

## Jouni Frilander

### Director de gestión digital, preservación y acceso, YLE radio

Finlandia. Primero que nada quisiera dar las gracias a nuestros anfitriones por la invitación, es realmente todo un honor y me da mucho gusto poder compartir mis perspectivas con el público.

Les voy a hablar sobre algunos elementos que hemos abordado desde mi institución y de las estrategias para elegir los sistemas de almacenamiento masivo; así es que voy a tratar de ser lo más explícito.

Primero les hablaré de las finalidades del almacenamiento, su características, cuáles son los tipos de almacenamiento alternativos, cómo se clasifican los almacenamientos, según su dimensión, y de un caso con costos, ya que muchas veces no conocemos cuánto cuesta un MAM, y también de una lista de identificación para todas aquellas personas que quieran comprar un almacenamiento; y posteriormente ahondaré en las conclusiones.

Para empezar, todos estarán de acuerdo en que los archivos audiovisuales, incluyendo una fonoteca o una videoteca, necesitan información estandarizada para elegir la solución tecnológica más correcta para almacenar de manera masiva su tecnología.

Estos almacenamientos de tecnologías de la información pueden gestionarse de una manera bastante fácil y realmente no tienen por qué ser tan caros si vemos cuánto tiempo van a durar. Hay muchos socios tecnológicos para los archivos y compañías de TI que producen servicios específicamente para ellos y cada sistema se adaptará a las condiciones de cada una de las instituciones. No hay un sistema que esté listo de fábrica.

El propósito principal de tener un almacenamiento es salvaguardar todos los objetos, ya sean de video, audio y todos los materiales correlacionados y, a partir de esta conservación, dar acceso a ellos. Los almacenamientos también ofrecen la posibilidad de tener acceso en línea a todos los objetos y se pueden utilizar con una finalidad intermedia; es decir, para poder tener objetos de producción y almacenar todo lo que producen las compañías difusoras. También es una buena opción para respaldar todos los objetos, en muchas ocasiones no sólo necesitamos guardar el original, sino tener un respaldo.

Ahora, hablemos de las características de los almacenamientos, podemos tener disco duros, cintas de datos, discos ópticos, hoy, hasta tenemos distintas memorias portátiles y la capacidad del almacenamiento se basa en terabytes en horas, dependiendo de cuánto material se quiera almacenar y la unidad de almacenamiento por la que se opte, la velocidad del almacenamiento, puede ser la velocidad en la que se inyecte información y se saque del medio de almacenamiento. Hay muchas soluciones en donde variará la velocidad, ya sea si es audio o video o el tipo de archivo, habrá distintas velocidades según la versión del archivo que estemos manejando; por ejemplo, si se quiere sacar una copia del video que tenemos almacenado será muy distinto a los videos que se necesitan para producción que pueden estar en Beta y seguramente el acceso va a variar en velocidad.

Hay distintos sistemas de almacenamiento y cada uno tiene sus propios errores intrínsecos y modalidades para recuperar dichos errores. Una de las características que seguramente ustedes tendrán que tomar en cuenta antes de optar por un sistema de almacenamiento, será, qué tan compatible son con internet, con distintos formatos, si se podrán crear redes y almacenamientos alternativos; es necesario considerar antes, todo lo anterior.

Tenemos la cuestión de los costos, cuando uno se propone adquirir un almacenamiento habrán de contemplar la compra de software y hardware. Esta es una parte significativa de la inversión y además de esto, tenemos que contemplar los costos de operación y mantenimiento, además de los honorarios de mantenimiento y soporte. Especialmente después de que concluya el periodo de la garantía, ya que el equipo se compra y acaba la garantía en algún momento y ahí tenemos que asegurar que no incurramos en un costo oneroso para el mantenimiento y el soporte.

Hay servicios que prestan los proveedores TI, algunos serán directos, algunos tendrán que conseguirlos por otro lado, y todos estos costos hay que tomarlos en cuenta, además de las necesidades ambientales, como la climatización; otro elemento que hay que considerar es cuando se decide qué sistema usar, si necesitamos un sistema con operadores o vamos a utilizar un sistema que funcione totalmente de manera automatizada. Habrá partes del sistema que se deban manejar de manera manual y otras no. Acuérdense que un almacenamiento es un módulo que se combinará con otras partes del proceso, con el acceso a la información con los metadatos, con todos lo demás procesos que hacen que una entidad de almacenamiento funcione.

