

Белая книга

Экономическая ценность систем Data Domain и Integrated Data Protection Appliance (IDPA)

Подтверждение экономичности устройств Dell EMC для резервного копирования

Авторы: Винни Чоински, старший аналитик ESG Lab,
и Кристоф Бертран, старший аналитик
Июнь 2018 г.

Эта белая книга ESG составлена по заказу Dell EMC
и распространяется в соответствии с лицензией от ESG.

Содержание

Содержание	2
Введение	3
Преимущества архитектуры Data Domain и IDPA	4
Экономические преимущества систем Data Domain и IDPA	6
Эффективность дедупликации имеет большое значение	7
Скорость и масштабируемость	9
Факторы сокращения расходов	10
Выводы	11

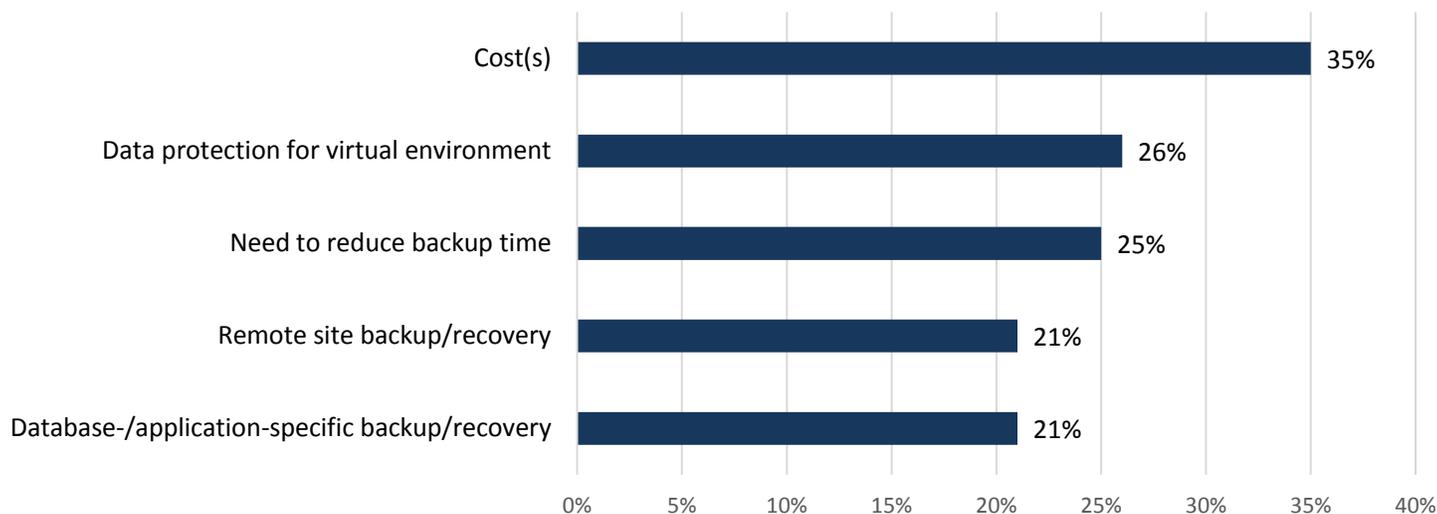
Введение

Чтобы опережать конкурентов по темпам развития и инновациям, организации должны постоянно поддерживать инфраструктуру на современном уровне. Перед ИТ-службами постоянно ставятся задачи оптимизировать инфраструктуры для новых бизнес-инициатив и вспомогательных приложений, при этом удерживая на неизменном уровне или даже сокращая затраты. Большинство респондентов, принявших участие в текущем исследовании ESG, называют необходимость в сокращении затрат одним из главных факторов, влияющих на расходы их компании на ИТ. Отвечая на вопрос о том, как их организации планируют справиться с ростом расходов в 2017 году, 27% респондентов заявили, что они приобретут новые ИТ-решения с лучшей окупаемостью инвестиций.¹

В частности, 35% ИТ-менеджеров, отвечающих за внедрение процессов и технологий для защиты данных, обычно называют рост расходов одной из основных сложностей (см. рис. 1). Кроме того, серьезные сложности у таких специалистов часто вызывают проблемы, связанные с рабочими нагрузками, особенно в плане защиты виртуализированных сред и удаленных офисов. Однако анализ общих оценок сложностей, связанных с защитой данных, открывает более широкую картину. Сокращение расходов и виртуализация — это главные проблемы, но следующие три сложности, о которых сообщают респонденты, относятся к *операционным* и *функциональным*. Это сложности, связанные с производительностью, резервным копированием и восстановлением в распределенной архитектуре, а также с защитой баз данных и приложений.²

Рисунок 1. Современные сложности, связанные с процессами и технологиями защиты данных

Which of the following would you characterize as challenges with your organization's current data protection processes and technologies? (Percent of respondents, N=387)



Источник: Enterprise Strategy Group

Эти сложности также соответствуют основным требованиям ИТ-руководителей, связанным с повышением скорости работы, оперативности и надежности, а также с контролем расходов. Все эти проблемы неизменно фигурировали в предыдущих отчетах ESG о модернизации защиты данных.

