Обеспечение отказоустойчивости и резервное копирование

# **Класстеризация**

**Кластеризация** (кластерный анализ) — это метод анализа данных, при котором объекты разделяются на группы (кластеры) по значимым критериям. Внутри каждого кластера располагаются только схожие компоненты, а элементы разных групп должны максимально отличаться друг от друга.

Кластеризация выполняется со следующими целями:

* Сжатие информации. Кластеризация актуальна, если исходные выборки слишком объёмные. В результате от каждого кластера остаётся по одному типичному представителю.
* Поиск паттернов в информационном пространстве. Разбиение объектов на кластеры позволяет добавить дополнительный признак каждому объекту.
* Обнаружение аномалий. При кластеризации объектов удаётся выделить нетипичные элементы, которые не подходят ни к одному сформированному сегменту.

Настройка отказоустойчивых кластеров (кластеров высокой доступности) включает несколько этапов: подготовку инфраструктуры, настройку кластера в конкретной операционной системе и тестирование.

## Подготовка инфраструктуры

Для корректной работы кластера необходимы:

* **Избыточность ресурсов**. Чтобы перенести задачу с аварийного сервера на работоспособный, нужны свободные вычислительные ресурсы.
* **Совместный доступ к данным**. Ключевой принцип работы — распределение задач вышедшего из строя сервера между оставшимися, поэтому необходимо резервировать каналы к данным и конечным пользователям.
* **Соответствие ПО**. Приложения, которые планируется разворачивать в кластере, должны уметь использовать общее хранилище данных и перезапускаться на другом узле в том же состоянии, в котором они были в момент сбоя первоначально используемого сервера.

## Настройка кластера в Windows Server

В Windows Server для настройки отказоустойчивого кластера используется диспетчер отказоустойчивости кластеров. В некоторые шаги для выполнения настройки:

* **Установить компонент** отказоустойчивой кластеризации на всех серверах, которые будут узлами кластера.
* **Настроить хранилище** (например, по технологии iSCSI).
* **Создать кластер** в диспетчере отказоустойчивости кластеров, добавив серверы.
* **Настроить кластерный кворум** — определить, какой диск будет использоваться в качестве диска-свидетеля.

## Тестирование и оптимизация

После завершения настроек важно протестировать систему на устойчивость к различным сценариям сбоев. Для этого можно имитировать отключение отдельных узлов и проверить реакцию кластера.

Также необходимо настроить мониторинг состояния узлов и оповещения о сбоях. Для этого используются инструменты, например:

* **Performance Monitor (PerfMon**) — для отслеживания производительности.
* **System Center Operations Manager (SCOM)** — для мониторинга и управления, предоставляет подробные отчёты и уведомления о состоянии системы.

# Автоматическое копирование

Для автоматического копирования конфигурации сетевых устройств используются специальные инструменты, такие как RANCID и Oxidized.

## Really Awesome New Cisco config Differ

**RANCID (Really Awesome New Cisco config Differ)** — система мониторинга конфигураций сетевых устройств, которая автоматически загружает файлы с устройств и сравнивает их с предыдущими версиями.

Некоторые особенности RANCID:

* поддерживает оборудование разных брендов, например Cisco, Juniper, Catalyst;
* позволяет отслеживать изменения в конфигурации и хранить её историю;
* поддерживает оповещение по электронной почте при изменении конфигурации.

Для хранения конфигураций RANCID использует системы контроля версий — Subversion или CVS (Concurrent Version System).

## Настройка RANCID

Некоторые шаги для настройки RANCID:

* **Создать файл конфигурации**. В нём указать группы устройств, список устройств в группе, тип устройства, учётные данные для доступа и протоколы для получения конфигураций.
* **Настроить права на файл конфигурации**. Например, установить права 640 для файла конфигурации и назначить пользователя RANCID его владельцем.
* **Протестировать работу**. Запустить команду sudo /usr/lib/rancid/bin/rancid-run, которая загрузит файлы конфигураций с устройств, перечисленных в файле конфигурации.

## Oxidized

**Oxidized** — инструмент для автоматического резервного копирования конфигураций сетевых устройств и контроля версий.

Некоторые особенности Oxidized:

* поддерживает более 130 типов операционных систем;
* позволяет автоматически создавать копии конфигураций и версий программного обеспечения для роутеров, коммутаторов, файрволов и других устройств;
* хранит данные конфигурации и версий в репозитории Git или RCS;
* имеет встроенный веб-интерфейс для просмотра истории конфигурации и сравнения изменений файлов.

## Настройка Oxidized

Некоторые шаги для настройки Oxidized:

* **Создать файл конфигурации**. По умолчанию он находится в каталоге ~/.config/oxidized/config и имеет формат YAML.
* **Настроить источник списка устройств**. Например, можно использовать CSV-файл ~/.config/oxidized/router.db, в котором указать устройства в формате hostname:model[:username:password:enable\_password].
* **Настроить аутентификацию устройств**. Oxidized поддерживает несколько методов аутентификации: имя пользователя и пароль, SSH-ключи, пароль для привилегированного режима.
* **Запустить Oxidized**. Для производственного использования рекомендуется запускать его в качестве сервиса.

