Russ Holden: Los Mejoramientos de la Base de Datos Domino de la Versión 5.0

Entrevistado por: Betsy Kosheff
(Traducido por: Anita L. Paci)

This article resides in "Iris Today", the technical Webzine located on the http://www.notes.net Web site produced by Iris Associates.

[Nota del Editor: Iris ha estado trabajando por mas de dos años en perfeccionar la base de datos mejorada que se va a debutar con la nueva versión 5.0 del Servidor Domino. Para el líder del equipo Russ Holden, se trata de la empresa, haciendo Domino cada vez mejor en fiabilidad, administración, escalabilidad, y rendimiento, rendimiento, rendimiento.]

En la versión 5.0 del servidor Domino cuales son las áreas de mejoramiento de mayor importancia para la base de datos ?
En general, hemos logrado unos mejoramientos muy significativos en la escalabilidad y rendimiento para los clientes de correo de las empresas. También existen muchas funciones nuevas tales como el archivo de control de transacciones (transactional logging), la compactación en línea, y el alto rendimiento de hacer las copias de seguridad y restaurarlas en línea. Y, los clientes verán mejoras significativas en la actualización y reconstrucción de las vistas -- de 3 a 5 veces mas rápido para las actualizaciones y de 3 a 10 veces mas rápido para la reconstrucción.

Que es lo que ha empujado su enfoque con respecto al rendimiento de la base de datos de correo?
La gente quiere aprovechar sus inversiones de añadir mas memoria o otro procesador. Anteriormente, no podíamos aprovechar bien los multiprocesadores, por lo tanto hemos trabajado mucho en eso porque los clientes están tratando de consolidar sus servidores y bajar el costo global de la administración de su correo.

Existen otras funciones nuevas para la mensajería?
Además de las mejoras fundamentales de la base de datos, hemos hecho mucho trabajo sobre el rendimiento del router, especialmente donde utiliza la base de datos. Muchas de esas mejoras son gratis, pero también hemos añadido algunas opciones para las empresas grandes. Una de ellas es la capacidad de tener múltiples buzones de correo, lo cual reduce los conflictos entre los servidores de correo donde hay tráfico excesivo.

Como funciona?
La forma como trabaja Domino hoy en día es que tenemos una base de datos llamado mail.box en el cual se deposita todo el correo. Esto puede llegar a ser un cuello de botella, especialmente si tienen muchos usuarios utilizando el correo o si quiere compactar la base de datos. Ahora, se puede tener cualquier numero de bases de datos mail.box y se puede distribuir la carga entre tantos que uno desea. También se pueden bajar

individualmente. Esta funcionalidad realmente nos ha permitido superar la escalabilidad y explotar los benchmarks.

Que es lo que el archivo de control de transacciones le brinda al usuario final?
Es una técnica estándar en la industria y realmente la ultima tecnología para la fiabilidad del almacenaje de los datos. Un archivo de control de transacciones provee información secuencial registrada por todos quienes escriben en él -- la escritura secuencial en disco es mucho mas rápido que escribir en varios sitios distintos en el disco.

Cuando se realiza una operación sobre la base de datos, esa operación se termina o completamente con éxito o sin éxito, pero nada entre medio. Además, cuando ha terminado de ejecutarse, y el usuario recibe un respuesta, tiene que ser garantizado que la operación fue registrada en disco, no solamente en memoria. El archivo de control de transacciones es lo que registra las operaciones realizadas. Se pudiera simplemente efectuar la grabación directamente en disco mediante la base de datos, pero eso termina siendo muy costoso porque hay muchas cosas que tienen que ser actualizadas.

Cuales son otros beneficios que se obtienen con el archivo de control de transacciones?
Le permite al usuario arrancar el sistema sin ejecutar el DbFixup en sus bases de datos. Esto es una ganancia importante en el rendimiento.

