

VERITAS™

ÍNDICE GENÓMICO DE DATOS

Un informe realizado por los expertos en datos no estructurados sobre la verdadera composición de los entornos de almacenamiento.

2016



Detalles del informe

El Proyecto Genómico de Datos (Data Genomics Project) es una iniciativa diseñada para cambiar la manera en la que concebimos el manejo de la información. Veritas creó esta iniciativa con el propósito de entender mejor la verdadera naturaleza de los datos no estructurados que cotidianamente estamos creando, almacenando y manejando. Para esto, Veritas reunió una comunidad de científicos, expertos y líderes innovadores del sector de la información. Nuestra primera contribución al proyecto fue este informe de referencia preliminar sobre la composición real de los entornos de almacenamiento – el Índice Genómico de Datos (Data Genomics Index).

Al trabajar con el 86% de las empresas de la Fortune 500, realizando backups, archivando y analizando exabytes, Veritas cuenta con una posición única para analizar las características principales del entorno de una organización. Actualmente, la característica en la que nos enfocamos es en los metadatos. Haciendo uso del acumulado de estos metadatos del cliente - mediante nuestros productos para el análisis de archivos - Veritas es capaz de identificar los elementos exactos de los que está compuesto un entorno real.

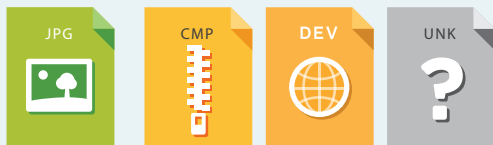
El entorno de datos actual



41%

del entorno total no ha sido modificado en los últimos

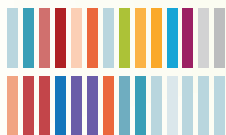
3 años



Los principales tipos de archivos son más extraños de lo previsto.



Los tipos de archivo office tradicionales están sobrecargando excesivamente su entorno.

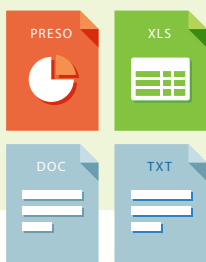


Datos Orfãos



Los datos huérfanos son desproporcionadamente pesados y obsoletos.

O arquivamento de



Reduz o custo em

50%

Las organizaciones deberían priorizar su estrategia de supresión y almacenamiento a fin de obtener ganancias excepcionales en términos de espacio de almacenamiento.



El primer Índice Genómico de Datos de Veritas

En 2015, y con el fin de entender mejor cómo estaba compuesto el entorno de datos de muchos de nuestros clientes, Veritas analizó decenas de miles de millones de archivos - y sus propiedades - provenientes del entorno de datos no estructurados. Más de 8.000 de las extensiones de archivo más comunes fueron consideradas en el análisis. Generalmente, esta información es una muestra representativa de todo el entorno del sistema de archivos de un cliente determinado.

