

WHITE PAPER

La compresión en tiempo real promueve la optimización del almacenamiento.

Con el auspicio de IBM

Laura DuBois
Mayo de 2012

RESUMEN EJECUTIVO

No debería sorprender que los presupuestos de almacenamiento estén constantemente sometidos a presiones de fuerzas opuestas. Por un lado, las fuerzas económicas impulsan a los presupuestos a mantenerse planos o en muchos casos, a comprimirse como porcentaje de los ingresos de una compañía. Por otro lado, la infraestructura lucha por mantenerse al ritmo del crecimiento de los datos, presionada por muchas variables tanto sociales como económicas. Las empresas no tienen otra opción más que adaptar su Infraestructura de Almacenamiento a los niveles del crecimiento de datos.

En su mayor parte, la ineficiencia de la infraestructura de cómputo ha sido abordada a través de la virtualización de servidores. Racks y más racks de servidores ociosos fueron reemplazados sin prisa y sin pausa por entornos de cómputo virtualizados ultradensos que ocupan una fracción del espacio, delegando en los proveedores de almacenamiento la responsabilidad de abordar el desafío de hacer que el almacenamiento de datos sea eficiente y económicamente sustentable.

Los proveedores de almacenamiento han respondido desarrollando un conjunto de soluciones destinadas a desacelerar el consumo de almacenamiento. Las Tecnologías de Optimización del Almacenamiento, buscan reducir la huella del almacenamiento eliminando la redundancia y el desperdicio y optimizando la colocación de datos. Tradicionalmente, las ofertas de optimización del almacenamiento incluían aprovisionamiento ligero, de-duplicación y compresión post-proceso. El aprovisionamiento ligero apunta al almacenamiento desperdiciado que está asignado pero no es utilizado debido al sobre-aprovisionamiento. De igual modo, la de-duplicación apunta a datos redundantes o duplicados, creando instancias únicas de datos. Para los datos que no pueden "aligerarse" ni "des-duplicarse", la compresión apunta a la reducción de la huella total bajando el tamaño de los datos. La promesa ha sido que al implementar una o más de estas tecnologías, las empresas pueden aspirar a reducir su gasto en almacenamiento, diferir compras para el futuro y mejorar el rendimiento de su inversión.

Sin embargo, cada optimización supone algo de carga (overhead) para el sistema, en términos de rendimiento o eficiencia. Históricamente, el peor componente de esta combinación ha sido la compresión, debido a su inherente naturaleza post-proceso intensiva en cómputo. Esto a menudo ha llevado a las empresas a implementar aprovisionamiento ligero y/o de-duplicación pero a dejar la compresión de lado cuando se trataba de almacenamiento primario activo.

EN ESTE ARTÍCULO

En este artículo examinamos cómo la tecnología IBM Real-time Compression está cambiando el juego al ofrecer una solución de optimización robusta, eficiente y eficaz en función de costos. Más aún, IBM está demostrando que la compresión sí tiene un lugar en la optimización del almacenamiento y es un buen complemento para otras técnicas que hoy están en juego. IBM estima que implementando soluciones habilitadas para Real-time Compression tales como IBM Storwize V7000, las empresas pueden ahorrar un 50% o más en la cantidad de espacio físico requerido y reducir el crecimiento total del almacenamiento en casi 30%. Según IBM, esto se traduce en una reducción del 30% al 40% en costos por gigabyte (GB) para configuraciones de almacenamiento primario, sin comprometer la capacidad ni el rendimiento.

IBM Real-time Compression es un componente central de IBM Smarter Storage, que a su vez forma parte de la estrategia Smarter Computing de IBM, un enfoque construido a lo largo de años de liderazgo en la industria del almacenamiento y tecnologías innovadoras y previsoras que ayudan a guiar el diseño y la implementación de sistemas de almacenamiento. Smarter Storage permite a las organizaciones tomar el control de su almacenamiento para que puedan enfocarse en extraer conocimientos más valiosos de sus datos y ofrecer más valor al negocio.