¹ Источник: отчет об исследовании ESG «2017 IT Spending Intentions Survey» ([Планируемые затраты на ИТ в 2017 г.](#)), март 2017 г.

² Источник: исследование ESG «Data Protection Modernization Trends» ([Тенденции в модернизации защиты данных](#)), декабрь 2016 г.

Анализ ESG подтверждает, что несоответствие сложностей, с которыми сталкивались ответственные за внедрение, и требований ИТ-руководителей часто приводило к весьма скромным масштабам модернизации и трансформации. Ответственные за внедрение не могли выполнять требования руководителей из-за технических проблем в средах защиты данных.³

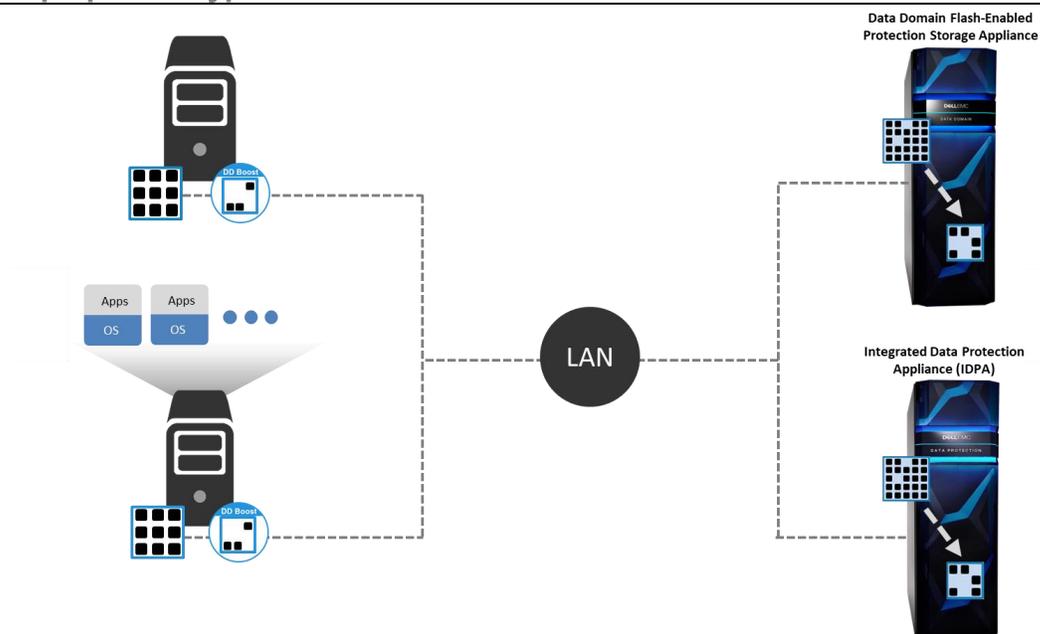
В этой белой книге описано, каким образом системы Dell EMC Data Domain и конвергентные решения на базе архитектуры Data Domain, например Dell EMC Integrated Data Protection Appliance (IDPA), помогают обеспечить оперативность, которая требуется ИТ-специалистам по внедрению. Благодаря этим решениям ИТ-службы могут трансформировать инфраструктуры в соответствии с требованиями ИТ-руководителей, устранить современные сложности в защите данных и, что важнее всего, удешевить хранение защищаемых данных. Результаты исследования ESG основаны на аудите и анализе ключевых индикаторов производительности (KPI) — фактических данных, полученных от развернутых систем. К таким индикаторам относятся первоначальная стоимость приобретения, расходы на обслуживание среды, емкость, коэффициент использования ресурсов хранения и производительность.

Преимущества архитектуры Data Domain и IDPA

В 2006 г. специалисты EMC решили не внедрять код дедупликации данных в существующие решения для защиты данных, что стало принципиальным изменением стратегии EMC в сфере защиты данных. Это изменение началось с использования технологий от двух приобретенных компаний. Первая из них предоставила технологию дедупликации на стороне источника. Второй приобретенной компанией была Data Domain, предоставившая технологию дедупликации на стороне целевой системы. Технологии, полученные от этих двух компаний, стали базовыми компонентами решений EMC для защиты данных. А сейчас, уже под брендом Dell EMC, системы Data Domain, в том числе IDPA, представляют собой полностью интегрированные специализированные устройства для защиты данных с поддержкой флэш-накопителей, уменьшающие объем дискового пространства СХД, необходимый для длительного хранения и защиты данных. В архитектуру этих систем интегрированы технологии дедупликации данных на стороне источника и на стороне целевой системы. В результате эти решения могут выполнять больше операций резервного копирования за меньшее время, обеспечивать более быстрое и надежное восстановление, а также сокращать емкость СХД, необходимую для защиты данных. В новое поколение систем Data Domain и IDPA внедрены флэш-накопители для метаданных. Они гарантируют высокую производительность в любом масштабе и позволяют в 20 раз ускорить доступ к виртуальным машинам и их восстановление непосредственно из Data Domain и IDPA по сравнению с системами предыдущих поколений.

