



## Maximizar los Beneficios de la Virtualización con la Compresión en Tiempo Real

por Dan Kusnetzky, Analista Reconocido

### RESUMEN EJECUTIVO

Las organizaciones utilizan cada vez más datos. Gran parte de este crecimiento puede atribuirse a la creciente colección de datos analíticos, documentos, mensajes, presentaciones y otros datos no estructurados. Los requisitos para la retención de datos y la necesidad de una copia de seguridad y de planes de recuperación en caso de desastre han llevado a las organizaciones a enfrentar el hecho de que ambos procedimientos absorben mucho tiempo.

Con la dificultad que actualmente conlleva responder a las necesidades de mantenimiento de copias de seguridad, es probable que la introducción de servidores y entornos virtuales sólo agrave el problema. La proliferación de entornos virtuales significa que existen más fuentes de datos para incluir en una copia de seguridad de la empresa y en un plan de almacenamiento de archivos.

Los encargados de tomar las decisiones deben considerar el rápido crecimiento de los activos de información de la compañía, los desafíos de realizar copias de seguridad para estos datos y desarrollar planes concretos para resolver este problema.

Un abordaje sólido sería reducir el tamaño de los datos que deben ser almacenados, comenzando con los datos principales. A pesar de su solidez, este abordaje solo puede utilizarse exitosamente si se centra adecuadamente la atención en no impactar el rendimiento y compatibilidad con las aplicaciones y procesos actuales. La compresión en tiempo real puede ser el mejor abordaje, ya que no solo contrae los datos principales en tiempo real, sino que también reduce los requisitos para el almacenamiento nuevo y existente durante todo el ciclo de vida de los datos sin impactar en el rendimiento de las aplicaciones ni forzar a la compañía a cambiar aplicaciones, procesos, procedimientos o reentrenar el personal comercial o de TI.

### ¿CUAL ES EL PROBLEMA?

Las compañías enfrentan una explosión en la necesidad y utilización de datos. Al mismo tiempo, estas mismas organizaciones también enfrentan la necesidad de reducir los gastos generales de TI. Estas tendencias divergentes se encuentran en la base de los diversos problemas que la gestión de TI enfrenta actualmente.

Documento #20110519

El Kusnetzky Group © 2011

El Kusnetzky Group es un proveedor independiente de servicios de marketing para proveedores y organizaciones de usuarios, y proveedores en el área de sistemas, virtualización, computación en nube y mercados de tecnología de código abierto. Los proveedores de tecnología se encuentran entre esas organizaciones. Las opiniones que aparecen en este documento se basan en nuestras investigaciones, en nuestra experiencia personal y en el uso real de la tecnología independientemente de si este documento o la investigación de soporte fue patrocinada por uno o más de los clientes del Kusnetzky Group. Este documento no debe ser copiado parcial ni totalmente sin la autorización por escrito del Kusnetzky Group. Este documento informativo del Kusnetzky Group fue encargado por IBM y se encuentra bajo la licencia del Kusnetzky Group

## FACTORES QUE GENERAN EL IMPORTANTE CRECIMIENTO EN EL ALMACENAMIENTO

Observemos algunos de los factores que generan este importante crecimiento de los datos.

- ☒ Para reducir los costos relacionados con el personal, las compañías están, de manera creciente, automatizando los procesos que antes eran manuales. Estos nuevos procesos automáticos producen datos estructurados y no estructurados adicionales. En ambos casos, es necesario una capacidad de almacenamiento adicional.
- ☒ A fin de mejorar los procesos de venta y marketing, así como reducir los costos de adquisición y retención de clientes, las compañías están comenzando a utilizar la inteligencia de negocios y el software de proceso analítico en tiempo real para develar los requisitos, antes ocultos, de estos clientes. Esto significa recolectar y almacenar grandes cantidades de datos de la gestión de relaciones con el cliente y de los sistemas de ventas y marketing.
- ☒ Las compañías observan un crecimiento continuo en el uso de datos no estructurados, tales como mensajes, documentos, presentaciones, hojas de cálculo y similares, como parte de sus operaciones. Independientemente de donde se utilicen, ya sea en una computadora, en un smartphone o en una computadora tablet, estos datos deben ser respaldados en una copia de seguridad.
- ☒ Los entornos virtuales están aumentando la cantidad de datos que se procesan. A medida que las compañías comienzan a encapsular la carga de trabajo de servidores en máquinas virtuales, existen cada vez más fuentes de datos para incluir en sus copias de seguridad y en sus planes de almacenamiento de archivos.