Situación del mercado

Vivimos en el mundo post-PC. Los dispositivos de computación móvil siguen proliferando por doquier, incluso en las empresas. En el 2015, en el mercado de computación móvil habrá más de 243 millones de tablets y más de 1,5 mil millones de smartphones. Según IDC, aproximadamente 9,57 zettabytes (ZB) de información se consumen por año, de los cuales 1,2 ZB son datos de medios no estructurados que se desarrollan a una tasa de crecimiento anual compuesta del 62%. A medida que los aspectos demográficos de la empresa, el "prosumidor" y el consumidor se inclinan por la creación y el acceso a los datos en movimiento, se moviliza también la demanda del almacenamiento. Investigaciones de IDC muestran que los ingresos por sistemas de almacenamiento en disco externo en el mundo registraron un crecimiento interanual del 7,7%, por un total que roza los US\$ 6,6 mil millones, en el cuarto trimestre de 2011. La capacidad total de los sistemas de almacenamiento en disco distribuidos en 2011 llegó a 6.279, con un crecimiento del 22,4% año a año.

Se proyecta que la demanda empresarial de capacidad de almacenamiento en el mundo crecerá a una tasa anual superior al 43% de 2008 a 2013. Por consiguiente, la demanda de datos sobrepasará holgadamente la oferta de almacenamiento. Además, a causa del ritmo rezagado de la economía, las empresas ya no pueden darse el lujo de dar cabida a este crecimiento de datos utilizando enfoques tradicionales. Atrás quedaron los días en que el almacenamiento se trataba como una capa boba con mecanismos suficientes de protección y rendimiento como grupos RAID y administradores de volúmenes. Los costos por GB para disco están bajando, pero los costos de infraestructura se han disparado, forzando a las empresas a descartar enfoques tradicionales del almacenamiento de datos, a favor de enfoques nuevos.

Los proveedores de almacenamiento han seguido el ejemplo dado por los proveedores de virtualización de servidores, al crear un ecosistema de tecnologías que mejoran la utilización de la infraestructura de almacenamiento, tornándola más eficiente en el proceso. La prueba de la mayor adopción de tales tecnologías en el mercado es el hecho de que hoy casi todas las soluciones de almacenamiento vienen pre-construidas con una o más tecnologías de optimización listas para implementar. Mientras que medir el tamaño del mercado de optimización del almacenamiento puede ser complejo, la tasa de adopción de estas tecnologías indica que con el tiempo todo el almacenamiento instalado será optimizado de una u otra manera.

¿Por qué la optimización del almacenamiento ya no es óptima?

Si bien el determinante primario de la optimización del almacenamiento puede ser el crecimiento vertiginoso de los datos, la optimización del almacenamiento también está influida, en gran medida, por ineficiencias introducidas en la manera en que los datos se crean, acceden y/o almacenan. Algunas de estas ineficiencias tienen un elemento humano, pero otras son subproductos de la proliferación tecnológica.

- ☒ La cantidad de datos estáticos (datos a los que rara vez se accede en forma continua) sigue creciendo a un ritmo mucho más rápido que la cantidad de datos activos (datos a los que se accede todo el tiempo).
- ☒ Hay una duplicación creciente en los datos que se generan gracias a las tecnologías nuevas. Por ejemplo, la virtualización de servidores puede crear imágenes de sistemas operativos que son casi idénticas entre sí. Otro ejemplo son los usuarios que generan copias duplicadas de imágenes, archivos y otros datos.
- ☒ Algunos de los formatos utilizados para crear contenido estático son ineficientes (es decir, el formato crea mucho espacio vacío). El resultado es espacio superfluo en el disco.
- ☒ El almacenamiento de datos estructurados, como bases de datos, puede ocasionar ineficiencias en el almacenamiento debido a los elementos de metadatos que dan soporte a diversos atributos para estos datos.
- ☒ Las empresas pueden tener activos de almacenamiento disímiles (y a veces, multiproveedor) que generaron silos en la infraestructura. Debido a pautas de interoperabilidad o consideraciones del desempeño, a las empresas puede parecerles un gran desafío implementar tales activos en forma homogénea. El resultado es la gran subutilización de los activos.
- ☒ Las capas de almacenamiento parecen ser el plan ideal para racionalizar la infraestructura, pero en la práctica la falta de mecanismos de capas automatizadas hace que sea un desafío para los administradores del almacenamiento mover datos de una capa a la otra. El resultado es el uso ineficiente de capas de almacenamiento y una reducción de los beneficios de las capas.