³ Источник: отчет об исследовании ESG [«2015 Trends in Data Protection Modernization» \(Тенденции в модернизации защиты данных в 2015 г.\)](#), сентябрь 2015 г.

Рис. 2. Обзор архитектуры Data Domain и IDPA



Источник: Enterprise Strategy Group

Ниже описаны ключевые особенности архитектуры Data Domain и IDPA.

- **SISL (Stream Informed Segment Layout).** SISL позволяет системам Data Domain и IDPA выполнять 99% операций дедупликации с помощью ЦП и ОЗУ. То есть повышение производительности систем не зависит от количества дисков.
- **Использование сегментов переменной длины.** Чтобы получить оптимальные коэффициенты дедупликации, системы Data Domain и IDPA используют сегменты переменной длины, разделяя потоки на основе естественной структуры данных. В результате системы могут определять уникальность сегментов перед их сжатием и сохранением.
- **Дедупликация на лету.** Системы Data Domain и IDPA выполняют дедупликацию в ЦП и памяти по мере того, как система получает поток данных резервного копирования. То есть на диск отправляются для сохранения только уникальные данные. Это устраняет необходимость создавать область для переноса данных на диск и выделять ресурсы для дедупликации после обработки.
- **ПО Data Domain Boost.** ПО Data Domain Boost делегирует часть процесса дедупликации клиентам приложений или серверу резервного копирования. Благодаря DD Boost с сервера резервного копирования или клиента в систему Data Domain и IDPA перемещаются только уникальные данные. Это также помогает освободить ресурсы в сети заказчика и системах Data Domain и IDPA, чтобы повысить производительность дедупликации на стороне целевой системы.
- **Архитектура неустойчивости данных Data Invulnerability Architecture.** Архитектура Data Domain и IDPA гарантирует надежное восстановление данных. Архитектура Data Invulnerability Architecture обеспечивает проверку записи и чтения на лету, которая защищает от нарушений целостности данных и автоматически исправляет данные во время получения и извлечения. Функции непрерывного обнаружения ошибок и самовосстановления гарантируют возможность восстановления данных на протяжении всего их жизненного цикла в системе Data Domain и IDPA.

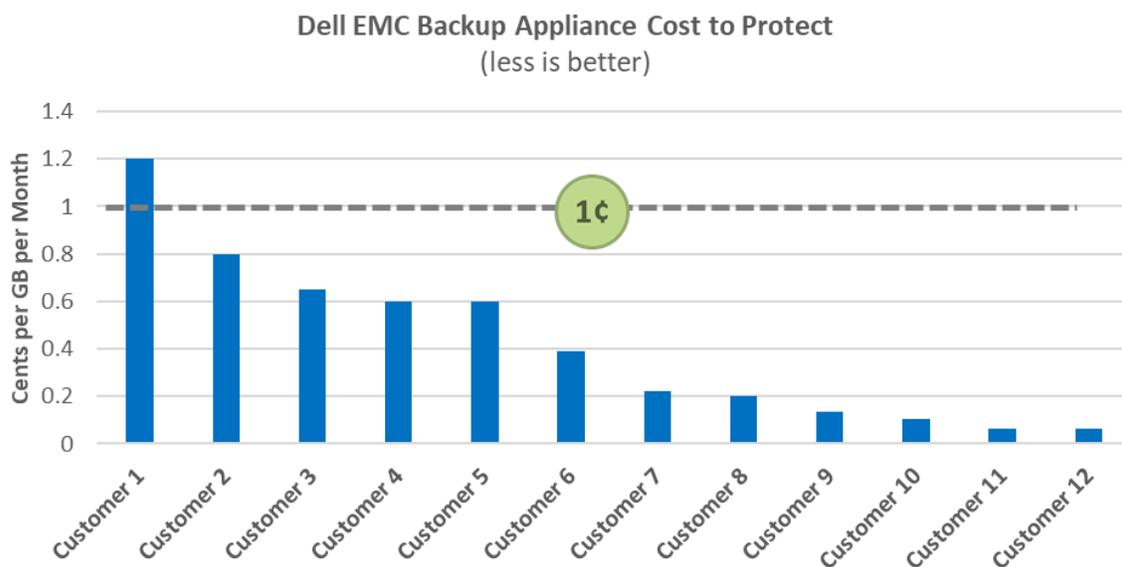
Экономические преимущества систем Data Domain и IDPA