Ahora, hablemos de las alternativas o los distintos tipos de almacenamiento. En mi opinión, los ópticos no son medios para archivo, pero realmente los voy a mencionar por no dejarlos fuera; los almacenamientos ópticos son discos de CD (discos compactos), DVD (disco versátil digital) y Blue Ray, y su ventaja es que son realmente accesibles, no cuestan mucho, tenemos la posibilidad de tener equipo que escriba, sobre escriba o lea los medios y se pueden intercambiar entre usuarios con relativa facilidad. Pero, de alguna manera, los medios ópticos no son tan confiables, ya que están sujetos a muchos problemas de escritura y lectura, además de los problemas de compatibilidad y muchos de ustedes seguramente están considerando los medios ópticos para sus archivos. Les recomiendo que lean el informe de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (Unesco) por sus siglas en inglés, sobre los medios ópticos, donde claramente se especifican las ventajas y desventajas de los mismos; también pueden ser difíciles de gestionar si no tienen éstos un sistema automatizado en su biblioteca. El futuro de los medios ópticos es muy difícil de predecir, seguramente sí tendrán una vida útil en los años venideros, pero seguramente serán reemplazados por otros medios portátiles, tarjetas de memoria y servicios en internet. Se pueden meter a Amazon.com y comprar un espacio de almacenamiento a un precio bastante accesible.

Después, tenemos los discos de información: externos o internos; se pueden almacenar en un sistema de disco duro independiente, las ventajas de estos sistemas es que tienen la posibilidad de enviar y recuperar información y son realmente medios bastante confiables. Si se utiliza un RAID con las protecciones adecuadas que son fáciles de conectar y trabajan con

redes y computadoras independientes, no tendrán mayor dificultad. Lo único es que los sistemas de disco duro siguen siendo muy caros.

Por otro lado, se puede comprar un disco externo de 1 terabyte que cuesta aproximadamente US\$100, En cinco años seguramente el precio bajará un 40% para un sistema como éstos que no son tan confiables, ya que se necesita de alguna manera vigilar el nivel de calidad de los objetos guardados aquí. Hay sistemas de respaldo que no son muy comunes y si los hay, son muy caros, tienen sus limitaciones, se necesitan sistemas operativos o sistemas exclusivos para su funcionamiento y un sistema de respaldo independiente, que será igual de caro que el primero y además se basará en discos o en cintas.

El futuro de estos mecanismos es bastante estable y seguramente van a tener una presencia mucho mayor en el mercado. Los sistemas de cintas también son independientes, se pueden conformar en una biblioteca de cintas: son una combinación de memoria caché en disco y un sistema de almacenamiento por jerarquía HSM. La ventaja es que son más baratos, fáciles de usar que un sistema de disco duro cuando se necesite almacenar gran cantidad de archivos, los costos ambientales van a variar en cinco años. Las cintas son realmente confiables y es más fácil sacar copias y respaldos en esta tecnología; en este sentido hablamos de mucho mayor confiabilidad, las desventajas de las cintas son que no duran para siempre y se necesita hacer una revisión de la calidad e implementar un sistema para hacerlo, ya que hay que tratar de conservarlas de la mejor manera para que no pierdan el estado en el que se encuentran. También hay sistemas HSIM y AMS, en esta modalidad, que representan un costo extra, sube la complejidad y además podrían representar problemas en un futuro.

La entrada y salida de las cintas podría realizarse a una velocidad mucho menor que en una unidad de disco duro, el futuro de los sistemas de cintas serán bastante estables. Por ejemplo, la tecnología LTO con el LTO6 es la nueva generación después de otras dos, es bastante estable pero no sabemos qué va a venir después, así es que tendremos que esperar a ver qué pasa.

La combinación más alternativa actualmente es usar sistemas AM. Las ventajas son que los sistemas son más baratos, ya que solamente tienen una solución de disco duro y combinan las bondades del disco duro y las tecnologías en cinta; además, permiten una vida útil combinada de cintas y disco duro. Las desventajas son que el tiempo de acceso a los archivos se basa en lo que dure la cinta y, por lo tanto, es un sistema mucho más complejo que tener nada más un dispositivo de disco duro.

¿Podríamos clasificar los sistemas de almacenamiento por su dimensión? Los sistemas de bajo costo sirven para instituciones de baja escala que no tienen la posibilidad de gestionar a un nivel mucho más profundo y todo se maneja de manera manual.

Podríamos tener un sistema empresarial con la posibilidad de componer errores de alguna manera y se puedan recuperar archivos perdidos en casos de desastre y se podrá transferir parte del archivo a otro punto.

Vamos hablando de sistemas que varíen según la dimensión de la institución y del desempeño que se quiera, aquí tenemos los detalles de los sistemas de bajo costo: tienen un par de discos

duros con varias USB o Firewire conectadas a medios externos y son bastante viables para aquellas instituciones que no tienen la necesidad de sistemas sofisticados y que se basan más en la funcionalidad y el desempeño de los objetos que tengan almacenados. No representan muchos errores y el costo de adquisición varía de US\$100 por terabyte, es muy difícil hablar de los costos de operación o costos de mantenimiento porque éstas son operaciones manuales y están sujetas a cada institución y sus necesidades.

