

2. Устав (Конституция) Всемирной организации здравоохранения. 1946 // apps.who.int constitution-ru.pdf
3. Чарлтон Э. Основные принципы обучения здоровому образу жизни // Вопросы психологии №2, 1997. — С. 3-14
4. Социальная модель благополучия [Электронный ресурс] = Social welfare model / Л. И. Иванкина, О. Г. Берестнева // Современные проблемы науки и образования : научный журнал. — 2015. — № 1.
5. Казначеев В.П. Здоровье нации – феномен экологии XXI века / В.П. Казначеев Материалы I Всероссийского форума «III тысячелетие. Пути к здоровью нации». Москва, 15-17 мая 2001. – 194 с.
6. Андреева О.Н. Количественные показатели уровня и качества жизни населения Приморского края // Качество и уровень жизни населения современной России: состояние, тенденции и перспективы. Сборник материалов Международной научно-практической конференции - ОАО ВЦУЖ, ИСЭПН РАН. М.: ООО "М-Студио", 2012. С. 9-18.
7. Андреева О.Н. Уровень и качество жизни: содержание понятий и их составляющие // Ойкумена. Регионоведческие исследования. 2012. № 4. С. 6877.
8. Бобков В.Н., Мстиславский П.С. Качество жизни: сущность и показатели // Человек и труд. - 1996. - №6, — С. 76-79.
9. Финогенко И.А., Дьякович М.П., Блохин А.А. Методология оценивания качества жизни, связанного со здоровьем // Вестник Тамбовского университета. Серия: естественные и технические науки. № 1. 2016. С 121–130.
10. Блохин А.А. Информационная система для поддержки мультицентровых исследований качества жизни // Вестник Иркутского государственного технического университета. 2016. № 8 (115). С. 56–64. DOI: 10.21285/1814–3520–2016–8–56–64.
11. Массель Л. В., Блохин А. А. Метод когнитивного моделирования индикаторов качества жизни с учетом внешних факторов // Наука и образование. Научное издание МГТУ им. Баумана. №4. 2016. С. 65–754.
12. Блохин А. А., Массель Л. В. Когнитивное моделирование индикаторов качества жизни: предлагаемый подход и пример использования // Вестник НГУ. Серия: Информационные технологии. Т.14. №2. 2016. С. 72–79.

ЭФФЕКТИВНОЕ ИСПОЛЬЗОВАНИЕ ДАННЫХ: ТЕХНОЛОГИЯ ДЕДУПЛИКАЦИИ

А.А. Протасова

*(г. Москва, Московский финансово-промышленный университет «Синергия»,
AProtasova@synergy.ru)*

EFFECTIVE USE OF DATA DEDUPLICATION TECHNOLOGY

A.A. Protasova

*(Moscow financial and industrial university "Sinergiya")
e-mail: AProtasova@synergy.ru*

Abstract. Today the deduplication experiences a rebirth, becoming an integral part of the technologies applied in processing of modern data arrays. However existence of this function not always provides to customers notable advantages, its efficiency in many respects is defined by features of data. Administrators should remember that possibilities of a deduplication aren't boundless. To estimate expediency of her use for your concrete needs, it is worth studying nuances of process attentively. The deduplication is surrounded by a set of myths which dethronement will help to avoid a set of mistakes during the work with your SSD.

Keywords: Storage system of Data (SSD), Hard Drive, Deduplication.

Функция дедупликации доступна для массивов, содержащих и продуктивные данные, и резервные копии. В зависимости от типа данных коэффициенты их сокращения могут значительно отличаться. Ошибочно полагать, что показатель, достигнутый в рамках обработки архивного массива, можно применить и к продуктивному хранилищу.

Правильный расчет коэффициента позволит не ошибиться при определении нужного размера хранилища, поэтому следует проводить предварительные тестирования перед покупкой. Такой подход поможет избежать лишних трат и рассчитать оптимальную стоимость покупки, удовлетворяющую всем конкретным нуждам.

Определение правильного коэффициента дедупликации – важнейший этап организации мультизадачных конфигураций массивов. Анализ основных мифов поможет избежать досадных просчетов и осознанно подойти к покупке хранилища оптимального объема.

Разберем основные заблуждения, связанные с функцией дедупликации в массивах. Чем больше коэффициент дедупликации, тем больше вы получаете преимуществ. При сравнении предложений разных вендоров обращайте внимание не только на величину коэффициента сокращения данных, но и совокупную стоимость владения. Это поможет оценить реальную экономию объема в условиях сокращения требований к ресурсам. На практике коэффициент 50:1 помогает на 98% уменьшить объем данных, что лишь на 8% больше показателя коэффициента 10:1.