Cuando cae el servidor hoy en día, nosotros ejecutamos el Fixup en todas las bases de datos que se estaban modificando en el momento que cayó el sistema. En la versión 5 de Domino, utilizamos los archivos de control para volver a ejecutar todos los cambios que no fueron realizados a las bases de datos, en vez de ejecutar el Fixup en cada una de ellas.

El resultado de esto es que el arranque es dramáticamente mas rápido porque no tienes que ejecutar el proceso de fixup en las bases de datos grandes, lo cual puede ser muy lento. Simplemente, repetimos el proceso utilizando el archivo de control para actualizar todo.

Entonces, arrancar el servidor va a ser mas rápido?
Definitivamente -- es mas, el arranque del servidor con el archivo de control de transacciones habilitado debe ser con una rapidez de una magnitud sumamente grande -- arbitrariamente mas rápido.

El otro beneficio grande que obtenemos con el archivo de control de transacciones, es el mejoramiento con las copias de seguridad, porque ahora tenemos un conjunto completo de interfaces del programador de aplicaciones (API) para respaldar a las bases de datos. La versión 4 de Domino no tiene un API de respaldo y restauración, aunque existen varias herramientas para esto, hechas por otras compañías. En vez de hacer un producto que hace las copias de seguridad -- que ya existen muchos productos que lo hace muy bien -- vamos a proveer un conjunto de APIs muy poderosos que permiten al vendedor integrar sus productos. Esto le va a dar al cliente los programas en línea de alto rendimiento de respaldo y restauración que necesitan para Domino y Notes.

Como esto cambiaría el proceso de realizar las copias de seguridad?
Inicialmente, los administradores harían sus copias completas de seguridad. Luego, utilizarían el archivo de control para registrar los cambios. Entonces, por ejemplo, una vez a la semana puedes registrar la base de datos completa y luego solamente necesitarías utilizar el archivo de control. Esto provee un mecanismo muy útil para asegurar que no se ha perdido nada, sin tener que copiar todos los gigabytes de las bases de datos todas las noches.

Van a trabajar con algunos productos específicos de administración de sistemas?
Estamos trabajando con aproximadamente siete empresas que ya tienen productos de respaldo para Notes. Ya estamos entregándoles los interfaces que necesitan para entrar al servidor, para que el servidor trabaje mejor. Nuestra fecha de lanzamiento de la versión 5.0 va a coincidir con el lanzamiento de tantos de estos productos que sea posible. Cada uno se conformará con nuestro API.

Han modificado el límite de tamaño de la base de datos?
Sí. La nueva arquitectura de la base de datos incluye un cambio con respecto a como se guarda los apuntadores dentro de ella, y esto nos permitirá proveer bases de datos individuales mayor de un terabyte, en el futuro muy cercano. Para la versión 5, vamos a probar y certificar bases de datos hasta por lo menos 16 gigabytes, versus el límite de 4 GB en la versión 4. Ese tamaño aumentará muy rápidamente.

Que hay del nuevo utilitario de compactación en línea?
El propósito de compactación es para reclamar espacio en las bases de datos -- por ejemplo, los documentos que han sido eliminados. La manera en que trabaja Domino es que en vez de achicar al archivo en el momento que se elimina algo, el servidor espera un rato e intenta utilizar ese espacio mas tarde. Por consiguiente, algunas veces existe una acumulación de espacio libre. Entonces, para poder reclamarlo se requiere la compactación. En la versión 4, la compactación requería una nueva copia de la base de datos. Si tenías una base de datos de cuatro gigabytes, necesitabas otros cuatro gigabytes de espacio libre en ese mismo dispositivo. Eso era muy lento. En la versión 5, una vez que tienes la base de datos en el nuevo formato, la compactación se ejecuta sobre la misma. Nosotros hacemos todo el trabajo de intercambiar todo dentro de los límites del archivo. Así que no necesitas todo ese espacio demás. Además, es sumamente mas rápido.

Sin embargo, debes estar consciente que cuando se ejecuta la compactación por primera vez, tienes que copiar las bases de datos, pero una vez que hayas hecho esto, obtendrás todo la nueva funcionalidad y rendimiento a cambio de esa inversión.

