

Технология оптимизации данных — IBM Real Time Compression

Обзор функциональных особенностей технологии IBM RACE (Random Access Compression Engine), позволяющей в сочетании с использованием аппаратного ускорителя компрессии, основанного на технологии Intel quick assist, достигать гарантированной компрессии данных.



Дмитрий Агеев — High-End Storage Brand Manager, IBM East Europe/Asia Ltd.

Технологии и стратегии повышения эффективности датацентров

Разговорами о том, что объем корпоративных данных постоянно растет, а бюджеты остаются прежними или увеличиваются, но незначительно, уже никого не удивишь. Основными драйверами хранения данных сегодня являются аналитика (большие данные), соответствие законодательным нормам. При этом такие технологии и стратегии защиты данных как репликация и создание локальных копий (снимки и клоны томов), как минимум, удваивают их количество.

Данные опроса — что является основной причиной закупок систем хранения данных, — дают следующий результат: 25% заказчиков заявили, что это потребность в новой емкости и невозможность расширения установленных систем (рис. 1). В топе остальных факторов — это обновление технологической составляющей, производительность и консолидация инфраструктуры.

В серверной части неэффективность использования вычислительных мощностей привела к их виртуализации. Огромное количество физических серверов размещают в новые ультра компактные решения, которые стали занимать меньше места и потреблять меньше ресурсов.

Следующие шаги для повышения эффективности хранения данных за счет снижения занимаемого объема в датацентрах были сделаны производителями систем хранения, постепенно представляющими новые технологии (каждая технология оптимизации требует ресурсы от СХД):

- thin provisioning — виртуальное выделение дискового пространства;
- дедупликация данных — удаление повторяющихся объектов и компрессия данных — уменьшение объема хранимого объекта.

Основным назначением технологии thin provisioning является слежение за выде-

ленным, но неиспользуемым дисковым пространством. Дополнительный функционал этой технологии — более быстрый возврат неиспользуемого пространства после удаления путем обнуления блоков. Несмотря на то, что это требует дополнительной процессорной мощности (overhead, в среднем, составляет до 7%), это достаточно эффективно, так как не изменяет сами данные. Но есть и нюансы: использовать эту технологию нужно аккуратно, так как не все файловые системы и приложения годятся для этого из-за их особенностей использования дискового пространства. Правильно используемый и хорошо контролируемый процесс приводит к лучшей утилизации пространства и снижению издержек на администрирование большого количества томов.

Дедупликация потребляет несколько больше процессорной емкости, но требует достаточно большого количества оперативной памяти (кэш-памяти дискового массива). На каждый терабайт данных может потребоваться до 16 Гбайт памяти. Это нужно для хранения данных и для операций над базой данных о хранимых объектах для их сравнения. Основная проблема в том, что, как только база не помещается в кэш-памяти, ее приходится хранить и частями загружать и выгружать для операций сравнения. Это пагубным образом влияет как на производительность системы в целом, так и на приложения. Поэтому, как правило, данный функционал рекомендуется использовать или реализовать как пост-процесс. Это ведет к сокращению эффективности техники снижения объема, так как физически приходится иметь, как минимум, два объема для хранения сравниваемой информации. Онлайн-процесс дедупликации, конечно же, реализован, но в основной массе — для устройств резервного копирования, таких, как виртуальные ленточные библиотеки, где вся процессорная мощь и ресурсы направлены на единственную задачу — последовательную запись. Для этой задачи подобная технология подходит лучше, чем для транзакционных нагрузок, где операции, в основном, произвольные и производятся небольшими блоками.

У компрессии данных требования к процессорным ресурсам такие же, как и у дедупликации, так как в их основе лежат одинаковые алгоритмы сжатия/уплотнения (например, LZ). Но компрессия не требовательна к оперативной памяти