Un crecimiento del

39%

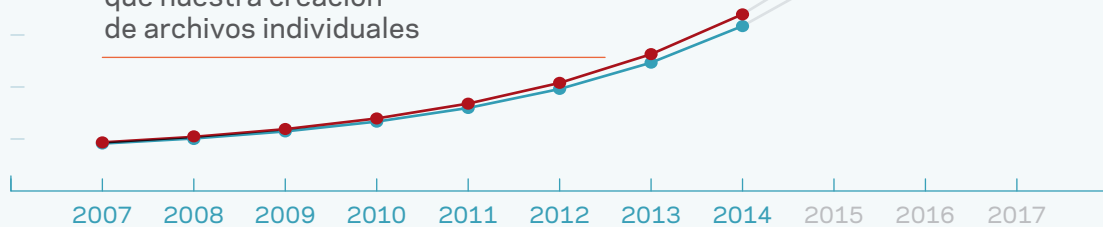
promedio
de los datos
anual



Una velocidad de almacenamiento

9% más rápida

que nuestra creación
de archivos individuales



Los datos se disparan

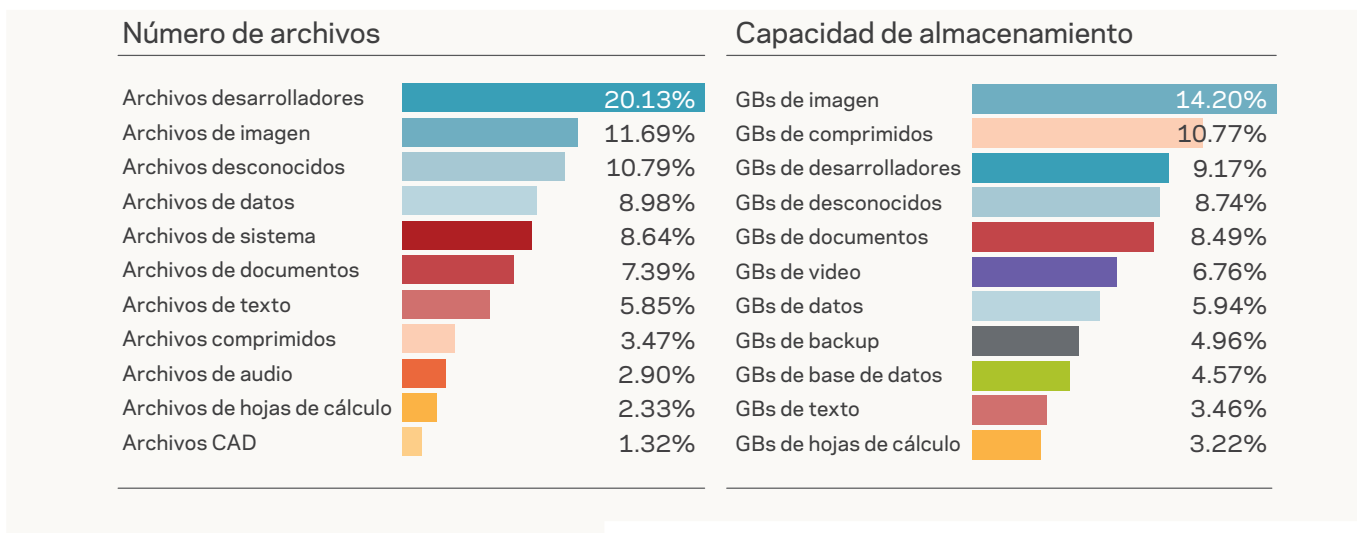
Durante los últimos siete años, la verdadera velocidad a la que creció la información al nivel de archivo fue del 39.2481189%, año tras año.

Este requisito de capacidad de almacenamiento está creciendo 9% veces más rápido que nuestra creación de archivos individuales. Aunque un cambio de comportamiento sin duda ayudaría a frenar una parte de este crecimiento, el problema real está en el manejo mismo del almacenamiento.

Contener la capacidad del almacenamiento no es únicamente un problema de almacenamiento. El entorno de almacenamiento está saturado, con el promedio de los PB de información conteniendo 2,312,000,000 archivos.

Los peores delincuentes: las imágenes y los archivos desarrolladores

Revisando el total de los datos, podemos observar importantes diferencias entre los tipos de archivos que representan el mayor número vs. los que ocupan el mayor espacio. Usted puede concebir esto como la saturación vs. los costos de almacenamiento.

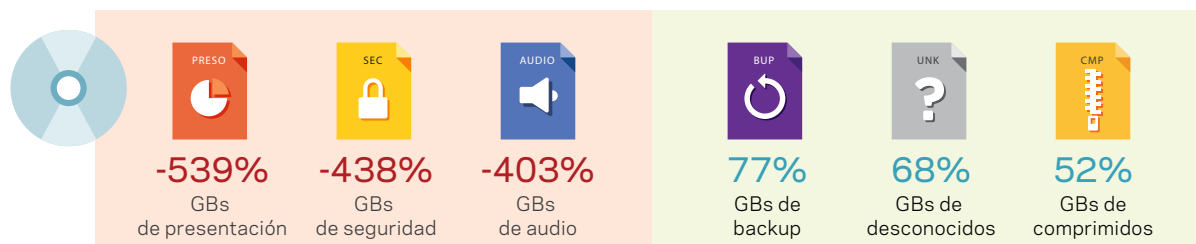


Los mayores motores y accionadores de datos

Esta composición ha ido cambiando con el tiempo. En los últimos diez años, los mayores motores de los datos relacionados con otro tipo de archivos son:



