

Удаленная репликация: критерии выбора

В статье рассматриваются особенности построения отказоустойчивых решений на основе синхронной и асинхронной репликации, реализуемых средствами систем хранения. Приводятся основные критерии выбора отдельных компонент и всего решения в целом.

Введение

Технологии удаленной репликации на базе FC систем хранения на текущий момент остаются основным инструментом обеспечения катастрофоустойчивости центров данных. Все ведущие поставщики SAN ориентированных дисковых систем предлагают их в различных вариациях в своих портфелях предложений. Примерами являются:

- синхронная удаленная репликация — Hitachi TrueCopy Remote Replication, IBM Metro Mirror (Synchronous PPRC), EMC MirrorView/Synchronous, EMC SRDF Synchronous;
- асинхронная удаленная репликация — IBM XRC, IBM Global Copy (PPRC-XD), IBM Global Mirror (Asynchronous PPRC), IBM Metro/Global Copy (Asynchronous Cascading PPRC), Hitachi Compatible Replication for IBM® XRC (for z/OS® u OS/390®), Hitachi TrueCopy Asynchronous, Hitachi Universal Replicator for TagmaStore Universal Storage Platform.

Эти технологии, являясь ядром непрерывности бизнеса, при современных темпах внедрения широкополосной связи уже в ближайшей перспективе могут стать “массовым” явлением не только в секторе high-end рынка, но и для среднего и малого бизнеса.

Современность и репликация данных

Основная технология, используемая для организации устойчивых к единичным сбоям вычислительных систем, — репликация информации в реальном времени. Практически все производители систем хранения данных сегодня предлагают технологии для организации репликации между удаленными узлами, восстановления и перезапуска системы после серьезных аварий. Эти решения позволяют максимально оперативно принять меры по восстановлению работоспособности системы. Они дают возможность в течение нескольких минут с момента аварии запустить базу данных и приложения на удаленном сайте и восстановить подключения пользователей. Репликацию можно также использовать для быстрого создания на аппаратном уровне надежных реплик базы данных, расположенных на другом узле, в другом городе или на другом континенте. Механизм репликации можно использовать, к примеру, в приложениях для создания отчетов. Репликация позволяет обновлять данные, выполнять операции выгрузки данных при резервном копировании в безопасное место, импортировать информацию в хранилище данных и т.д.

Если расстояние между центрами данных невелико, существует возможность реализовать схему синхронной реплика-

ции, обеспечивающую “зеркальное” копирование данных с локального узла на удаленный в режиме реального времени. При увеличении расстояния асинхронная схема, с учетом соображений минимизации задержки и обеспечения достаточной производительности, становится более жизнеспособной.

Основная проблема заключается в том, чтобы обеспечить логическую целостность данных на удаленном узле. Как синхронный, так и асинхронный механизмы репликации обеспечивают последовательный порядок записи данных, при котором целостность данных поддерживается благодаря синхронизации каждой последовательной операции ввода-вывода (в случае синхронной репликации) или синхронизации циклических сегментов данных (в случае асинхронной репликации). Кроме того, в определенных случаях, например, если основной центр обработки данных не имеет физических повреждений, исходная сторона операции копирования данных может работать с данными, хранящимися на удаленном зеркальном узле, без переноса рабочей нагрузки на удаленный узел. Если же основному узлу причинен реальный физический ущерб, операции перезапуска баз данных на удаленном узле выполняются без задержек, связанных с поиском и сортировкой в правильной последовательности магнитных лент. Благодаря тому, что на удаленном узле реплики

данных хранятся на физических дисках, процедура возвращения рабочей нагрузки на основной узел также значительно упрощается и ускоряется.

Продукты репликации средствами систем хранения реализуют следующие функции:

- независимое от сервера удаленное реплицирование данных в режиме реального времени;
- восстановление после плановых отключений;
- перезапуск после аварии;
- тестирование перезапуска после аварии;
- поддержка параллелизма и совместимости с подавляющим большинством промышленных операционных систем и Linux;
- поддержка миграции центра данных;
- технология обеспечения целостности для повторного использования приложений;
- интеграция с семейством продуктов для локальной репликации.

