СИСТЕМЫ, УСТОЙЧИВЫЕ К КАТАСТРОФАМ: ПРАКТИЧЕСКИЙ ОПЫТ РЕАЛИЗАЦИИ

Геннадий Карпов
Корпорация Квазар-Микро

Аннотация: **Изложен опыт, накопленный в процессе создания устойчивой к катастрофам системы, предназначенной для применения в банковской сфере. Система представляет собой кластер, узлы которого разнесены на значительное расстояние (географический кластер). Рассмотрены существующие технологии удаленной репликации данных. Показана важность разработки сценариев реакции на критические ситуации, учитывающих местные условия.**

Summary: **Experience of implementation of the disaster tolerant system for financial organization was described. The system is cluster with nodes placed on geographically distant location (geographical cluster). Available technologies of remote data replication were considered. Importance of development of localized disaster recovery plan was shown.**

Ключевые слова: **Катастрофоустойчивая система, географический кластер, удаленная репликация.**

I Введение

По мере повышения роли электронной обработки информации во всех областях человеческой деятельности интерес к вычислительным системам высокой надежности постоянно возрастает. Последние 5-7 лет наблюдается широкое использование кластерных технологий в различных информационных системах. Прежде всего это относится к высокодоступным кластерам (High Availability, HA), обеспечивающим устойчивость к отказам как отдельных аппаратных компонент, так и узлов (серверов) целиком.

Однако существует иной класс угроз, для защиты от которых применение традиционных кластерных решений является неэффективным. К указанному классу следует, прежде всего, отнести локальные и глобальные катастрофы, вызванные причинами природного либо антропогенного характера: пожары, затопления, нарушения электропитания, авиакатастрофы, военные действия и т. д. Неэффективность высокодоступных кластеров в указанных случаях объясняется их географической компактностью: как правило, вся система располагается в пределах одного помещения или, реже, здания. При этом в случае катастрофы одновременно разрушаются все узлы кластера и хранилище данных, что приводит к невозможности дальнейшего функционирования информационной системы.

Долгое время наличие специальных средств обеспечения целостности и доступности данных в условиях катастроф являлось прерогативой комплексов военного назначения. Большинство гражданских систем либо вообще не имело защиты от подобных ситуаций, либо довольствовалось периодическим резервным копированием с последующим перемещением съемного носителя в удаленное безопасное место. Однако в последнее время непрерывность функционирования информационной системы стала настолько критичным фактором поддержания нормальной работы во многих областях деятельности (прежде всего, финансы, телекоммуникации, системы массового обслуживания), что появилась реальная потребность в применении специальных средств обеспечения устойчивости к катастрофам.

Практически единственным приемлемым методом обеспечения катастрофоустойчивости гражданских информационных комплексов является размещение их элементов на различных площадках (site), разнесенных на достаточное расстояние. Величина указанного расстояния определяется типами аварийных ситуаций, к которым должна быть устойчива система. Например, для защиты от локального пожара или затопления достаточно разместить площадки в различных зданиях в пределах одного города. В то же время, угроза военных действий или крупных стихийных бедствий потребует разнесения площадок между различными городами или даже странами.

Подобные распределенные системы, по своей сути, являются кластерными и часто именуются в литературе географическими кластерами или геокластерами. Одним из наиболее существенных отличий геокластера от классического HA-кластера является методика обеспечения доступа различных узлов к одним и тем же данным. В HA-кластерах для достижения этой цели традиционно используется разделяемый дисковый массив. По понятным причинам, подобный подход в случае геокластера неприменим. Вместо этого применяются средства удаленной репликации данных, при этом узлы на каждой площадке обладают собственной копией данных той или иной степени актуальности. Тип удаленной репликации (синхронная или асинхронная), а также требования к пропускной способности и латентности каналов связи между площадками в огромной степени зависят от используемого приложения.

II Постановка задачи

Наша задача заключалась в проектировании и создании катастрофоустойчивой системы для одного из крупнейших украинских банков. Построенную систему предполагалось использовать как платформу для развертывания критичного финансового приложения, не допускающего потери данных даже в условиях массового отказа оборудования, вызванного внешними деструктивными факторами.