Los mayores accionadores en términos de capacidad de almacenamiento son:

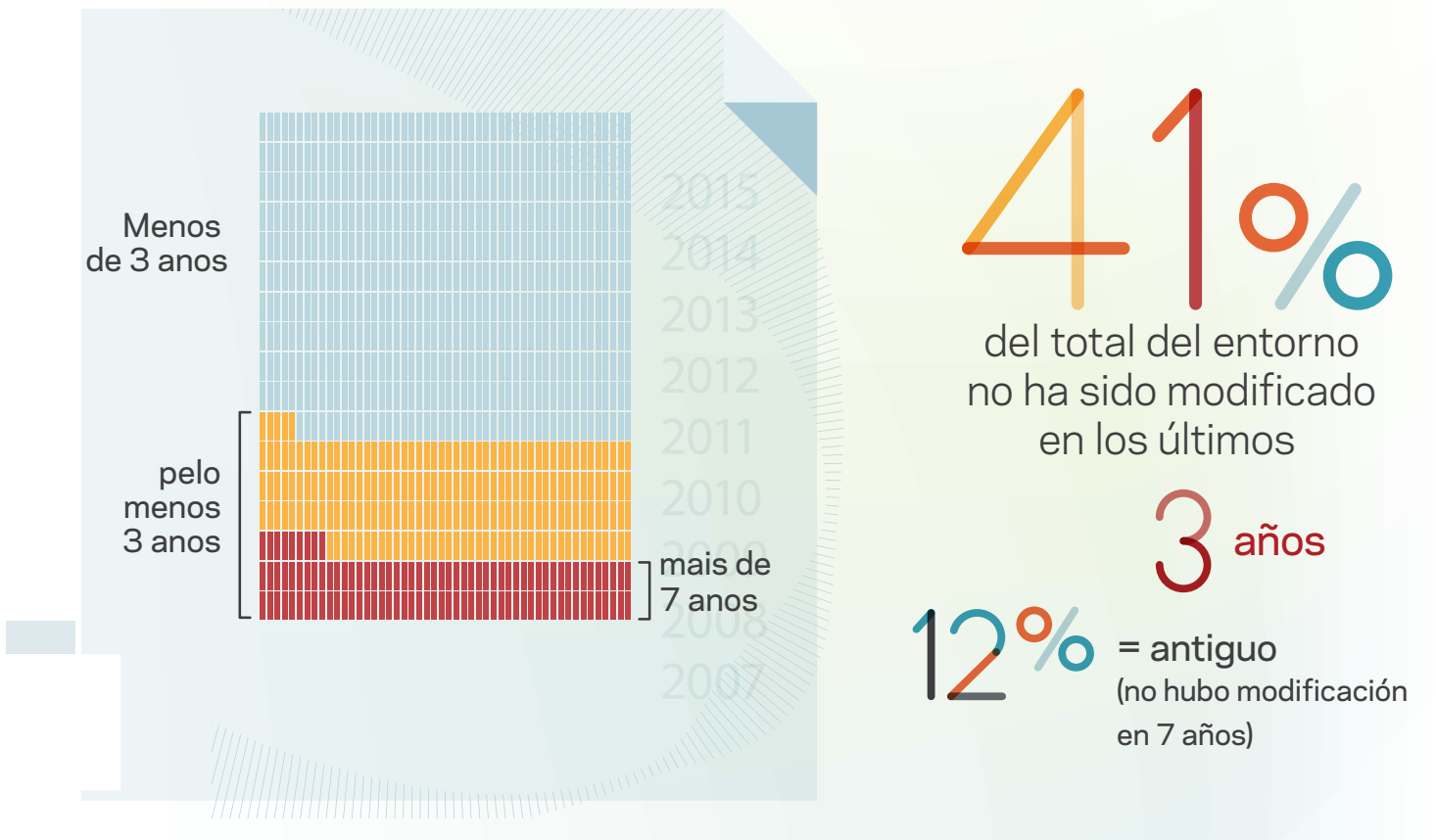




El crecimiento de los datos. Siempre en temporada

El otoño es la temporada en la que se crea la mayor cantidad de archivos. Creamos un 91% más de archivos de texto durante el otoño que en cualquier otra estación, así como 48% más de hojas de cálculo y 89% más de sistemas de archivo de información geográfica.

