Управление доступностью данных

Излагается методология подразделения Global Services компании Symantec управления доступностью (защитой) данных и приложений в соответствии с их ценностью и требованиями бизнеса. Данная публикация является продолжением серии материалов, посвященных современным методам защиты данных (SN № 2/27, 2006 – "Методы защиты данных: обзор решений", "Правильное архивирование данных: теория и практика", "Удаленная репликация: критерии выбора" и др.).

Введение

Обеспечение защиты данных и, соответственно, поддержание доступности приложений – одно из основных условий успешности бизнеса любой компании.

Со временем требования к непрерывности только возрастают. Так, по данным одного из ведущих авиаперевозчиков США, только 30-минутный простой ИС оказывает минимальное влияние, но уже при простоях от 4 до 10 часов убытки возрастают до $100 млн/час, а при простоях ИС более 10 часов возникает угроза банкротства. В целом, влияние простоев ИС (для западных компаний) по отраслям промышленности можно оценить из рис. 1.

Еще лет 10 назад технологии поддержания доступности данных имели очень узкий спектр решений. В основном, это было резервирование данных на ленту и кластеризация работы приложений в горячем или активном режимах. Современные технологии поддержания доступности данных имеют гораздо более широкое семейство решений, среди которых: глобальная дедупликация данных; CDP-решения; виртуальные библиотеки; синхронная/асинхронная репликация, реализованная на сетевом уровне, уровне систем хранения, программном уровне и многое другое. И если лет 10 назад постановка задачи выбора системы резервного копирования для корпоративной инфраструктуры была вполне корректна, то сейчас она во все большей степени трансформируется в концепцию/корпоративную стратегию управления доступностью данных/приложений на основе целого семейства (в ряде случаев – гетерогенного) технологий и решений.

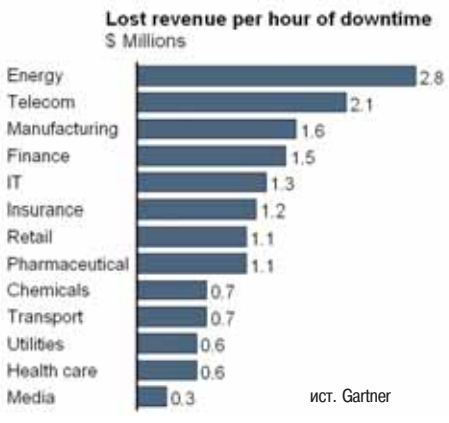

Рис. 1. Возможные потери оборота в $млн/час по отраслям для западных компаний *(ист. Gartner)*.

Методология управления доступностью данных

Основой методологии управления доступностью данных является классификация существующих технологий (рис. 2) и приложений по трем показателям: 1) RPO (recovery point objective) – средний период времени, в течение которого можно позволить потерю данных, или как часто должны выполняться резервные копии/снимки работающих приложений; 2) RTO (recovery/recall time objective) – максимально допустимое время восстановления работоспособности приложения; 3) стоимости.

Современные корпоративные ИС, как правило, имеют в своем составе множество приложений с разной степенью критичности данных. В подавляющем большинстве случаев не удается всю защиту данных реализовать на базе одного решения. Это происходит по ряду причин: из-за невозможности поддержки различных платформ на базе одного решения; неоправданно излишними затратами; эксплуатационными особенностями и др. Примерное распределение требований по доступности данных с учетом их стоимости в среде корпоративных приложений дано на рис. 3. Аналогичное распределение требований по доступности данных, но уже между отраслевыми приложениями, приведено на рис. 4.

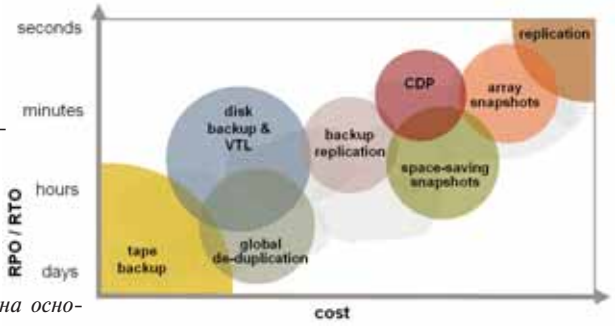

Рис. 2. Классификация основных решений обеспечения доступности по трем основным показателям: RPO/RTO и стоимости.

