

Управление доступностью данных в концепции EMC

Статья — продолжение темы, начатой в SN № 4/33, 2007. В данной публикации излагается видение компании EMC на проблему управления доступности данных/приложений на основе развиваемых ею решений.

Введение

Системы мониторинга/управления в решениях поддержания доступности данных выполняют в них функцию обратной связи — важнейший компонент реального контроля современных высокодоступных IT-инфраструктур (отсутствие которой, например, по одной из гипотез, привело к вымиранию динозавров). И чем сложнее IT-инфраструктура, чем больше приложений она обслуживает, тем больше требований и функций она должна выполнять, как с точки зрения реального мониторинга процессов/сервисных уровней, так и прогнозирования их поведения в будущем.

Важнейшие функции систем мониторинга это фильтрация, консолидация и установление отношений между событиями IT-инфраструктуры и бизнес-процессами, на которые они влияют.

В зависимости от сложности IT-инфраструктуры и предъявляемых к ней требований по поддержанию бизнес-приложений системы мониторинга доступности данных могут быть представлены тремя типами. В первом случае — наиболее простые топологии — бывает достаточно базовых средств управления/контроля систем резервного копирования/восстановления (РКВ). Во втором случае (более 25–50 серверов приложений), как правило, уже необходимо иметь специальные средства мониторинга/управления доступности данных, которые (через агентов) интегрируются с системой РКВ и компонентами IT-инфраструктуры. В третьем случае (при наличии сотен серверов и приложений, требующих поддержания доступности на уровне трех девяток и более), разрабатываются сложные интегрированные системы из ряда решений, которые наиболее трудоемки при создании, но зато обеспечивают контроль максимального числа параметров/SLA и требуемую глубину анализа причины их отклонения (при необходимости), а также максимальную эффективность контроля IT-инфраструктуры с точки зрения поддерживающего персонала.

Классификация данных по доступности и поддерживающих их технологий

Методология классификации данных по доступности корнями уходит в концеп-

цию ILM, которую EMC начала активно продвигать около 5 лет назад. Впервые термин ILM был использован компанией StorageTek при управлении информа-

цией на ленточных накопителях. EMC взяла на вооружение эту концепцию и развила ее в ином направлении, а именно: ввела уровни обслуживания инфор-

Табл. 1. Классификация данных по доступности и в соответствии со стадиями их жизненного цикла (модель обслуживания данных).

Функция	Параметры сервиса		Уровень 1	Уровень 2	Уровень 3	Уровень 4
	Характеристика	Спецификация				
Оперативные данные	Гарантированная производительность	Пропускная способность (I/O сек)	10,000+	до 10,000	до 3,000	до 500
		Время отклика для OLTP (мс)	< 8 мс	7-14 мс	12-50мс	до 50мс
	Готовность	Максимальное время простоя в год (мин.)	< 5.26 мин/год (99.999%)	< 26 мин/год (99.995%)	до 52.6 мин/год (99.99%)	до 263 мин/год (99.95%)
Архивные данные	Производительность	Время отклика	< 5 сек	> 10 мин	> 24 часов	
		пропускная способность	>200 МБ/сек	<200 МБ/сек	<50 МБ/сек	
	Готовность	макс допустимое время простоя в год (мин) или % от рабочего времени	< 5.26 мин/год (99.999%)	до 263 мин/год (99.95%)	до 175.2 часов	
	Хранение и уничтожение	Период хранения	> 30 лет	< 10 лет	< 3-х лет	
		Гарантированное уничтожение данных после окончания периода хранения	Да	Нет	Нет	
	Доступ	частота доступа к данным	ежечасно	ежедневно	ежемесячно	
Целостность	Гарантированная аутентификация	Да	Нет	Нет		
Оперативное восстановление	Восстановление	Класс восстановления	Полное восстановление приложения	Восстановление файлов или файловой системы	Восстановление файлов или файловой системы	
		Допустимая потеря данных (ORPO)	1 час	~24 часа	7 дней	
		Целевое время восстановления (ORPO)	< 30 мин	4 часа	24 часа	
	Хранение	Период	2 часа	4 недели	15 мес	
Аварийное восстановление	Восстановление	Допустимая потеря данных (DRPO)	0 мин	1 час	>24 час	
		Целевое время восстановления (DRPO)	< 5 мин	< 1 час	> 24 час	

мации. Информация распределялась по разным уровням сервиса и была доступна для использования в соответствии с согласованными с бизнесом параметрами качества обслуживания (производительность, доступность, обеспечение целостности, масштабируемость, восстанавливаемость и пр.). В результате, это потенциально могло привести к экономии стоимости хранения данных или, по крайней мере, к обоснованным затратам на хранение.