Los backup y los documentos son los únicos dos tipos de archivo que tienden a incrementarse del otoño al invierno. Con todas las operaciones de backup durante el año, los backup se ensanchan, en términos de tamaño, en un 756%. El 68% de los videos de todo un año es producido durante el verano y el otoño. Las imágenes, por su parte, presentan una caída durante el invierno, con una disminución del 63%. Aunque la tendencia de los emails (pst) también es predecible, su desviación estándar entre las estaciones es de tan solo 0.7%.

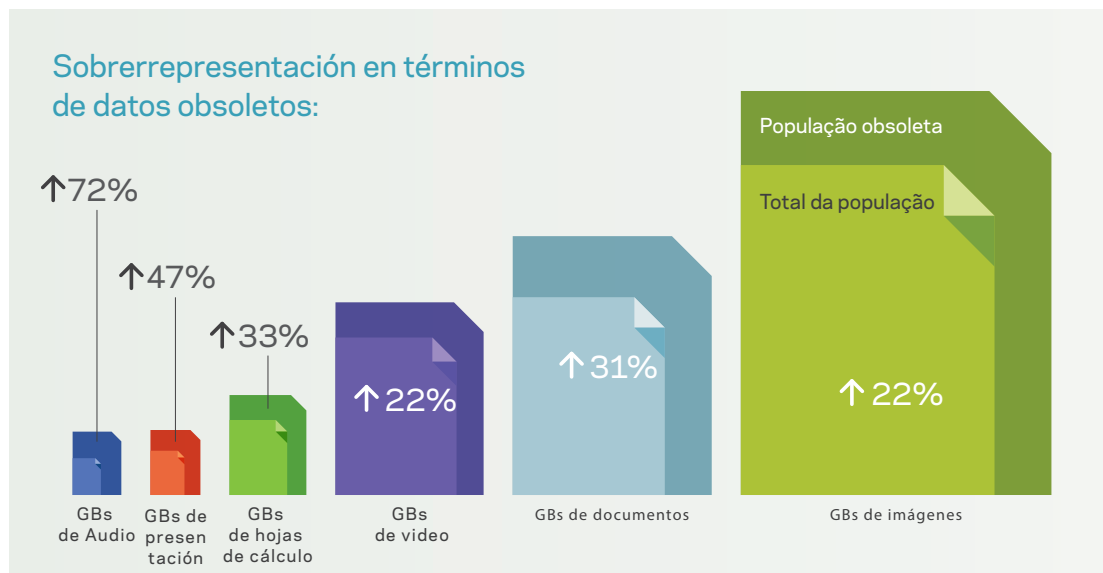


¿Por qué continúa guardando aquello?

En los negocios de hoy en día, la información lo es todo. No obstante, ésta viene siendo creada a una tasa tan abrumadora que la utilidad de cada elemento individual es más bien efímera.

El remedio: tipos de archivos en el blanco

Al momento de enfrentarse a una cantidad abrumadora de información obsoleta, así como de tener que tomar posibles decisiones para enmendar dicha situación, siempre es útil darle mayor prioridad a aquellos lugares donde usted puede jugar sus “mejores cartas” de gestión de la información. Revisando cuáles tipos de archivo están sobrerrepresentados, al establecer una proporción de los datos obsoletos frente a los datos totales, los resultados demuestran que los archivos “office” tradicionales representan una gran carga:



Si usted quiere saber en qué lugar la corrección de archivos individuales se traduce en una ganancia mayor en cuanto a espacio de almacenamiento se refiere, enfóquese en los siguientes cinco formatos, los cuales le brindarán el mayor retorno de GB por archivo:

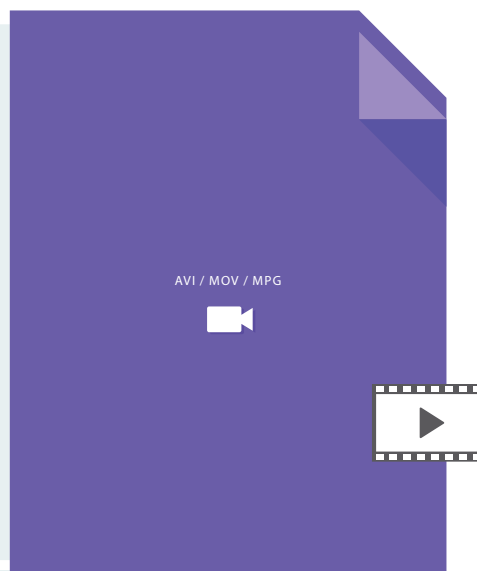
1. Tipos de archivo de Máquina Virtual
2. Tipos de archivo de seguridad
3. Tipos de archivo de juego
4. Tipos de archivo científico
5. Tipos de archivo de sistemas de información geográfica