Рассмотрим подробнее, какие экономические преимущества для бизнеса предоставляют архитектуры Data Domain и IDPA. Используя дедупликацию данных, заказчики могут хранить больше данных на том же пространстве физических дисков. Это снижает требования к емкости СХД и позволяет сократить расходы. Дедупликация данных, используемая на стороне источника или целевой системы, также помогает экономить емкость, при этом повышая производительность резервного копирования. При дедупликации на стороне источника только уникальные блоки данных отправляются из источника в целевую систему во время операции резервного копирования, что существенно сокращает объем сетевого трафика. Повышение эффективности сети позволяет поддерживать увеличение объемов данных резервного копирования, используя существующую сетевую инфраструктуру. Таким образом можно устранить необходимость в дорогостоящей модернизации сети или отсрочить такую модернизацию. Разумеется, чем меньше данных требуется передавать, тем быстрее выполняется резервное копирование. Ускоренное резервное копирование также позволяет заказчикам повысить частоту создания резервных копий, чтобы уменьшить риск потери данных, которая может повлечь за собой серьезные убытки для организации. Благодаря ПО DD Boost системы Data Domain и IDPA поддерживают дедупликацию как на стороне источника, так и на стороне целевой системы. Таким образом, заказчики могут выполнять дедупликацию в тех областях ИТ-среды, где она больше всего необходима.

Как показано на рис. 3, результаты проведенного компанией ESG анализа фактических данных (включая данные об оборудовании, программном обеспечении, энергопотреблении, охлаждении и дедупликации) демонстрируют, что системы Data Domain и IDPA могут с легкостью выделять ресурсы хранения в средах защиты данных при очень низкой ежемесячной стоимости защиты одного гигабайта.

Обратите внимание, что стоимость защиты одного гигабайта данных у двенадцати заказчиков, опрошенных ESG, составляет от 0,062 до 1,2 цента. Этот относительно широкий диапазон экономии связан прежде всего с размером сэкономленной емкости для дедупликации в инфраструктурах заказчиков Dell EMC. В этом отчете более подробно рассматриваются расходы на защиту данных. При этом особое внимание уделяется экономичности дедупликации, которую можно повысить, используя комплексное сочетание ПО Dell EMC Data Protection Suite и оборудования Data Domain и IDPA.

Рис. 3. Экономическая ценность Data Domain и IDPA

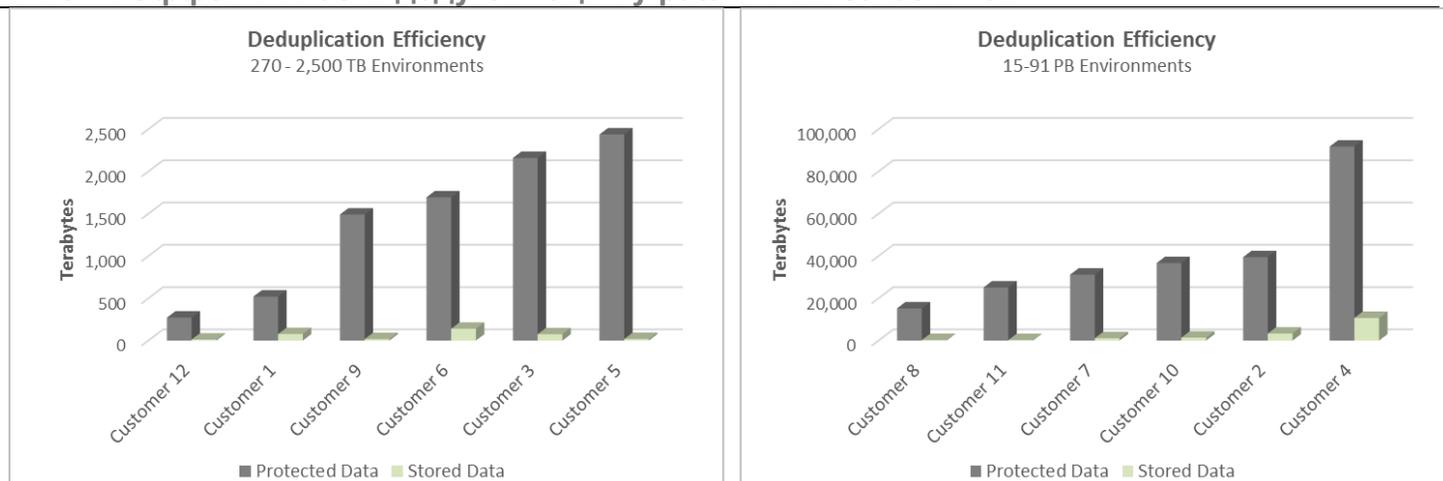


Источник: Enterprise Strategy Group

Эффективность дедупликации имеет большое значение

Компания ESG начала исследовать экономическую ценность сокращения емкости для дедупликации в Data Domain и IDPA, проведя аудит и анализ данных, полученных из центров удаленной поддержки 12 активных заказчиков Dell EMC. Как показано на рис. 4, объем данных в средах заказчиков, которые требовалось защитить, составлял от около 270 Тбайт до свыше 90 Пбайт. Заказчики представляли различные отрасли, в том числе ИТ-отрасль, производство, сферу страхования и отрасль здравоохранения. Компания ESG выбрала заказчиков из разных отраслей, поскольку ей требовалось получить показатели дедупликации наборов данных различных типов. Серые столбцы на рис. 4 обозначают объем данных, которые необходимо защитить в каждой среде. Зеленые столбцы — это объем данных, хранимых после дедупликации.