Первые применения ИЛМ показали успешность этой модели. EMC продолжала развивать эту модель, придя в итоге к трехфазной модели ИЛМ Программы, которая стала стандартом де-факто в индустрии.

Первая фаза: группировка приложений и их данных по уровням сервиса многоуровневой СХД.

Вторая фаза: индивидуальная ИЛМ оптимизация данных приложений, особо значимых для бизнеса и оказывающих наибольшую нагрузку на СХД.

Третья фаза: создание интегрированной ИЛМ среды для большинства централизованных данных организации.

ИЛМ методология EMC учитывает индивидуальные для организации требования бизнеса, особенности ИТ-инфраструктуры и основана на лучших международных практиках оценки и построения процессов управления ИТ-сервисами (ITIL/ITSM v.3.0, Cobit 4.0, международном стандарте ISO 20000-1:2005). Требования, принимаемые во внимание, включают производительность, готовность и требования резервирования/восстановления, длительность архивного хранения, защиту от изменений и пр.

Методология такого проекта для первых двух фаз ИЛМ программы состоит из ниже перечисленных этапов:

- анализ бизнес- и ИТ-требований репрезентативной выборки 20–30 бизнес-приложений, которые отбираются согласно критериям: критичность для бизнеса, объем данных, динамика роста данных;
- группировка требований приложений и их данных в соответствии со стадиями жизненного цикла (модель обслуживания данных – табл. 1):

– **оперативные данные** – данные, которые оперативно используются бизнесом;

– **архивные данные** – данные, которые перестали активно использоваться бизнесом. Однако они должны сохраняться в неизменном виде определенное время для удовлетворения целей бизнеса и требований внешнего регулирования и законодательства;

– **данные оперативного восстановления**. Копии оперативных данных, необходимые бизнесу для восстановления бизнес-среды на определенные даты в прошлом по требованию бизнеса;

– **данные аварийного восстановления**. Копии оперативных и архивных данных, которые необходимы для восстановления работоспособности бизнеса после серьезных аварий. Эти данные

представляют собой минимальный объем критически важной информации, необходимой бизнесу для поддержки критических бизнес-процессов в аварийных условиях функционирования;

- построение оптимизированной модели обслуживания данных, включая согласованные с бизнесом политики перемещения данных по уровням сервисов хранения в соответствии с фазами их жизненного цикла. Эта модель не привязана к конкретным технологиям и вендорам, что позволяет на ее основе объективно сформировать структурированные тендерные задания для каждого уровня сервиса хранения;
- соотнесение требований приложений с классами сервисов оптимизированной модели обслуживания данных;
- анализ текущих возможностей инфраструктуры хранения с учетом изменения требований на 24–36 месяцев. Также на этом этапе проводится исследование существующих затрат и бизнес-рисков, связанных с использованием имеющихся решений;
- построение каталога сервисов СХД (нейтрального по отношению к технологиям и вендорам) для существующих и новых приложений.

Хотя технологии, обеспечивающие доступность данных, и технологии, позволяющие оптимизировать стоимость хранения данных в соответствии с их ценностью (ИЛМ), являются достаточно жестко связанными, первые, как правило, более приоритетные по сравнению со вторыми. И в дальнейшем основное внимание будет уделяться именно первым.

