

**ФЕДЕРАЛЬНОЕ АГЕНТСТВО СВЯЗИ**  
**Федеральное государственное образовательное бюджетное**  
**учреждение высшего профессионального образования**  
**«САНКТ-ПЕТЕРБУРГСКИЙ**  
**ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ТЕЛЕКОММУНИКАЦИЙ**  
**им. проф. М. А. БОНЧ-БРУЕВИЧА»**

---

**А.Н.Губин**

## **Сети хранения данных**

**Краткий конспект лекций**

**Раздел 7. Перспективные направления развития технологий хранения  
данных и управления информацией**

**САНКТ-ПЕТЕРБУРГ**  
**2017**

## Содержание

### **Раздел 1. Содержание курса, цели и задачи дисциплины**

- 1.1. *Общее содержание курса, цели и задачи изучения дисциплины.*
- 1.2. *Структура дисциплины и ее связь с другими курсами.*
- 1.3. *Хранение информации как основная задача информационных центров.*

### **Раздел 2. Общие характеристики процессов хранения и управления данными**

- 2.1. *Основные технологии хранения данных.*
- 2.2. *Структурированные и неструктурированные данные.*
- 2.3. *Общие характеристики процессов хранения и управления данными.*
- 2.4. *Общая структура информационных центров.*
- 2.5. *Управление хранением данных. Многоуровневое хранение данных.*

### **Раздел 3. Основные компоненты среды хранения данных**

- 3.1. *Физические компоненты. Хост, дисковое устройство.*
- 3.2. *Логические компоненты. Файловые системы.*
- 3.3. *Производительность дисковых устройств.*

### **Раздел 4. Защита данных. RAID-массивы**

- 4.1. *Распределение данных. Зеркалирование данных.*
- 4.2. *Контроль четности.*
- 4.3. *Основные конфигурации RAID-массивов.*
- 4.4. *IOPS-операций и конфигурация дисков.*
- 4.5. *Влияние наличия RAID на производительность дисковых устройств.*

### **Раздел 5. Системы хранения данных. Интеллектуализация систем хранения данных**

- 5.1. *Основные компоненты интеллектуальных систем хранения информации.*
  - 5.2. *Операция чтения и записи данных с использованием КЭШ-памяти.*
- Защита КЭШ- данных*
- 5.3. *Основные компоненты системы хранения данных Symmetrix.*

### **Раздел 6. Сети хранения данных**

- 6.1. *Архитектура сетей хранения данных. Особенности архитектуры СХД Fiber- Channel. Зонирование данных.*
- 6.2. *Контекстная адресация данных.*
- 6.3. *Виртуализация систем хранения данных.*

### **Раздел 7. Перспективные направления развития технологий хранения данных и управления информацией**

- 7.1. *Внеполостная и внутриполостная виртуализация систем данных.*
- 7.2. *Основные проблемы виртуализации систем хранения данных.*

## Раздел 7. Перспективные направления развития технологий хранения данных и управления информацией

### 7.1. Внеполосная и внутриполосная виртуализация систем данных

Сетевая виртуализация систем хранения данных осуществляется с использованием внеполосной и внутриполосной технологий. При использовании внеполосной технологии конфигурация виртуализированной среды хранится в отдельном устройстве вне каналов передачи данных. Эта конфигурация называется также системой с передачей по разделенному каналу. В этом случае происходит разделение канала управления и канала передачи данных (канал управления проходит через устройство виртуализации, а канал передачи данных – не проходит).

