

CommVault Simpana — единая платформа для защиты и управления информацией

В конце августа 2014 г. компания CommVault анонсировала новую версию 10 R2 своего основного продукта — ПО Simpana, расширенную по функционалу и поддержке облачных сред. Также в 2014 г. четвертый раз подряд Commvault становится лидером магического квадранта Gartner в исследовании Enterprise backup/Recovery/Integrated Appliances как по присутствию на рынке, так и стратегии развития.



Евгений Пухов — технический консультант, Commvault Россия и СНГ.

Введение

Рынок корпоративных решений по управлению информацией — один из самых быстрорастущих в ИТ. Gartner в своем исследовании “Magic Quadrant for Enterprise Backup Software and Integrated Appliances” (16 June 2014, ID:G00262146) отмечает:

- к 2018 г. 40% организаций дополнят или заменят свои существующие приложения для резервного копирования;
- к 2016 г. 20% организаций (по сравнению с 7% в настоящее время) будут использовать только моментальные снимки и методы репликации, отказавшись от традиционного резервного копирования / восстановления;
- к 2016 г. менее чем 30% всех больших данных будут резервироваться;
- к 2017 г. 70% (по сравнению с 30% в настоящее время) организаций заменят свои технологии резервного копирования на ленту, используемые для удаленных офисов, решениями на базе жестких дисков (которые также будут включать и репликацию).

Работа любой крупной организации невозможна без единого подхода к управлению данными как одним из самых главных бизнес-активов компании. Это позволяет снизить риски от любых программных и аппаратных сбоев и всегда гарантировать восстановление данных с заранее заданными уровнями обслуживания.

Вот наиболее распространенные вопросы, связанные с управлением данными, которые стоят перед ИТ-директорами:

- как правильно выстроить систему защиты неоднородной информации в сложной ИТ-среде;

- как эффективно интегрировать точные библиотеки, облачное хранение, СХД различных производителей, множество распределенных по удаленным ИТ-островам данных в один репозиторий;
- как контролировать безопасность и конфиденциальность данных, хранимых на мобильных компьютерах пользователей;
- как сделать гарантированное восстановление данных максимально быстрым и прозрачным для бизнес-приложений.

CommVault предлагает решение этих задач в единой интегрированной платформе Simpana. Благодаря такому комплексному подходу достигается автоматизация и унификация механизмов защиты данных. За счет этого снижаются риски, повышается операционная эффективность, снижаются затраты на управление.

Архитектура CommVault Simpana

CommVault Simpana — полностью программно- и аппаратно-независимая платформа для защиты и управления данными самого широкого круга приложений, операционных систем (включая Open Source), систем хранения и облачных провайдеров.

Часто платформа Simpana ассоциируется с ПО резервным копированием данных (далее РК). Действительно, 70% инсталляций по всему миру, прежде всего, реализуют функции бэкапа. На самом деле, CommVault — это намного больше. С помощью Simpana реализуются:

- резервное копирование данных, включая виртуальные среды;
- автоматический контроль виртуальной инфраструктуры на предмет активности ВМ, перенос ВМ между различными гипервизорами, перенос физических машин на виртуальную среду;
- архивирование электронной почты, файловых серверов, виртуальных машин, баз данных;
- защита данных рабочих компьютеров персонала с защитой от потери или кражи ноутбука;
- создание индексируемого хранилища с полнотекстовым поиском;
- контроль содержимого на предмет несанкционированного контента;

- портал самообслуживания для самостоятельного восстановления данных пользователями;
- соответствие требованиям регуляторов. Compliance и Legal Hold.

Основные компоненты CommVault Simpana:

- **CommServe** — управляющий сервер Simpana;
- **ContentStore** — единый набор ресурсов хранения данных. Это могут быть СХД (в том числе программно-определяемые СХД), любые backup appliance, файловые серверы, облачные хранилища Amazon и Microsoft и др., ленточные библиотеки;
- **Data Agent** — агент для приложений и файловых систем. Выполняет также функции компрессии, дедупликации и шифрования на источнике (source-side dedup compression);
- **Media Agent** — медиа-агент Simpana, принимающий поток данных с Data Agent, реализующий выбранную политику хранения и записывающий данные на соответствующие носители в ContentStore. Непосредственно выполняет процесс дедупликации данных «на получателе» (target-side dedup).