Рис. 4. Эффективность дедупликации у различных заказчиков



Источник: Enterprise Strategy Group

- **Дедупликация.** Самый высокий коэффициент дедупликации, полученный по итогам анализа данных, составил 126:1. Этот результат был получен в среде заказчика. Он указан как крайний

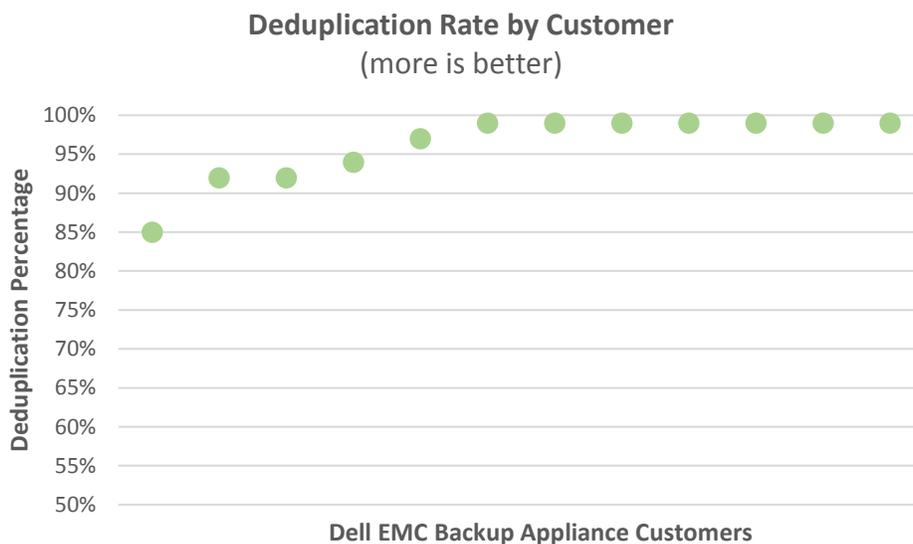
показатель слева на рис. 4. Средний коэффициент дедупликации для всех проанализированных сред заказчиков составил 41:1.

- **Защита данных.** Даже заказчик с самым низким коэффициентом дедупликации (примерно 9:1) смог обеспечить защиту почти 92 Пбайт данных, используя лишь 10 Пбайт емкости. Показатели этого заказчика приведены последними справа на рис. 4. Этот заказчик работает в обрабатывающей промышленности, а его ИТ-среда отличается высокой степенью распределения.

На рис. 5 проиллюстрирован еще один аспект дедупликации в Data Domain и IDPA. На этом рисунке указаны процентные показатели дедупликации в ИТ-среде каждого из вышеупомянутых 12 заказчиков. Коэффициенты дедупликации составляют от 85% до 99%. Обычно чем дольше решение Data Domain и IDPA принимает данные в ИТ-среде, тем выше коэффициент дедупликации. Это связано с тем, что со временем комбинации данных повторяются и сохраняются только уникальные данные.

Возможность обрабатывать сегменты данных переменной длины в архитектурах Data Domain и IDPA играет ключевую роль в достижении такого уровня дедупликации. Благодаря использованию сегментов данных переменной длины решения Data Domain и IDPA проще адаптируются к стандартным сочетаниям в структуре данных (включая временные метки баз данных), отправляемых в устройство для защиты. При использовании сегментов данных переменной длины достигаются значительно более высокие уровни дедупликации, чем в архитектурах с сегментами фиксированной длины. Кроме того, такой подход обеспечивает более эффективную масштабируемость в одном пуле хранения данных. Пользователям также потребуется управлять меньшим числом устройств, что позволит снизить операционные издержки.

Рис. 5. Коэффициент дедупликации заказчиков (в процентах)



Источник: Enterprise Strategy Group

Еще одно преимущество дедупликации сегментов переменной длины проявляется, когда данные реплицируются между устройствами Data Domain и IDPA. Из-за сокращения общего объема требуется реплицировать меньше данных. В результате не только повышается эффективность репликации, но и снижаются требования к полосе пропускания глобальной сети для процесса репликации.

- **Коэффициенты дедупликации.** По результатам аудита и анализа, коэффициенты дедупликации составляли от 85% до 99% (в среднем 96%). Это означает, что лишь малую часть

производственных данных (в лучших случаях 1% или менее) необходимо было сохранять в системах Data Domain и IDPA для защиты.