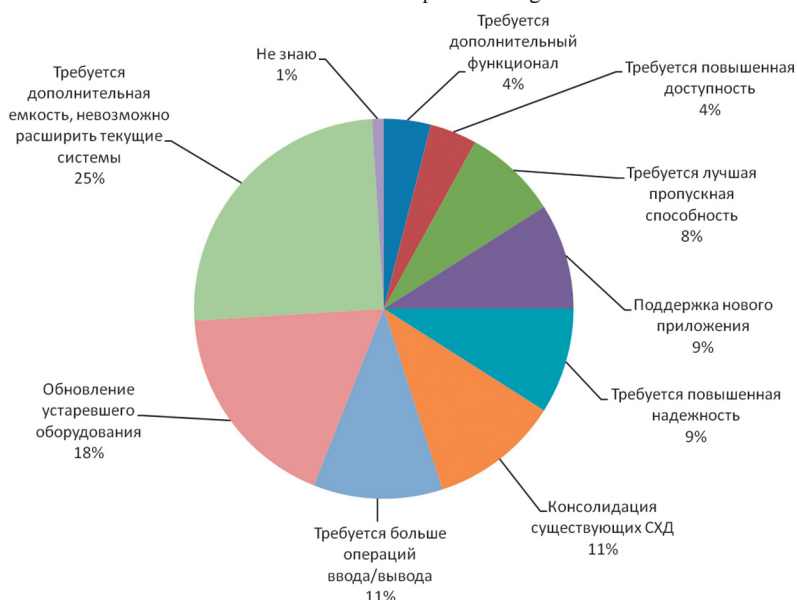


Рис. 1. Распределение ответов заказчиков на вопрос об основной причине закупок систем хранения данных (ист. Enterprise Strategy Group 2013 г.)

Табл. 1. Возможная степень сжатия в зависимости от типа данных.

Базы данных		До 80%
Серверная виртуализация	Linux virtual OSes	До 70%
	Windows virtual OSes	До 55%
Документы	Office 2003	До 75%
	Office 2007-2013	До 25%
CAD/CAM Проектирование и дизайн		До 75%

(кэш-памяти дискового массива), так как работает над уплотнением одного объекта. Эта техника является крайне эффективной. Но с ней тоже не все так просто: очень важна сама технология компрессии, которая влияет на ее возможность для использования в онлайн-режиме. Большинство производителей предлагает для компрессии пост-процесс на манер дедупликации, что значительно снижает эффект от ее использования.

Важным фактором применения технологий оптимизации данных является понимание уровня эффективности их применения. В табл. 1 приведена информация о возможной степени сжатия – в зависимости от типов данных. Не стоит применять технологии оптимизации к данным, к которым они уже были применены приложениями, а также и к зашифрованным данным. Это, как правило, различные видео и изображения, различные форматы архивов – zip, rar, tgz и т.д. Это приводит к трате мощности оборудования впустую.

IBM Random Access Compression Engine

В дисковых системах от IBM используется собственная технология компрессии данных в реальном времени – RACE (Random Access Compression Engine), в которой содержатся 35 патентов IBM. Это единственная технология на рынке дисковых массивов, которая позволяет сжимать данные на лету, и они попадают на носитель уже в сжатом виде, позволяя приобретать меньше начального физического объема.

Каковы же главные отличия технологии IBM от остальных, представленных на рынке?

Для начала нужно понимать, как работает стандартный алгоритм компрессии (рис. 2). Рассмотрим это на примере текстового файла: используется техника “сдвигающегося окна” (обычно это

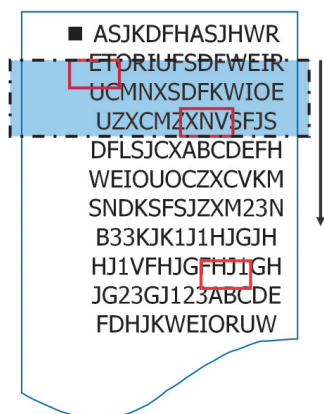


Рис. 2. В стандартном алгоритме компрессии используется техника “сдвигающегося окна” (обычно это 32 Кбайт).

32 КБ). На повторения непопадающие в окно ссылки не создаются. Размер блока это компромисс, больше окно – лучше компрессия, но потребление процессорной мощности и памяти больше, ну и, соответственно, меньше окно – хуже компрессия и лучше производительность.

Подобным образом это работает в СХД. Данные при записи разделяются на фиксированные куски, каждый из которых сжимается и разжимается независимо друг от друга. Основная проблема состоит в том, что последующие блоки опираются на предыдущие, и чтобы прочитать данные в определенном месте, нужно сделать декомпрессию предыдущих, а чтобы обновить данные нужно сделать компрессию всех последующих, что является неэффективным для больших файлов и блочного доступа.