Los sistemas de mediana escala utilizan una solución de cintas y discos duros y pueden almacenar varias decenas de terabytes en datos y ya tienen herramientas para gestionar la entidad; me refiero al sistema y al desempeño de entrada y salida, en promedio es bastante bueno y el rango de error no es tanto, hay posibilidad de componer los errores y hay distintos costos. Si uno quiere comprar 50 terabytes con un sistema de disco duro, tendrán que gastarse alrededor de US\$175 y estaríamos hablando de US\$125 por el soporte, utilizando la tecnología SATA; si quisieran 200 terabytes tendrían que contemplar una erogación de US\$640,000 y US\$510,000 para todos los High Drives.

Los costos ambientales de este tipo de sistemas de almacenamiento en disco duro ascenderían un 30 a 40% de la compra en un periodo de 5 años, y lo mismo sería para una biblioteca de cintas, estaríamos hablando de un 10%, los costos de mantenimiento para un sistema de 50 terabytes sería alrededor de US\$12,000 por el hardware por año, y US\$8,500 por el software, y si se opta por un sistema un poco mayor, aquí tenemos los costos.

Si hablamos de un sistema más sofisticado en donde tenemos disco duro y cintas, una vez más, un sistema que pueda almacenar cientos de terabytes de datos, hasta petabytes, sería un sistema en donde ya cuentan con herramientas para gestionar el sistema con muy buen desempeño de entrada y salida y se puedan almacenar hasta petabytes, con la capacidad de recuperarse de los errores y no hay tiempos muertos; además, en estos sistemas hay posibilidades de recuperarse en caso de desastres, van a poder replicar su almacenamiento a otra ubicación geográfica. Los costos de adquisición para un sistema de disco duro de 500 terabytes ascenderían a US\$1,700,000 y tendrían que gastarse US\$1,300,000 en los hardware con los drives SATA y si quisieran tener una biblioteca de cintas de 500 terabytes costaría US\$140,000 y dedicarían entonces US\$185,000 para comprar las cintas, los soportes y también deben tomar en cuenta que van a tener que contemplar, dependiendo de sus necesidades, la adquisición de los dispositivos para las cintas que cuestan alrededor de US\$20,000 a la fecha.

En cuanto a los costos ambientales, no hay mucho qué decir en este rango, es básicamente lo mismo al anterior. Los costos de mantenimiento para un sistema de disco duro de 500 terabytes ascienden alrededor de US\$140,000 al año para el hardware, y US\$52,000 al año para el software, y para la biblioteca de cintas de 500 terabytes, baja un poco para el hardware, serían US\$46,000 al año y para el software sólo US\$1,500 al año.

Hablemos de los costos, hace algunos meses trabajamos en esta propuesta para hacer una estimación de las capacidades de un sistema de MAM y cómo lo proyectamos a los próximos cinco años, tratamos de hacer una aproximación de cuánto dinero tenemos que gastar para éste y cuáles son los costos relacionados, y nos dimos cuenta que tendríamos que utilizar cinco terabytes de imágenes y materiales diversos cada año y añadir 10 terabytes de audio lineal

cada año y 25 terabytes de video previamente revisado en cuanto a su calidad, cada año y todo este material estaría en nuestro almacenamiento y tendríamos que tener aproximadamente 900 terabytes a 1.2 petabytes de video de calidad para producción; esto lo almacenaríamos en cintas con una copia en la biblioteca y probablemente otra copia en línea en otra ubicación, y nuestro costo de inversión para esta propuesta para los próximos cinco años sería alrededor de 2.3 millones de dólares y, a razón de esto, tendríamos que gastarnos alrededor de ocho millones de dólares para operar el sistema, y al gastarnos tal cantidad de dinero tendríamos un costo de operación para mantener nuestros datos y serviríamos a unos 200 usuarios para nuestra organización, nosotros hemos decidido que nuestra propuesta es quedarnos con los soportes viejos durante un periodo de tiempo, entonces tenemos aquí, costos que se empalman debido a esta decisión.

Aquí tengo un diagrama de las mismas cifras, las columnas en verde indican la inversión al año y en rojo tenemos los gastos anuales por el servicio de las operaciones, estamos hablando de un dineral. Estas cifras van a utilizar material en calidad DV25, si estuviéramos almacenando la información con formatos comprimidos, realmente creo que bajaríamos hasta ocho veces la cifra que tenemos aquí, pero ¿cómo vamos a dar soporte a este sistema? Estos son elementos que no hay que subestimar, porque el sistema debe ser lo suficientemente flexible para poder continuar con las operaciones;

es decir, si queremos hacer la conversión a otro formato o queremos tener la posibilidad de agrandarlos y tenemos un sistema que es fácil de gestionar y, aunque se puede automatizar en ciertos puntos, sabemos que vamos ahorrar dinero a futuro y nos ayudará mucho en todas las labores de migración al futuro.

