Министерство образования Республики Беларусь

Учреждение образования

БелорусскиЙ государственный университет

информатики и радиоэлектроники

Факультет компьютерных систем и сетей

Кафедра электронных вычислительных машин

Дисциплина: Интерфейсы и устройства вычислительных машин

КОНТРОЛЬНАЯ РАБОТА

по теме

|  |  |  |
| --- | --- | --- |
| Дисковые массивы | | |
| Выполнил  студент гр. 250541 |  | Бобрик В.Ю. |
|  |  |  |
| Проверил |  |  |
| ст. преподаватель каф. ЭВМ |  | Куприянова Д.В. |

Минск 2025

Содержание

|  |  |
| --- | --- |
| Введение…………………………………………………………………………… | 3 |
| 1 Классификация дисковых массивов…………………………………………… | 4 |
| 2 Уровни RAID…………………………………………………………………… |  |
| 3 Практическое применение дисковых массивов……………………………… |  |
| Заключение………………………………………………………………………… |  |

# Введение

Современные информационные системы невозможно представить без надежных и производительных средств хранения данных. Объемы информации растут экспоненциально, а требования к ее доступности и сохранности становятся критически важными как для крупных компаний, так и для индивидуальных пользователей. Особое значение приобретают дисковые массивы – технологии, позволяющие объединять несколько физических накопителей в единую логическую систему.

Под дисковым массивом понимается совокупность жестких дисков, управляемых специализированным контроллером, который обеспечивает их согласованную работу. В простейшем случае массив может представлять собой набор независимых дисков (JBOD), однако в большинстве практических сценариев используется более сложная логика объединения, направленная на решение двух ключевых задач: повышение скорости доступа к данным и обеспечение отказоустойчивости. Эти цели достигаются за счет распределения информации между накопителями, применения методов зеркалирования, чередования и избыточного хранения контрольных кодов.

Исторически развитие дисковых массивов тесно связано с технологией RAID, предложенной в конце 1980‑х годов исследователями Калифорнийского университета в Беркли. Первоначально RAID рассматривался как способ снизить стоимость систем хранения путем использования относительно дешевых дисков, однако со временем акцент сместился на обеспечение надежности и производительности. RAID-массивы стали стандартом в серверных решениях и системах хранения данных, где требуется баланс между скоростью и безопасностью информации.

Актуальность изучения дисковых массивов обусловлена не только их широким распространением, но и необходимостью понимания принципов их работы для грамотного выбора архитектуры хранения данных. Различные уровни RAID обладают своими достоинствами и ограничениями, и именно от правильного выбора зависит эффективность функционирования всей информационной системы. Современные тенденции, такие как виртуализация, облачные вычисления и развитие сетевых систем хранения, делают тему дисковых массивов особенно значимой.

Цель данной контрольной работы сводится к рассмотрению основных принципов построения дисковых массивов, их классификации, анализа уровней RAID и оценки практических областей применения. Работа позволит систематизировать знания о современных технологиях хранения данных и подчеркнуть их роль в обеспечении надежности и производительности информационных систем.

# 1 Классификация дисковых массивов

Прежде чем переходить к классификации дисковых массивов, целесообразно дать определение основным терминам, используемым при работе с ними.

Дисковый массив – объединение нескольких физических накопителей в единый логический диск, управляемый контроллером.

Контроллер массива – специализированное устройство или программный модуль, обеспечивающий распределение данных, контроль целостности и восстановление массива.

По способу организации дисковые массивы можно разделить на два основных типа: JBOD и RAID-массивы.

JBOD (Just a Bunch of Discs) – простое объединение дисков без избыточности и повышения производительности. В данном случае контроллер воспринимает несколько физических дисков как единый логический том, последовательно соединяя их пространство. Такой подход позволяет получить большой виртуальный диск, что бывает полезно при работе с файлами, превышающими размер одного накопителя. Таким образом, JBOD решает исключительно задачу увеличения объема, не затрагивая вопросы надежности или скорости.

RAID (Redundant Array of Independent Discs) – технология организации массива с использованием методов чередования, зеркалирования и избыточности для повышения скорости и надежности. Здесь данные распределяются между дисками по определенным алгоритмам, что позволяет достичь двух ключевых целей: повышения производительности и обеспечения отказоустойчивости. В основе работы RAID лежат такие методы, как:

* чередование (striping) – метод повышения производительности за счет параллельной записи и чтения блоков данных на нескольких дисках;
* зеркалирование (mirroring) – метод повышения надежности за счет создания копии данных на другом диске;
* четность (parity) – дополнительная информация, вычисляемая по данным и позволяющая восстановить их при отказе одного из дисков.

По способу реализации выделяются такие варианты, как аппаратные и программные массивы. Аппаратная реализация предполагает использование выделенного RAID‑контроллера, который может быть выполнен в виде отдельной платы расширения, встроенного модуля в сервер или даже внешнего устройства в составе системы хранения данных (NAS, SAN). Такой контроллер содержит собственный микропроцессор, память для кэширования и специализированные алгоритмы для обслуживания массива. Программная реализация RAID осуществляется средствами операционной системы. В этом случае логика распределения данных и контроль избыточности выполняются центральным процессором компьютера, а диски подключаются напрямую без специализированного контроллера.

По назначению дисковые массивы распределяются следующим образом:

* массивы для повышения производительности. Главная цель таких массивов – ускорение операций чтения и записи. Это достигается за счет параллельной работы нескольких дисков. Наиболее подходящим примером является RAID 0 (четность), где данные разбиваются на блоки и записываются одновременно на несколько накопителей. Появляется значительный прирост скорости, особенно при работе с большими файлами или интенсивных потоках данных. Однако проблемой является полное отсутствие отказоустойчивости – выход из строя одного диска приводит к потере всего массива. Такие массивы применяются там, где важна скорость, а данные легко восстановимы: видеомонтаж, обработка изображений, временные файлы;
* массивы для повышения надежности. Приоритетом является сохранность информации даже при отказе одного или нескольких дисков. Простейший пример – RAID 1 (зеркалирование), где данные дублируются на двух накопителях. Преимуществами являются высокая отказоустойчивость и возможность продолжать работу при выходе из строя диска. К недостаткам следует отнести снижение доступного объема и иногда падение производительности при записи. Такие массивы востребованы в банковских системах, корпоративных базах данных, финансовой отчетности – там, где потеря информации недопустима;
* комбинированные массивы. Это сочетание преимуществ первых двух подходов. Пример – RAID 10, где зеркалирование обеспечивает надежность, а чередование – скорость. Это баланс между производительностью и отказоустойчивостью. Имеет высокую стоимость, так как требуется большее количество дисков. Комбинированные массивы используются в дата‑центрах, облачных сервисах и других критически важных системах, где важны и скорость, и надежность.