Общие требования, предъявленные к системе, были следующими:

- обеспечение целостности и полной актуальности данных как в случае отказа отдельных аппаратных компонент, так и полного физического разрушения географически компактно расположенного оборудования;

- возможность восстановления доступности данных (возобновление обслуживания клиентов) в случае возникновения аварийной ситуации за приемлемое время (не более нескольких часов);
- возможность использования для внутрикластерных коммуникаций существующей IP-сети, связывающей принадлежащие заказчику удаленные площадки;
- минимизация времени простоя системы в случае одиночных отказов аппаратуры;
- возможность полного контроля и управления системой силами специалистов заказчика.

Следует отметить, что требование *автоматического* восстановления обслуживания клиентов в аварийных ситуациях не предъявлялось. Это, прежде всего, объясняется особенностями приложения, требующего для своей работы специализированного периферийного оборудования, активация которого в резервном центре требует обязательного участия человека. Кроме того, приемлемое время восстановления сервиса также не вступает в конфликт с необходимостью вмешательства оператора.

III Выбор технологии удаленной репликации данных

Как уже подчеркивалось, удаленная репликация данных является основой геокластерной системы, во многом определяющей ее эксплуатационные характеристики, а также оборудование и программное обеспечение (ПО), необходимое для построения системы. По этой причине выбор технологии репликации исходя из предъявляемых требований является одним из важнейших этапов проектирования.

Анализ литературы, а также спектра имеющегося на рынке оборудования и ПО показал, что в настоящее время могут быть реализованы три метода удаленной репликации (рис. 1). Указанные методы отличаются структурным уровнем вычислительной системы, на котором выполняется создание удаленной зеркальной копии: контроллер дисковой подсистемы, операционная система, приложение.

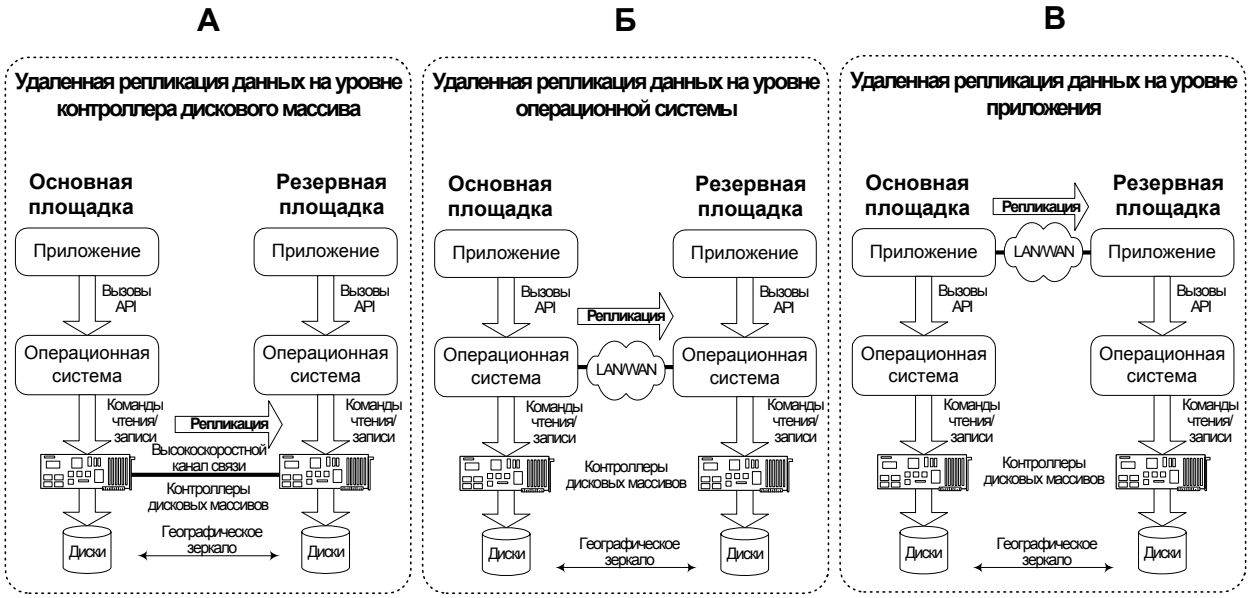

Рисунок 1 – Методы удаленной репликации данных

Несмотря на то, что все перечисленные методы решают одну задачу, каждый из них имеет ряд особенностей (табл. 1). Возможность и степень эффективности применения той или иной технологии определяется исходными условиями задачи.