Основой для построения “модели обслуживания данных” по поддержанию доступности данных на уровне организации служит группировка требований приложений по ряду параметров: производительности, доступности, восстанавливаемости, архивному хранению. Формализуя эти группы, можно предложить адекватные технологические решения по обеспечению удовлетворения требований, например, допустимой по времени потери данных/времени по восстановлению данных (RPO/RTO). Третьим

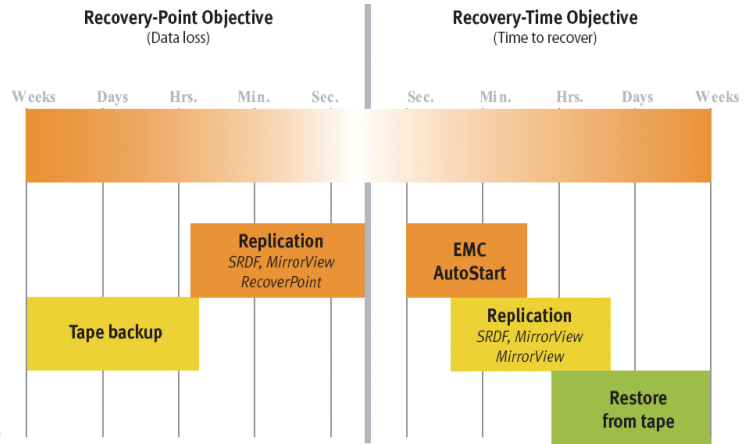


Рис. 1. Классификация технологий по поддержанию уровня доступности данных по показателям RPO/RTO.

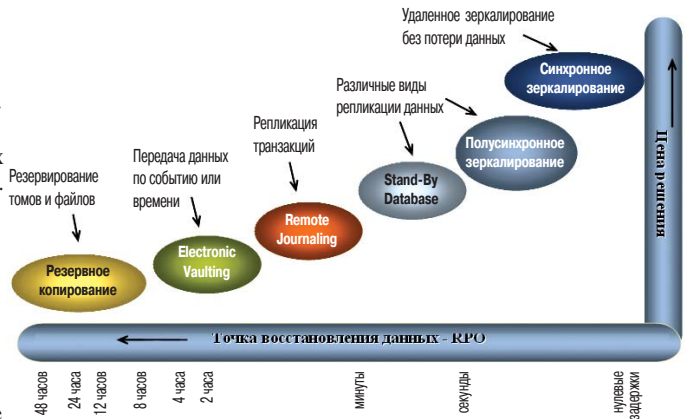


Рис. 2. Классификация технологий по поддержанию уровня доступности данных по показателям RPO/RTO и цене решения.

из наиболее важных критериев, определяющих выбор технологий, является, безусловно, цена решения (рис. 1, 2).

Необходимо отметить, что под терминами “резервное копирование”, “tape backup”, “restore from tape” в настоящее время “скрывается” целое семейство решений, которые объединяются только методом защиты данных. Так, сейчас, например, активно развиваются решения по дедупликации данных при резервном копировании, среди которых выделяются “решения по дедупликации на источнике” и “решения по дедупликации на целевом устройстве”. Большое развитие получили дисковые библиотеки, эмулирующие ленточные, в состав которых также активно интегрируются технологии дедупликации. При этом на рынке эти технологии предлагаются в виде интегрированных программно-аппаратных решений, чисто программных, различных специализированных appliance и др. для самых различных схем использования в составе полного интегрированного решения. Следует также не забывать и о различных политиках лицензирования. В последние 1,5 года активно продвигается т.н. технология continuous data protector (CDP, в EMC это решение RecoverPoint), позволяющее сделать показатель RPO в пределах до минут и даже секунд доступным для “средних” компаний и с возможностью поддержания удаленной репликации.

Очень быстро в последнее время развиваются технологии поддержания доступности данных в виртуальных средах на

базе стандартных серверов, как со стороны разработчиков ПО виртуализации (например, анонсирование опции VMware Site Recovery Manager), так и со стороны разработчиков СХД, которые уже сейчас дают возможность выполнения на их базе критичных приложений, ранее устанавливаемых только на UNIX-серверы.

В целом, в ряде случаев, новые технологии дают возможность полностью пересмотреть архитектуру решений по поддержанию доступности данных и вывести ее на новый уровень, ранее доступный исключительно в составе корпоративных инфраструктур.

Компоненты системы EMC для мониторинга доступности данных

В общем случае “полная” система мониторинга/управления доступностью данных EMC строится на базе двух ключевых компонент: решения Backup Advisor и семейства продуктов Smarts (рис. 3). При этом, как уже упоминалось, потребность в таких законченных системах возникает, в основном, только для больших корпоративных IT-инфраструктур.