Esta encuesta rápida cubre solo una parte de la creciente cantidad de datos que el personal y los sistemas de una organización están creando. Lo que queda claro es que la cantidad de datos que las compañías necesitan guardar en copias de seguridad está creciendo rápidamente. También queda claro que es cada vez más difícil realizar copias de seguridad de todo.

### ¿QUÉ DIFICULTADES CAUSA ESTE PROBLEMA?

La mayor cantidad de requisitos para el almacenamiento de datos causa una gran cantidad de problemas que incluyen mayores costos, dificultad para encontrar un lapso de tiempo para realizar con éxito una copia de seguridad para los datos y el estar preparado ante una auditoría o algún desastre.

### MAYORES COSTOS AL MISMO TIEMPO EN QUE SE REDUCEN LOS PRESUPUESTOS

El crecimiento exponencial de datos estructurados y no estructurados, con frecuencia, conduce a la adquisición de dispositivos de almacenamiento, equipos para la red de almacenamiento y software para almacenamiento adicionales. Estos equipos adicionales requieren energía y enfriamiento, lo que solo aumenta el alto precio de adquirir y mantener el almacenamiento. Además, los requisitos para realizar una copia de seguridad de todos los datos también representa un incremento en la cantidad de personal necesario. Generalmente, esto conduce siempre a aumentar los costos de TI.

Debido a que las organizaciones buscan reducir los presupuestos de TI, el crecimiento de los datos se presenta como un serio problema.

¿CUÁNDO ES EL MOMENTO PARA REALIZAR UNA COPIA DE SEGURIDAD EN UN MUNDO QUE FUNCIONA LAS 24 HORAS DEL DÍA, LOS 7 DÍAS DE LA SEMANA, LOS 365 DÍAS DEL AÑO?

Algunas organizaciones se dan cuenta que, simplemente, no pueden realizar una copia de seguridad de todos los datos que sus procesos y sistemas crean. Actualmente, tanto el personal como los clientes del mundo entero acceden a los sistemas operativos. Por lo tanto, los sistemas operativos y de soporte deben estar disponibles las 24 horas del día, los 7 días de la semana, los 365 días al año.

Anteriormente, el personal de TI solía realizar copias de seguridad de los sistemas luego del horario comercial, de acuerdo con su propio huso horario. Esto ya no funciona así, puesto que los sistemas están siempre en uso. Ya no existe más una hora específica que pueda considerarse como oportuna para realizar la copia de seguridad de los datos que una empresa está utilizando.

LA VIRTUALIZACIÓN EXACERBA EL PROBLEMA

Este problema se ve agravado por la creciente utilización de entornos virtuales.

- ☒ Cada servidor virtual consume y produce datos. A medida que los servidores virtuales crecen en número, también lo hacen los datos. Cada servidor virtual posee su propia copia del sistema operativo, del software de aplicaciones, del software que dan soporte a estas aplicaciones y de los datos y software de las bases de datos. A medida que la compañía confía cada vez más en cientos o, tal vez, miles de servidores virtuales, la cantidad de datos producida crece a un ritmo paralelo.
- ☒ Las aplicaciones de Comercio Electrónico de Internet, con frecuencia basadas en tecnología de virtualización para los procesos y aplicaciones, aumentan bruscamente la cantidad de datos que una organización debe almacenar. Además, los datos cambian lo suficientemente rápido como para desbordar las posibilidades de muchos sistemas de almacenamiento.
- ☒ Las organizaciones utilizan cada vez más aplicaciones de modelado para gestionar riesgos, realizar pruebas no destructivas, desarrollar procesos de fabricación y, por supuesto, como parte de su desarrollo de arquitectura. Estos modelos son, frecuentemente, archivos muy grandes.

¿CUÁL SERÍA LA SOLUCIÓN IDEAL?

Consideremos las características de una solución ideal para estos problemas. La solución debe poseer los siguientes atributos para tener éxito:

- ☒ La solución debe obtener el control del crecimiento explosivo de datos y reducir los costos de los centros de datos al disminuir el almacenamiento de espacio físico en el piso, la energía consumida y el calor producido por los dispositivos de almacenamiento.
- ☒ En vez de enfrentar la constante instalación de nuevos dispositivos de almacenamiento, la solución debe posibilitar que la organización obtenga una mayor utilización de sus sistemas de almacenamiento actuales. También debe reducir la necesidad constante de dispositivos de almacenamiento adicionales.