Продукты семейства репликации позволяют защитить данные, хранящиеся в центре обработки данных, равно как и в распределенных системах при возникновении непредвиденных аварийных ситуаций, а также в случае запланированных отключений системы, например, при проведении профилактических работ.

Традиционные методы защиты информации, обычно основанные на резервном копировании данных на магнитную ленту, обеспечивают адекватный уровень защиты для некоторых видов информации. Однако критичные для бизнеса базы данных, информационные хранилища и информационные ресурсы масштаба предприятия имеют настолько важное значение — не только для обеспечения жизнеспособности бизнеса, но и для самого его существования — что требует более эффективных методов защиты. Hi-End решение для безотказной работы с данными, обеспечивающее не только сохранность данных и возможность восстановления после сбоев, но и предлагающее средства поддержания непрерывности деловых операций, становится необходимостью во многих отраслях бизнеса. Решая эту задачу, частные компании и государственные учреждения все чаще прибегают к решениям, базирующимся на использовании технологий удаленного зеркалирования. Удаленное зеркалирование позволяет экономить время, деньги и силы, затрачиваемые на восстановление данных, по сравнению с традиционными методами восстановления после аварий. При использовании удаленного зеркалирования время восстановления работоспособности системы значительно сокращается; при этом предотвращаются потери информации, т.к. происходит оперативное восстановление данных с откатом к заранее заданному, максимально близкому к точке аварии, моменту времени. Пользователи получают возможность продолжать деловые операции, а негативный финансовый эффект от

запланированных и незапланированных отключений значительно уменьшается или вообще сводится к нулю.

Решения репликации средствами систем хранения помогают компаниям решать проблемы, связанные с непредвиденными происшествиями и запланированными отключениями. Каждое решение репликации позволяет организовать постоянно доступное, независимое от типа хост-системы, информационное хранилище с дублированием производственных данных на одном или нескольких физически разделенных целевых системах хранения. Эти системы могут быть сколь угодно удалены друг от друга — они могут находиться в одной комнате или в разных частях света. Но, независимо от расположения систем хранения, решения репликации позволяют бизнесу функционировать в режиме 24x7 и поддерживать высокий уровень обслуживания клиентов, дающий дополнительные преимущества перед конкурентами.

Синхронная репликация средствами систем хранения

Синхронная репликация — это решение для организации постоянно доступного, независимого от хост-системы хранилища с зеркальным резервированием, обеспечивающее дублирование производственных данных на одной или нескольких физически разделенных целевых системах хранения. Она является наиболее высокопроизводительным и устойчивым решением для защиты данных из имеющихся сегодня на рынке. Это решение идеально подходит для тех случаев, когда сервисные уровни требуют 100%-й защиты данных (полного отсутствия потерь).

Синхронная репликация имеет следующие основные преимущества:

- способность быстро и эффективно реплицировать критически важные данные с одного узла на другой при 100%-й защите данных;
- уменьшение негативного эффекта от запланированных эксплуатационных процедур;
- возможность тестирования процесса восстановления работоспособности после сбоев;
- легкий и удобный перезапуск бизнес-процедур после незапланированных отключений;
- дополнительные функции, например, динамический перенос рабочей нагрузки с исходного узла на целевой и обратно, возможность одновременной поддержки нескольких переключаемых конфигураций;
- одновременное функционирование (на одних и тех же томах)

синхронной и асинхронной схем репликации;

- одновременное функционирование (в рамках одной системы) разных схем репликации;
- дополнительные режимы синхронной репликации (полусинхронная репликация и репликация с адаптивным копированием) предлагают пользователям дополнительные возможности выбора.

Перезапуск системы после аварий

Синхронная репликация позволяет возобновить бизнес-операции после аварии гораздо быстрее, чем любое другое решение репликации. При синхронной репликации пользователи имеют возможность просто перезапустить операционные системы вместо того, чтобы заниматься их восстановлением. Это означает, что приложения можно перезапустить сразу же после аварии; при этом будут сохранены все данные, записанные на момент последней подтвержденной операции ввода-вывода перед аварией. Традиционные методы восстановления данных и приложений часто предусматривают довольно долгую процедуру.