При оценке преимуществ, которые дает тот или иной коэффициент дедупликации следует опираться на закон убывающей доходности. Чем выше коэффициент – тем меньше прирост реальных преимуществ. Это объясняется неизменным объемом одного фактора (объема данных) в условиях дополнительного увеличения затрат другого фактора. Это утверждение легко продемонстрировать на примере одного офиса. Увеличивая число работников, но не меняя площадь помещения, вы с каждым новым сотрудником ухудшаете условия труда, снижая прирост эффективности бизнес-процессов.

Механизм сжатия, лежащий в основе дедупликации, заключается в удалении повторяющихся последовательностей данных. Процесс осуществляется на нескольких уровнях – файла, блока, контента или приложения. С целью повышения эффективности сжатия большинство решений сочетают дедупликацию и компрессию. Некоторые вендоры отождествляют эти термины, другие же выделяют особую технологию, предлагая продукты с функцией «уплотнения» (дедупликация+сжатие). Единого определения дедупликации не существует, но ее роль в процессах обработки массивов данных четко определена. Она помогает заказчикам экономично использовать пространство систем хранения и резервного копирования.

Линейка продуктов от НРЕ отличается функциональностью. Для СХД в массивах 3PAR создан программный комплекс под названием НРЕ 3PAR Thin Technologies, представленный следующими механизмами.

- Thin Deduplication. Решение, позволяющее выполнить дедупликацию продуктивного массива в режиме онлайн без ущерба для показателей производительности.
- Thin Provisioning. Технология для систем хранения 3PAR с функцией виртуализации дискового пространства, использующая внутреннюю карту хранимых блоков. Администратор получает к освобожденным ресурсам мгновенный доступ, так как массиву не требуется проводить ревизию.
- Thin Persistence и Thin Copy Reclamation. Технологии, позволяющие массиву 3PAR при освобождении физических ресурсов переводить блоки в соответствующий пул.
- Thin Conversion. Решение, облегчающее процесс трансформации томов со старых массивов данных различных вендоров в «тонкие» тома для экономии ресурсов целевого хранилища.

Технологии можно использовать бесплатно без ограничений по функциональности и времени для всех СХД 3PAR.

Коэффициенты дедупликации продуктивных массивов и массивов с резервными копиями одинаковы

Алгоритмы дедупликации разных типов массивов отличаются. Некоторые задачи требуют большего числа ресурсов, поэтому коэффициенты сокращения данных сильно варьируются. Наибольшее влияние на этот показатель оказывает объем повторяющихся данных, который в каждом конкретном случае уникален. По этой причине системы резервного копирования данных, содержащие большой объем повторяющихся данных, имеют наиболее высокий коэффициент дедупликации по сравнению с оперативными массивами.

В продуктах от HPE для оперативных массивов ZPAR используется особый сверхбыстрый алгоритм поиска повторяющихся данных с использованием микросхемы ASIC в каждом контроллере. Такой подход разгружает процессоры, делая их доступными для других важных задач, что обеспечивает непрерывность бизнес-процессов.

Такое мнение ошибочно – все данные разные, что и определяет вариативность коэффициента дедупликации. На его значение оказывают влияние следующие факторы.

- Тип данных. Данные, прошедшие программное сжатие, зашифрованные, мета- и потоковые данные отличаются минимальным коэффициентом дедупликации.
- Срок хранения. Чем дольше хранится массив, тем больше в нем повторяющихся фрагментов данных, что обеспечивает наиболее высокий коэффициент дедупликации.
- Степень изменяемости данных. Чем чаще в течение дня изменяются и обновляются данные массивов, тем меньше в нем повторяющихся фрагментов и ниже коэффициент дедупликации.
- Механизм резервного копирования. Массивы с полными копиями сжимаются легче, чем массивы с дифференцированными бэкапами. Коэффициент дедупликации на резервном массиве всегда выше, чем на основном.

Смешивая данные разного типа в общем секторе хранения, администраторы пытаются расширить набор уникальных данных, чтобы повысить количество повторяющихся фрагментов. Однако для несвязных данных такой подход не работает, ведь повторяющиеся фрагменты заключены в разных форматах. В этом случае пул становится слишком массивным, что лишь затрудняет процесс поиска повторяющихся данных. Лучшее решение для повышения производительности процесса дедупликации – разделение данных на сектора по типу.