По сфере использования дисковые массивы можно разделить так:

* персональные системы, которые используются в настольных компьютерах и рабочих станциях. Обычно реализуются средствами встроенных контроллеров материнских плат и поддерживают базовые уровни RAID (0, 1, иногда 5 или 10);
* корпоративные серверы применяются в бизнес‑среде, где критически важны надежность и производительность. Здесь используются аппаратные RAID‑контроллеры с поддержкой «горячей замены» дисков, резервных накопителей и сложных уровней RAID (5, 6, 10, 50).
* системы хранения данных (NAS, SAN) – это специализированные решения для дата‑центров и облачных сервисов. Отличаются масштабируемостью, расширенными функциями мониторинга и управления, возможностью интеграции в распределенные сети хранения.

# 2 Уровни RAID

Под уровнем RAID понимают определенный способ организации данных в массиве, то есть набор правил, по которым контроллер распределяет информацию между дисками. Каждый уровень описывает:

* как именно данные делятся на стрипы и формируются в страйпы;
* используется ли зеркалирование (создание копий данных);
* применяется ли четность (контрольные суммы для восстановления информации);
* какие цели достигаются: повышение производительности, надежности или их комбинация.

Чтобы понимать работу различных уровней RAID, необходимо ввести несколько базовых понятий.

Стрип (strip) – это минимальный блок данных, который контроллер массива записывает на один диск. Размер стрипа задается в настройках и может варьироваться (например, 16 КБ, 64 КБ или 128 КБ). От выбранного размера зависят характеристики массива: мелкие стрипы повышают эффективность при работе с мелкими файлами, а крупные – при последовательной обработке больших объемов данных.

Страйп (stripe) – совокупность стрипов, расположенных на всех дисках массива. Это «полоса данных», которая распределена по всем накопителям. Например, если массив состоит из четырех дисков, то один страйп будет включать четыре стрипа – по одному на каждый диск. При чтении или записи страйпа контроллер может задействовать все диски параллельно, что обеспечивает прирост производительности.

Для упрощения понимания уровней RAID опишем JBOD как нулевой уровень, хоть и не относящийся к RAID напрямую. В данном случае данные просто пишутся подряд на несколько дисков в один сплошной страйп, что проиллюстрированно на рисунке 2.1.

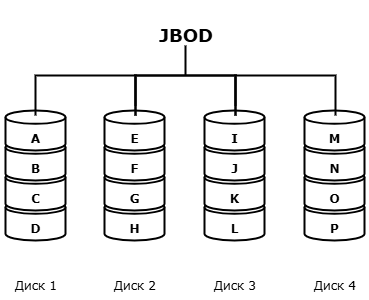


Рисунок 2.1 – Схема данных JBOD

JBOD не обеспечивает ни повышения производительности, ни отказоустойчивости. Его единственное преимущество – возможность получить большой виртуальный диск, размер которого превышает объем любого из отдельных накопителей. Это может быть полезно, например, при хранении крупных файлов, которые не помещаются на один физический диск.

Однако выход из строя любого диска приводит к потере части данных, и восстановить их средствами массива невозможно. Поэтому JBOD применяется редко и в основном в тех случаях, когда требуется исключительно увеличение объема, а надежность и скорость не являются критически важными параметрами.

RAID 0 – это уровень массива, при котором данные разбиваются на стрипы и записываются параллельно на несколько дисков. Совокупность стрипов, распределенных по всем накопителям, образует страйп. Таким образом, при чтении и записи контроллер может задействовать все диски одновременно, что обеспечивает значительный прирост производительности. Схема хранения данных RAID 0 показана на рисунке 2.2.

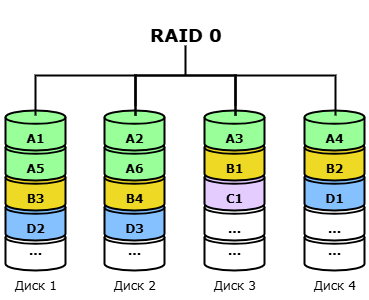


Рисунок 2.2 – Схема данных RAID 0

Для RAID 0 минимальным требованием является наличие как минимум двух дисков. Это связано с тем, что принцип работы массива основан на чередовании: данные разбиваются на блоки и записываются параллельно на разные накопители. Если использовать только один диск, распределять стрипы будет просто некуда, и массив не сможет существовать как RAID 0.

Полезный объем массива в RAID 0 равен сумме всех дисков, а производительность растет пропорционально их количеству. Весь объем накопителей используется без потерь на избыточность, что делает RAID 0 привлекательным для задач, где важен каждый гигабайт.

Однако проблемой этого уровня является полное отсутствие отказоустойчивости. Если выходит из строя хотя бы один диск, данные всего массива становятся недоступны, так как каждый файл «разрезан» на части и распределен по всем накопителям. Восстановление массива RAID 0 как таковое невозможно, поскольку этот уровень не предусматривает избыточности данных. В случае сбоя остается лишь заменить поврежденный диск и заново создать массив, восстановив данные из резервных копий или исходных источников.

Поэтому RAID 0 применяют в тех случаях, когда приоритетом является скорость, а сохранность данных не критична. Это может быть обработка видео и графики, работа с большими временными файлами или игровые системы, где потерянную информацию легко восстановить из исходных источников.