В большинстве случаев резервные копии, обычно хранящиеся на магнитных лентах с достаточно большим временем доступа, необходимо передать на узел, на котором будет происходить восстановление данных, либо по электронным каналам, либо на физических носителях. Следующий шаг — довольно длительный процесс восстановления данных и анализа логов. Стандартная практика предусматривает создание резервных копий один или несколько раз в неделю. Для современных компаний, действующих в глобальном масштабе, эта практи-



NVisionGroup ЭНВИЖН ГРУП Новое время – новое видение

Искусство организации информационных потоков

- ИТ-консалтинг
- ИТ-инфраструктура
- Телекоммуникации
- Системы управления (OSS/BSS)
- Информационные системы (ERP, SRM, CRM, SCM, BI, EAM)
- Информационная безопасность
- Видеоконференцсвязь
- Сервисная поддержка

115054, Москва, ул. Дубининская, д. 53, стр.5
телефон: (495) 641-1212, факс: (495) 641-1211
e-mail: info@nvisiongroup.ru, http://www.nvisiongroup.ru

ка превратилась в настоящую головную боль. Приложения теперь занимают больше места, времени для резервного копирования все меньше, а любые простои приводят к потере дохода или ухудшению качества сервиса. Синхронная репликация помогает избавиться от этих проблем и обеспечить непрерывность деловых операций. Запланированные операции резервного копирования, загрузка данных в информационное хранилище или тестирование новых приложений могут выполняться с данными, находящимися на целевом узле СХД, в то время как на исходном узле СХД обработка данных осуществляется в обычном режиме. При этом в любой момент времени данные надежно защищены, доступны и могут обрабатываться бизнес-приложениями.

Асинхронная репликация средствами систем хранения

Асинхронная репликация представляет собой наиболее производительное решение для репликации данных на большие расстояния среди представленных на рынке. Она позволяет снизить совокупную стоимость владения (ТСО) за счет уменьшения общей требуемой пропускной способности и является оптимальным решением для репликации данных на большие расстояния в тех ситуациях, когда соображения экономии и производительности приложений имеют более важное значение для обеспечения должного уровня сервиса, чем 100%-я сохранность данных. Использование технологии репликации на основе Delta Set обеспечивает по-настоящему высокопроизводительную асинхронную репликацию данных между системами хранения, находящимися на значительном удалении друг от друга. Основные преимущества асинхронной репликации следующие:

- высокопроизводительная репликация, не оказывающая негативного воздействия на скорость работы приложений в хост-системе;
- расстояние между системами может быть сколь угодно большим;
- наличие в любой момент времени целостных удаленных копий, позволяющих осуществить перезапуск системы;
- эффективное использование имеющейся полосы пропускания благодаря технологии Delta Set и механизмам усреднения операций ввода-вывода;
- поддержка практически всего спектра операционных систем.

Современные схемы асинхронной репликации используют инновационный подход, как для открытых систем, так и для мэйнфреймов, позволяющий постоянно поддерживать доступную целостную копию производственных данных на удаленном узле, который может находиться на сколь угодно большом расстоянии от основного узла. При этом асинхронная репликация не оказывает отрицательного воздействия на работу приложений в хост-системе. В отличие от традиционных асинхронных архитектур, основанных на принципе упорядоченной записи по временным меткам, современные технологии, например SRDF и TrueCopy, используют новую ар-

хитектуру дельта-фрагментов (Delta Set) с кэшированием для отправки данных на удаленный узел, позволяющую значительно снизить требования к пропускной способности, необходимой для дельта-вида репликации. Технология дельта-фрагментов все время поддерживает атомарную копию данных на удаленном узле, обеспечивая постоянную доступность целостных копий данных, позволяющих выполнять перезагрузку системы.

Опция работы с согласованными группами (consistency group) — это опция репликации, позволяющая обеспечить целостность данных на уровне предприятия, гарантирующую успешный перезапуск приложений после аварии. Приложениям корпоративного масштаба, таким, как ERP, CRM и web-приложения для обработки транзакций, требуется все более высокий уровень координации программной обработки в условиях гетерогенных информационных сред. Корпоративные приложения предъявляют высокие требования к доступности средств координированного восстановления данных во всех информационных хранилищах, с которыми эти приложения работают. Период восстановления должен длиться недолго, и все хранилища данных должны быть восстановлены на один и тот же, синхронный момент времени, соответствующий одинаковому циклу обработки деловой информации, в противном случае восстановленные данные будут непригодны для использования.