Ключевая роль в такой системе в настоящее время отводится решению Backup Advisor и не только как средству мониторинга параметров и процедур РКВ, но как инструменту для их проактивного анализа.

EMC Backup Advisor

Сервисы РКВ остаются одним из наиболее важных направлений поддержания доступности данных, “вокруг” которых активно развиваются современные технологии с ними связанные и все еще остающиеся одним из наиболее ненадежных компонентов систем поддержания доступности данных (согласно оценке Enterprise Storage Group, почти 25% всех операций резервного копирования завершаются с ошибкой).

Backup Advisor имеет богатый функционал, позволяющий ему выполнять множество функций, основные среди них:

- *создание отчетов обо всех аспектах работы системы РКВ* – наличие гибких средств составления отчетов и механизма их периодической публикации;
- *мониторинг* – сбор сведений и информирование о проблемах, связанных с производительностью, неправильной конфигурацией, нетипичных объемах или других отклонениях в работе системы резервного копирования;
- *устранение неполадок* – возможность сопоставления различных компонент среды РКВ, обнаружения потенциальных узких мест и детализация данных для определения причины проблем;
- *управление производительностью* – отображение широкого спектра показателей (системной производительности серверов резервного копирования и устройств, на которые производится копирование);
- *планирование емкостей* – подсчет числа доступных лент/виртуальных картриджей и экстраполяция их числа с точки зрения потребностей;
- *отчетность по сервисным соглашениям* – отчеты об оказанных услугах и их стоимости; предупреждение о возможных будущих нарушениях соглашений.

Главное достоинство решения Backup Advisor состоит в его возможности непрерывно собирать большие объемы сведений о среде резервного копирования, анализировать их и представлять в виде легко читаемых таблиц, графиков и диаграмм. Проводимый глубокий анализ информации и сопоставление определенных показателей позволяет не только отобразить то, что происходит с операциями резервного копирования и восстановления, но и подготовить возможность быстрой идентификации и реакции на любые будущие события (рис. 4).

В ограниченных серверных инфраструктурах (с числом серверов от 30 до 200 и соответствующим числом приложений) Backup Advisor может выступать в качестве достаточного решения мониторинга доступности данных по каждому приложению, позволяя отслеживать состояние сервисных соглашений, относящихся к работе системы резервного копирования, для различных бизнес-приложений предприятия, которые выступают в роли потребителя услуг защиты информации. В качестве параметров для соглашений можно использовать основные метрики системы резервного копирования, такие как параметр RPO и RTO, определяющие допустимые потери данных и скорость их восстановления, а также задать расписание действия этих соглашений. Backup Advisor позволяет отслеживать выполнение существующих соглашений и предоставляет отчетность о качестве их выполнения и статистику по ним.

Backup Advisor интегрирован с рядом программных и аппаратных решений, поддерживающих сервисы РКВ, как от EMC, так и других производителей, в частности, с: EMC NetWorker, EMC Avamar, EMC Disk Library, Symantec Veritas NetBackup/Backup Exec, CommVault Galaxy, Tivoli Storage Manager, HP Data Protector, CA ARCserve, Oracle RMAN.

В целом, применение таких специализированных средств сбора и анализа данных о событиях сервисов РКВ как EMC Backup Advisor позволяет:

- выявить риски в существующей стратегии резервного копирования;
- сократить время, необходимое на устранение проблем резервного копирования;
- увеличить производительность и сократить периоды резервного копирования, снизить затраты на программное и аппаратное обеспечение, а также на носители;