Las Tecnologías de Optimización del Almacenamiento apuntan directamente a mejorar la utilización de los activos cambiando la ubicación y organización de los datos. Como capa adicional de tecnología, la optimización del almacenamiento al fin de cuentas está diseñada para superar las ineficiencias introducidas por humanos o por la misma tecnología:

- ☒ Capas automatizadas es la capacidad de mover solo porciones activas de datos a capas más altas (y más costosas) mientras que la parte menos activa de datos permanece en capas más bajas. El principio subyacente es que mientras que los datos crecen, un porcentaje cada vez mayor de estos datos son estáticos o fríos que no necesitan residir en capas de alto rendimiento en todo momento.

- ☒ Virtualización del almacenamiento es la capacidad de reunir en un pool recursos de almacenamiento diferentes y a menudo multiproveedor bajo una capa de gestión y presentación común.
- ☒ Aprovisionamiento ligero es la capacidad de definir una unidad de almacenamiento (sistema completo, pool de almacenamiento o volumen) con un tamaño de capacidad lógica mayor que la capacidad física asignada a esa unidad. El host o dispositivo que accede a esa unidad de almacenamiento ve la capacidad lógica y no la capacidad física.
- ☒ De-duplicación es la capacidad de examinar un único bloque de datos o archivo o una serie de bloques de datos o archivos para detectar patrones comunes y reemplazarlos, dirigiéndolos a una única instancia de ese patrón, reduciendo así la duplicación de esos patrones en el marco de almacenamiento. Debido a la naturaleza del acceso de bloques, la de-duplicación se ofrece principalmente en el almacenamiento basado en archivos.
- ☒ Compresión es la capacidad de comprimir los datos para que los bloques sean más pequeños. El uso de compresión permite que estos datos consuman mucho menos almacenamiento en comparación con otras tecnologías de optimización, tales como la de-duplicación o el aprovisionamiento ligero. Tal como lo demuestran los factores determinantes y las técnicas de optimización del almacenamiento que responden a esos factores, la optimización del almacenamiento no es un enfoque del tipo 'talla única', a diferencia de algunas otras tecnologías de almacenamiento, como la replicación, los clones in-array, etc.

La naturaleza diversa de los datos hace que algunas técnicas de optimización del almacenamiento sean más adecuadas para ciertos tipos de datos que otras. Sin embargo, al implementar Tecnologías de Optimización del Almacenamiento como un paquete en una infraestructura de almacenamiento compartida, las empresas pueden obtener importantes beneficios tangibles, sin comprometer la calidad del servicio.

Compresión en tiempo real

Antes de resaltar los beneficios de Real-time Compression, examinemos por qué la compresión a menudo se excluye de la lista a la hora de optimizar el almacenamiento. Muchos proveedores no dudarán en señalar que, a diferencia del aprovisionamiento ligero o la de-duplicación, la compresión puede tener resultados mixtos cuando no adversos. Esto se debe parcialmente a que el enfoque tradicional de la compresión fuerza a los proveedores a alcanzar un delicado equilibrio entre compresibilidad y la penalidad del rendimiento, lo cual da resultados subóptimos. Por ese motivo, en general se implementa para datos que se usan con poca frecuencia.

Desafíos en el uso de la compresión

El enfoque tradicional y comúnmente adoptado es comprimir los datos que no están activos. Este tipo de compresión, que también se conoce como compresión post-proceso, comienza una vez que los datos ya se escribieron en el disco. Es muy similar a varios utilitarios basados en un host que comprimen archivos y carpetas una vez que se crearon. Este método es inherentemente ineficiente porque sus algoritmos consumen importantes ciclos de cómputo, son intensivos en disco y requieren un espacio de "salto" adicional para almacenar los datos no comprimidos. En un host, en la mayoría de las circunstancias, uno puede sortear esta penalidad agregada, pero en un array de almacenamiento construido con una finalidad, un overhead agregado de

este tipo puede producir una degradación considerable en la calidad de servicio, que automáticamente se percibe en el servidor y las aplicaciones. Como resultado, los proveedores rápidamente restan importancia a la compresión en estos escenarios a menos que el conjunto de datos sea significativo o el Sistema de Almacenamiento como un todo esté subutilizado.