RAID 1 – это уровень массива, в основе которого лежит принцип зеркалирования. При такой организации каждая запись данных дублируется на два и более диска. В результате все накопители содержат идентичные копии информации, и массив продолжает работать даже в случае выхода из строя одного из них. Для операционной системы RAID 1 выглядит как единый диск, но фактически данные хранятся параллельно на нескольких носителях. Для организации RAID 1 необходимо как минимум 2 диска. Полезный объем массива равен объему одного диска, так как остальное пространство используется для хранения копий. Схема хранения данных RAID 1 показана на рисунке 2.3.

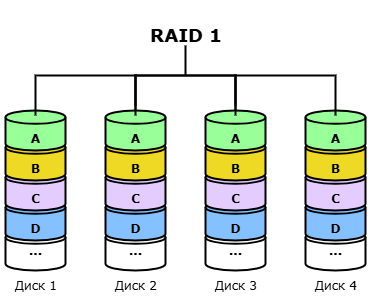


Рисунок 2.3 – Схема данных RAID 1

Главное достоинство RAID 1 – высокая надежность. Если один диск перестает функционировать, система автоматически переключается на его копию, и пользователь не замечает сбоя. Благодаря этому RAID 1 широко применяется там, где критична сохранность информации: в бухгалтерских системах, базах данных, корпоративных серверах. Дополнительным преимуществом является ускорение операций чтения, так как контроллер может обращаться к любому из зеркальных дисков, распределяя нагрузку. Главным недостатком является низкая эффективность использования дискового пространства: половина объема массива или более уходит на хранение копий. Прироста скорости записи не наблюдается, так как данные должны быть записаны на все зеркала одновременно.

Восстановление массива RAID 1 несложно: при выходе из строя одного из накопителей достаточно заменить поврежденный диск на новый, после чего контроллер автоматически запускает процесс ребилда. В ходе этой процедуры выполняется поблочное копирование данных с исправного диска на новый, и по завершении синхронизации массив возвращается в полноценное рабочее состояние без вмешательства пользователя.

RAID 1 можно рассматривать как противоположность RAID 0: если в RAID 0 приоритетом является максимальная производительность при полном отсутствии защиты, то в RAID 1 акцент сделан на надежности и сохранности данных ценой потери половины или более полезного объема.

RAID 5 – это уровень массива, при котором данные также разбиваются на стрипы и распределяются по нескольким дискам, но в каждом страйпе дополнительно хранится блок четности. Эти блоки четности не сосредоточены на одном накопителе, а распределяются по всем дискам массива, что позволяет сохранить баланс между производительностью и отказоустойчивостью. При выходе из строя одного диска контроллер может восстановить недостающие данные, используя информацию о четности и оставшиеся стрипы. Схема хранения данных RAID 5 показана на рисунке 2.4.

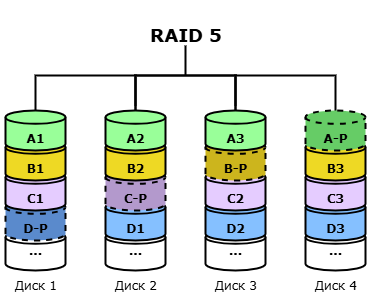


Рисунок 2.4 – Схема данных RAID 5

Для RAID 5 минимальным требованием является наличие как минимум трех дисков. Это связано с тем, что для вычисления четности необходимо иметь хотя бы два диска с данными и один для хранения контрольной информации. Полезный объем массива равен сумме всех дисков за вычетом одного, так как пространство, эквивалентное одному накопителю, используется для хранения распределенных блоков четности.

RAID 5 сочетает высокую производительность при чтении и достаточную надежность: массив способен выдержать отказ одного диска без потери данных. При этом эффективность использования пространства выше, чем в RAID 1, так как на избыточность уходит лишь часть массива, а не половина или больше.

У массива RAID 5 заметно снижение скорости записи, так как при каждой операции необходимо вычислять и записывать блок четности. Кроме того, процесс восстановления после сбоя (ребилд) может занимать значительное время и сопровождаться высокой нагрузкой на оставшиеся диски, что увеличивает риск выхода из строя еще одного накопителя.

Если в страйпе потерян блок данных, контроллер считывает все остальные блоки данных и соответствующий блок четности. Так как четность в RAID 5 вычисляется по принципу XOR, недостающий блок можно восстановить: контроллер выполняет XOR над всеми доступными блоками и блоком четности, в результате чего получается исходное содержимое утраченного блока. Если в страйпе отсутствует блок четности, восстановление еще проще: контроллер просто пересчитывает его заново, выполняя XOR над всеми блоками данных в этом страйпе. Таким образом, массив всегда может восстановить либо данные, либо четность, пока потерян только один диск.

Весь процесс ребилда заключается в том, что контроллер последовательно проходит по всем страйпам массива, вычисляет недостающие блоки (данные или четность) и записывает их на новый диск. При больших объемах это может занимать много времени и сопровождаться высокой нагрузкой на оставшиеся накопители, но по завершении синхронизации массив возвращается в полноценное рабочее состояние.

RAID 5 применяется в тех случаях, когда требуется компромисс между производительностью, надежностью и эффективностью использования дискового пространства. Он широко используется в файловых серверах, системах хранения корпоративных данных и других приложениях, где важен баланс между скоростью и сохранностью информации.

RAID 6 – это уровень массива, при котором данные, как и в RAID 5, разбиваются на стрипы и распределяются по нескольким дискам. Отличие заключается в том, что в каждом страйпе вычисляются не один, а два блока четности, которые размещаются на разных накопителях. Такая схема обеспечивает повышенную отказоустойчивость: массив способен продолжать работу даже при выходе из строя сразу двух дисков. Схема хранения данных RAID 6 показана на рисунке 2.5.