- ☒ La solución debe pasar por la actual infraestructura de TI sin causar un impacto negativo en el rendimiento de la aplicación. Las aplicaciones deben, como mínimo, tener el mismo rendimiento que antes.
- ☒ La solución debe reducir la cantidad de datos que se almacena y, por lo tanto, el tiempo que consume la realización de la copia de seguridad de los datos y el soporte para los planes de recuperación en caso de desastres. Este cambio debe realizarse con facilidad y sin requerir extensas y costosas modificaciones en las configuraciones de almacenamiento.
- ☒ La solución no debe forzar cambios en los procesos, procedimientos o en el personal de la organización.
- ☒ Las aplicaciones y bases de datos, independientemente de que tengan soporte físico o de servidores virtuales, deben continuar funcionando como antes, sin necesidad de reestructuraciones o nuevas copias.
- ☒ La solución ideal debe ser un buen vecino y funcionar con otras tecnologías eficientes de almacenamiento, tal como la eliminación de datos duplicados.

IBM Real-time Compression puede resolver estos problemas, fácil y rápidamente, sin excederse del presupuesto.

#### ¿QUE ES UNA COMPRESION EN TIEMPO REAL Y DE QUE MANERA PUEDE AYUDAR?

La compresión de datos es una herramienta que los desarrolladores de TI utilizan para reducir el tamaño de los objetos de datos.

Tradicionalmente, esto se logra mediante la codificación de datos para disminuir el almacenamiento físico.

Los métodos de compresión se basaban en encontrar patrones repetidos de ceros y unos que forman un archivo digital y reemplazarlos por una única copia de ese patrón y la cantidad de veces que ese patrón debe repetirse cuando el archivo es expandido para su utilización.

Antiguamente, la compresión requería una gran cantidad de energía para su proceso y, por ende, no era adecuada para su utilización con propósitos generales. Los desarrolladores de TI tenían requisitos para la integridad y disponibilidad de los datos que desafiaban los abordajes anteriores para la compresión de datos. Aún con estas limitaciones, este método encontró sólidos partidarios en las comunidades de creación de contenido digital y redes.

Mejorías recientes en el rendimiento de los microprocesadores y en los algoritmos de compresión de datos (Lempel-Ziv) permitieron que la compresión de datos se lograra rápidamente y con un nivel de confiabilidad que hizo posible su utilización como una herramienta para la optimización del almacenamiento principal. Los Real-time Compression Appliances de IBM son un ejemplo de la moderna tecnología actual.

Los IBM Real-time Compression Appliances fueron diseñados para mejorar significativamente la eficiencia del NAS. Algunos de los beneficios que las empresas pueden obtener:

- ☒ Reducir los datos principales de NAS y de la copia de seguridad en hasta un 80%, sin degradación del rendimiento. Los Real-time Compression Appliances de IBM permiten que los datos de NAS sean manejados de forma comprimida, lo que conduce a mejoras en las operaciones finales de los sistemas de almacenamiento (por ejemplo, copia de seguridad y recuperación, archivo, etc.). La menor cantidad de almacenamiento para realizar una copia de seguridad y almacenar representa mayor rapidez en estos procesos y en el de recuperación, mayor disponibilidad de datos, mejor RPO/RTO, así como menores costos. La compresión funciona con datos y sistemas de almacenamiento del NAS y puede lograr una eficiencia hasta cinco veces mayor en el almacenamiento.
- ☒ Los archivos de la máquina virtual y los datos producidos por los entornos virtuales también pueden verse significativamente reducidos en tamaño. En entornos VMware, IBM Real-time Compression reduce el almacenamiento principal en hasta un 72%.
- ☒ IBM Real-time Compression mejora la eliminación de datos duplicados para mejorar la reducción general de datos.
- ☒ Hay un efecto de eficiencia en cascada. IBM Real-time Compression permite importantes reducciones en la CPU de almacenamiento y en la utilización de discos, así como menos tráfico de almacenamiento en la red.
- ☒ La reducción en la extensión de los datos redundante en menos equipos de almacenamiento para gestionar, enfriar y proveer energía.
- ☒ Los IBM Real-time Compression Appliances no ocasionan interrupciones en los sistemas, aplicaciones y procesos existentes.
- ☒ Los IBM Real-time Compression Appliances son fáciles de instalar y de utilizar.
- ☒ Las mejoras en todo el ciclo de vida de la información pueden ayudar a disminuir aún más los gastos de gestión.