Перезапуск приложений, защищенных с помощью опции согласованных групп, можно выполнить практически сразу после аварии, обеспечив при этом 100%-ю сохранность данных. Это обеспечивает синхронный перезапуск всех баз данных и приложений на резервном узле, вместо того, чтобы восстанавливать все многочисленные информационные компоненты по одному.

Данная опция позволяет группировать воедино различные объекты, принадлежащие как одной, так и разным системам, таким образом, что для внешней среды они предстают единым объектом или группой. Если данные с одного или более исходных устройств, входящих в согласованную группу, не могут попасть на соответствующие целевые устройства, весь поток данных, адресованный целевому устройству согласованной группы, приостанавливается. Такой подход гарантирует целостность базы данных на удаленном узле (на момент возникновения проблем с передачей данных) и, соответственно, обеспечивает возможность использования этих данных приложениями.

Выбор транспортной сети

При реализации решения на основе удаленной репликации факторы стоимости транспортной сети и технические особенности доступа к ней часто бывают или решающими, или оказывающими большое влияние на конфигурацию и построение системы. Хотя последние стандарты Fibre Channel SAN и способны к передаче данных до 500 км, реальное расстояние определяется длительностью задержки в сети, которое не приводит к сбоям в работе приложений. Для синхронного реплицирования этот предел составляет 35–50 км.

Хотя синхронное реплицирование обеспечивает самые быстрые средства восстановления приложений, на практике при больших расстояниях оно редко используется из-за высокой стоимости сетевого обеспечения связи. Для асинхронной передачи используется 2 вида сетевого транспорта: темное оптоволоконно и SONET (Synchronous Optical NETwork).

Темное стекловолокно является неразделяемым ресурсом — это частный транспорт, в отличие от общедоступных колец SONET. Из-за технических ограничений темное стекловолокно может использоваться только для относительно коротких расстояний (особенно в сравнении с SONET) и, в основном, в конфигурациях “точка-точка”.

SONET (SDH в Европе) — стандарт для транспортировки данных в публичных телекоммуникационных кольцах. Полоса пропускания в SONET определяется договором с поставщиком услуг и, как правило, допускает масштабирование. Пример кольца SONET, объединяющего различные типы связей — OC1, OC3, Ethernet и др., — представлен на рис. 1.

Основные характеристики канала связи: 1) пропускная способность (Mbps или Gbps) связи определяет, какое количество

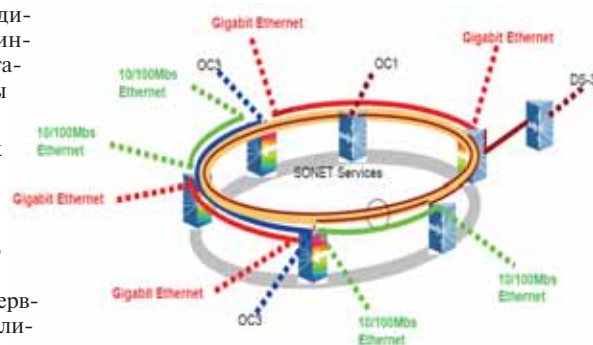


Рис. 1. Пример кольца SONET, объединяющего различные типы соединений.

данных можно транспортировать и сколько времени займет эта передача; 2) время ожидания (мс), требуемое для прохождения сигнала от одного конца до другого.

Скорости различных соединений представлены в табл. 1. Время ожидания определяется скоростью света, которая одинакова для всех типов связи и составляет примерно 1 мс на каждые 100 миль.