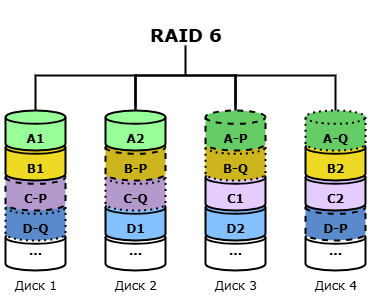


Рисунок 2.5 – Схема данных RAID 6

Для RAID 6 минимальным требованием является наличие как минимум четырех дисков, так как необходимо хранить два блока данных и два блока четности. Полезный объем массива равен сумме всех дисков за вычетом двух, поскольку пространство, эквивалентное двум накопителям, используется для хранения распределенных блоков четности.

В плюсы RAID 6 записывается высокая надежность: массив способен выдержать отказ двух дисков без потери данных. Это делает его особенно ценным для корпоративных систем хранения, где критична непрерывность работы и сохранность информации. При этом эффективность использования пространства выше, чем в RAID 1, хотя и ниже, чем в RAID 5.

Недостатком RAID 6 является снижение производительности записи, так как при каждой операции необходимо вычислять и записывать два блока четности. Кроме того, процесс восстановления после сбоя (ребилд) занимает больше времени и создает еще большую нагрузку на оставшиеся диски, чем в RAID 5.

Процесс ребилда в RAID 6 зависит от того, какие блоки утрачены. Если потерян один диск с данными, контроллер восстанавливает недостающий блок, используя оставшиеся данные и оба блока четности. Если отсутствует один блок четности, он пересчитывается заново по данным страйпа и второму блоку четности. При потере двух дисков возможны разные сценарии: если отсутствуют два блока данных, они восстанавливаются с помощью всех оставшихся данных и обоих блоков четности; если потерян блок данных и блок четности, сначала восстанавливаются данные, затем пересчитывается четность; если отсутствуют оба блока четности, они пересчитываются заново по данным страйпа.

Разница между двумя блоками четности заключается в методе их вычисления. Первый блок (P) формируется по принципу XOR, аналогично RAID 5, и обеспечивает восстановление при потере одного блока. Второй блок (Q) рассчитывается по более сложным алгоритмам, например, с использованием кодов Рида–Соломона, что позволяет восстановить данные при одновременной потере второго блока. Совместное использование P‑ и Q‑четности обеспечивает RAID 6 двойную защиту и делает его одним из самых надежных уровней массива.

RAID 6 применяется в тех случаях, когда приоритетом является максимальная сохранность данных при сохранении приемлемой производительности и эффективности использования пространства. Он широко используется в дата‑центрах, корпоративных системах хранения и инфраструктуре, где недопустимы потеря информации или длительный простой.

После рассмотрения основных и наиболее применяемых уровней RAID следует упомянуть и другие варианты организации массивов, которые описаны в классификации, но практически не используются на практике. Они демонстрируют разные подходы к распределению данных и хранению четности, однако в современных системах оказались вытеснены более эффективными решениями.

RAID 2 – это уровень массива, при котором данные разбиваются на отдельные биты и распределяются по дискам, а для контроля используются специальные диски с кодами Хэмминга. Такая организация позволяет исправлять ошибки на уровне битов, но требует большого количества накопителей и сложной логики контроллера. Схема хранения данных RAID 2 показана на рисунке 2.6.

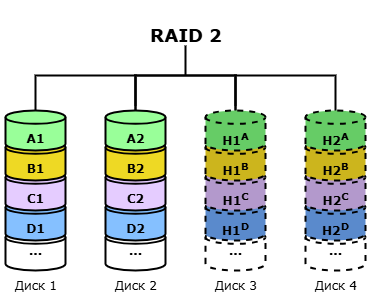


Рисунок 2.6 – Схема данных RAID 2

Ощутимым минусом RAID 2 является чрезмерная избыточность и высокая стоимость реализации. Современные диски и контроллеры уже имеют встроенные механизмы коррекции ошибок, поэтому необходимость в RAID 2 отпала.

RAID 3 – это уровень массива, при котором данные разбиваются на байты и записываются параллельно на несколько дисков, а один выделенный диск хранит четность для всех страйпов. Такая схема обеспечивает высокую скорость последовательного чтения, так как все диски работают синхронно. Схема хранения данных RAID 3 показана на рисунке 2.7.

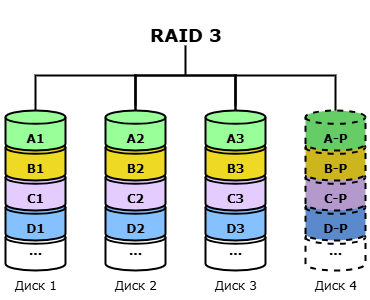


Рисунок 2.7 – Схема данных RAID 3

Основным недостатком RAID 3 является наличие выделенного диска четности, который становится узким местом при записи: все операции обращаются к нему, что снижает производительность.

RAID 4 – это уровень массива, при котором данные разбиваются на блоки (стрипы) и распределяются по дискам, а один выделенный диск хранит четность для всех страйпов. В отличие от RAID 3, здесь используется блочная организация, что повышает гибкость доступа к данным. Схема хранения данных RAID 4 показана на рисунке 2.8.

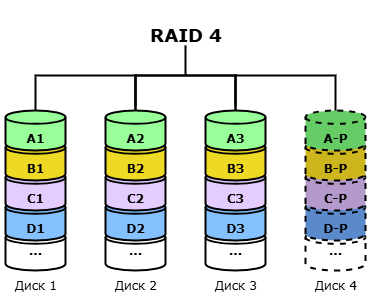


Рисунок 2.8 – Схема данных RAID 4

Недостатком RAID 4 также является выделенный диск четности, который перегружается при каждой операции записи. Эта проблема была решена в RAID 5 за счет распределения четности по всем дискам.

RAID 7 – это проприетарное развитие идей RAID 3/4, в котором применялся выделенный диск для четности, кэширование и специализированный контроллер. Схема хранения данных RAID 7 показана на рисунке 2.9.