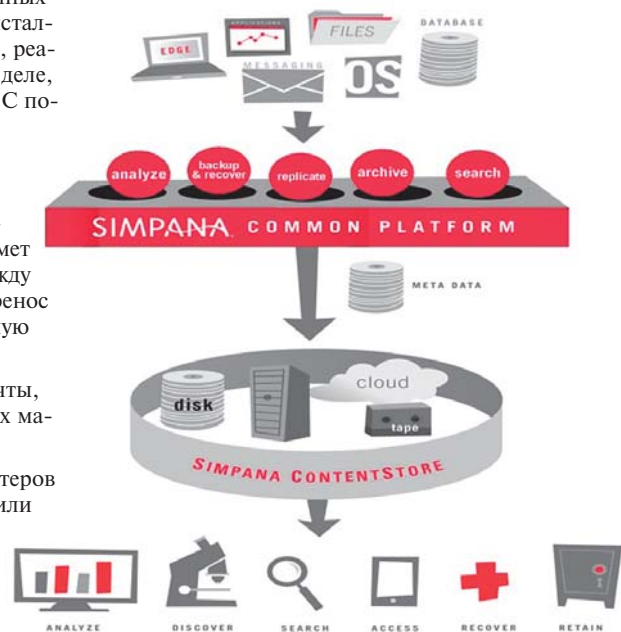


Рис. 1. Simpana представляет собой единую платформу для защиты и управления всей корпоративной информацией.

Архитектурно Simprana скорее похожа на OLAP-платформу, чем на систему резервного копирования или архив (рис. 1). С помощью Data Agents данные собираются со всех приложений, файловых систем, индексируются затем с помощью Media Agent попадают в ContentStore — некое подобие Data Warehouse. В ContentStore задаются автоматические политики перераспределения данных и систему аналитических отчетов для принятия решения ИТ-руководителем (например, по изменению SLA того или иного набора данных). Такая схожесть с OLAP и обуславливает главное отличие Simprana от альтернативных предложений по резервному копированию. Simprana — унифицированная горизонтально масштабируемая платформа, поэтому добавление нового функционала — без изменения конфигурации системы и потерь производительности.

В какой-то мере функционал Simprana можно покрыть 2–4-мя сторонними продуктами, но это будут разные продукты с разными интерфейсами управления, ресурсами хранения, что усложняет и удорожает обслуживание.

ContentStore: это единое дедуплицированное хранилище. База данных дедубликации (DDB) надежно защищена, и при размещении на быстрых носителях (SSD) включение дедубликации очень слабо влияет на общую производительность системы.

В отличие от других разработок, Simprana хранит индексы на медиа-агентах (которых может быть множество для балансировки нагрузки и отказоустойчивости), а не на управляющем сервере. Наличие такого распределенного индексирования позволяет обеспечивать горизонтальное масштабирование системы РК. Это особенно важно для удаленных офисов, в которых каждый локальный медиа агент имеет копию индексной базы.

Другой функционал для территориально распределенных систем — репликация с дедубликацией **DASH copy**, позволяющая связывать «острова данных» по медленным каналам.

Следующий важный компонент — это набор правил по хранению того или иного набора данных в соответствии с заданным SLA — Storage Policy. В Storage Policy задается:

- когда делается копирование;
- где это хранить;
- как долго хранить.

Правила могут комбинироваться в любом порядке в рамках даже одной Storage Policy, например, файлы одного типа могут храниться на дисковых носителях для быстрого восстановления, файлы другого типа — на медленных — ленте и в облачных провайдерах. Можно также задавать различные retention-периоды, многоуровневое хранение, ролевое администрирование с привязкой ролей к Active Directory, разную видимость ресурсов. Мощный web-интерфейс позволяет реализовать РК и архивирование под девизом: «каждому пользователю — своя Simprana».

Важно, что в Simprana клиенты, устанавливаемые на приложения (Data Agent), отвечают только на вопрос: какие данные защищать? Сами же политики защиты predeterminedляются в Storage Policy на этапе первоначальной настройки. Все, что нуж-

но сделать при установке нового сервера, это проассоциировать Data Agent с той или иной политикой хранения.