Укрупненно задачу управления доступностью данных/приложений корпоративной инфраструктуры можно разбить на 3 этапа: 1) определение уровней доступности данных; 2) "привязка" к ним соответствующих технологий защиты данных и приложений; 3) консолидированное управление и мониторинг заданных уровней доступности данных. Третий этап, как показывают, например, исследования Enterprise Strategy Group, становится крайне важен. В соответствии с ними, около 35% всех операций резервного копирования/восстановления (РКВ) по разным причинам не завершаются успешно. Другие исследова-

Рис. 3. Примерное распределение требований по доступности данных с учетом их стоимости в среде корпоративных приложений *(по данным Symantec)*.

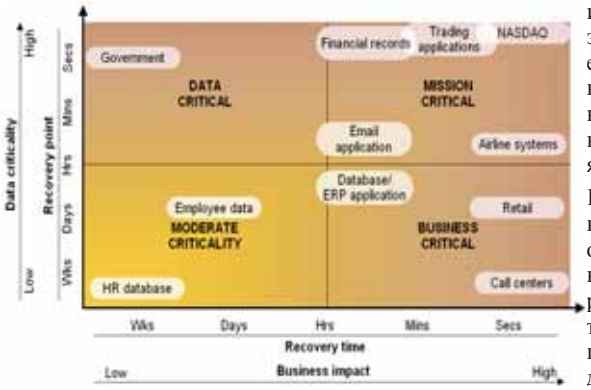

Рис. 4. Распределение требований по доступности данных, но уже между отраслевыми приложениями *(по данным Symantec)*.

ния, проведенные InfoStor, выявили следующую классификацию проблем, связанных с РКВ. 36% респондентов заявили, что они достоверно не знают, копировали ли они соответствующие данные. Для 24% опрошенных восстановление было самой большой проблемой, для 25% – большой трудностью было управление множественными устройствами резервного копирования/восстановления, а 15% заявили, что им трудно было "вписаться" в выделенное окно для ре-

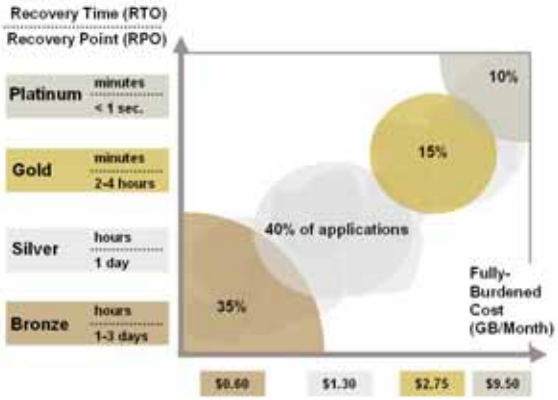

Рис. 5. Усредненное распределение четырех классов SLA ("платиновый", "золотой", "серебряный", "бронзовый") в корпоративной ИС между приложениями *(по данным Symantec)*.

зервного копирования. Помимо этого, консолидация управления РКВ позволяет снизить общие эксплуатационные расходы на РКВ, повысить управляемость операций РКВ и их прогнозируемость.

Пример задания уровней доступности данных дан на рис. 5, на котором выделено 4 уровня поддержания доступности данных: "платиновый", "золотой", "серебряный" и "бронзовый" в соответствии с требованиями доступности на основе показателей RPO/RTO и соответствия ценности информации стоимости ее хранения.

После определения уровней доступности данных к ним "привязываются" соответствующие технологии, их поддерживающие. В табл. 1 приведены типовые рекомендуемые технологии/подходы, которые могут использоваться при определении стратегии защиты данных для каждой из конкретных систем в рамках того или иного класса восстановления.

Рекомендации, приведенные в данной таблице, не означают невозможность использования других вариантов защиты данных. Кроме того, "нет" против той или иной позиции означает "по умолчанию – нет", а не "ни при каких условиях – нет"[1].

Кроме этого, необходимо отметить, что несколько вариантов подходов по защите данных могут быть использованы параллельно для наибольшей эффективности.