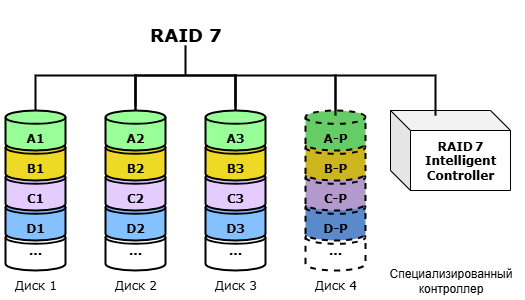


Рисунок 2.9 – Схема данных RAID 7

Главным недостатком RAID 7 стала закрытость технологии и высокая стоимость реализации. Отсутствие стандартизации и совместимости с другими решениями привело к тому, что RAID 7 не получил распространения и остался редким экспериментом.

Таким образом, уровни RAID 2, RAID 3, RAID 4 и RAID 7 представляют собой скорее исторические этапы развития технологий. Они демонстрируют поиск баланса между скоростью, надежностью и эффективностью, но в современных системах вытеснены RAID 5, RAID 6 и комбинированными уровнями, которые обеспечивают более оптимальные характеристики.

Комбинирование уровней объединяет сильные стороны базовых уровней. Такие массивы, как RAID 10, RAID 01, RAID 50 и RAID 60, стали востребованы в реальных системах: они позволяют одновременно повысить производительность и надежность, сохраняя при этом гибкость конфигурации.

RAID 10 – это комбинированный уровень массива, сочетающий принципы RAID 1 и RAID 0. В основе лежит зеркалирование пар дисков (RAID 1), поверх которого выполняется чередование данных (RAID 0). Таким образом, данные сначала дублируются на два и более накопителя, а затем распределяются по зеркальным парам. Для операционной системы RAID 10 выглядит как единый диск, но фактически информация хранится в виде зеркал, объединенных в страйпы. Для организации RAID 10 необходимо как минимум 4 диска. Полезный объем массива равен половине суммарной емкости накопителей, так как половина пространства используется для хранения копий. Схема хранения данных RAID 10 показана на рисунке 2.10.

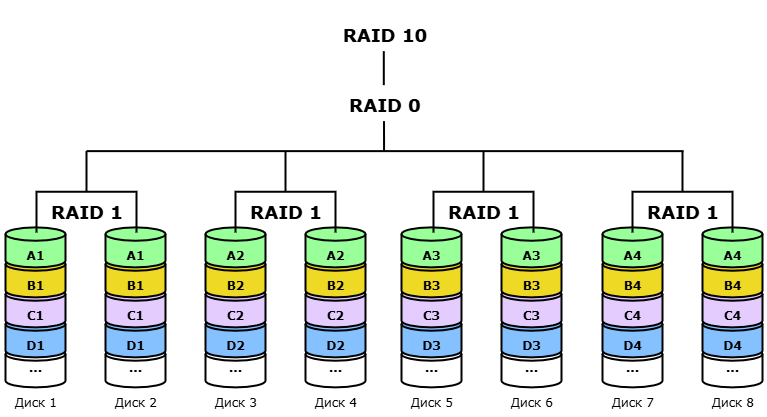


Рисунок 2.10 – Схема данных RAID 10

RAID 10 сочетает высокую надежность и высокую производительность. Благодаря зеркалированию массив способен выдержать отказ одного диска в каждой паре без потери данных. При этом страйпинг обеспечивает ускорение операций чтения и записи, так как нагрузка распределяется между всеми зеркальными парами. RAID 10 часто применяется в системах, где одновременно важны и скорость, и отказоустойчивость: в высоконагруженных базах данных, виртуализационных платформах, корпоративных приложениях.

Недостатком RAID 10 является низкая эффективность использования пространства: половина емкости массива уходит на хранение копий. Кроме того, стоимость реализации выше, чем у массивов без зеркалирования, так как требуется больше дисков.

Восстановление массива RAID 10 относительно простое: при выходе из строя одного накопителя контроллер использует его зеркало. После замены неисправного диска запускается процесс ребилда, в ходе которого данные копируются с исправного зеркала на новый диск. Благодаря страйпингу восстановление может выполняться параллельно для разных пар, что ускоряет процесс.

RAID 01 – это комбинированный уровень массива, сочетающий принципы RAID 0 и RAID 1, но в обратном порядке по сравнению с RAID 10. В основе лежит страйпинг (RAID 0), поверх которого выполняется зеркалирование (RAID 1). Таким образом, данные сначала распределяются по нескольким дискам в виде блоков, а затем весь получившийся страйп целиком дублируется на другую группу дисков. Для операционной системы RAID 01 выглядит как единый диск, но фактически информация хранится в виде зеркал, каждое из которых состоит из страйпированных дисков. Для организации RAID 01 необходимо как минимум 4 диска. Полезный объем массива равен половине суммарной емкости накопителей, так как половина пространства используется для хранения копий. Схема хранения данных RAID 01 показана на рисунке 2.11.

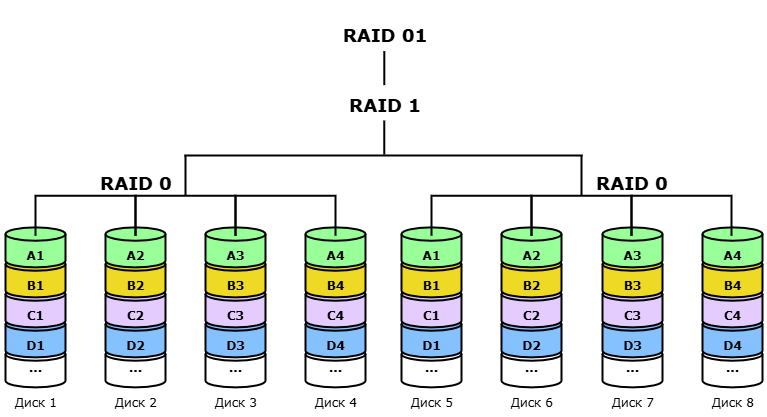


Рисунок 2.11 – Схема данных RAID 01

Достоинство RAID 01 – высокая скорость работы при чтении и записи благодаря страйпингу, а также повышенная надежность за счет зеркалирования. Однако отказоустойчивость у RAID 01 ниже, чем у RAID 10: выход из строя одного диска в одном из страйпов делает недоступной всю группу, и массив продолжает функционировать только за счет зеркала. Поэтому RAID 01 менее популярен и используется редко, уступая RAID 10, который обеспечивает более гибкую защиту от сбоев.