- **Менее 90%.** По результатам анализа, только у одного заказчика коэффициент дедупликации данных был меньше 90%. Этот заказчик опять-таки имеет ИТ-среду с высокой степенью распределения и распределенными пулами защиты данных.
- **90% и выше.** У остальных клиентов коэффициенты дедупликации превышали 90%, а в большинстве сред заказчиков коэффициент дедупликации составлял 99%.

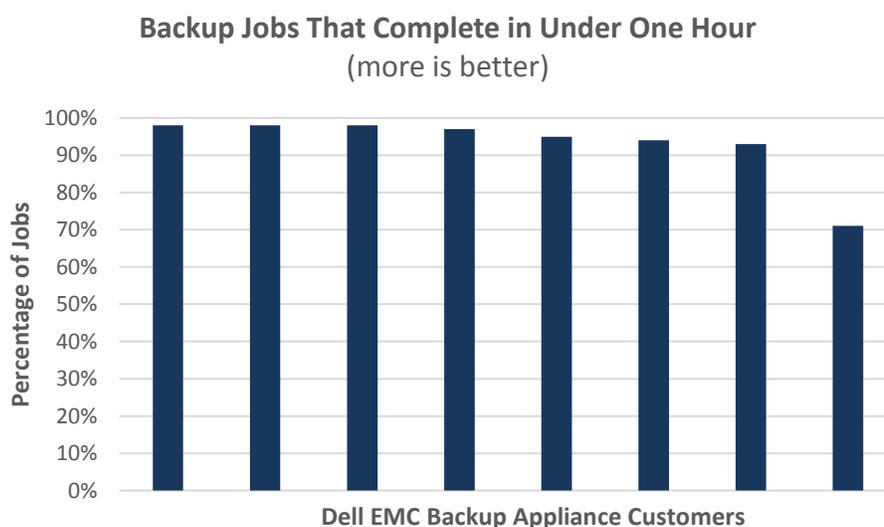
Скорость и масштабируемость

Увеличение емкости СХД с помощью дедупликации не всегда гарантирует высокую производительность. Data Domain и IDPA устраняют эту проблему, используя два ключевых элемента для повышения производительности: Stream-informed Segment Layout (SISL) и программное обеспечение DD Boost.

SISL — это элемент архитектуры платформ Data Domain и IDPA, благодаря которому 99% дедупликации может выполняться в ЦП и ОЗУ. Это значит, что данному устройству не требуется много дисков, чтобы обеспечить высокую производительность. В результате более компактные решения могут иметь производительность на уровне конфигураций, включающих в себя больше дисков. Кроме того, поскольку системы Data Domain и IDPA используют ЦП Intel, усовершенствования каждого нового выпуска процессоров Intel повышают их производительность.

ПО DD Boost также помогает повысить производительность, распределяя процесс дедупликации между клиентами приложений и сервером резервного копирования. При использовании DD Boost с сервера резервного копирования или клиента в систему Data Domain и IDPA перемещаются только уникальные данные. DD Boost может сократить объем переносимых данных на 99%. Компания ESG провела аудит, чтобы определить чистую выгоду от ускорения резервного копирования. Как показано на рис. 6, большинство заказчиков обычно выполняют задачи резервного копирования за один час или быстрее.

Рисунок 6. Производительность резервного копирования



Источник: Enterprise Strategy Group

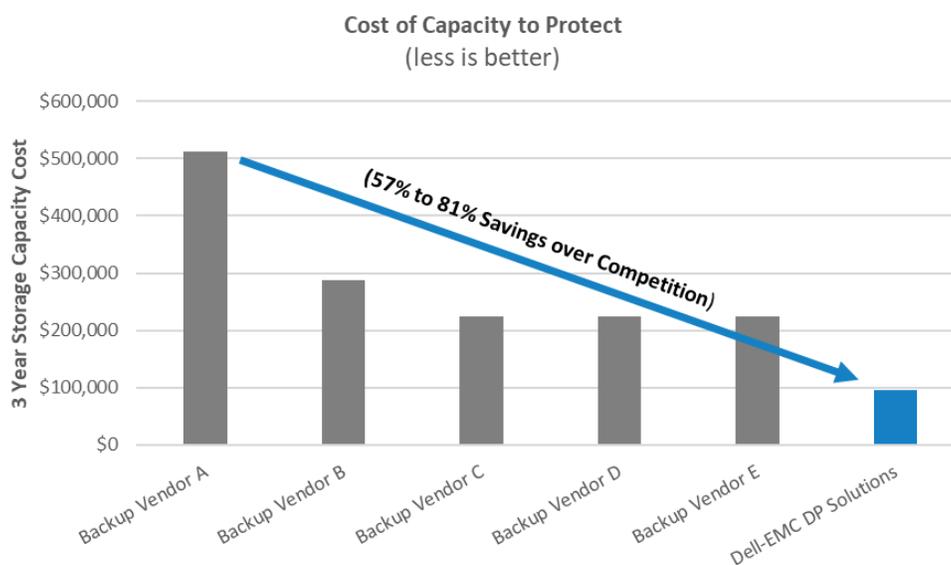
- **Менее 15 минут.** В ИТ-среде одного заказчика из ИТ-отрасли 76% всех задач резервного копирования выполнялись менее чем за 15 минут.