Ожидаемый коэффициент дедупликации можно получить уже после первого резервного копирования

Это ошибочное мнение формируется по результатам сравнения коэффициентов на оперативном массиве и системах резервного копирования. Если вы храните всего одну копию данных, то сможете получить показатель дедупликации, незначительно превышающий единицу. По мере роста количества резервных копий похожих данных вы сможете наблюдать увеличение показателя коэффициента дедупликации.

При первом резервном копировании вы сможете достигнуть определенного показателя дедупликации, но по мере роста числа копий в пуле коэффициент сокращения данных будет увеличиваться. Значительный его рост наблюдается уже после появления второй копии, ведь первые 2 бэкапа очень похожи на блочном уровне.

Искусственно увеличить уровень дедупликации в маркетинговых целях предельно просто – достаточно хранить большое количество копий одних и тех же данных.

Заказчик преследует иные цели и фиктивно высокий показатель дедупликации его не устроит, ему нужна эффективно работающая схема резервного копирования. Сравнение политик показывает, что для достижения наиболее высокого коэффициента дедупликации необходимо использовать механизм создания полных дневных копий. Однако следует понимать, что изменения в политике резервного копирования в значительной мере повлияют на фактический объем хранимых данных.

Заранее спрогнозировать коэффициент дедупликации невозможно. На практике это действительно непростая задача, однако решить ее помогут наборы утилит от различных

производителей. Наборы программ для резервного копирования и основных систем, которые помогают получить подробную информацию о типе данных в массиве, сроке их хранения и используемой политике резервного копирования. На основании этих данных можно получить весьма четкие представления об ожидаемом коэффициенте сжатия данных. Для более точного представления прогноза производители также используют данные, полученные от заказчиков при работе в похожей среде и условиях. Несмотря на объективные основания для прогноза, гарантировать, что на практике вы получите эквивалентный показатель дедупликации, вендоры не могут.

Снижение стоимости хранения данных на каждом Гб Flash-накопителей и стремительный рост объемов SSD снижает актуальность использования технологий дедупликации в условиях оперативных хранилищ. Однако для систем резервного копирования эта функция становится все более актуальной. Компания HPE разработала утилиты, помогающие рассчитать необходимый объем хранилища, что поможет рационально использовать ресурсы бизнеса.

- Утилита, позволяющая оценить текущую утилизацию оперативного хранилища данных и оценить преимущества перехода на ZPAR.
- Программы для оценки утилизации систем резервного копирования и построения прогноза роста объема данных. Программа бесплатна, достаточно разрешить отправку данных о состоянии массива в службу технической поддержки HPE.

Существует иное видение рационального использования ресурсов в будущем, отличное от метода дедупликации – удаление копий одних и тех же данных. Такой подход потребует кардинальных изменений существующей инфраструктуры с последующим внедрением технических мер и новых правил для приложений с целью минимизации снижения их производительности и защищенности. Только время покажет, какое направление одержит победу в существующей на сегодняшний день дискуссии об эффективном использовании данных.

ЛИТЕРАТУРА

1. LAP Lambert Academic Publishing «Методы дедупликации данных» (2017).
2. Сергей Симонов «Семь тысяч двести», «Компьютерра», №39 (2016).
3. Сергей Леонов «Винчестер будущего», «Компьютерра», (2015).
4. Протасова А.А., Бурханова Ю.М. «BIG DATA», Сборник трудов III-Международного конгресса по проблемам развития "умной" страны SMART RUSSIA 2016

ПРИМЕНЕНИЕ МЕТОДОЛОГИИ ИЕРАРХИЧЕСКИХ ГИБРИДНЫХ КОГНИТИВНЫХ КАРТ ПРИ РАЗРАБОТКЕ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ДЛЯ НАУЧНО-МЕДИЦИНСКИХ ЦЕНТРОВ

С. В. Романчуков, С. В. Аксенов, О.Г. Берестнева

*(Томск, Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский Томский политехнический университет»)
e-mail: inoy@vtomske.ru, axoenows@tpu.ru, ogb2004@mail.ru*

HIERARCHICAL HYBRID COGNITIVE CARDS METHODOLOGY IN THE DECISION SUPPORT SYSTEM DEVELOPMENT FOR SCIENTIFIC AND MEDICAL CENTERS

*S. Romanchukov, S. Aksenov, O. Berestneva
(Tomsk, Tomsk Polytechnic University)*

Abstract. In this article, the authors of the ongoing study appeal to the problem of IT products development and implementation for improving the effectiveness of health institutions and organizations, namely