Восстановление массива RAID 01 происходит путем замены неисправного диска и синхронизации его содержимого с зеркала. Однако при выходе из строя целого страйпа нагрузка на оставшуюся группу возрастает, что делает процесс ребилда более уязвимым.

RAID 50 – это комбинированный уровень массива, сочетающий принципы RAID 5 и RAID 0. В основе лежат несколько независимых массивов RAID 5, объединенных в единый страйп (RAID 0). Внутри каждой группы дисков реализуется распределенная четность, обеспечивающая отказоустойчивость, а поверх этих групп выполняется страйпинг, повышающий производительность. Для операционной системы RAID 50 выглядит как единый диск, но фактически данные хранятся в нескольких подмассивах RAID 5, объединенных в чередование. Для организации RAID 50 необходимо как минимум 6 дисков (два массива RAID 5 по три диска), однако на практике чаще используют 8 и более накопителей. Полезный объем массива равен суммарной емкости всех дисков за вычетом одного диска в каждой группе, так как это пространство используется для хранения четности. Схема хранения данных RAID 50 показана на рисунке 2.12.

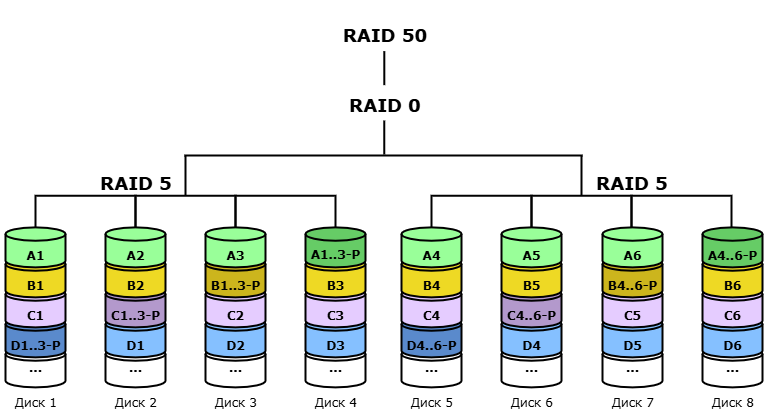


Рисунок 2.12 – Схема данных RAID 50

Главное достоинство RAID 50 – сочетание высокой производительности и повышенной надежности. Благодаря страйпингу массив обеспечивает ускорение операций чтения и записи, а распределенная четность внутри каждой группы защищает данные от выхода из строя одного диска в подмассиве. Недостатком RAID 50 является усложненная структура и необходимость большого числа дисков. Кроме того, при выходе из строя более одного накопителя в одной группе массив теряет работоспособность.

Восстановление массива RAID 50 происходит в два этапа: сначала контроллер восстанавливает данные внутри поврежденного подмассива RAID 5 с использованием четности, затем синхронизирует страйп. При больших объемах дисков процесс ребилда может занимать значительное время, однако благодаря страйпингу нагрузка распределяется между группами, что ускоряет восстановление по сравнению с одиночным RAID 5.

RAID 60 – это комбинированный уровень массива, сочетающий принципы RAID 6 и RAID 0. В основе лежат несколько независимых массивов RAID 6, объединенных в единый страйп (RAID 0). Таким образом, внутри каждой группы дисков реализуется двойная распределенная четность, позволяющая выдержать отказ двух накопителей, а поверх этих групп выполняется страйпинг, повышающий производительность. Для организации RAID 60 необходимо как минимум 8 дисков (два массива RAID 6 по четыре диска), однако на практике используют 12 и более накопителей для повышения эффективности. Полезный объем массива равен суммарной емкости всех дисков за вычетом двух дисков в каждой группе, так как это пространство используется для хранения двойной четности. Схема хранения данных RAID 60 показана на рисунке 2.13.

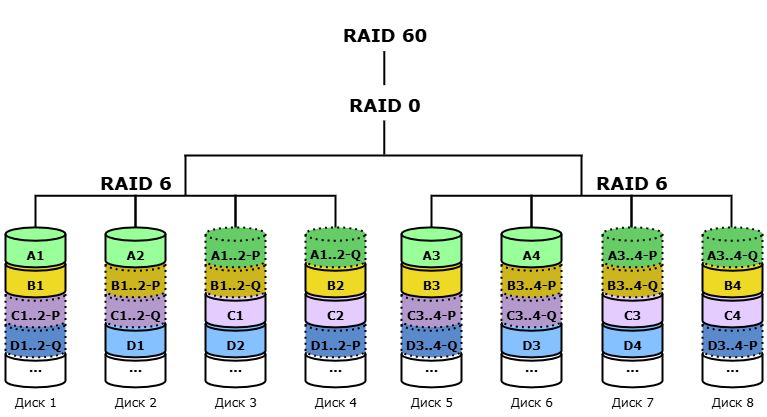


Рисунок 2.13 – Схема данных RAID 60

RAID 60 – это сочетание высокой надежности и производительности. Благодаря страйпингу массив обеспечивает ускорение операций чтения и записи, а двойная четность внутри каждой группы защищает данные от выхода из строя сразу двух дисков в подмассиве. RAID 60 применяется в системах, где критичны и скорость, и максимальная отказоустойчивость при больших объемах хранения: в корпоративных дата‑центрах, облачных инфраструктурах, системах резервного копирования и архивирования. RAID 60 имеет высокую стоимость реализации и необходимость большого числа дисков. Кроме того, вычисление двойной четности замедляет операции записи по сравнению с RAID 50.

Восстановление массива RAID 60 происходит в два этапа: сначала контроллер восстанавливает данные внутри поврежденного подмассива RAID 6 с использованием двойной четности, затем синхронизирует страйп. Благодаря страйпингу нагрузка распределяется между группами, что ускоряет процесс ребилда, однако при очень больших объемах дисков он может занимать значительное время.