- **Менее 1 часа.** В ИТ-средах восьми из двенадцати проанализированных заказчиков от 70% до 98% всех задач резервного копирования выполнялись менее чем за час (как показано на Рисунок 6).
- **Менее 4 часов.** В ИТ-среде одного заказчика из сферы страхования все задачи резервного копирования выполнялись менее чем за четыре часа.
- **Более 4 часов.** В двух самых крупных средах (по количеству и размеру приложений) на выполнение всех задач резервного копирования уходило более четырех часов.

Факторы сокращения расходов

Компания ESG завершила изучение экономической ценности дедупликации Data Domain и IDPA, проведя аудит и анализ данных, полученных от свыше 15 000 устройств резервного копирования Dell EMC, развернутых по всему миру в центрах удаленной поддержки заказчиков. Представленные ниже данные собраны из систем, фиксирующих используемое ПО для резервного копирования и полученные коэффициенты дедупликации. Эти данные показывают, что эффективность дедупликации устройств резервного копирования Dell EMC можно дополнительно повысить, если внедрить комплексное решение Dell EMC для защиты данных. Как показано на рис. 7, сочетание оборудования устройства Dell EMC с ПО Data Protection Suite помогало заказчикам обеспечить более экономичную дедупликацию и снизить расходы на емкость СХД в течение трех лет на 57–81% по сравнению с решениями от других вендоров.

Рис. 7. Повышение экономичности с помощью Data Domain и IDPA при использовании Data Protection Suite



Источник: Enterprise Strategy Group

Решения Data Domain и IDPA предоставляют и другие экономические преимущества, помимо экономии емкости СХД, как показано на рис. 7. Проведенный ESG анализ фактических данных заказчиков также выявил перечисленные ниже преимущества.

- **Сокращение нагрузки на полосу пропускания.** Как и большинство корпоративных приложений, решения для резервного копирования и восстановления используют сетевые ресурсы, чтобы функционировать должным образом. То есть им необходимы ресурсы локальной сети для

локального резервного копирования и восстановления, а также ресурсы глобальной сети для аварийного восстановления и обеспечения непрерывности бизнеса. Неэффективные решения для защиты данных могут быстро потреблять ценные сетевые ресурсы и даже негативно влиять на возможности подключения пользователей и производительность их работы. Используя дедупликацию в Data Domain и IDPA, международная производственная компания сократила нагрузку на локальную и удаленную полосы пропускания для защиты данных на 98%.

- **Производительность.** Эффективная дедупликация данных может значительно повлиять на повышение общей производительности средств защиты данных. Внедрив ПО DD Boost, сокращающее объем данных, которые необходимо передавать между клиентом и системой Data Domain, производитель тяжелого оборудования ускорил резервное копирование и восстановление на 50%. Кроме того, этот заказчик повысил готовность своей системы к аварийному восстановлению на 90%.
- **Уменьшение занимаемой площади центра обработки данных.** Недвижимость, особенно достаточно крупная для соответствия требованиям современного ЦОД к площади, электропитанию и охлаждению, стоит недешево. Физическое расширение центра обработки данных без перемещения оборудования может быть сложной или даже невыполнимой задачей. Например, национальная сеть универмагов, перейдя на Data Domain, отказалась от использования физической ленты в среде защиты данных и освободила три полных ряда ценного стоечного пространства ЦОД.

Выводы

Основные требования к защите данных, предъявляемые ИТ-руководителями, сосредоточены на улучшении ключевых показателей доступности и оперативности используемых решений. Еще одно важное требование — сокращение расходов, которое также считают основным приоритетом ИТ-специалисты, внедряющие средства защиты данных. Эти требования не следует считать противоречащими друг другу или взаимоисключающими. На самом деле их можно выполнить, используя улучшенные решения для защиты данных, разработанные как для повышения эффективности и надежности, так и для расширения функциональности.

Эффективность может проявляться в самых разных аспектах. Часто она заложена в экономичности ведения бизнеса и функциональной совместимости решений. Некоторые аспекты эффективности приведены ниже.

- Возможность создать платформу защиты данных с минимальными расходами, сократив издержки на обслуживание приложения для защиты данных.
- Функциональная совместимость программных решений для резервного копирования. Ее важно обеспечить, так как только четверть организаций используют одно приложение резервного копирования.⁴
- Эффективность масштабирования и дедупликации платформы защиты данных, с помощью которой можно сократить число развертываемых и управляемых систем, а также уменьшить расходы на них.
- Возможность использовать один пул платформы защиты данных с оптимизированной емкостью для резервного копирования и архивирования данных.