Общая рекомендация CommVault: создавайте наименее возможное количество Storage Policies, которое может обеспечить заданный уровень SLA всех приложений (что упростит управление), например:

одна политика для баз данных, вторая — для почты, третья — для файловых серверов и рабочих компьютеров пользователей.

Архив + РК для файловых и почтовых серверов

Самая распространенная задача для Simprana — резервное копирование и архивирование почты и файлов.

Для файловых и почтовых серверов есть специализированный режим OnePass, который позволяет объединить резервное копирование, архивирование и построение отчетов о содержимом в одну операцию. Обычно для РК и архивирования используют два отдельных продукта. Но в Simprana OnePass агент собирает данные в один проход, помещает в единый репозиторий, затем использует для обеих целей (рис. 2). Это существенно ускоряет процесс.

Гранулярное восстановление данных из архива или из резервной копии

В большинстве случаев восстановление сервера или виртуальной машины целиком не требуется. Как правило, потеря данных носит локальный характер: файлы и сообщения теряются по причине проблем с рабочим компьютером, ПО, или же по банальной ошибке. В этом случае логично предоставить пользователю возможность удобного доступа к своим резервным копиям и архиву и научить его выполнять элементарное пофайловое восстановление самостоятельно. Это радикально упростило жизнь как пользователя, так и администратора. Пользователю это сокращает время решения проблемы, а администратор разгружается от рутинных операций.

CommVault предлагает портал самообслуживания в виде web-консоли или outlook плагина, в которых пользователь мо-

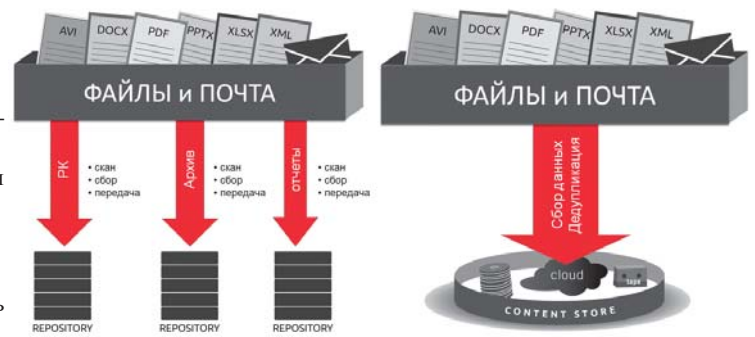


Рис. 2. Сравнение традиционного (слева) и однопроходного (OnePass — справа) способов резервного копирования и архивирования данных, а также подготовки отчетности.

жет восстановить из архива или резервной копии выбранные сообщения или файлы (рис. 3).

Традиционные способы восстановления (подключение базы целиком) могут работать параллельно с гранулярным OnePass (который работает по MAPI). Возможно восстановить ящик в виде pst файла.

В версии Simprana 10 R2 появилась возможность видеть архив почты ContentStore непосредственно из Outlook как дополнительный почтовый ящик Inbox. Соответственно, все вложения, письма и другие функции Outlook теперь работают с архивом, как с обычными почтовыми ящиками Exchange. Появилась возможность архивации также PST-файлов, находящихся на локальных рабочих станциях пользователей.

Управление мгновенными снимками (снапшотами) СХД

Как правило, в ИТ-инфраструктурах крупных компаний используются СХД нескольких производителей, которые применяют собственные не совместимые с другими технологии защиты и управления данными, особенно это относится к мгновенным снимкам.

Simprana умеет управлять мгновенными снимками практически любой современной СХД, включая программно-определяемые СХД.

Функционал называется **Intellisnap** и является частью общей схемы резервного копирования, архивирования и управления данными, хотя при желании его можно внедрить независимо, не реализуя другие функции CommVault.

