

## ПРОГРАММНО-ОПРЕДЕЛЯЕМЫЕ СИСТЕМЫ ХРАНЕНИЯ ДАННЫХ

*Камышев Ю.С.*

*Институт информационных технологий БГУИР,  
г. Минск, Республика Беларусь*

*Скудняков Ю.А. – доцент каф. ПЭ, к.т.н., доцент*

Рассматриваются возможности современных программно-определяемых систем хранения данных.

Программно-определяемая сеть хранения данных (СХД) (англ. software-defined storage, SDS) — программное решение, обеспечивающее создание СХД на неспециализированном оборудовании массового класса, как правило, группе серверных узлов архитектуры x86-64 под управлением операционных систем общего назначения (Linux, Windows, FreeBSD). Основная отличительная возможность — виртуализация функции хранения, отделяющая аппаратное обеспечение от программного, которое управляет инфраструктурой хранения и в этом смысле является развитием концепции программно-определяемой сети, специализированном для систем хранения [1].

Аппаратное обеспечение в такой сети хранения обычно без какой-либо аппаратной агрегации или защиты предоставляет доступные накопители в программную часть, которая, как правило, объединяет их в пулы, и уже в рамках агрегированных пулов реализуются необходимые функции выделения томов, их презентации, ведения ограничений, управления производительностью, отработки отказов. Среди возможных функций программного уровня — кэширование, дедупликация, репликация, мгновенные снимки, резервное копирование, тонкое резервирование.

Центральную роль программно-определяемые сети хранения играют в гиперконвергентных системах, где они обеспечивают отказоустойчивую среду хранения томов виртуальных машин в предконфигурированных системах на базе серверного оборудования массового класса, выполняющих функции одновременно узлов сети хранения и узлов виртуализации вычислительных ресурсов. Среди серийно используемых продуктов для построения программно-определяемых сетей хранения являются vStorage (Virtuozzo), vSAN (VMware), GlusterFS и Ceph (обе — Red Hat), Storage Spaces Direct (Microsoft).

Отличительная особенность систем хранения SDS (Software-defined storage) состоит в том, что управление их многочисленными сервисными функциями (дедупликация, репликация, создание снапшотов, бэкапов и пр.) осуществляется на основе бизнес-ориентированных правил.

Благодаря широкому проникновению виртуализации программное управление инфраструктурой СУБД отделяется от аппаратного обеспечения. Это обеспечивает высокую гибкость в работе, упрощается масштабируемость. Но есть и минусы: поиск нужного сервера в кластере значительно усложняется, найти место физического размещения искомого фрагмента данных оказывается непросто. Считается, что это приведет в итоге к формированию абсолютно автономного программного стека, который будет полностью отвечать за решение любых функциональных задач СХД на уже имеющемся оборудовании.

Фрагментацию сопровождает явление, называемое «теневыми ИТ» (shadow IT). Оно проявляется в том, что отдельные сотрудники, а иногда и целые департаменты или филиалы компании, «забирают» в свои руки часть информационных ресурсов, делая их недоступными для других. Это выглядит как собственный набор серверов, СХД и систем резервирования, функционирующих автономно от общей ИТ-инфраструктуры компании. Нередко встречаются случаи переноса ответственных данных в публичные облака, сохранение их на USB-носителях. Возможности для надежного контроля за сохранностью и доступностью данных снижаются, а подчас просто выходят из-под контроля.

Естественная фрагментация возникает при расширении бизнеса и выстраивании филиальной сети. Тогда на каждом новом участке создается собственный сегмент ИТ-инфраструктуры, который оптимизируется под определенные, локальные условия. Установка и обслуживание сети, серверов и СХД связываются прежде всего с надежностью внутреннего доступа. Взаимодействие с другими точками локации ИТ-инфраструктуры предприятия часто оказывается в условиях реальной эксплуатации затруднительным. Поддержка этой возможности требует дополнительных ресурсов, поэтому их выделение подчас откладывается на неопределенный срок.

Но при нарушении связей (обрыве доступа) часть важных корпоративных данных оказывается в изоляции. Некоторые пользователи не могут получить к ним доступ, даже если потребность велика. Это может отрицательно сказаться на принимаемых решениях. Проблема может затронуть руководителей любого уровня. Будущее компании становится зависимым от степени полноты доступных данных. В этих случаях приходится возвращаться к принятию решений на интуиции. В условиях развитого рынка и больших рисков цена принимаемых решений может быть очень высокой.

Нарастающая фрагментация также может стать причиной ряда других, не менее существенных проблем:

1. Сохранение целостности данных. Когда данные поступают из разных мест, необходимо аккуратно применять механизмы контроля их целостности. Но если появляется несколько авторизованных копий общих данных, то могут возникнуть проблемы.

2. Снижение уровня защищенности данных. Когда к обслуживанию данных и их резервированию получают доступ пользователи или местные, неавторизованные службы, то каждый раз возникает риск потери контроля над собранной информацией и даже возможности ее частичной или полной утери.

3. Неэффективное использование ИТ-ресурсов. Серверы и СХД, выделенные под специализированные приложения, часто оказываются недозагруженными. Но при этом повышенное внимание приходится уделять каждому ресурсу, учитывать его специфические требования. В результате растут капитальные и операционные расходы, требуется индивидуальная поддержка для специализированных участков ИТ-инфраструктуры компании.

4. Неэффективное использование трудовых ресурсов. В случае возникновения фрагментации данных приходится наращивать штат сотрудников или привлекать на аутсорсинг внешние компании. Это необходимо для поддержки всех участков хранения данных в рабочем состоянии.

В заключение необходимо отметить, что программно-определяемые СХД являются актуальным и перспективным направлением развития систем хранения и обработки больших объемов данных, поскольку использование SDS позволяет обеспечить более высокие уровни гибкости, производительности и экономичности информационно-вычислительного процесса по сравнению с существующими.

**Список использованных источников:**

1. Крупнейший в Европе ресурс для IT-специалистов «Хабр» [Электронный ресурс]. – 2006. – Режим доступа : – <https://habr.com/company/croc/blog/278929/>. – Дата доступа : 09.11.2018.