En la compresión tradicional, cuando las aplicaciones realizan múltiples actualizaciones, tales cambios se escriben en el disco de manera no comprimida. Posteriormente, debe programarse una operación separada que realiza un muestreo y una compresión en base a su ubicación física en un volumen, independientemente de su relación con otros bloques de datos que la aplicación pueda estar accediendo.

Los resultados producidos por motores de compresión tradicionales también presentan amplias variaciones. Esos motores de compresión ingieren una porción pre-programada fija de datos y producen una salida variable, según la compresibilidad de esos datos. Además, los índices de compresión se basan en tamaños de porciones: cuanto mayor la porción, mayor el índice de compresión y mayor el overhead de rendimiento. Los tamaños de porciones menores dan como resultado índices de compresión más bajos. El punto justo de equilibrio entre tamaños de porciones y overhead de rendimiento no es tan justo después de todo.

Un efecto colateral de la compresión es que causa fragmentación. Debido al post-proceso y a la condición variable de los motores de compresión tradicionales, la compresión “degenera” la continuidad de los datos con el tiempo. Los datos comprimidos se distribuyen en porciones a lo largo del volumen y requieren de una recolección de residuos frecuente. El impacto de este efecto es una disminución del rendimiento con el tiempo.

La compresión en tiempo real revitaliza el rol de la compresión en la optimización del almacenamiento

Hoy la mayoría de los proveedores ofrecen la compresión como un atributo licenciable opcional pero recomiendan fuertemente que los administradores lo habiliten solo en conjuntos de datos seleccionados. En la mayoría de las situaciones, la adopción de la compresión para datos primarios activos en la Infraestructura de Almacenamiento sigue siendo mayormente limitada. Como cada penalidad del rendimiento en el Sistema de Almacenamiento puede tener un efecto expansivo al resto de la pila de aplicaciones, la mayoría de las empresas elige dejar la compresión inhabilitada.

Un enfoque más novedoso, conocido como ‘Real-time Compression’ o compresión en tiempo real, puede traer de nuevo al centro de la escena el papel de la compresión en la optimización del almacenamiento primario. Real-time Compression hace que el uso de la compresión en configuraciones para propósitos generales sea factible sin penalidades operativas. Promete múltiplos de compresión de hasta cinco veces para datos primarios en línea. Más aún, al achicar los datos primarios, también comprime todas las copias derivadas de esos datos, tales como copias de seguridad, archivos, instantáneas y réplicas. IBM ha incorporado esta tecnología al Sistema de Almacenamiento Storwize V7000, eliminando la necesidad de appliances externos.

IDC proyecta que otros proveedores seguirán el ejemplo de IBM, alejándose de la compresión tradicional durante momentos de inactividad y desarrollando tecnologías de compresión similares pero competidoras. Esto validará y reforzará el rol de la compresión como componente central de la eficiencia. De hecho, conforme los procesadores de almacenamiento crezcan en velocidad y potencia, los proveedores de almacenamiento podrían descargar los ciclos de compresión / rehidratación, cargándolos a núcleos o procesadores dedicados en el controlador del almacenamiento, además de los ciclos de de-duplicación.

¿Qué es Real-time Compression y qué beneficios tiene?

A diferencia de las tecnologías tradicionales de compresión, que comprimen datos como una operación post-proceso, Real-time Compression funciona sobre los datos activos primarios mientras se está teniendo acceso a ellos. Esto expande el dominio de la compresión a un conjunto mucho más amplio de cargas de trabajo con resultados predecibles y mensurables. Más aún, esta compresión está “siempre activada”, lo cual significa que puede habilitarse para cargas de trabajo activas y no requiere períodos programados para post-procesamiento, a diferencia de sus antecesores.