Часто снапшот — это промежуточный этап перед попаданием на ленту или дру-

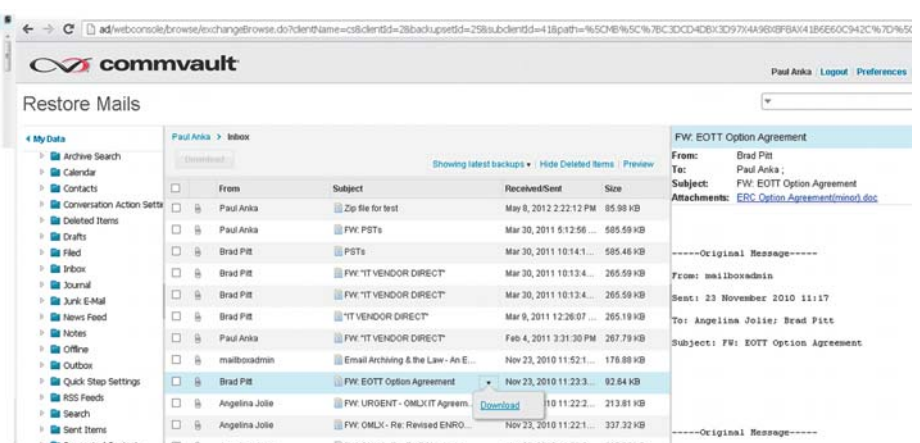


Рис. 3. Персонализированные интерфейсы по восстановлению файлов и почты.

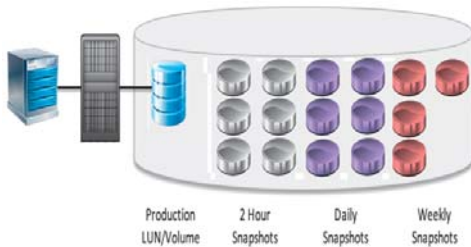


Рис. 4. Пример многоуровневой иерархии снимков.

ное бэкап-устройство, но некоторые производители (например, NetApp) рекомендуют рассматривать снимshots как конечную цель РК, а для защиты — просто переносить всю иерархию снимshots на другой такой же массив, который становится по сути дисковым устройством РК (функционал называется NetApp SnapVault). Для таких случаев есть отдельная схема лицензирования, которая, в том числе, продается через OEM-канал NetApp под названием SnapProtect.

Как найти самую актуальную версию какого-то утерянного файла? Можно, конечно, вручную перебирая все снимshots, а если их десятки или сотни? Благодаря индексированию всего содержимого мгновенных снимshots (опционально) можно находить и восстанавливать файлы определенной версии. Поиск по файлам может быть полнотекстовым, для этого нужно дополнительно установить *Simpana Search Engine* (подробнее — <http://www.commvault.ru/simpana-software/features/search-and-ediscovery>).

Intellisnap управляет временем жизни снимshots — понятно, что для снимshots месячной давности уже не важна гранулярность в 10 минут — нужно минимизировать количество с течением времени. *Simpana* имеет возможность создавать отдельные расписания для снимshots разных retention периодов. И, конечно же, периодичность снимshots можно привязать к критичности приложений и соблюсти самые строгие SLA (рис. 4).

Защита виртуальных инфраструктур

CommVault 100% совместим со всем функционалом vSphere 5.5 и Hyper-V. В *Simpana 10 R2* возродилась поддержка гипервизора Xen. Поддерживаются версии XenServer до 6.2.

Основа успешной системы РК виртуальных машин — правильное планирование ресурсов и предопределение политик хранения. Но перед этим необходимо классифицировать VM.

CommVault использует следующую классификацию:

- **GOLD.** Высоконагруженные 24x7 VM, с высоким процентом изменяемых данных практически без окна РК;
- **SILVER.** Низконагруженные VM, с невысоким процентом изменяемых данных, допускающие небольшое окно РК;
- **BRONZE.** Малонагруженные и практически неизменные VM, допускающие выключение в нерабочие часы или остановку на длительный срок.

Тип Silver может быть защищен с помощью обычного VADP streaming backup.

Для виртуальных машин с высоким процентом изменяемых данных использование обычного VADP с CBT (Change Block Tracking) неэффективно, так как во время копирования данных с программного снимshots VADP накапливается большой процент измененных данных, что, в конечном счете, приводит к задержкам при «докатке» обновлений CBT. Поэтому для таких высоконагруженных машин типа GOLD необходимо использование аппаратных снимshots СХД. Здесь CommVault уникален благодаря совместимости практически со всеми современными СХД.

Что же делать, если виртуальные машины постоянно добавляются и удаляются? Как ИТ-администратору успеть задать каждой вновь появившейся виртуальной машине ту или иную политику защиты? Спасает мониторинг виртуальных машин VM autodiscovery. Каждой VM, удовлетворяющей определенной набору свойств, привязывается та или иная политика защиты данных автоматически (рис. 5).