Выбор типа соединения обуславливается многими факторами, в частности, расстоянием, типом и объемом трафика, скоростью, типом доступа, стоимостью и т.д. На практике соединение T3 рассматривают как минимально необходи-

Табл. 1. Пропускная способность различных типов соединений

Тип связи	Пропускная способность
Copper Telco Links	
T1 / DS1	1.544 Mbps
T2	6.312 Mbps
T3 / DS3	44.736 Mbps
Optical Telco Link s	
OC-1	51.84 Mbps
OC-3	155.52 Mbps
OC-12	622.08 Mbps
OC-24	1.244 Gbps
OC-48	2.488 Gbps
OC-192	9.6 Gbps
Fiber Channel / FICON	1/2/4 Gbps
Ethernet Links	
10Base-T	10 Mbps
100Base-T	100 Mbps
Gigabit Ethernet	1 Gbps

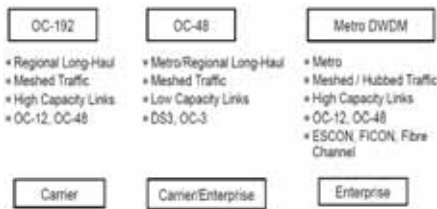


Рис. 2. Характеристики трех сетевых технологий, классифицированные по сценарию их использования и сетевому провайдеру.

мое для реплицирования данных. Самые большие компании требуют значительно большей полосы пропускания — OC3, OC12, OC48 или OC192 (с пропускной способностью от 155 Mbit/sec — OC3 — до 9,6 Gbit/sec — OC192). На рис. 2 представлено “резюме” для трех сетевых технологий, а также сценарий их использования и сетевой провайдер.

До последнего времени даже не столь высокопроизводительные сети типа T3 требовали достаточно больших арендных платежей, в зависимости от поставщика, в диапазоне \$12–21 тыс., что являлось весьма сдерживающим фактором для средних и небольших компаний.

В настоящий момент в качестве решения такой “транспортной проблемы” стали предлагать решения на основе широкополосной беспроводной связи в нелицензируемом радиочастотном спектре (5,725–5,825 GHz), работающие в диапазоне до 80 км со скоростью до 48 Мбит/с (что превосходит T3 “наземного” исполнения) и значительно более экономичнее, чем T3. Примером такого решения может служить совместная разработка для удаленной синхронной репликации данных для SMB-рынка компаний Hitachi Data Systems, McDATA и Redline Communications (рис. 3). В состав решения входят:

- модульные дисковые системы HDS AMS-типа и ПО Hitachi TrueCopy (для удаленной синхронной репликации);
- McDATA Eclipse 1620 SAN Router, позволяющий двум FC дисковым системам использовать Ethernet-сеть (Fast/Gigabit Ethernet) для репликации;
- оборудование Redline Communications — BFW Microwave — для широкополосной передачи данных.

Большое значение при доступе к транспортной сети имеют процедуры компрессии и оптимизации трафика, которые могут значительно улучшить использование доступной сетевой полосы пропускания. Наилучшим способом это достигается аппаратно, в частности, в рассмотренном выше решении (см. рис. 3)

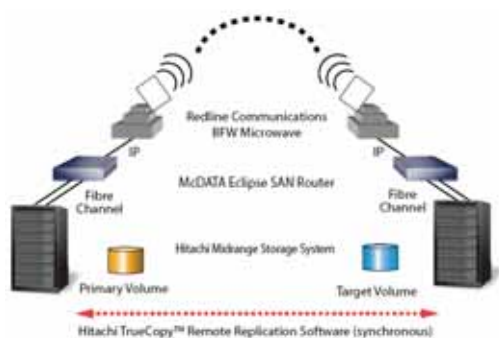


Рис. 3. Решение для удаленной синхронной репликации с использованием широкополосной радиочастотной связи.

это достигается встроенными средствами SAN-маршрутизатора, который использует интеллектуальные подпрограммы сжатия с перестраиваемой конфигурацией для динамической оптимизации трафика на порту Ethernet с целью обеспечения максимального использования полосы пропускания. Это достигается как компрессией данных, так и за счет пакетирования данных в большие фреймы и устранения излишних команд подтверждения.