Puedo realizar la compactación mientras los usuarios están accesando las bases de datos?
Sí. Ya no se requiere acceso exclusivo a las bases de datos para poder compactarlas. Anteriormente, nadie mas podía estar leyendo o actualizando un base de datos. Pero con la versión 5.0, los usuarios pueden continuar sus operaciones mientras el proceso de compactación se esta ejecutando. Los administradores no tienen que bajar al servidor o buscar la forma de tener acceso exclusivo a las bases de datos.

También existe una nueva estructura On-Disk Structure (ODS)?
Sí, y va a traer unos mejoramientos de rendimiento muy grandes para los usuarios que actualizan. De la misma manera, todos los formatos anteriores seguirán trabajando con la versión 5.0 de Domino.

Cuales fueron las consideraciones atrás del ODS?
Iterativamente hemos estado cambiando el formato de la base de datos para optimizar el I/O y los CPUs de correo y de otras cargas de aplicación. El idea es que cuando por ejemplo, se añade correo, tenemos que escribir menos cosas en disco. Entonces cuando escribimos en disco, escribimos en un numero menor de sitios y realizamos mucho mas escritura secuencial. Lo que obtenemos con esto es aproximadamente la mitad de las operaciones de I/O que con las bases de datos de la versión 4. Y, en algunos casos es mucho menos que eso aun. El nuevo ODS es lo que permite también la compactación sobre la misma base de datos y la recuperación por el archivo de control de transacciones.

Existen otras características nuevas de la base de datos?
Existen varias características nuevas que hacen que ciertas funciones sean opcionales, particularmente aquellas que pueden consumir muchos recursos del sistema.

Por ejemplo, ahora tenemos la capacidad de crear bases de datos que no mantienen las marcas de documentos no leídos (unread marks). Es importante tener la capacidad de hacer seguimiento a los documentos que no han sido leídos, pero para algunas bases de datos, tal como el Registro de nombres y direcciones (NAB) o los archivos de control (log), típicamente no se requiere estas marcas. Entonces, ahora existe un suiche para desactivar estas marcas. Esto también nos proporciona un mejoramiento adicional de rendimiento.

También existe la capacidad de desactivar el blanqueo de los datos borrados (overwrite), si la seguridad física ya es adecuada. La versión 4.x del servidor permitía que todos los datos borrados podían ser sustituidos en disco, lo cual tiene un costo de rendimiento. Ahora va a ser opcional.

Cuantos usuarios puede soportar un servidor?
La mayoría de los clientes hoy en día tienen entre 100 a 1,000 usuarios activos en un solo servidor. Pero tenemos muchos sitios grandes que soportan números mucho mas grandes -- hay un caso que es de 10,000 usuarios en un solo servidor. Con la versión 5, nuestro análisis actual indica que muchos clientes entrarán en el rango de 500 a 5,000 usuarios por servidor, aunque esperamos poder manejar números mucho mayores todavía, probablemente en el rango de 10,000.

Por supuesto, las medidas verdaderas de escalabilidad no son solamente el numero de usuarios por servidor, sino las necesidades totales de infraestructura de una organización grande. Esto significa directorios múltiples, administración distribuida, autentificacion/codificación (encryption) flexible, y las tecnologías de replicación/encaminamiento (routing) que se mantienen antibalas ante cualquier limitación de topología de red. Estos son cosas de igual importancia que la personas deben tomar en cuenta cuando hacen comparaciones.

Que mas mejoraron en la versión 5.0?
Hemos estado trabajando y haciendo mejoras a la base de datos por unos 27 meses, entonces hemos tenido mucho tiempo y tenemos confianza de que por el trabajo que hemos realizado vamos por delante de la competencia en varios áreas. En general, hemos reconstruido la estructura interna de la base de datos, con lo cual se ha logrado mejoras mayores tanto en la utilización de CPU y I/O, como en la integridad de la base de datos.