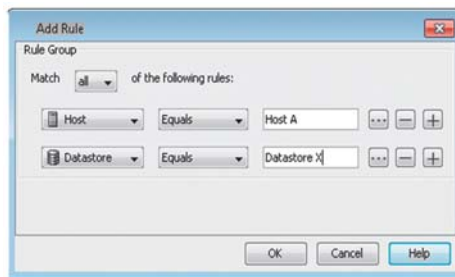


Рис. 5. Создание политики защиты для всех VM, которые одновременно выполняются на хосте Host A и хранилище Datastore X.

Управление жизненным циклом VM и архивирование: VM lifecycle management

Данный функционал полностью реализует концепцию самостоятельного управления жизненным циклом виртуальных машин. Используя портал самообслуживания пользователь может создавать, изменять настройки и восстанавливать свои виртуальные машины из шаблонов, заданных администратором.

Огромное количество VM типа BRONZE (см. выше), занимая оперативную память, вычислительные ресурсы, при этом используются редко, однако данные на них также нужно защищать. Для отслеживания таких машин нужен отдельный сервис, так как подобный функционал нереализуем стандартными средствами VMware. На помощь опять приходит *Simpana*, а точнее, новый компонент *VM Archiving*, появившийся в версии 10 R2.

VM Archiving позволяет выбирать малоиспользуемые VM по критериям низкой загрузки CPU, Disk, Network на протяжении длительных периодов (длительность можно задавать). Если VM попала под выбранные правила, можно выполнить следующие действия (рис. 6):

- перенести VM на другой менее производительный Datastore или на менее производительный хост (с помощью vMotion), при этом VM продолжает работать;

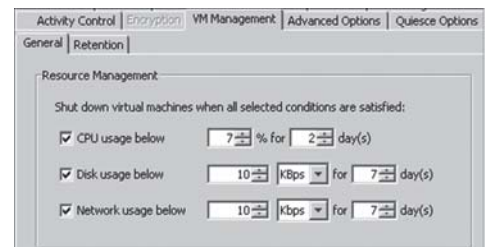


Рис. 6. Настройка политики выключения VM по уровню нагрузки.

- выключить и заархивировать VM, перенести его в долговременное дедуплицируемое хранилище *Simpana Contentstore*. vCenter продолжает видеть виртуальную машину и позволяет ее восстановить по требованию пользователя. Восстановить VM также можно через панель самообслуживания CommVault;
- удалить VM по прошествии определенного Retention period.

В *Simpana 10 R2* появилась группа функционала гранулярного восстановления данных под общим названием *Live recovery*.

Live file recovery — возможность восстановить отдельный файл виртуальной машины без необходимости сбора гранулярных метаданных на этапе РК. Выбранная версия VM просто монтируется как NFS datastore, и пользователь может восстанавливать информацию через файловый браузер. *Live file recovery* работает также с аппаратными снимshots (интеграция с *intellisnap*).

Multi-VM recovery — возможность восстанавливать несколько VM в один проход.

Live Recovery — возможность запустить VM напрямую из бэкап-копии, параллельно запустив восстановление на продуктив в фоновом режиме.

Резервную копию физического сервера можно восстановить в виртуальную среду VMware или Hyper-V. Функционал называется **Virtualize me**. Если в организации одновременно используются два типа гипервизоров — VMware vSphere и Hyper-V, то с помощью *VM conversion* можно конвертировать виртуальные машины из VMware в Hyper-V: Виртуальную машину VMware можно восстановить, как Hyper-V из РК, или же сделать прямую миграцию (создать копию) VM из vSphere в Hyper-V.

Аналитика

Не всегда легко подобрать правильную политику хранения резервных копий и архивов. Для того чтобы принимать подобные решения, необходимо ознакомиться со структурой данных. Сделать это помогает модуль **File Analytics**. Отчеты, наглядно демонстрирующие потребление ресурсов хранения данными в зависимости от их атрибутов и характеристик. На рис. 7 приведен пример, демонстрирующий, что большую часть доступного пространства используют файлы с датой последнего изменения более трех лет.