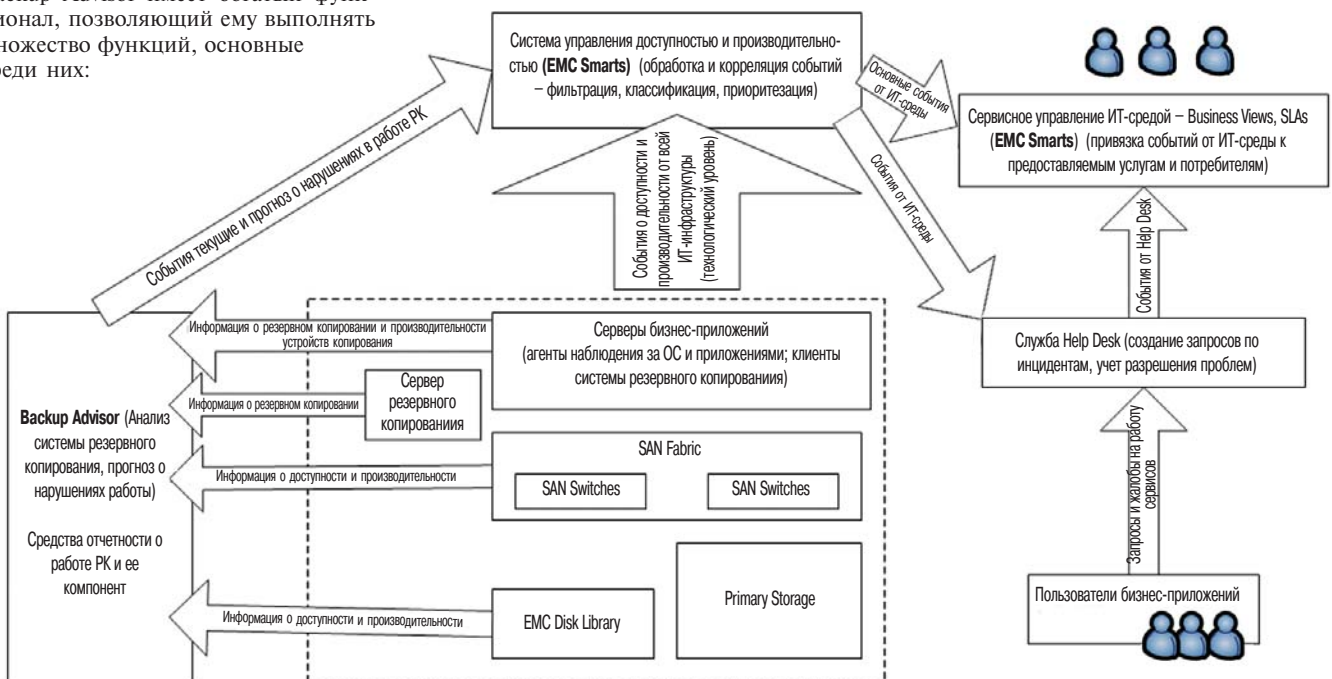


Рис. 3. Архитектура комплексной системы мониторинга/управления доступностью данных EMC.

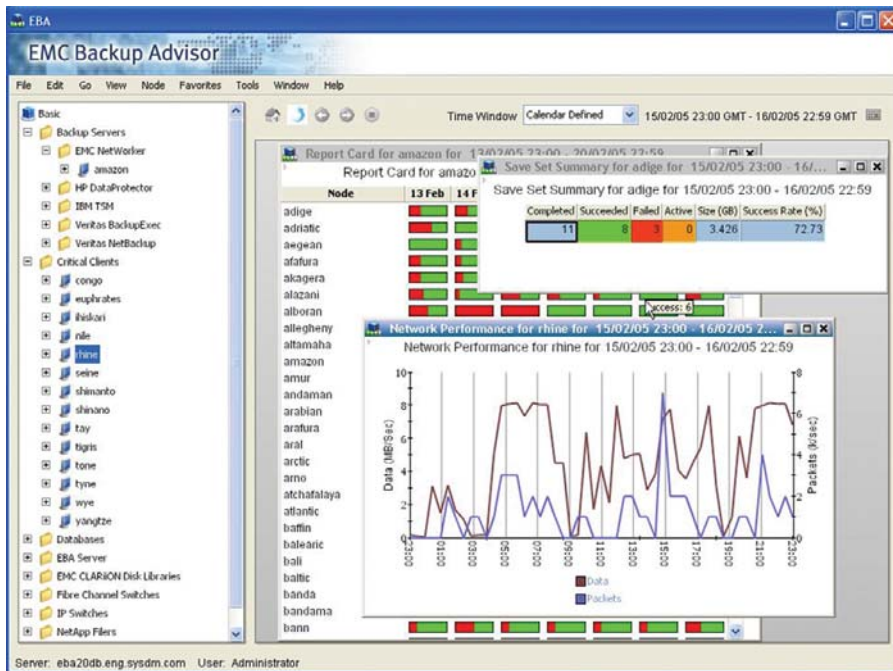


Рис. 4. Backup Advisor непрерывно собирает большие объемы сведений о среде резервного копирования, анализирует их и представляет в виде легко читаемых таблиц, графиков и диаграмм.