#### HISTORIAS DE LOS USUARIOS

Examinemos un ejemplo de cómo una organización de TI que desplegó un entorno de virtualización se encontró frente a un explosivo crecimiento de datos, y cómo luego se benefició de la utilización de la Real-time Compression de IBM para gestionar ese crecimiento.

##### UNIVERSIDAD BEN-GURION

La Universidad Ben-Gurion, en Israel, es uno de los principales centros de enseñanza e investigación, con más de 19.000 estudiantes matriculados en las Facultades de Ciencias de Ingeniería, Ciencias de la Salud, Ciencias Naturales, Humanidades, Ciencias Sociales, como también en la Guilford Glazer Faculty of Business and Management y en la Kreitman School of Advanced Graduate Studies.

El departamento de TI de la Universidad Ben-Gurion es responsable por todos los proyectos y problemas relacionados con la informática, incluyendo las matriculas, la investigación académica, las clases de computación y los programas de soporte. Además de los programas de la universidad, algunos selectos grupos de investigación externos

pueden utilizar los entornos de computación de alto rendimiento en el campus.

La Universidad Ben-Gurion utiliza VMware para la virtualización de los servidores.

#### DESAFÍO

Todos los grupos son añadidos al principal clúster de datos de la universidad para proporcionar acceso continuo a los recursos y la posibilidad de utilizar los recursos de almacenamiento disponibles asignados a otros grupos de la misma universidad. Al utilizar un modelo de "almacenamiento por demanda" que permite a los investigadores, laboratorios y a departamentos enteros adquirir espacio de disco gestionado, de acuerdo con sus necesidades, la Universidad Ben-Gurion permite a los grupos crecer esos espacios a lo largo del tiempo, al mismo tiempo que garantiza que los datos tengan una copia de seguridad completa en un archivador de copias de seguridad, para luego ser archivados en una cinta. Ben-Gurion comenzó a experimentar un crecimiento exponencial en el almacenamiento de datos con una cantidad creciente de grupos de investigación, administración de la universidad y departamentos académicos que exigían cada vez mayor capacidad de almacenamiento.

El departamento de TI actualmente cuenta con una agrupación central de servidores VMware que funciona en más de 50 servidores virtuales con 1,5 TB de espacio de disco cada uno, desde un servidor de archivo NetApp® centralizado. Con la demanda de almacenamiento de datos que continuaba creciendo rápidamente, el departamento de TI se encontraba en la búsqueda de una nueva tecnología que pudiera mejorar su eficiencia en almacenamiento. Necesitaba una solución que le permitiera utilizar la capacidad existente de disco con mayor eficiencia sin afectar el rendimiento ni crear una capa adicional de complejidad en la gestión.

#### SOLUCIÓN

El departamento de TI de la Universidad Ben-Gurion desplegó los IBM Real-time Compression Appliances para comprimir automáticamente los datos de Microsoft Office, los datos de investigación y las bases de datos de Oracle almacenados en sus soluciones virtuales NAS y Sun® Solaris®. El departamento de TI logró comenzar a comprimir los datos inmediatamente. Todos los datos que ya estaban almacenados en el sistema de la universidad fueron comprimidos, reduciendo cada volumen existente de 300 GB a menos de 50 GB. El tamaño reducido de los volúmenes los hizo más manejables para realizar funciones tales como la duplicación y la realización de copias de seguridad, y el entorno de almacenamiento total se ejecuta ahora con mucha más rapidez que antes.

Ben-Gurion logró una reducción del 65% al 83% en todos sus datos de Microsoft Office, datos de investigación y bases de datos de Oracle. Al comprimir directorios de inicio en el momento, los administradores de TI fueron capaces de crear nuevos volúmenes y de copiar a través de datos con mayor rapidez y facilidad. Todos los datos son comprimidos sin ninguna dificultad y no es necesario detener el entorno de producción, lo que facilita la copia de datos en el momento necesario y mejora el rendimiento general del sistema.

Con la habilidad para comprimir datos sin dificultades y sin afectar el rendimiento del entorno de producción, el departamento de TI de la



“En los últimos tres años, hemos visto una tasa de crecimiento exponencial de los requisitos para el almacenamiento de datos. Esto no ha sido un incremento lineal del crecimiento del almacenamiento. Los datos crecen tan rápidamente en el campus y el almacenamiento posee una demanda tan grande que tuvimos que encontrar una solución que nos permitiera utilizar mejor nuestra infraestructura ya existente. Estamos sorprendidos por la cantidad de compresión que podemos lograr al utilizar el IBM Real-time Compression.”