Таблица 1 – Сравнительная характеристика методов удаленной репликации

Уровень репликации	Достоинства	Недостатки
Контроллер дискового массива (метод А)	не требует поддержки со стороны операционной системы и приложения; репликация возможна в синхронном и асинхронном режимах; высокая производительность	высокая стоимость оборудования; требуется наличие выделенного высокоскоростного канала связи; наличие ограничений на расстояние между площадками (как правило, не более 10 км.); при синхронизации зеркальных копий возможна избыточная репликация

Продолжение таблицы 1

Операционная система (метод Б)	не требует поддержки со стороны приложения и аппаратуры; репликация возможна в синхронном и асинхронном режимах; позволяет использовать существующие LAN/WAN сети; при синхронизации зеркальных копий реплицируются только измененные данные	поддерживается не всеми операционными системами; производительность достаточна не для всех приложений
Приложение (метод В)	не требует поддержки со стороны операционной системы и аппаратуры; позволяет использовать существующие LAN/WAN сети; при синхронизации зеркальных копий реплицируются только измененные данные	поддерживается не всеми приложениями (как правило, только промышленными СУБД); асинхронность репликации может приводить к потере последних изменений; производительность не всегда приемлема

Вполне очевидно, что в нашем случае репликация на уровне контроллера неприемлема, поскольку не позволяет использовать существующую IP-сеть, связывающую удаленные площадки. Метод "В" также оказался нереализуемым, поскольку приложение не обладает необходимой функциональностью. Следовательно, мы пришли к необходимости организации удаленного зеркального копирования на уровне операционной системы.

IV Решение

Исходя из имеющихся у заказчика ресурсов, было принято решение о создании двухузловой геокластерной системы, узлы которой расположены на площадках в различных зданиях. Исходя из предъявленных к системе требований, был выбран режим работы узлов "активный/резервный", т.е. в штатном режиме работы системы резервный узел не несет полезной нагрузки и выполняет только функции удаленной репликации данных. Упрощенная логическая схема построенного географического кластера показана на рис. 2.

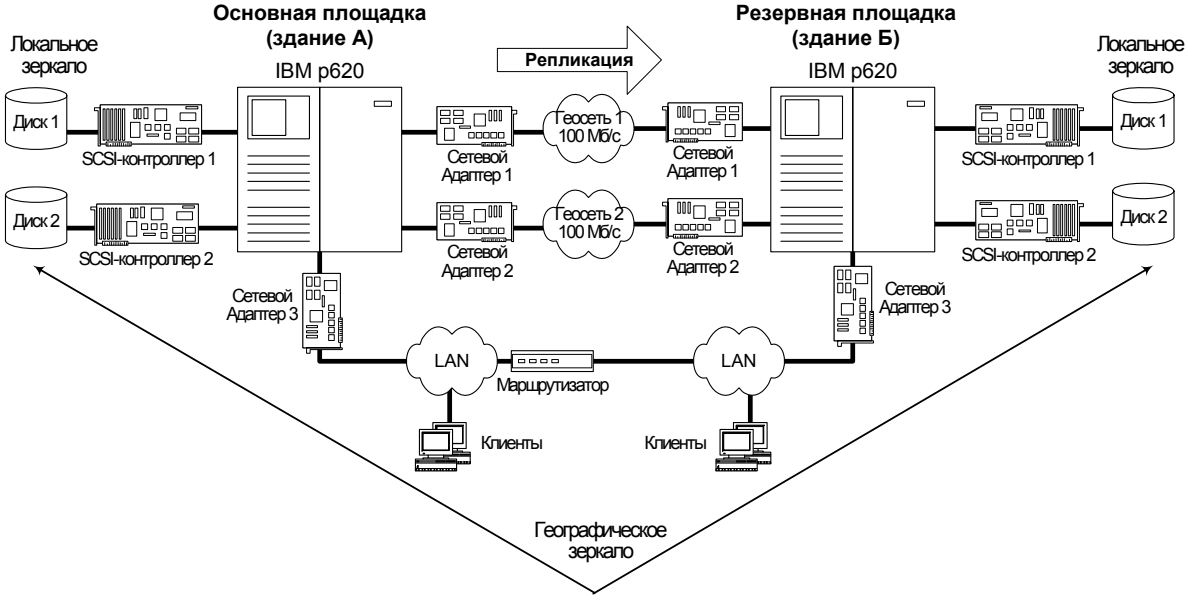

Рисунок 2 – Упрощенная логическая схема созданного географического кластера

Проанализировав совокупность поддерживаемых приложением операционных систем, мы остановились на ОС IBM AIX по следующим причинам: а) компания IBM предлагает ПО Geographic Remote Mirroring (GeoRM) [1], добавляющее в AIX функции удаленной репликации; б) версия приложения для ОС AIX является одной из наиболее активно поддерживаемых его производителем; в) AIX зарекомендовал себя как один из наиболее стабильных и хорошо масштабируемых вариантов UNIX. ПО GeoRM поддерживает режим