El diferenciador primario es que en la tecnología Real-time Compression el motor de compresión procesa un flujo de datos variable sobre la base de los patrones de datos que realmente se escriben. La tecnología Real-time Compression aprovecha la “localidad temporal” y no la ubicación física: los datos a los que se acceden juntos se comprimen juntos, independientemente de su ubicación física. De modo que cuando las aplicaciones realizan actualizaciones relacionadas a distintas partes del volumen se manejan juntas en forma contigua. El método es similar a la operación del sistema real, pues aprovecha la estructura de los datos y la relación de los datos con otros datos. Esta conciencia de la carga de trabajo minimiza la cantidad de operaciones de compresión y/o descompresión. El resultado es menor cantidad de IOPS de disco y menos overhead en el controlador de almacenamiento. Como consecuencia, se elevan los índices de compresión y la eficiencia, sin sacrificar el rendimiento.

Implementación de IBM Storwize V7000

IBM adquirió la tecnología Real-time Compression cuando obtuvo Storwize en 2010. Inicialmente, IBM ofrecía esta tecnología a través de appliances en banda construidas para tal fin, para la compresión transparente de datos NAS primarios. IBM sigue vendiendo estos appliances actualmente, pero también ha dado el paso audaz de llevar esta tecnología a su plataforma Storwize V7000. La plataforma Storwize V7000 utilizará la misma tecnología Random Access Compression Engine (RACE) que los appliances.

IBM planifica introducir la compresión en el Storwize V7000 como componente licenciable opcional a través de una actualización de firmware. Así, los usuarios existentes de Storwize V7000 podrán seleccionar la compresión como atributo (preconfigurado) al crear nuevos volúmenes. Los administradores también tendrán la opción de convertir volúmenes existentes utilizando el espejado de volúmenes para comprimir en forma ligera y eliminar el espacio no utilizado en el proceso. IBM inicialmente planea dar soporte de hasta 200 volúmenes comprimidos pero puede aumentar este límite con releases posteriores. La interfaz gráfica de usuario del Storwize V7000 presentará información de rendimiento relacionada con la compresión y permitirá a los administradores administrar y monitorear la compresión desde una

única consola. La tecnología Real-time Compression mejora la funcionalidad sin crear overhead adicional.

¿Cómo se compara la compresión en Storwize V7000 con el enfoque tradicional?

Tal como ya se mencionó, los enfoques tradicionales de la compresión han recibido bastantes críticas debido a su rendimiento impredecible, overhead y engorrosa condición de post-proceso. Real-time Compression por el contrario, cambia la ecuación al crear una capa de optimización en línea y eficiente. El impacto inmediato de implementar Real-time Compression es que el Sistema de Almacenamiento se entrega con menos IOPS durante la rutina de compresión. Las cargas de trabajo activas, como bases de datos y sistemas de e-mail, a menudo realizan pequeñas actualizaciones a los datos existentes. El análisis de IBM sugiere una mejora significativa en comparación con los enfoques tradicionales para estas cargas de trabajo (ver la Tabla 1).

TABLA 1

Comparación de la compresión tradicional con IBM Real-time Compression

Compresión tradicional	IBM Real-time Compression
"Porción"1 MB	100 Byte actualizar
1 MB leer	0 MB leer
1 MB descomprimir	0 MB descomprimir
100 Byte actualizar	0 Byte actualizar
1 MB comprimir	100 Byte comprimir
1 MB escribir	< 100 Byte escribir
Total IO por operación de compresión: 2 MB	Total IO por operación de compresión: < 100 Bytes

Fuente: IBM, 2012

Storwize V7000 Real-time Compression logra índices de compresión similares a los de los appliances IBM Real-time Compression (ver la Tabla 2). Como Real-time Compression puede implementarse con una gama más amplia de datos que la compresión tradicional, los beneficios potenciales pueden ser mayores. Los índices de compresión predecibles facilitan a las empresas planificar los presupuestos en forma acorde.