Выбор технического решения

При выборе конкретного технического решения для организации Disaster Recovery между двумя площадками (ЦОД), как правило, необходимо иметь в виду следующие аспекты:

- наличие детальной схемы организации связи между площадками: какие используются физические каналы связи, какая скорость передачи данных по ним возможна, какая возникает задержка в распространении сигнала;
- проведение оценки совместимости оборудования и прикладного/системного ПО, используемого на серверах, работающих с СХД;
- необходимость определения того, какие данные будут реплицироваться; какие приложения следует иметь в качестве резерва для второй площадки; все ли серверы будут дублироваться; будут ли дополнительные серверы на резервной площадке, данные которых надо будет дублировать на основную площадку;
- использование для отказоустойчивости приложений (по крайней мере Exchange, Oracle, Lotus) кластерного софта, например (Veritas Cluster Server, EMC Legato AutoStart, IBM HACMP);
- ожидаемый уровень нагрузки на дисковую подсистему (I/O, количество операций, Мбайт/с — по каждому серверу (чтение и запись отдельно).

Решения репликации для обеспечения непрерывности бизнеса позволяют подобрать наиболее адекватный сервисный уровень защиты информации, соответствующий деловым целям организации. Такой подход минимизирует общий объем затрат благодаря многократному использованию оптимальных методик, включающих полный набор таких компонентов, как хранилище данных, сетевое и коммуникационное оборудование, ПО, элементы внедрения и администрирования, необходимые для обеспечения заданных показателей точек восстановления (RPO) и времени восстановления (RTO). Решения для удаленной репликации покрывают весь спектр сервисных уровней RPO. Асинхронная репликация прекрасно подходит для тех случаев, когда репликация осуществляется на больших расстояниях, а требования к восстановлению данных допускают некоторый

Табл. 2. Обобщенное сравнение различных вариантов репликации

Вариант	Описание
Синхронная удаленная репликация Решения для RPO, равного нулю	<ul style="list-style-type: none"> • Отсутствие потерь данных, ограниченные расстояния • Наиболее высокий сервисный уровень защиты информации • Производительность решения зависит от расстояний и объемов изменений данных • Работает при максимальной пропускной способности коммуникационных каналов
Асинхронная репликация Решение для RPO, составляющего от нескольких секунд или минут до нескольких часов	<ul style="list-style-type: none"> • Неограниченное расстояние между центрами данных • Промежуточный сервисный уровень защиты информации • Производительность не зависит от расстояния • Работает при средней пропускной способности коммуникационных каналов
Асинхронная удаленная репликация в режиме создания “снимков” Решения для RPO, составляющего от 4 до 24 часов	<ul style="list-style-type: none"> • Восстановление данных с откатом от 4 до 24 часов, неограниченные расстояния • Более низкий сервисный уровень защиты информации • Если решение правильно спроектировано, его производительность не зависит от расстояния или скорости изменения данных • Работает при низкой пропускной способности коммуникационных каналов

откат во времени (точки восстановления RPO больше нуля). Обобщенное сравнение различных решений дано в табл. 2.

Заключение

Репликация средствами систем хранения является наиболее надежным методом организации зеркалирования данных. Основными ее преимуществами являются:

- прозрачность данного метода для серверов — нет необходимости проводить дополнительные настройки на серверах, использующих зеркалирующиеся данные;
- отсутствие дополнительной загрузки серверных ресурсов — все операции выполняются средствами систем хранения.

Основными недостатками репликации средствами систем хранения являются:

- невозможность репликации между массивами разных производителей — каждый производитель использует собственные протоколы репликации, не совместимые с протоколами других производителей;
- относительно высокая стоимость в сравнении с программными методами репликации.

Для выбора конкретного механизма репликации в распределенных серверных системах наиболее важными исходными данными являются:

- самая подробная информация о связи между площадками;
- ожидаемый уровень нагрузки на дисковую систему (особенно операции записи, так как передаются именно записываемые на диск данные);
- используемое оборудование и прикладное ПО — решение должно быть совместимо с ним.

Наиболее важным шагом является выбор типа репликации: синхронной или асинхронной. Синхронная репликация — более надежная, но работает только на быстрых каналах и относительно небольших расстояниях. Как правило, лучше стремиться сделать систему с синхронной репликацией, и только в случае невозможности по каким-либо причинам ее использования — выбрать асинхронную.

Александр Лебедев,
зам. руководителя отдела инфраструктурных решений, “Энвижн Груп”