Shmuel Gruber, Director del Centro de Informática de la Universidad Ben-Gurion

Universidad Ben-Gurion ahora puede ofrecer a los usuarios capacidad adicional de almacenamiento al mismo costo. Esto también permite a los gestores de TI a utilizar mejor las soluciones de almacenamiento ya existentes en la universidad.

Con el ahorro de espacio adicional, el departamento de TI de la universidad logró ampliar el ciclo de vida de la infraestructura actual sin tener que adquirir continuamente capacidad adicional de almacenamiento basado en disco y utilizada de manera ineficaz. Al ofrecer almacenamiento por demanda y de capacidad optimizada a las salas de computación, a los departamentos y a los grupos de investigación de la universidad, el departamento de TI puede brindar mayor espacio utilizable a los usuarios finales y obtener un mejor retorno de sus adquisiciones para almacenamiento.

#### BENEFICIOS CLAVES

El departamento de TI de Ben-Gurion no solo logró su ROI en el primer año de utilización, sino que la gestión de menos almacenamiento también permitió a los gestores de TI centrarse en otras tareas importantes. Además de los datos estructurados ya comprimidos, la universidad pronto comenzará a comprimir datos no estructurados que se encuentran almacenados en el entorno informático de alto rendimiento.

**Eficiencia en el almacenamiento:** la Universidad Ben-Gurion logró reducir de 300 GB a 50 GB cada uno de sus volúmenes de almacenamiento en el primer año del despliegue de IBM Real-time Compression, lo que, a su vez, redujo los requisitos de gestión y permitió la utilización eficiente de sus servicios de almacenamiento por demanda de un 65% a un 83%.

**Ahorro en los costos:** los aparatos de IBM Real-time Compression permiten a la Universidad Ben-Gurion ampliar la vida útil de la capacidad de sus discos existentes y reducir el almacenamiento, lo que da como resultado un mejor ROI en almacenamiento para los departamentos individuales y los grupos de investigación en menos de un año. El departamento de TI es capaz de ofrecer a los laboratorios de investigación y a los departamentos mayor capacidad de almacenamiento al mismo costo, y permite ampliar la vida útil de su infraestructura ya existente a pesar del crecimiento de los requisitos de almacenamiento.

#### R E S U M E N

Las organizaciones utilizan cada vez más datos. Gran parte de este crecimiento puede atribuirse a la creciente colección de datos analíticos, documentos, mensajes, presentaciones y otros datos no estructurados. Los requisitos para la retención de datos y la necesidad de una copia de seguridad y de planes de recuperación en caso de desastre han llevado a las organizaciones a enfrentar el hecho de que ambos procedimientos absorben mucho tiempo.

Con la dificultad que actualmente conlleva responder a las necesidades de mantenimiento de copias de seguridad, es probable que la introducción de servidores y entornos virtuales sólo agrave el problema. La proliferación de entornos virtuales significa que existen más fuentes de datos para incluir en una copia de seguridad de la empresa y en un plan de almacenamiento de archivos.

La mayor cantidad de requisitos para el almacenamiento de datos causa una gran cantidad de problemas que incluyen mayores costos, dificultad para encontrar un lapso de tiempo para realizar con éxito una copia de



seguridad de los datos y el estar preparado ante una auditoría o algún desastre.

La reducción de datos principales del NAS permite a las compañías resolver los problemas de eficiencia en el almacenamiento en su mismo origen. Los beneficios de la compresión en tiempo real se extienden en cascada hacia las operaciones finales de almacenamiento, tales como la realización de la copia de seguridad, recuperación y el archivado, además de optimizar la eficiencia de otras tecnologías de reducción de datos, como la eliminación de datos duplicados. En los entornos de virtualización, la compresión en tiempo real complementa los requisitos de reducción del sistema y del almacenamiento, y contribuye a mejorar el costo general de ejecución de la infraestructura de TI al disminuir la cantidad de equipos de almacenamiento necesarios, a los cuales se debe gestionar, enfriar y proveer energía. Los IBM Real-time Compression Appliances pueden mejorar este alto nivel de optimización de almacenamiento en hasta cinco veces, sin degradar el rendimiento ni forzar a la compañía a cambiar aplicaciones, procesos o procedimientos.

IBM Real-time Compression debe ser considerado como una parte de cualquier gran iniciativa empresarial para la optimización de la capacidad de almacenamiento en grandes centros de datos privados virtuales y entornos de gestión en nube.

Para obtener más información, visite <http://www.ibm.com/storage/rtc>