Como lo han dicho ya otros oradores, en algún momento van a tener que cambiar todo su contenido a otro medio, habrá que migrarlo cada cinco años aproximadamente y si no tienen un sistema bien estructurado, van a tener que hacer las cosas de manera manual y eso representaría un problema. Imagínense una situación en donde el sistema no pueda ser gestionado de la manera más adecuada o si no pueden recibir apoyo profesional por la versión que manejan, entonces estaríamos hablando de que en sus manos tendrían una catástrofe.

Habrá cambios en el futuro, el precio por gigabyte en los discos duros va a seguir bajando más rápido de lo que van a bajar las tecnologías de cintas y las compañías consultoras como CTel, y hasta nos dicen que sus precios van a bajar hasta un 15% cada año y seguramente habrá nuevas tecnologías como el almacenamiento holográfico; el almacenamiento DNA seguramente dejará de ser un programa piloto en los laboratorios y serán propuestas, pero no tenemos claro cuál es la perspectiva a futuro, de qué es lo que vendrá a reemplazar los dispositivos de discos duros, de discos de datos y qué es lo que le dará más valor a las aplicaciones de documentación. En el futuro estoy seguro de que será posible seguir comprando servicios profesionales de almacenamiento y esto nos va ayudar a contemplar los costos totales de una operación de almacenamiento y no vamos a tener que pensar separadamente. Los costos seguramente tendrán más claridad en el futuro; si acudimos hoy a un proveedor y le decimos: “mira quiero comprar un petabyte de almacenamiento, ¿cuánto me va a costar?”. Nos darán un precio en este momento pero en un futuro las cosas van a

cambiar y, si me permiten, les voy a compartir una pauta: primero hay que asegurarnos que estemos conscientes de todos los requisitos, de la necesidad de capacitación, de exportación, conectividad, procedimientos de respaldo, de todas las necesidades de mantenimiento y el personal; primero hay que entender cuál es el modelo de precio de los proveedores de software y hardware, cuál es el soporte, y no olviden, deben incorporar todos los servicios IT, los costos de respaldo, e incorporarlos en sus cálculos; acuérdense del costo total de la propiedad del sistema; es decir, cuáles son los componentes que tendrán que adquirirse y operar y cuáles son los costos totales de los componentes adicionales.

Cuando hay más competencia, los precios bajan, si compran un sistema al cierre del año fiscal, seguramente les van a dar un mejor precio. Asegúrense de que tengan un sistema que pueda desempeñarse bien cuando haya que hacer una migración de un medio a otro y que funcione automáticamente. Acuérdense que la próxima migración probablemente incluya el cambio de todo el sistema de almacenamiento y no sólo del reemplazo de un disco duro o un cartucho de cinta en sus soportes. Es necesario decidir cuál será el ciclo de vida del sistema; hay que mantenernos en términos prácticos, no se olviden de los honorarios derivados del soporte y que en cuanto termine la garantía, la situación cambia y habrá sistemas que después de cinco años ya serán sistemas viejos y será muy difícil encontrar soporte técnico.

Este no es un comercial, pero aquí tienen a las compañías que nos ayudaron a establecer nuestros sistemas y además me dieron la oportunidad de obtener información para preparar esta presentación. Para concluir, la migración es una labor muy compleja y para establecer una propuesta de compra para estos sistemas, hay que acudir a un proveedor competente, si es que lo llegan a encontrar en donde estén ubicados, y si van a comprar un sistema de almacenamiento con una capacidad de alrededor de 100 terabytes ó 150, o menos, probablemente sea más práctico tener una solución que se base en un sistema de disco duro. La migración de datos de un medio de almacenamiento a otro será un proceso constante, y así será el futuro; por lo tanto, todo el tiempo se repetirá este paso. Si tienen cinco terabytes, a lo mejor tendrán que migrarlos constantemente en el futuro y asegúrense de que sus sistemas puedan manejar este tipo de migraciones automáticamente en el futuro. Hay que tener una herramienta de virtualización que nos haga pasar de este entorno al que sigue, y primero que nada, hay que estudiar cuáles son las necesidades reales, si solamente después de entender claramente qué es lo que se necesita, podremos entonces acudir a un proveedor para comprar un solución que realmente satisfaga sus necesidades particulares, ya que las necesidades son múltiples.

Las Fonotecas Nacionales tienen necesidades muy distintas a la de una televisora, las distinciones son múltiples.