Подобная отчетность крайне полезна для принятия решений в части организации хранения и позволяет сэкономить за счет автоматического переноса файлов на более экономичные уровни хранения с помощью агентов CommVault *Simpana*.

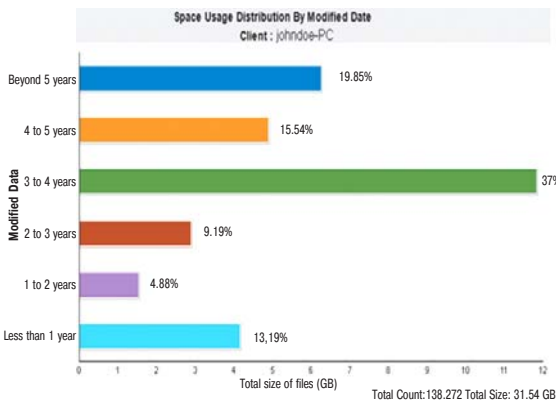


Рис. 7. Пример работы модуля File analytics.

Защита мобильных устройств

Защита мобильных устройств **CommVault EDGE** — еще один востребованный функционал Simrana, который был существенно дополнен в 2014 г.

Для пользователя защита мобильных устройств в EDGE выглядит как удобный облачный бэкап с минимумом настроек (все настройки предопределяются на уровне Storage Policy, как и с бэкапом любых других данных). Но только в отличие от интернет-компаний, которые предлагают похожие сервисы, с CommVault достигается полный контроль над контентом, доступностью, конфиденциальностью данных. Как и все остальные сервисы, CommVault EDGE интегрируется в хранилище ContentStore.

Что же предлагает CommVault EDGE?

Во-первых, инкрементальный бэкап с возможностью дедупликации на источнике. Гибкая настройка автоматических сценариев РК: бэкап может запускаться только при достаточно широком канале, только при низкой активности процессора. Есть отдельные политики для ноутбука, в котором EDGE может останавливать РК, если ноутбук работает от батареи, а также останавливать бэкап при низком ее заряде. Можно даже сделать так, чтобы бэкап работал только в защищенных сетях и запретить бэкап по общественным WiFi-сетям, что очень сильно повышает безопасность.

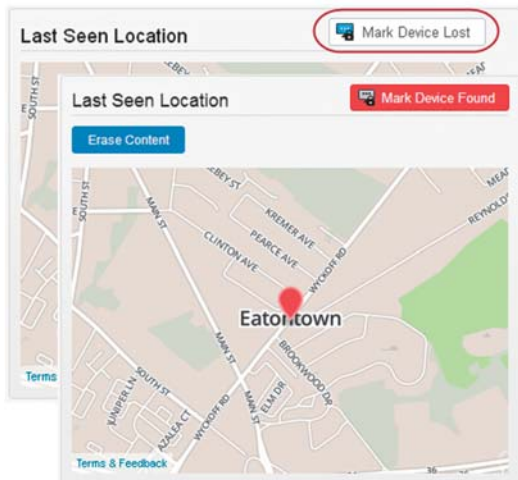


Рис. 8 Поиск ноутбука через портал самообслуживания CommVault (можно через мобильное приложение iOS/ Android).

Что произойдет при утере ноутбука? Все данные — в резервной копии, мы их сможем восстановить! А как быть с данными, которые хранятся на ноутбуке и могут быть доступны злоумышленникам? Их защитой занимается модуль **DLP** (Data Loss Prevention).

Пользователь, обнаружив пропажу ноутбука, заходит на портал самообслуживания CommVault (можно через мобильное приложение iOS/ Android, помечает устройство утерянным (рис. 8). В случае если утерянный ноутбук после этого подключится в интернет, данные окажутся заблокированными и не доступными для чтения. Но что будет, если ноутбук утерян/ украден и не подключается в сеть? Для защиты от этого можно создать режим, при котором данные будут блокироваться, даже когда устройство офлайн: есть настройка, которая позволяет автоматическую блокировку определенных данных если ноутбук долгое время «не выходит на связь». Понятно, что злоумышленник попытается прочитать данные с извлеченного жесткого диска компьютера, но здесь нас спасет шифрование (возможно выборочное) файловой системы.