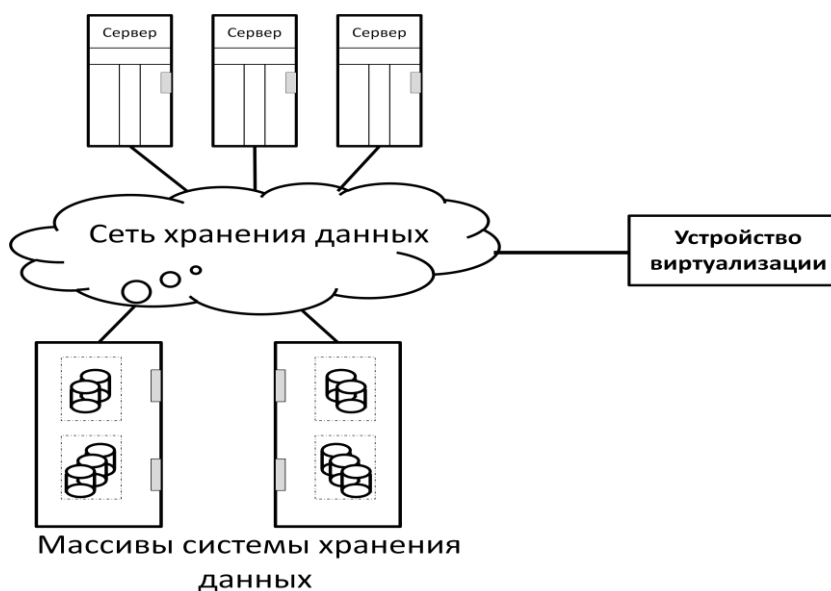


Рис.7.1. Конфигурация внеполосной виртуализации системы хранения данных

Такая конфигурация (рис. 7.1) позволит обрабатывать данные с сетевой скоростью при минимальной задержке, необходимой для преобразования информации о виртуальной конфигурации в информацию о физическом устройстве.



Рис.7.2. Конфигурация внутриполосной виртуализации системы хранения данных

При внутриполосной конфигурации виртуализация осуществляется в канале данных, серверами и устройствами общего назначения (рис. 7.2). В процессе обработки данных в этом случае пакеты часто кэшируются, чтобы затем после преобразования виртуальной информации их можно было отправить в соответствующую область физических устройств. Возникает дополнительная задержка времени отклика для приложения, связанная с пребыванием данных в Кэш-памяти перед их отправкой на физический диск.

В качестве примера реализации внеполосной виртуализации можно привести устройство EMC Invista, функционирующей на базе интеллектуальных маршрутизаторов SAN. Основные компоненты Invista и способ подключения к SAN приведены на рис. 7.3.

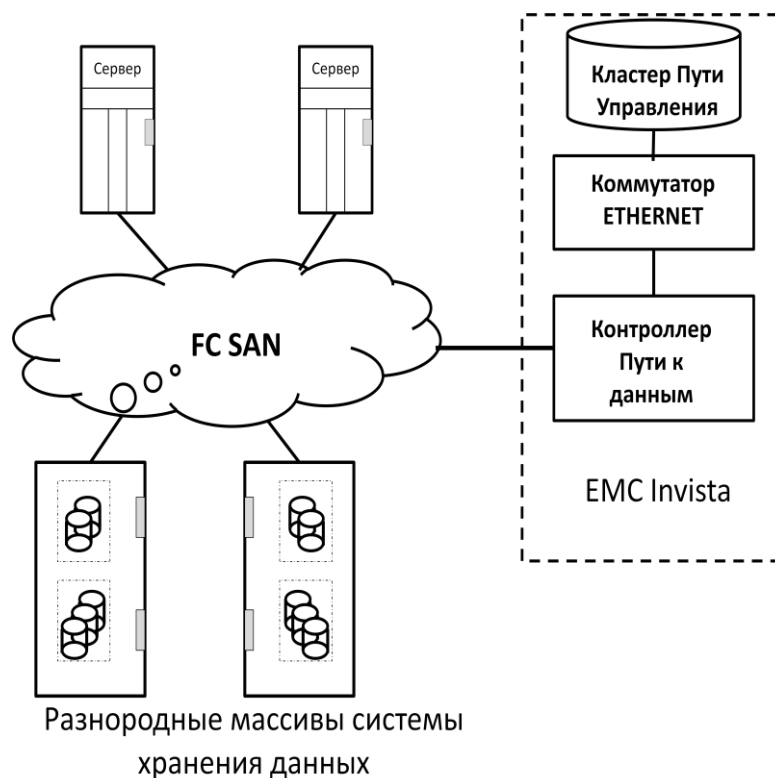


Рис.7.3. Основные компоненты устройства виртуализации EMC Invista

Кластер пути управления представляет собой специальное запоминающее устройство, обеспечивающее запуск программы Invista и содержащее информацию о параметрах конфигурации Invista. Кроме того блок кластера пути обеспечивает выполнение основных функций контроля и управления виртуальными запоминающими устройствами.

Коммутатор сети Ethernet обеспечивает доступ к основным устройствам Invista для их конфигурирования и контроля трафика.

Контроллер пути к данным выполняет функции маршрутизатора для системы SAN, обеспечивая корректное выполнение операций ввода/вывода данных.