синхронной репликации, что, в соответствии с предъявляемыми требованиями, обеспечивает полную актуальность данных на удаленной площадке. Для обеспечения внутрикластерных коммуникаций может использоваться любая IP-сеть (LAN/WAN).

В качестве аппаратной платформы были выбраны серверы среднего уровня p620 семейства IBM pSeries (новое название серии RS/6000). Указанные сервера обладают умеренной стоимостью, достаточной для работы приложения производительностью и функциональностью, хорошо масштабируются, а также имеют встроенные средства мониторинга и резервирования аппаратных компонент.

Наличие резервной площадки позволяет восстановить сервис не только в случае катастроф, но и при отказах отдельных аппаратных компонент основного сервера (например, дискового контроллера). Однако мы пришли к выводу, что подобная защита оказывается неэффективной: активация резервного центра может занять достаточно продолжительное время и требует присутствия квалифицированного персонала. Кроме того, после восстановления работоспособности основного центра потребуется вновь его активировать (и, соответственно, деактивировать резервный), что также связано с перерывом в обслуживании. Подобный сценарий вполне приемлем при катастрофе, однако не оправдывает себя в случае локальных отказов аппаратуры.

По указанным причинам следует стремиться к обеспечению максимальной отказоустойчивости узлов (серверов) комплекса. Мы отказались от варианта использования классических HA-кластеров в качестве узлов геокластера по причинам неоправданного удорожания системы и указанных выше особенностей приложения, не допускающих полную автоматизацию процедуры восстановления. В нашем случае задача была решена путем выбора соответствующей конфигурации применяемых серверов.

Все наиболее уязвимые (имеющие механически движущиеся части, находящиеся под воздействием высокого напряжения) компоненты, такие как блоки питания, вентиляторы, диски, SCSI-контроллеры, были продублированы. Фактически, в рамках каждого узла были реализованы независимые альтернативные маршруты доступа к локальным зеркальным копиям данных по цепочке шина PCI $\Rightarrow$ SCSI-контроллер $\Rightarrow$ дисковый контейнер $\Rightarrow$ диск. Отказ любого из элементов данной цепочки незаметен для приложения и не приводит к прекращению обслуживания. Кроме того, SCSI-контроллеры и диски поддерживают режим горячей замены, позволяя устранить неисправность без остановки узла.

Использование двух геосетей, созданных на основе уже существующих коммуникаций, позволяет предотвратить нарушение процесса репликации в случае отказа сетевого оборудования. Кроме того, подобное решение позволяет увеличить производительность системы за счет распределения трафика между сетями. Следует отметить, что в случае недоступности обеих геосетей первичный центр продолжает функционировать, однако репликация данных прекращается. После восстановления работоспособности геосети производится синхронизация удаленных локальных копий, причем синхронизируются только автономно измененные в первичном центре данные. Процедура синхронизации происходит в фоновом режиме и существенно не сказывается на производительности всего комплекса.

В процессе реализации проекта особое внимание было уделено обеспечению возможности управления системой непосредственно силами специалистов заказчика. При этом мы столкнулись со следующей проблемой.

ПО GeoRM поставляется с комплектом управляющих утилит, сложность использования которых вполне адекватна сложности геокластерной системы. Однако уровень квалификации, необходимый для безошибочного выполнения процедур управления комплексом (особенно в критических ситуациях), в большей степени характерен для опытных системных инженеров компаний-интеграторов, чем для специалистов службы поддержки информационной системы. Например, активация (деактивация) одной из площадок требует углубленного анализа текущего состояния системы и последующего выполнения (возможно, на различных узлах) 4…7 команд в строгой последовательности и с неочевидными аргументами. Возможные ошибки чреваты увеличением времени восстановления или даже разрушением данных.