# Хранение данных

Вопрос хранения данных становится всё более актуальным. Выбор между хранением данных в облаке и на локальном сервере зависит от конкретных задач и потребностей.

## Облачные сервисы

**Облачные сервисы** – это онлайн-платформы, предоставляющие пространство для хранения данных на удалённых серверах. Наиболее известные из них: Google Drive, Dropbox, Яндекс.Диск, iCloud и др.

**Преимущества облачного хранения:**

* **Доступность в любой точке мира**. Достаточно иметь интернет-подключение и ввести пароль — ваши данные доступны с любого устройства.
* **Автоматическая синхронизация**. Файлы и изменения мгновенно передаются на все ваши устройства.
* **Безопасность при поломке ПК**. Даже если домашний компьютер сломается или будет украден, данные останутся в целости.
* **Резервное копирование**. Многие облачные сервисы автоматически делают резервные копии, что минимизирует риск потери информации.
* **Легко делиться файлами**. Передать большой объём документов или фото друзьям и родственникам – вопрос пары кликов.

**Недостатки облачных сервисов:**

* **Зависимость от интернета**. Без подключения доступ к данным ограничен.
* **Ограниченное бесплатное пространство**. Бесплатный объём обычно не превышает 5–15 ГБ, за увеличение придётся платить.
* **Вопросы приватности**. Сервисы могут анализировать ваш контент, а при взломе аккаунта ваши личные данные могут оказаться под угрозой.
* **Скорость доступа**. Работа с очень большими файлами (например, видео в 4K) может быть неудобной из-за ограничений по скорости интернет-соединения.
* **Проблемы с загрузкой/выгрузкой больших архивов**. Загрузка и скачивание больших архивов – процесс не самый быстрый.

## Локальное хранение

**Локальное хранение** – это сохранение данных непосредственно на встроенном жестком диске, SSD, внешних накопителях или NAS-серверах в пределах домашней сети.

**Преимущества локального хранения:**

* **Полный контроль над данными.** Вы сами определяете, кто имеет доступ к вашим файлам.
* **Мгновенный доступ**. Нет зависимости от скорости интернета, файлы доступны моментально.
* **Большие объёмы хранилища**. Жёсткие диски и SSD быстро дешевеют, что позволяет хранить терабайты информации без абонентской платы.
* **Конфиденциальность**. Никто, кроме вас, не имеет доступа к содержимому вашего накопителя.
* **Возможность использования специализированных решений**. Для дома доступны различные решения: NAS, RAID-массивы, которые позволяют создать свою собственную систему резервного копирования и хранения данных.

**Недостатки локального хранения:**

* **Уязвимость к поломкам и утере.** Повреждение накопителя или кража устройства = потеря данных (если нет копии).
* **Нет удалённого доступа**. Чтобы получить доступ к файлам вне дома — необходимы дополнительные настройки (например, создание домашнего облака).
* **Затраты времени и сил на обслуживание**. Требуется самостоятельная организация резервных копий, настройка оборудования.
* **Устаревание оборудования**. Через 3–5 лет диски приходится менять, переносить данные.

## Рекомендации по выбору

Некоторые факторы, которые стоит учитывать при выборе способа хранения информации:

* **Размер и потребности данных**. Облачные решения подходят для больших объёмов, локальные — для конфиденциальных данных или аналитики с низкой задержкой.
* **Бюджет**. Нужно оценить первоначальные и постоянные затраты на оба варианта.
* **Требования к безопасности**. Для конфиденциальных данных может понадобиться более высокий уровень контроля, который предлагает локальное хранение.
* **Технические ресурсы**. Локальное решение может быть предпочтительным, если есть доступ к квалифицированному IT-персоналу.

Также можно рассмотреть гибридный подход: хранить конфиденциальные данные локально, а облачные — для масштабируемости и аналитики.

# Данные пользователей

**Планирование бэкапов** (резервного копирования) необходимо для обеспечения восстановления данных и минимизации рисков их потери.

**Полный бэкап** — это полная резервная копия диска со всеми данными. Такой тип бэкапа выбирают в случае высокой значимости информации, например, потребности постоянного обновления клиентской базы. Однако полный тип связан с нагрузками на физическое оборудование и во многом тормозит производительность внутри локальной сети.

**Инкрементальный бэкап** — это копия изменений между текущим состоянием диска и предыдущим бэкапом, созданным по плану. Такой бэкап создаётся быстрее, чем полный, и занимает меньше места, потому что хранит только изменения на диске.

При выборе схемы резервного копирования следует учитывать множество факторов:

* объём данных, с которыми оперируют в процессе резервного копирования;
* в течение какого времени требуется восстановить данные при сбое оборудования или других происшествиях;
* за какой период допустимо потерять данные в результате инцидента;
* тип хранилища — локальное или облачное;
* на каких устройствах хранятся бэкапы — диски, ленты, NAS и т. д.;
* для каких объектов защиты выполняется бэкап — операционные системы, приложения, платформы виртуализации и т. д.;
* в какой инфраструктуре осуществляется резервное копирование: пропускная способность сети, расположение серверов и хранилищ резервных копий и т. п..