Следует отметить, что комбинированные уровни RAID строятся только на основе тех базовых конфигураций, которые доказали свою эффективность и получили широкое распространение: RAID 0, 1, 5 и 6. Попытки объединить менее популярные уровни не нашли практического применения вследствие их ограничений и низкой востребованности.

Комбинируют только те уровни, которые дополняют друг друга разными принципами: RAID 0 (страйпинг) + RAID 1 (зеркалирование), либо RAID 0 (страйпинг) + RAID 5/6 (распределенная четность). Комбинации «четность + зеркалирование» или «четность + четность» не дают практического выигрыша, а лишь усложняют архитектуру и снижают эффективность. Следовательно, теоретически возможные комбинации вроде RAID 15, RAID 51 или RAID 65 оказываются нецелесообразными. Расчеты показывают, что они обеспечивают крайне низкий коэффициент полезного использования дискового пространства (часто менее 50 %), усложняют архитектуру и существенно снижают производительность записи из‑за двойной обработки четности или избыточного зеркалирования. При этом уровень отказоустойчивости у таких схем не превосходит уже существующих решений.

Таким образом, в настоящее время закрепились лишь RAID 10, RAID 01, RAID 50 и RAID 60, которые обеспечивают оптимальный баланс между скоростью, надежностью и эффективностью использования ресурсов.

# 3 Практическое применение дисковых массивов

Дисковые массивы применяются в самых разных масштабах – от домашних компьютеров и рабочих станций до корпоративных дата‑центров и облачных инфраструктур. Реализация RAID может осуществляться как программными средствами операционных систем и файловых систем, так и аппаратными контроллерами или готовыми сетевыми хранилищами (NAS). В зависимости от задач и бюджета выбирается оптимальный вариант: программные решения обеспечивают гибкость и доступность, аппаратные контроллеры – высокую производительность и расширенные функции, а специализированные устройства и серверные платформы – комплексную интеграцию в инфраструктуру предприятия.

Программные реализации RAID разворачиваются на базе уже имеющегося оборудования – персонального компьютера или сервера. Это означает, что их производительность и надежность напрямую зависят от характеристик хоста: мощности центрального процессора, объема оперативной памяти, скорости шин и подсистемы ввода‑вывода. При недостатке ресурсов программный RAID может создавать дополнительную нагрузку на систему, снижая общую производительность.

Одним из наиболее показательных примеров программной реализации RAID является файловая система ZFS (Zettabyte File System), разработанная компанией Sun Microsystems. В отличие от классических подходов, где RAID реализуется отдельным модулем или контроллером, ZFS объединяет в себе функции файловой системы и менеджера томов. Такой подход позволяет управлять массивами напрямую и обеспечивает высокий уровень защиты данных.

Ключевая особенность ZFS заключается в использовании концепции пулов хранения. Вместо работы с отдельными дисками или разделами администратор объединяет устройства в единый пул, внутри которого ZFS самостоятельно распределяет данные и четность. Это упрощает администрирование и позволяет динамически расширять хранилище без сложных операций по переразметке.

ZFS поддерживает несколько уровней отказоустойчивости, аналогичных классическим схемам RAID: RAID‑Z1 (аналог RAID 5), RAID‑Z2 (аналог RAID 6) и RAID‑Z3, уникальная конфигурация с тремя независимыми четностями. Такой набор возможностей делает ZFS особенно востребованной при работе с большими объемами данных.

Важным преимуществом ZFS является использование механизма копирования при записи, а также встроенной системы контрольных сумм, что обеспечивает защиту от «тихого повреждения данных». Эти механизмы выгодно отличают ее от многих других файловых систем.

Наряду с ZFS существуют и другие программные решения, позволяющие организовать RAID‑массивы. В Linux широко применяется mdadm, представляющий собой классический инструмент для создания и управления программными RAID‑массивами на уровне ядра. Он проще по архитектуре, чем ZFS, но обеспечивает базовую функциональность и гибкость конфигурации. В экосистеме Linux также используется Btrfs, файловая система, которая, подобно ZFS, сочетает в себе функции RAID, снапшотов и проверки целостности данных. В среде Windows аналогичные задачи решаются с помощью Storage Spaces, позволяющих объединять диски в пулы и создавать отказоустойчивые массивы.

В зависимости от задач и платформы администратор может выбрать как легкие и универсальные инструменты (mdadm), так и более сложные файловые системы нового поколения (Btrfs, Storage Spaces). Общим для всех программных решений остается то, что их эффективность определяется ресурсами хоста, на котором они развернуты.

Широкое распространение получили аппаратные решения для организации RAID‑массивов, которые в отличие от программных не зависят от ресурсов хоста и обеспечивают стабильную производительность. отдельные системы хранения данных (СХД). В отличие от встроенных решений, такие устройства представляют собой самостоятельные аппаратно‑программные комплексы, которые подключаются к сети и предоставляют доступ к данным множеству пользователей или серверов. Наиболее распространенными архитектурами являются NAS и SAN.

NAS (Network Attached Storage) – это сетевое хранилище, работающее на уровне файловой системы. Оно подключается к локальной сети по протоколам SMB, NFS или AFP и предоставляет пользователям доступ к файлам так же, как к общим папкам. NAS‑устройства, такие как QNAP или Synology, сочетают в себе RAID‑контроллер, дисковую подсистему и специализированную операционную систему. Они удобны в администрировании, поддерживают различные уровни RAID (0, 1, 5, 6, 10 и др.) и часто дополняются функциями резервного копирования, медиасервера или облачной синхронизации. NAS широко применяется в домашних условиях, малом и среднем бизнесе, где требуется централизованное хранение данных при умеренной стоимости.

SAN (Storage Area Network) – это сеть хранения данных, работающая на уровне блочного доступа. В отличие от NAS, SAN не предоставляет файловый доступ, а эмулирует для серверов «удаленные диски», подключаемые по протоколам Fibre Channel или iSCSI. Такая архитектура позволяет объединять десятки и сотни серверов с единой системой хранения, обеспечивая высокую скорость, масштабируемость и отказоустойчивость. SAN‑решения применяются в корпоративных дата‑центрах, где RAID используется совместно с другими технологиями (кластеризация, репликация, снапшоты) для обеспечения непрерывной работы критически важных сервисов.