Los videos utilizan

15.8x

veces más del total de la capacidad de almacenamiento obsoleto (de lo que representan sobre el total del número de archivos obsoletos)



7.3x



6.4x



2.2x



Archivos desmesurados

Si usted quiere priorizar unos tipos de archivo específicos, fíjese en qué lugar tanto el número de archivos, como su respectivo tamaño en el disco, se encuentran desmesurados. Por ejemplo, los videos ocupan 15.8 veces más del total de la capacidad de almacenamiento obsoleto de lo que representan sobre el total del número de archivos. Por su parte, las Máquinas Virtuales ocupan 7.3 veces más espacio, las presentaciones 6.4 veces más y los emails 2.2 veces más, erigiéndose así como las opciones más relevantes en cuanto a la priorización de tipos de archivo se refiere.



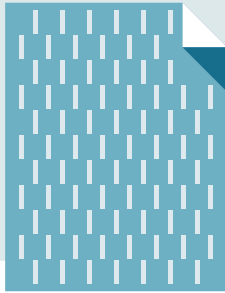
Ellos limpiaron sus escritorios. Eso fue todo.

Un dato huérfano es un dato que no tiene un dueño asociado. Con la renovación de personal, el cambio de funciones y el caos general del Directorio Activo, es fácil imaginar la dificultad que tiene el hacerle seguimiento al historial del entorno. Esto le genera costos a las organizaciones.

Una forma en la que los costos se manifiestan es en el hecho de que los datos huérfanos estén ocupando una parte desproporcionada de la capacidad de almacenamiento. Aun cuando los datos huérfanos representan únicamente un 1.6% del total de archivos, éstos ocupan el 5.1% del total de la capacidad de almacenamiento. Los datos huérfanos también están siendo desmesuradamente canalizados hacia los tipos de archivo con abundante información. Así, las imágenes ocupan un 88% más del espacio normal y los videos y las presentaciones un 165% y 229%, respectivamente.

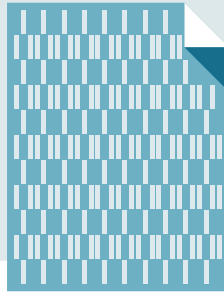
Igualmente, podemos ver que las tendencias de contratación tienen un posible impacto en los entornos de almacenamiento. Los archivos huérfanos son 222% más grandes que un archivo promedio. Los gerentes han llegado a creer que, cuanto más grande el archivo, más importante es su contenido. Por consiguiente, al momento de ocurrir un cambio de personal, éstos han tendido a conservar únicamente los elementos más densos. Si usted quiere recuperar espacio de almacenamiento, el concentrarse en el contenido sin dueño puede llegar a ser un buen comienzo.

Tamaño promedio de los archivos



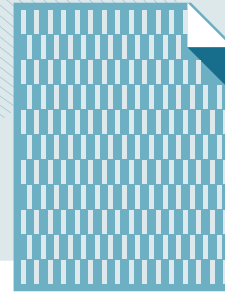
0.24 MB

en la última década



0.40 MB

5 años atrás



0.53 MB

modificados
en el último año

La densidad es un sutil indicador de la utilidad

No es algo extraño el que hoy en día estemos creando contenidos más densos. Sin embargo, puede sorprenderle el hecho de que, en los últimos 7 años, ha habido un crecimiento relativamente más lento del 10.3%. El tamaño promedio de un archivo:

