 0,026€, pasar a una solución HSM Getronics, el aumento de este almacenamiento se podrá realizar sobre discos de tecnología ATA con un coste por MB de 0,0009€. La reducción de costes al aumentar su capacidad supera el 96,5%, además de estar asegurando la disponibilidad de sus datos constantemente.

LAS SOLUCIONES HSM DE GETRONICS LE PERMITIRÁN REDUCIR SUS COSTES EN ALMACENAMIENTO HASTA UN 95%

Este sistema transparente de almacenamiento jerárquico actúa sobre tres elementos esenciales en toda organización:

- Personas. Generando cargas de trabajo menores y actividades mas eficaces y eficientes.
- Tecnología. Integración con el resto de sistemas existentes.
- Procesos. Centralizando su lógica de proceso a través de accesos rápidos y seguros.

Las ventajas de contemplar un Almacenamiento Jerárquico HSM que Getronics propone son:

- Mejora de la productividad(trabajo sin excepción y sin tiempos muertos).
- Mejora de calidad en el servicio al ciudadano.
- Reducción del coste administrativo.
- Reducción del coste de material(reducción del volumen de archivos, de numero de fotocopias, de duplicidades, de soporte papel).
- Reducción de riesgos (control sobre la información).
- Optimización de los procesos de servicio al ciudadano(acceso instantáneo y concurrente a la información en cualquier momento).

Y los resultados a obtener son:

- Eliminación de trabajo manual de reparto, firma, registro y archivo de documentos.
- Acceso inmediato a la información, que sea accesible por varias personas a la vez y desde cualquier lugar.
- Eliminación de la posibilidad de perdida de documentos.
- Reducción del tiempo de ejecución.(disminución de tiempos de acceso, recuperación y distribución de la información)
- Reducción de los tiempos de espera.
- Facilitar la toma de decisiones(persona adecuada con la información oportuna).
- Incremento en la transparencia de los procesos de información(mejora de la calidad de gestión, tramitación y consulta de la información).
- Racionalización y automatización de los procesos de negocio , permitiendo los trabajos por excepción.

todo lo anterior deberá conllevar a:

- Incremento de la productividad, mejorando los procesos.
- Reducción de los costes de las operaciones(tiempo, espacio y costes) .
- Mejora del servicio al ciudadano.

mejorando la imagen publica y visibilizando el trabajo en curso.

En definitiva la implementación de este sistema de almacenamiento jerárquico dotara a la Administración Pública de un sistema transparente que facilitara el acceso a los diversos documentos, optimizando su operativa habitual y aumentando la eficacia en las diferentes tareas administrativas y técnicas, de cara a conseguir en definitiva un mejor y más rápido servicio al ciudadano.

Los clientes potenciales son los Ayuntamientos, Comunidades Autónomas y la Administración del Estado. Las características de su actividad son procesos complejos /alto volumen, necesidad de manejar excepciones, entorno sometido a regulación y con contenidos que se consideran valiosos.

A día de hoy la Administración Pública necesita una actualización tecnológica, partiendo de sistemas con dificultades en el mantenimiento u obsoletos.

La propuesta de valor que plantea el sistema de almacenamiento jerárquico propuesto es alcanzar la eficiencia y calidad en el servicio al ciudadano a través de la automatización de los procesos y contenidos (información no estructurada) críticos que son gestionados y distribuidos a usuarios(externos/ internos) y comunidades de usuarios.

CONCEPTOS CLAVE EN HSM

- **Cliente de Migración.** Sistema en red que contiene la información susceptible de ser migrada, ya sea ahora o en el futuro. Básicamente consiste en múltiples filesystems o volúmenes.
- **Servidor de Migración.** Es el servidor que realiza las tareas de migración de los clientes hacia los soportes físicos definidos.
- **Migración de fichero.** La operación a través de la que los ficheros son “migrados” desde la localización física donde reside hacia otro soporte, basados en las políticas de migración definidas. Esta operación puede ser realizada de forma manual o automática.
- **Políticas de migración.** Criterios por los cuales se realizará la migración de la información de forma automática. Estas políticas pueden estar basadas en el uso de los ficheros (ficheros que no han sido accedidos o escritos durante 90 días), tamaño del fichero (ficheros mayores a 100MB) o propietarios de la información (migración de todos los ficheros pertenecientes a un usuario determinado)
- **“De-migración” de ficheros.** Cuando un usuario necesita acceder a un fichero migrado, este es llamado automáticamente desde el almacenamiento secundario (o de migración). Esta operación es transparente al usuario, y aparece como información local. En algunos casos el usuario puede apreciar un ligero retraso en el momento de presentarle dicha información si ésta se encuentra en soportes con tiempos de acceso altos, como dispositivos de cinta.