Табл. 1. Распределение технологий поддержания доступности данных по классам

Класс восстановления	Мгновенный снимок (Point In Time Snapshot)	Репликация на базе зеркалирования	РК по сети SAN (SAN Based Backup)	РК по сети LAN (LAN Based Backup)	РК данных на дисковую память	РК на ленты
1	Да	Да	Да	Да	Да[2]	Да
2	Нет	Нет	Нет	Да	Да	Да
3	Нет	Нет	Нет	Да	Да	Да
4	Нет	Нет	Нет	Да	Нет	Да

Табл. 2. Задание показателей RPO/RTO для "повседневного" и "аварийного" восстановления

Показатели восстановления		Класс восстановления			
		1	2	3	4
"Повседневное" восстановление	RTO	< 1 часа	< 6 часов	< 24 часов	< 72 часов
	RPO	< 1 часа	24 часа	24 часа	24 часа
"Аварийное" восстановление	RTO	4 часа	48 часа	5 дней	4 недели
	RPO	4 часа	24 часа	24 часа	72 часа

Помимо этого, как правило, определяются показатели RPO/RTO для "повседневного" и "аварийного" восстановления (табл. 2).

Управление и мониторинг заданных уровней доступности (или SLA – Service Level Agreement – соглашение о сервисных уровнях) в рамках продуктового семейства Symantec осуществляется на основе Veritas Backup Reporter (VBR).

VBR является частью платформы Veritas NetBackup и предлагает расширенную отчетность о резервном копировании для проверки соответствия требованиям и бизнес-планирования. Данное ПО предоставляет внутренним отделам (таким, как юридический или финансовый) и внешним объектам (таким, как внешний аудит, государственные регулирующие органы) отчеты, необходимые для проверки уровня защиты важных данных компании и их соответствия требованиям по показателям RTO, которые вычисляются на основе простых формул с учетом законченности операций РКВ, и RPO (мониторинг событий).

Информация, предоставляемая VBR, формируется на основе изменяемых панелей инструментов и ролевого принципа в соответствии с требованиями получателя (руководство, бизнес-пользователи, ИТ-менеджеры и др.).

В числе основных особенностей VBR можно отметить следующие:

- *подтверждение соответствия резервного копирования SLA-требованиям* – отслеживание завершенности выполнения заданий по резервному копированию для запланированных целей;
- *анализ рисков и выявление угроз* – оценка восстанавливаемости наиболее важных клиентов и приложений;
- *прогнозирование роста инфраструктуры резервного копирования* – сбор хронологических данных для определения тенденций роста и эффективного планирования ресурсов резервного копирования;
- *отслеживание использования ресурсов* – мониторинг использования ресурсов и определение ограничений, связанных с различным географическим положением и различными приложениями;
- *составление отчетов в бизнес-контексте* – изменение способа отображения данных и адаптация содержания к определенной аудитории, например, по сфере деятельности, резервному домену или приложению;
- *составление отдельных, основанных на пороговых значениях, отчетов* – осуществление контроля по методу исключения, с уведомлением только о нарушении определенных соглашений об уровне обслуживания;
- *упрощение составления отчетов о защите данных* – возможность создания сотен готовых отчетов, позволяющая начать использование функции составления отчетов немедленно, либо использование мастера создания отчетов для настройки параметров составления отчетов;
- *сбор данных без использования агентов* – обеспечение простоты установки и текущего технического обслуживания, а также того, что с важных серверов резервного копирования не снимаются "отпечатки";
- *поддержка множества приложений для резервного копирования* – централизация и нормализация процедуры составления отчетов в разнообразных средах резервного копирования, включая Veritas NetBackup, Veritas NetBackup PureDisk, Symantec Backup Exec, CommVault Galaxy, EMC NetWorker и IBM Tivoli Storage Manager.

Заключение

В современных корпоративных IT-инфраструктурах для поддержания доступности данных/приложений и выстраивания стратегии РКВ требуется консолидация многих технологий и интегрированный подход к мониторингу и управлению операциями РКВ с учетом их бизнес-критичности.

1) Для приложения, относящегося к Классу восстановления 2, чьи данные располагаются на дисковых ресурсах в сети SAN, может иметь смысл использовать технологии РК по сети SAN или Мгновенный снимок. Однако само по себе использование приложением ресурсов сети SAN не является достаточным условием для включения этого приложения в Класс 2.

2) Этот подход может быть не применим для приложений, имеющих значительные (более 1 Тбайт) объемы данных (распространяется на все Классы).