Utilizado en la última década o más: 0.24MB

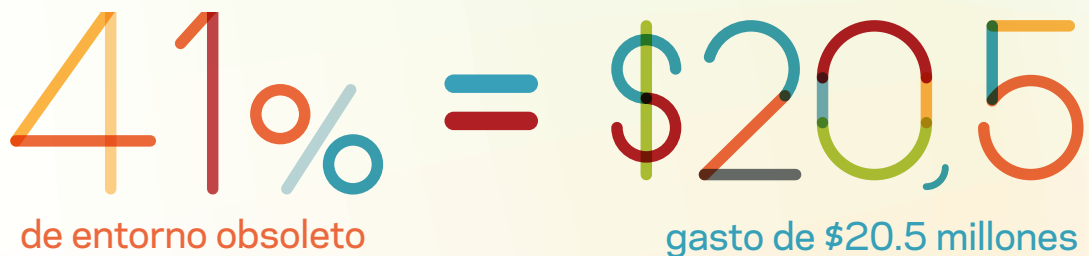
Utilizado en los últimos 5 años: 0.40MB

Modificado en el último año: 0.53MB

Los archivos que son clasificados como obsoletos son **33% más pequeños** que los archivos que han sido modificados en el último año.

Bueno, ¿y ahora qué hacemos?

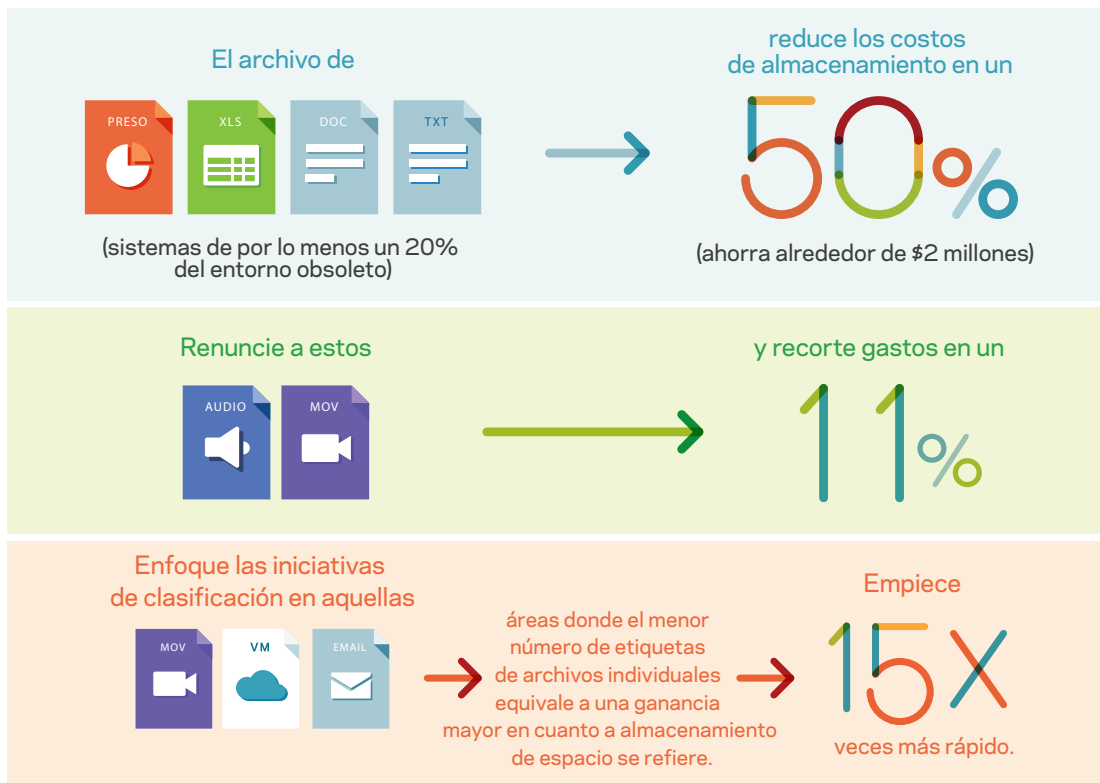
Si su entorno de almacenamiento se parece a los entornos aquí analizados, entonces usted tiene una gran oportunidad. Imagine el entorno promedio de 10PB...



Si el 41% del entorno es obsoleto, usted podría estar gastando hasta \$20.5 millones anuales por administrar una información que no ha sido utilizada en tres años. Limpiar el entorno no es un trabajo fácil, pues ese 4.1PB equivale a 9.479.200.000 archivos individuales que hay que clasificar, borrar o archivar.