- планировать рост инфраструктуры резервного копирования.

Компоненты для мониторинга доступности данных семейства Smarts

Семейство продуктов EMC Smarts в значительной степени позволяет расширить функциональные возможности Backup Advisor при построении полных сервис-ориентированных систем для мониторинга функционирования бизнес-процессов и связанной с ними ИТ-инфраструктуры в корпоративной среде.

Семейство продуктов EMC Smarts ориентировано для построения современ-

ных систем управления всеми аспектами современной ИТ-среды на основе лучших международных практик оценки и построения процессов управления ИТ-сервисами, начиная от наблюдения за технологическим уровнем (события от оборудования, операционных систем и платформ приложений) до построения качественного презентационного уровня, отражающего состояние бизнес-приложений и сервисов.

Информация, собираемая на технологическом уровне (состояние систем хранения и их производительность, состояние задач репликации, состояние высокодоступных систем – кластеры для бизнес-

приложений и др.), консолидируется, фильтруется и передается штатными средствами на следующий уровень управления – сервисный, где увязывается с состоянием бизнес-процессов.

Для отображения качественного состояния сервисов предназначены модули – Smarts Service Assurance Manager (SAM) и Smarts Business Impact Manager (BIM).

SAM является интеллектуальной шиной и управляющей системой (“менеджером менеджеров”) над всеми остальными модулями (подсистемами), работающими непосредственно с технологическими доменами. SAM самостоятельно не осуществляет процесс обнаружения (Discovery), агрегирует топологические, событийные и прочие данные из других модулей (подсистем) и является обязательной для установки.

SAM включает в себя важнейшие компоненты, ответственные за представление информации пользователям, такие как интегрированная подсистема Global Manager, которая, в свою очередь, поддерживает пользовательские консоли – EMC Smarts Global Console и/или Business Dashboard. Благодаря интегрированному модулю BIM, SAM дает возможность видеть сквозное влияние выявленных инфраструктурных проблем на реальные бизнес-сервисы, процессы и бизнес.

Вместо заключения

Развитие систем мониторинга/управления доступностью данных позволяет вывести бизнес даже относительно небольших и средних компаний на качественно новый уровень адаптивности постоянно возрастающим требованиям к присутствию на рынке, а также сделать бизнес значительно более устойчивым и предсказуемым.

“EMC & SAP Integration Day”

Май 2008 г. – Российское представительство корпорации EMC провело совместно с Fujitsu Siemens Computers “ежегодный технологический день” – “Решения EMC для оптимизации информационной инфраструктуры и расширения возможностей бизнес-приложений SAP”. EMC и SAP сотрудничают с 1997 года и имеют более 5 000 общих заказчиков.

Одна из тем, обсуждавшихся на конференции, – “Виртуальные ЦОД на базе решений VMware для SAP-решений”.

Принципиальные улучшения в составе продуктов VMware, произошедшие в последнее время, позволили на новом уровне строить SAP-системы на их основе. В частности, это: автоматический мониторинг ресурсов (динамическая балансировка, постоянная оптимизация без остановки SAP-приложений); постоянная доступность (на основе полной автоматизации и поддержкой катастрофоустойчивости на базе опции VMware

Site Recovery Manager и решений от EMC); производительность по требованию (масштабируемость, гибкость, реконфигурация). Другие причины для внедрения в SAP-решения виртуальных инфраструктур: возможность обслуживания в 3 раза больше серверов; в 5-10 раз меньшее потребление электроэнергии, “холода” и площади в ЦОД; она на

порядок быстрее в обслуживании, уже надежна почти как мэйнфреймы и упрощает расчет эффективности (chargeback); возможность продления жизни “старых” приложений на новом оборудовании (жизненный цикл приложений – 4–8 лет и более, ОС и драйверов устройств – 3–5 лет, оборудования – 2–4 года).

Виртуальный ЦОД SAP – с полной поддержкой от SAP