Кроме того, существует еще несколько факторов, отрицательно влияющих на степень готовности персонала к адекватной реакции на катастрофические события: а) критические ситуации возникают достаточно редко и, как следствие, необходимые навыки забываются; б) стрессовый характер ситуации мешает объективной оценке состояния системы и реализации сценария восстановления; в) при смене персонала передача необходимых знаний и навыков недостаточно эффективна.

С целью разрешения описанной проблемы нами были предприняты следующие действия: а) разработан комплект дополнительного ПО GeoTools, упрощающего и повышающего эффективность управления системой; б) разработаны и задокументированы подробные пошаговые сценарии восстановления в различных критических ситуациях.

GeoTools представляет собой программную надстройку над стандартными средствами GeoRM и операционной системы и выполняет следующие функции: а) позволяет администратору полностью

управлять системой с любого из узлов с минимальным риском выполнения ошибочных действий; б) производит автоматический мониторинг состояния геокластера с немедленным уведомлением администратора о любых отклонениях от штатного режима работы; в) поддерживает постоянство внутренней среды системы (ротация системных журналов, удаление "забытых" временных файлов, и т. д.); г) предоставляет администратору средства автоматического и ручного резервного копирования/восстановления данных приложения. Следует отметить, что система резервного копирования является необходимым компонентом системы любого уровня надежности, поскольку является практически единственным средствам восстановления в случае человеческих ошибок.

Разработанные сценарии восстановления предоставлены заказчику в виде документа "План восстановления в аварийных ситуациях". Документ регламентирует необходимую реакцию персонала в случае возникновения аварийных ситуаций различного характера, в том числе отказа отдельных аппаратных компонент, обратимого и необратимого отказа узлов, ошибок оператора. Для каждого случая предусмотрены три этапа реагирования: идентификация проблемы на основе имеющихся симптомов, восстановление сервиса, восстановление штатного режима работы. Эффективность разработанных процедур была подтверждена в ходе проведенных со специалистами заказчика практических занятий по отработке сценариев восстановления в критических ситуациях.

V Выводы

Практический опыт, накопленный в процессе создания катастрофоустойчивой системы, позволяет нам сделать следующие выводы.

1. В настоящее время технология создания устойчивых к катастрофам систем в достаточной степени востребована и, что не менее важно, доступна на отечественном рынке.
2. Средства и технологии, используемые при создании географических кластеров, могут варьироваться в широких пределах и определяются, прежде всего, предъявляемыми к системе требованиями и особенностями целевого приложения.
3. Активация резервного центра является неэффективными методом восстановления в случае одиночных отказов аппаратных компонент. Устойчивость системы к подобным авариям должна достигаться в рамках каждого центра (площадки) резервированием оборудования или применением классических высокодоступных кластеров.
4. Вне зависимости от применяемого решения требуется существенная привязка конфигурации геокластерной системы к местным условиям, а также разработка специфических сценариев восстановления в критических ситуациях. Указанные сценарии в сочетании с дополнительным управляющим программным обеспечением, разработанным с учетом местных условий, позволяют добиться необходимой степени эффективности реакции обслуживающего персонала на аварийные ситуации.

Литература: 1. Geographic Remote Mirror for AIX V2R3.0: Concepts and Facilities – IBM, 2001-44 p. 2. К. Вахрамеев. Защита данных от катастроф // Открытые системы, №3 – 2000.

УДК 681.321;322:621.395

АНАЛІЗ ЗАХИЩЕНОСТІ ІНФОРМАЦІЇ В АВТОМАТИЗОВАНИХ СИСТЕМАХ УПРАВЛІННЯ ПРИ МОДЕРНІЗАЦІЇ ТЕЛЕКОМУНІКАЦІЙНИХ СИСТЕМ

Микола Тардаскін, Володимир Кононович

Одеський регіональний центр технічного захисту інформації ВАТ "Укртелеком"

Анотація: **Аналізується проблема захисту інформації в автоматизованих систем управління, що впроваджуються при модернізації діючих систем телекомунікацій, Розглядається класифікація порушників та загроз інформації, порівнюються засоби захисту, рекомендується склад комплексу засобів захисту від несанкціонованого доступу.**

Summary: **An article the problem of security in automated management systems that are being incorporated in the acting systems of telecommunication in case of their enhancement is analyzed. Also, the classification of threats and users violator is considered. The protection facilities are compared and a list of such facilities from unauthorized access is recommended.**

Ключові слова: **Безпека інформації, автоматизована система управління, показники захищеності.**