## ***7.2. Основные проблемы виртуализации систем хранения данных***

В последнее время виртуализация систем хранения данных вышла за пределы простой агрегации ресурсов и базовых средств управления томами. Программно-определяемые системы (или SDS в англоязычной аббревиатуре) ставят своей целью предоставление ресурсов хранения в виде сервиса.

Программно-определяемое хранение добавляет такие мощные возможности, как защита данных, автоматическая организация многоуровневого хранения данных и обеспечение эффективности хранения. И, наконец, оно предоставляет возможность для таких расширений функционала, как объектное хранение.

Программно-определяемые среды хранения последнего поколения освобождают центр обработки данных от привязки к одному поставщику аппаратных и программных средств. Однако при этом выбор ПО может быть ограничен. По этой причине ЦОДами следует выбирать для поддержки сервисов данных единое мощное решение.

### **Изолированные зоны хранения — основная проблема традиционной виртуализации СХД**

Большинство систем хранения спроектированы для работы с определенными нагрузками. Одни подходят для баз данных, другие для систем аналитики, третьи — для вычислительных задач, четвертые — для резервного копирования. Существуют и универсальные СХД, которые могут работать с разными типами нагрузок, но, как правило, цена их достаточно высока. В итоге для разных задач экономически выгоднее приобретать разные комплексы.

В итоге быстро разрастаются изолированные зоны хранения в ЦОД. К тому же каждая из этих систем имеет свою консоль управления и применяет свои методы для распределения томов и для защиты данных.

Более того, организация становится зависимой от возможностей каждого поставщика по предоставлению необходимого набора базовых сервисов. У каждого поставщика имелись собственные технологии со своими компромиссами с точки зрения их эффективности и влияния на производительность.

Такие изолированные зоны, которые по-прежнему являются широко распространенным явлением в центрах обработки данных, существенно снижают уровень гибкости при эксплуатации инфраструктуры хранения. Например, если одно приложение исчерпывает ресурсы хранения или нуждается в доступе к более производительной системе, переход к ней может оказаться непростым мероприятием.

### **От традиционной виртуализации к программно-определяемым системам**

Виртуализация СХД позволяет объединить ресурсы хранения под единым интерфейсом управления, а также заимствовать ресурсы у других массивов или направлять задания на системы с большой свободной емкостью или быстродействием тогда, когда в этом возникает необходимость. Известно, что разные дисковые массивы даже в линейке одного поставщика хорошо подходят для определенных категорий задач и не очень — для других. И применение инструментов виртуализации существенно облегчает работу по проектированию комплексов хранения данных и их обслуживания в дальнейшем.

Одна из самых сложных задач, с которыми сталкивается администратор систем хранения данных — это перенос данных из старого массива на новый, или, по-другому, миграция данных. Усовершенствованные технологии виртуализации хранения (например, от IBM) допускают «живую» миграцию данных посредством переноса данных в новый массив хранения без прерывания обслуживания запросов приложений на ввод/вывод. Виртуальные комплексы способны напрямую отображать запросы данных на соответствующие физические системы хранения. После завершения переноса данных старый массив может быть выведен из эксплуатации или переориентирован на использование в качестве носителя данных при репликации или резервном копировании.

Следующее поколение средств виртуализации ресурсов хранения получило новое наименование — «программно-определяемые СХД». Это дальнейшее абстрагирование сервисов хранения данных от физических аппаратных средств.

Традиционная виртуализация СХД требовала, чтобы все сервисы данных предоставлялись одним поставщиком, использующим собственную платформу в виде «устройства» (appliance). При этом в пул можно было объединить только аппаратные средства. Программно-определяемая СХД позволяет сочетать сервисы данных от нескольких поставщиков, а также различные аппаратные устройства хранения.

Это концепция сводится к тому, что поставщики перемещают сервисы хранения со своих СХД на базе специфичных аппаратных средств к своим решениям, определяемым программным обеспечением. Другими словами, эти сервисы были абстрагированы от аппаратных средств и начали определять все возможности решения для виртуализации хранения.

До сих пор эти модульные дополнительные сервисы были направлены на повышение емкости и производительности. К примеру, IBM добавила к своей платформе Storwize поддержку технологий Realtime Compression и Easy Tier. Real-time Compression (сжатие в реальном времени) позволяет увеличить объем дисковых ресурсов без значительного снижения производительности в среднем вдвое. Технология Easy Tier ранее поддерживалась только на старших системах IBM DS8000, однако после реализации этой технологии на решениях семейства Storwize ее возможности — посредством виртуализации ресурсов хранения — стали доступными различным другим системам, в том числе устройствам, использующим только флэш-память. Это позволяет IBM демонстрировать, что платформа Storwize превратилась в программно-определяемое решение.