⁴ Источник: исследование ESG «Data Protection Modernization Trends» (Тенденции в модернизации защиты данных), декабрь 2016 г.

- Функциональная совместимость (и интеграция) платформы защиты данных и множества других платформ. Благодаря такой совместимости владельцы приложений (например, администраторы баз данных и ВМ) могут использовать собственные инструменты для дополнительной защиты и восстановления, при этом рационально расходуя ИТ-ресурсы и применяя эффективную централизованную платформу защиты данных.
- Возможность ускорить развертывание с помощью интегрированных решений для защиты данных, сокращая число систем, которые необходимо развернуть, упрощая лицензирование с помощью комплексного подхода и устраняя проблемы несовместимости.
- Возможность упростить общее администрирование сред защиты данных благодаря централизованному управлению и отчетности, а также одноэтапная модернизация и установка исправлений для всех компонентов защиты данных устройства позволяют сократить время обслуживания среды.
- Интеграция более широких возможностей, например облака (длительное хранение, аварийное восстановление в облаке), средств поиска и аналитики, также ускоряет развертывание и упрощает управление.

Dell EMC предоставляет решения, которые многие компании в отрасли считают эталонами платформ защиты данных и/или продуктов для дедупликации. Вместо простых компонентов защиты данных Dell EMC предлагает целую экосистему, которая включает в себя, среди прочего, производственную систему хранения, решения для резервного копирования и архивирования, а также подключаемые ускорители для приложений (например, DD Boost). Решение Dell EMC IDPA теперь стало частью этой экосистемы, охватывающей множество предварительно настроенных компонентов, среди которых программное обеспечение для защиты данных, серверы и хранилища, единый пользовательский интерфейс управления, облачные возможности, средства поиска и аналитики. Сегодня у нас есть весомый стимул по-прежнему внедрять инновации, опережая темпы роста рыночного спроса, чтобы еще больше повышать эффективность механизмов дедупликации, расширять возможности их интеграции и функциональность. В результате можно получить не только возможность локальной дедупликации, но и функции многоуровневого хранения в облаке и репликации данных в облачные устройства и из этих устройств, а также использовать виртуальные устройства, выполнять аварийное восстановление в облаке и т. д.

Одним из самых важных компонентов платформы дедупликации можно назвать гарантию целостности данных в репозитории, поскольку многое в работе ИТ-службы зависит от целостности каждого уникального блока в системе. Очень важно, чтобы организации пристально следили за тем, как часто и тщательно система проверяет целостность данных в базовой СХД. Подход Dell EMC к решению этой проблемы, выраженный в архитектуре Data Invulnerability Architecture (DIA), стал отличительной особенностью платформы устройств резервного копирования с момента ее создания.

По результатам аудита данных, полученных в рабочих условиях более чем 15 000 производственных сред, компания ESG подтвердила, что сочетание оборудования Dell EMC Data Domain и IDPA с программным обеспечением Data Protection Suite сокращает стоимость емкости для защиты данных на 57–81% по сравнению со средами устройств резервного копирования Dell EMC, в которых используется конкурентное ПО. Организации, которые сомневаются, стоит ли инвестировать средства в решения Data Domain или IDPA, ввиду «высокой стоимости» этих систем, могут со временем по-новому взглянуть на их экономические преимущества, в том числе на экономию благодаря сокращению простоев и потере данных, повышению производительности и выгодную ежемесячную стоимость защиты одного гигабайта данных.

Все наименования товарных знаков являются собственностью соответствующих компаний. Содержащаяся в этой публикации информация получена из источников, которые компания Enterprise Strategy Group (ESG) считает заслуживающими доверия, однако ESG не гарантирует ее достоверность. Данная публикация может содержать мнения компании ESG, которые могут со временем меняться. Авторские права на данную публикацию принадлежат компании Enterprise Strategy Group, Inc. Любое воспроизведение или предоставление данного документа или его части в печатном или электронном виде или другим способом лицам, не имеющим права на его получение, без явного согласия компании Enterprise Strategy Group, Inc. будет считаться нарушением законодательства США по защите авторских прав и станет поводом для возбуждения иска о возмещении ущерба и, если применимо, уголовного преследования. При возникновении любых вопросов обращайтесь в подразделение ESG по работе с заказчиками по телефону +1 (508) 482-0188.



Enterprise Strategy Group — это всемирно известная интегрированная фирма, занимающаяся исследованиями, анализом, валидацией и разработкой стратегий в области ИТ. Она собирает практически значимую аналитическую информацию и предоставляет ее международному ИТ-сообществу.

© The Enterprise Strategy Group, Inc., 2018 г. Все права защищены.