TABLA 2

Índices de compresión observados por IBM en entornos de clientes

Aplicación	Compresión observada	
Base de datos	Hasta 80%	
Servidores Virtuales (VMware)	Linux virtual OS	Hasta 70%
	Windows virtual OS	Hasta 50%
Office	2003	Hasta 75%
	2007 o posterior	Hasta 20 %
CAD/CAM	Hasta 70%	

Fuente: IBM, 2012

Al habilitar la compresión en volúmenes nuevos o existentes en Storwize V7000, las empresas pueden extraer una mayor capacidad utilizable de sus inversiones existentes. Ello permite a muchas empresas detener o reducir el gasto en almacenamiento para las configuraciones más comunes.

Un beneficio adicional del uso de Real-time Compression con Storwize V7000 es que es posible comprimir incluso los volúmenes extraídos del almacenamiento virtualizado externo. En muchos casos, esto presenta el doble de la capacidad utilizable para inversiones modestas y menores costos de gastos operativos.

Desafíos y oportunidades para Storwize V7000

En los tiempos que corren, cuando los presupuestos de almacenamiento se restringen o reducen, las organizaciones de TI siguen buscando formas de almacenar con eficiencia los volúmenes cada vez mayores de datos, a costos cada vez menores. La tecnología Real-time Compression, si bien es atractiva, aún se encuentra en la fase de adopción inicial. Muchas organizaciones siguen considerando la compresión como una tecnología post-proceso intensiva en cómputo.

Uno de los principales desafíos que las organizaciones enfrentan actualmente cuando tratan de habilitar Real-Time Compression es el riesgo versus la recompensa. Las organizaciones sin duda deberían exigir una respuesta a preguntas como: ¿Cuánto estoy pagando por comprimir mis datos y qué ahorros tendré como resultado? ¿Cuáles son los riesgos de comprimir mis datos? ¿Cuáles son los impactos en el rendimiento? ¿Podré mantener los índices de compresión o se degenerarán con el tiempo?

Poder cuantificar ahorros será un factor importante a la hora de decidir si vale o no la pena adquirir el software. Las organizaciones pueden utilizar la herramienta Compresstimator de IBM para modelar los beneficios esperados de la compresión en entornos específicos. Utilizando los resultados y las tasas de crecimiento esperadas, las organizaciones podrían obtener ahorros potenciales implementando Storwize V7000 con compresión o habilitando la compresión en volúmenes existentes. IBM podría apalancar proactivamente esta herramienta para ofrecer tanto a los clientes existentes como potenciales una vista de los ahorros que se pueden obtener.

La tarea que IBM tiene por delante consiste en ayudar a las organizaciones a dejar su percepción de la compresión como característica inherentemente post-proceso. IBM puede seguir ganando conciencia de marca o mindshare entre sus clientes transmitiéndoles mejores prácticas y casos de uso para implementar Real-time Compression y elaborando documentos que describan el uso de la compresión junto con otras Tecnologías de Optimización del Almacenamiento, como el aprovisionamiento ligero, las capas de almacenamiento y la de-duplicación.

CONCLUSIÓN Y GUÍA ESENCIAL

Las Tecnologías de Optimización del Almacenamiento han llegado para quedarse. Gracias a Real-time Compression, la compresión tiene un lugar en la cartera de optimización del almacenamiento. IBM ha dado un paso en la dirección correcta ofreciendo esta tecnología en uno de sus productos de almacenamiento más emblemáticos.

La naturaleza predecible, confiable y linealmente escalable de Real-time Compression alimentará su adopción en entornos con cargas de trabajo diversas. Esto a su vez alentará a las empresas a salir de su zona de confort y a hacer que Real-time Compression sea una de las tecnologías “imprescindibles” en sus entornos. La capacidad de cuantificar, predecir y medir los ahorros derivados de las Tecnologías de Optimización del Almacenamiento como Real-time Compression equipará mejor a las empresas para enfrentar el crecimiento explosivo de los datos.

Aviso de Copyright

Publicación externa de información y datos de IDC: toda la información de IDC que se va a utilizar en publicidades, comunicados de prensa o materiales promocionales requiere la aprobación previa por escrito del Vicepresidente o Gerente de País correspondiente de IDC. Dicha solicitud debe estar acompañada de un borrador del documento propuesto. IDC se reserva el derecho a negar su aprobación del uso externo por cualquier motivo.

Copyright© 2012 IDC. La reproducción sin permiso expreso queda terminantemente prohibida.