Cuales son algunas de las áreas especificas donde va a ser mejor?
Hemos ganado algunos mejoramientos por la nueva base de datos, mientras otros han sido funciones inherentes del servidor.

Específicamente, nuestros archivos de correo son fáciles de manejar como entidades individuales, lo cual es particularmente útil para los administradores cuando tienen que mudarlos y en la recuperación de ellos. Exchange requiere que uno restaura el contenedor de correo (mailstore) completo para poder extraer manualmente los datos de los archivos de correo individuales.

También, nuestra base de datos fue construida para datos auto-descriptivos (self-describing) y semi-estructurados, (semi-structured), entonces no tenemos algunas de las limitaciones de las bases de datos relacionales que tiene Exchange, tal como un directorio que tiene un límite de diez campos personalizados. Nuestros niveles de control de acceso sobre la base de datos y los mecanismos de replicación también son muy superiores, igual como es el poder de las funciones de agrupación de servidores (clustering) y los servidores compartimentados (partitioned servers). Exchange tiene solamente redireccionamiento (n) (función de) redireccionar de dos nodes (two node failover), mientras nosotros lo tenemos de dos a seis nodos, además de reparto de la carga, lo cual trabaja en múltiples plataformas, no solamente WindowsNT.

Finalmente, hemos casado la flexibilidad y poder de la base de datos con los beneficios de los archivos de control tradicionales de las base de datos y mecanismos de recuperación. El trabajo de la funcionalidad de los archivos de control y la recuperación ha sido realizado en el Almaden Research Center de IBM por un equipo de desarrolladores e investigadores altamente destacados. Esto es el sitio en donde las bases de datos relacionales fueron inventados y en donde nuevas tecnologías de mayor importancia para el data warehousing, data mining, multimedia y aplicaciones de la web se están perfeccionando. Entonces, estamos beneficiandonos de la experiencia de IBM y de muchas empresas del mundo real.

Existirá una diferencia entre las distintas plataformas con respecto a la escalabilidad de los servidores de la versión 5?
Ahora que hemos eliminado en gran parte los cuellos de botella en el mismo Domino, y se escala dramáticamente mejor en los multiprocesadores, estamos trabajando para explotar esos cambios para que el poder de algunos de los sistemas de grande escala, tal como Unix, AS400 y OS390, provee mejor valor por el nivel de precio y rendimiento, que los sistema NT.

Que deben saber las personas sobre la actualización de las bases de datos al nuevo formato?

El proceso de actualización es el mismo de siempre, lo cual es que uno compacta la base de datos al nuevo formato. Siempre permitimos que el cliente mantenga su versión vieja de las bases de datos. No tienen que compactarlas, las pueden dejar en el formato viejo y nosotros lo soportamos. Pero existen muchas razones funcionales por las cuales van a querer actualizarlas -- por mejor escalabilidad, por el nuevo archivo de control de transacciones, y, por supuesto, el mejor rendimiento.

Biografía:

Russ Holden está en el Grupo del Proyecto de Base de Datos en Iris Associates. Desde el lanzamiento de la versión 4.0, Russ ha estado trabajando en el rediseño del sistema de base de datos de Notes, el rendimiento en general, y los mejoramientos de escalabilidad. Actualmente hay aproximadamente doce personas trabajando solamente en el sistema de manejador de base de datos. Russ ha estado con Iris por casi cuatro años. Anteriormente, trabajó por siete años en la empresa Digital con los Sistemas de Manejadores de Bases de Datos Distribuidos (los proyectos de RdbStar y DB Integrator). Antes de eso estaba con la compañía CCA (Computer Corporation of America) trabajando también con el Sistema de Bases de Datos Distribuidos (Adaplex). Sus pasatiempos son, trotar, andar en bicicleta y raqueta (racquetball). Aparte del trabajo, Russ ocupa sus ratos libres casi completamente con su mejor amigo (por lo general!) -- su hijo de cuatro años Brandon.

Copyright 1998 Iris Associates, Inc. all rights reserved.