## **SSD и другие новшества**

Несмотря на то, что благодаря своей высокой производительности устройства на основе флэш-памяти привлекают внимание многочисленных создателей центров обработки данных, их внедрение - достаточно медленный процесс. Среди недостатков таких устройств можно отметить отсутствие

ряда функций, высокая стоимость единицы хранения и трудности при выявлении данных, подлежащих перемещению на эти устройства. В результате использование флэш-систем ограничено преимущественно крупными базами данных.

Концепция программно-определяемого хранения способна значительно ускорить внедрение флэш-технологий. Она позволяет интегрировать устройства на основе флэш-памяти в существующую архитектуру хранения без необходимости замены каких-либо активов. Она также помогает снизить стоимость единицы хранения, обеспечивая эффективный коэффициент сжатия 2:1 (и даже больше), и допускает перемещение данных между устройствами в автоматическом режиме. Ведь не все данные используются одинаково часто, а значит, требуют для работы производительности SSD-диска. К тому же со временем необходимо освобождать дефицитные ресурсы от данных, чья актуальность снизилась.

Программно-определяемое хранение повышает ценность SSD ресурсов посредством поддержки стандартных сервисов, таких как снимки и репликация данных на флэш-массивы. Эти возможности весьма важны для того, чтобы сделать флэш-системы более привлекательными для применения в интересах массовых приложений в ЦОД. Кроме того, потребители смогут опереться на хорошо изученные сервисы от поставщиков систем хранения корпоративного класса, а не продукцию малоизвестных компаний-стартапов.

Хотя возможности программно-определяемых систем хранения первого поколения увеличивают функциональность всех типов систем хранения, их потенциальное влияние является наиболее значимым для систем хранения на базе флэш-памяти. По существу SDS первого поколения сделают возможным широкое внедрение флэш-ресурсов с целью преодоления известных проблем производительности при хранении данных в центрах обработки данных. Они разовьют преимущества флэш-систем посредством поддержки для этих устройствах обширного набора хорошо протестированных сервисов данных.

Следует, однако, отметить, что ценность программно-определяемого хранения данных не ограничивается использованием флэш-памяти. При появлении очередного нового аппаратного средства хранения предприятия захотят внедрить его быстрее, чем поставщики смогут предложить соответствующее пакетное решение. По существу концепция программно-определяемого хранения позволяет поставщикам аппаратных средств сосредоточиться на своих аппаратных средствах и подключать их к существующей архитектуре хранения. А значит, потребители реализуют преимущества новых технологий быстрее и с минимальными отрицательными последствиями.



## **Проблемы организации сервисной платформы управления данными**

Обычно какая либо новая функция доступна только на совершенно новой платформе. Для систем хранения данных такая ситуация возникает особенно часто. Но для потребителя выгоднее и удобнее, когда ему достаточно добавить эту новую функцию к существующей у него платформе хранения, когда он избавлен от необходимости настройки другого устройства или установки новой системы хранения.

Следующий шаг в развитии программно-определяемого хранения данных состоит в интеграции программно-определяемых решений для хранения данных в платформу для сервисов данных всех типов. Конечная цель должна состоять в создании системы для ЦОД, которое позволяла бы выбирать сервисы данных от различных поставщиков в зависимости от конкретного сценария.

Но это потребует от SDS-поставщиков готовности к открытию своих платформ, чтобы другие поставщики смогли подключаться к ним свои новые сервисы. А для этого необходимо избавиться от отрицательного отношения к «чужим» изобретениям, которое сегодня превалирует.

Следующий шаг состоит в дальнейшем «открытии» программно-определяемой системы хранения и в превращении в SDS- платформу. Фактически речь идет о разрешении многочисленным независимым разработчикам программного обеспечения интегрировать свои приложения в данную SDS-платформу.

Предоставление такой компании разрешения на добавление ее сервисов к исходной платформе устранил возможные последствия, порождаемые добавлением новой системы хранения. Кроме того, это стимулирует дальнейшее инновации, поскольку нет необходимости повторно изобретать всю функциональность.

Еще один полезный момент для предприятия состоит в возможности добавления новых функций по мере необходимости к существующим средствам хранения без развертывания еще одной консоли управления. Это снижает потребности центра обработки данных в полезной площади и сокращает затраты времени на обучение персонала.