Кроме бэкапа, Simrana EDGE предлагает еще и совместную работу над документами: можно создавать общие папки, доступные рабочим группам, можно делать синхронизацию папок между несколькими устройствами одного пользователя. Благодаря мощному web-интерфейсу, вместо того, чтобы передавать

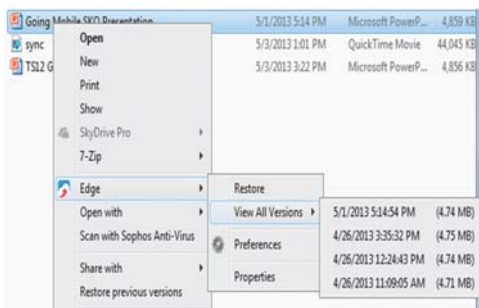


Рис. 9. Плагин EDGE для Explorer — восстановление файлов напрямую из проводника.

большие файлы по почте, можно обмениваться уникальными ссылками на файлы, которые затем загружаются получателем с хранилища Simrana.

Восстановление файлов можно делать напрямую из Explorer — часть общей концепции самообслуживания CommVault (рис. 9).

И неважно, из какой точки мира и с какого устройства вам понадобится доступ к данным — web-доступ всегда под рукой, и любая версия любого документа доступна для загрузки, в том числе через мобильное приложение для iOS/ Android. В сочетании с индексированием данных, гибким поиском, дедупликацией, файлобменным сервисом можно говорить о Simrana EDGE как о полноценном продукте по защите и организации совместной работы. И совершенно не важно, если в качестве основного ПО резервного копирования используется продукт другого вендора — Simrana полностью программно и аппаратно независимая среда, для работы которой требуется любое пространство хранения и небольшие вычислительные ресурсы x86.

Другие полезные инструменты

Simrana workflow tool — модуль, позволяющий создание автоматических сценариев с помощью наглядных диаграмм без необходимости написания сложных скриптов. Автоматизировать можно перемещение данных между уровнями хранения, регистрацию новых клиентов и другие задачи (рис. 10).

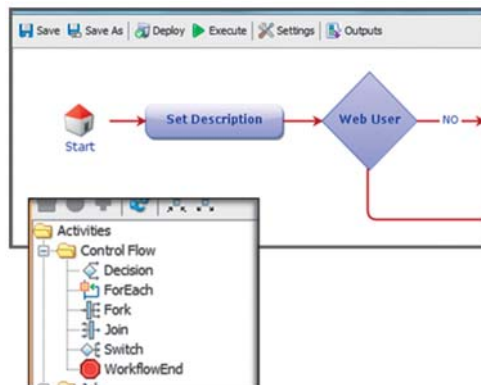


Рис. 10. Пример простейшей алгоритмизации процесса РК в Simrana и представление в виде блок-схем.

В дополнение к intellisnap в версии 10R2 появилась возможность упростить операции с мгновенными снимками в определенных ЦОД.

Значительно усовершенствовался функционал сбора информации и аналитики файловых серверов и почтовых серверов — **system discovery tools for Files and Exchange**. Эти утилиты позволяют ответить на вопрос: сколько дискового пространства на продуктивных СХД можно сэкономить, включив архивацию почты/файлов, и вообще — есть ли в этом смысл.

CommVault помогает мигрировать на Simrana с бэкап-инфраструктур других вендоров. Функционал называется **EDC** (External Data Collector). В версии R2 расширилось количество поддерживаемых вендоров.

Заключение

Simrana обеспечивает единую архитектуру с единой консолью администрирования и единым механизмом отчетности для всех типов резервного копирования; архивирования; синхронизации и разделения файлов и поиска. После улучшения масштабируемости в Simrana 10 она все чаще успешно используется в очень крупных, территориально распределенных компаниях.

Продукты и услуги CommVault будут полезны широкому кругу заказчиков, которые беспокоятся о надежном хранении информации, хотят снизить риски потери данных в случае сбоя любого программного или аппаратного компонента ИТ-инфраструктуры.

CommVault Simrana это намного больше, чем просто резервное копирование и архивирование. Это полноценное ПО по управлению жизненным циклом информации (включая контекстный поиск, удаленная репликация с дедупликацией, архивирование баз данных — то, что не попало в этот обзор), функционал которого постоянно расширяется.

*Евгений Пухов,
CommVault Россия и СНГ.*