Технология работы алгоритма в RACE построена противоположным образом. Данные из потока сжимаются в фиксированные, равные блоки, и карта ссылки на блоки хранится вместе с ними. В отличие от стандартного метода компрессии, каждый блок читается и обновляется независимо от другого. При стандартном методе, если имеется кусок данных в 32КБ и нужно обновить 4КБ, то придется: прочитать 32КБ, распаковать 32КБ, обновить 4КБ, запаковать 32КБ и записать обновленные 32КБ. В RACE это будут просто две операции – компрессия 4КБ и запись новых 4КБ данных, что намного эффективней и положительным образом сказывается на производительности. Движок RACE также применяет и использует дополнительное – временное измерение при операциях компрессии. Несколько операций записи объединяются в единый блок сжатия с общим индексом. Когда приложение записывает и обновляет данные, которые гомогенны и в них больше повторений и, как следствие, выше уровень компрессии. Этот аспект положительно влияет на упреждающее чтение, так как больше взаимосвязанных данных попадает в кэш.

Данная технология интегрирована как функционал в дисковый массив StorwizeV7000 и в виртуализатор систем хранения IBM SAN Volume Controller и не требует никаких изменений в имеющейся ИТ-инфраструктуре. Для клиентов, у которых уже есть данное оборудование, этот функционал доступен начиная с версии 6.4 как для блочного, так и для файлового (StorwizeV7000) доступа. Обычные тома можно конвертировать в сжатые. Для новых клиентов это тоже будет прозрачно, так как оба решения имеют функцию виртуализации и могут использовать существующие СХД как свои дисковые полки – данные можно будет сжать непосредственно на них или мигрировать на новые диски также в сжатый вид. В сочетании с технологией IBM Easy Tier, которая тоже присутствует в этих решениях, можно добиться значительной экономии такого дорогого ресурса как Flash-память, или на SSD-дисках массивов, или на специализированной All Flash СХД, как IBM Flash system. EasyTier – это автоматическое многоуровневое хранение данных, которое будет держать оптимальный процент данных на Flash-памяти, к которым очень часто обращаются. В дополнение к этому эти данные будут зани-

мать еще меньше места, что повысит эффективность системы хранения во всех аспектах и поможет размещать больше задач на ней. Кроме того, при транзакционной нагрузке в ряде случаев улучшается время отклика для сжатых данных в силу того, что приходится читать меньший объем, и в кэш дискового массива попадает больше уникальных блоков (рис. 3).

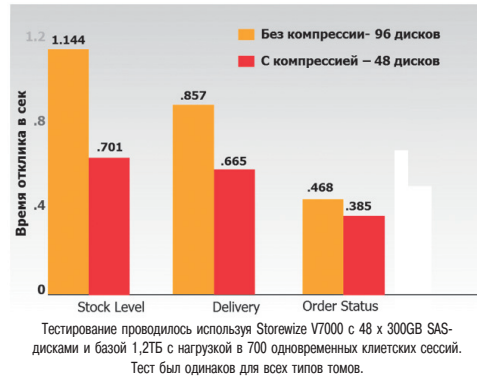


Рис. 3. Результаты тестирования совместного использования технологий RACE и IBM Easy Tier.

Новое поколение массивов IBM StorwizeV7000 G2 и SVC G7

В мае 2014 года компания IBM представила новое поколение дискового массива IBM StorwizeV7000 G2 и SVC G7. Во-первых, это новое “железо”. Новые мощные 8-ядерные процессоры и значительно увеличенный объем кэш-памяти. Во-вторых, был добавлен аппаратный ускоритель компрессии, основанный на технологии Intel quick assist. Он дает двойной прирост в производительности при работе со сжатыми томами. Таких плат можно установить до 2 штук. В Storwize V7000 один всегда установлен в базе. На программном уровне был изменен сам базовый алгоритм сжатия на LZ4, что дало рост производительности на операциях записи в 3 раза и на 15% при чтении. Это повышает количество параллельных сессий VMotion и позволяет создавать больше снимков на сжатых томах.

Как понять, повлияет ли технология компрессии на инфраструктуру и насколько сжимаются ваши данные. Для клиентов IBM, у которых уже есть выше упомянутое оборудование, любой функционал можно включить и опробовать когда угодно без установки каких-либо лицензий и на основании этого принять решения о полезности технологии для своего ИТ-окружения. Другой способ для тех, у кого пока нет СХД от IBM, – это воспользоваться бесплатной утилитой Comprestimator, которая ставится на сервер (это может быть Windows, Linux, AIX и VMWare ESX). Данная утилита не влияет на текущую производительность и на основе алгоритма RACE на ваших реальных данных покажет, как технологии компрессии и thin provisioning экономят ваши ресурсы.

В заключение хотелось сказать о том, что IBM гарантирует результат. Подтверждение этому – уже действующая программа, смысл которой в том, что если клиент не достигает заявленных значений, то корпорация вернет недостающий объем дисками.

Дмитрий Агеев,
IBM в России и СНГ.