Usted tiene que priorizar...



Los archivos ricos en contenido, tales como las presentaciones, las hojas de cálculo, los documentos y los archivos de texto representan el 20% del entorno obsoleto promedio. Estos mismos archivos pueden ser un excelente blanco de los proyectos para guardar sistemas de archivo, los cuales pueden reducir los costos de almacenamiento en un 50% o más -una ganancia de más de \$2 millones. Sólo los videos y el audio pueden implicar una ganancia del 11%. Las imágenes dominan el 18% del espacio de almacenamiento en las categorías antiguas, de archivos no utilizados en los últimos 7 años o más.

El enfocar los proyectos de clasificación en áreas donde el menor número de etiquetas de archivo individual equivale a una mayor ganancia en cuanto a espacio de almacenamiento se refiere (como por ejemplo, los videos, archivos VM y emails), puede ayudarlo a comenzar más rápido - de hecho, 15 veces más rápido. Reflexionar sobre cómo administrar la información que ha sido abandonada tras la salida de un empleado, o tras un cambio de funciones, puede conseguirle un 5%, o inclusive un hermoso millón...

Independientemente de su situación, con conocimientos como estos, hay muchas oportunidades de luchar contra la curva de crecimiento desmesurada y de recuperar su entorno.

VERITAS™

Veritas Technologies permite que las organizaciones aprovechen el poder de su información, promoviendo el éxito del negocio aun en los entornos más complejos. Trabajamos con organizaciones de todos los tamaños, incluyendo el 86% de las compañías de la lista Fortune 500 global. De hecho, por más de una década, Veritas ha sido reconocida como una empresa líder en el Gartner Magic Quadrant, tanto a nivel del Enterprise Backup Software y Integrated Appliances¹ como en el Enterprise Information Archiving².

De la mano con nuestra comunidad de aliados con experiencia, ayudamos a que nuestros clientes mejoren la disponibilidad de su información y obtengan conocimientos que los hagan más competitivos. Con más de 7,800 empleados en 58 países alrededor del mundo, Veritas es una compañía de \$2.5 mil millones que mantiene alianzas con algunos de los líderes más importantes del sector de tecnología. Entre estos aliados se encuentran Amazon, Cisco, Fujitsu, Google, Hitachi, HP, IBM, Microsoft, NetApp, OpenStack, Symantec y muchos otros.

Desde centros de datos tradicionales hasta nubes híbridas, públicas y privadas, Veritas - junto a su comunidad de aliados - ayuda a las organizaciones a proteger, identificar y administrar la información independientemente del entorno que posean, utilizando soluciones inteligentes de gestión de la información. Con Veritas, las compañías obtienen el conocimiento y la disponibilidad que necesitan para comprender con qué información cuentan, cómo hay que hacer para mantenerla protegida y cuál debe ser eliminada.

Para saber cómo podemos hacer lo mismo por usted, visítenos en veritas.com

Legal:

1 Fuente: Gartner, Inc., Magic Quadrant for Enterprise Backup Software e Integrated Appliances, Dave Russell, Pushan Rinnen, Robert Rhame, Junio 15 de 2015

2 Fuente: Gartner, Inc., Magic Quadrant for Enterprise Information Archiving, Alan Dayley, Garth Landers, Anthony Kros, Octubre 29 de 2015. Nota: entre 2005-2009 la investigación se llamó Magic Quadrant for Email Active Archiving.

Gartner no promociona a ningún proveedor, producto o servicio descrito en sus publicaciones de investigación, y no recomienda a los usuarios de tecnología que seleccionen exclusivamente a los proveedores con las calificaciones más altas u otras distinciones. Las publicaciones de investigación de Gartner se basan en las opiniones del departamento de investigación de Gartner, y no se deben interpretar como declaraciones de hecho. Gartner excluye toda garantía, expresa o implícita, relativa a esta investigación, incluyendo todas las garantías de comercialización o idoneidad para un propósito determinado.

© 2016 Veritas Technologies LLC. Todos los derechos reservados. Veritas y el logotipo de Veritas son marcas comerciales o marcas comerciales registradas de Veritas Technologies LLC o sus afiliadas en E.E. U.U y otros países.