В обоих случаях RAID является базовым механизмом, обеспечивающим отказоустойчивость и эффективность использования дискового пространства. Однако в отличие от программных решений, NAS и SAN выступают как самостоятельные устройства, независимые от ресурсов конкретного сервера, что делает их более надежными и предсказуемыми в эксплуатации. На рисунке 3.1 приведены изображения аппаратных RAID-контроллеров QNAP.

|  |  |
| --- | --- |
| https://www.pcplanet.ru/public_files/products/43/67/4367b4d30a85fb0f74f7467f4932906b/original.jpg | https://www.techpowerup.com/img/12-01-11/97e.jpg |
| https://m.media-amazon.com/images/I/51fDF9wkmWL.jpg | |
| Рисунок 3.1 – Аппаратные RAID-контроллеры QNAP | |

RAID может быть реализован непосредственно внутри серверов. В этом случае управление массивами осуществляется с помощью аппаратных RAID‑контроллеров, встроенных в архитектуру сервера или устанавливаемых в виде отдельных PCIe‑карт.

Такой контроллер не является отдельной «коробкой» с дисками, а представляет собой часть серверного оборудования, управляющую локальными накопителями, установленными в корпусе. Это позволяет объединять диски в массивы и обеспечивать их отказоустойчивость без привлечения внешних систем хранения.

Современные серверные RAID‑контроллеры, такие как Dell PERC, QCT, HPE Smart Array или LSI MegaRAID, оснащаются собственными процессорами и кэш‑памятью, что позволяет разгрузить центральный процессор и повысить общую производительность. Для защиты данных в кэше при сбое питания применяются батарейные или флэш‑модули, которые сохраняют незавершенные операции до восстановления работы системы. Поддержка различных уровней RAID обеспечивает гибкость конфигурации: администратор может выбрать оптимальное сочетание скорости, надежности и объема в зависимости от задач. Внешний вид встроенного RAID-контроллера показан на рисунке 3.2.

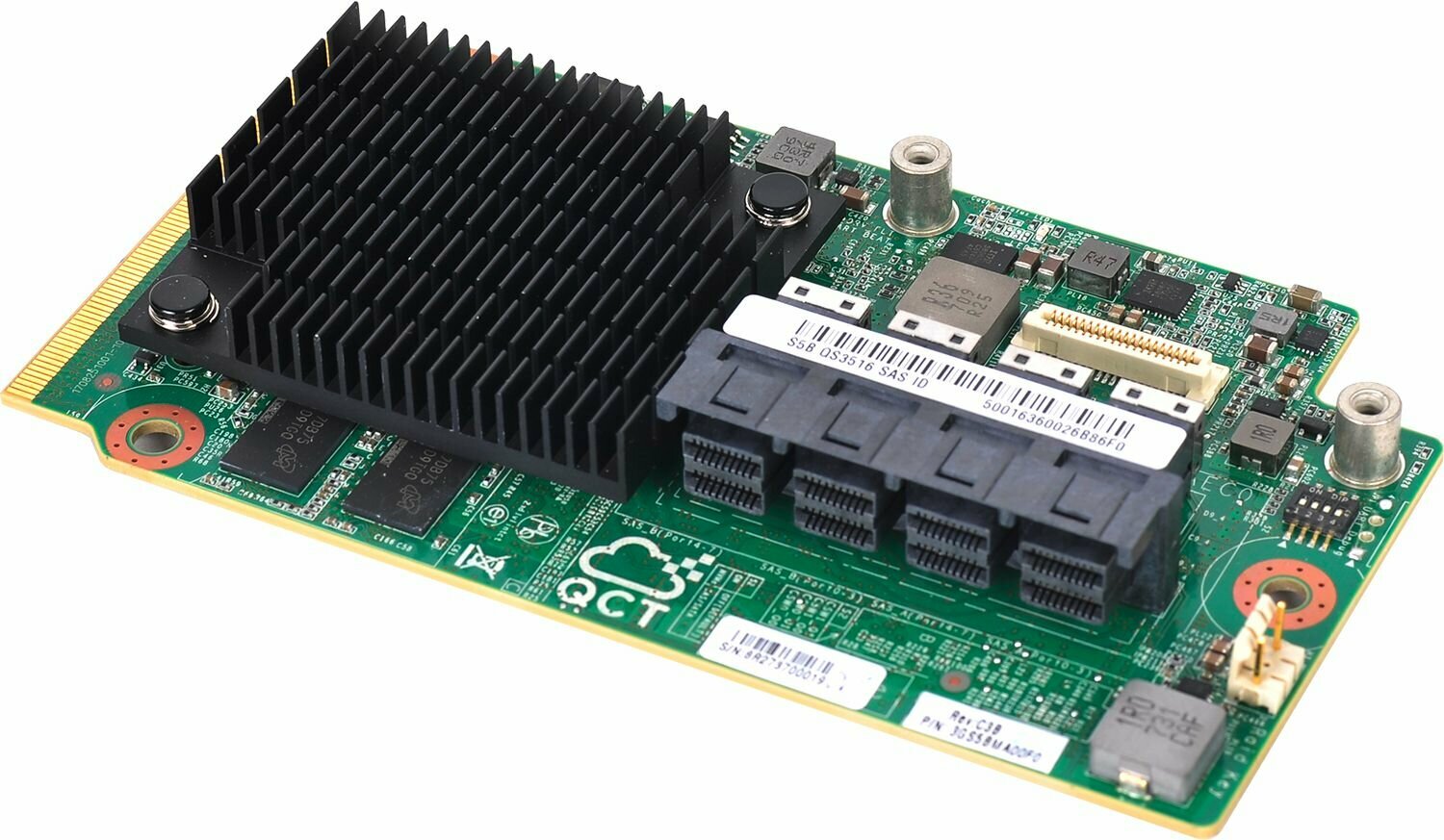


Рисунок 3.2 – Встроенный RAID-контроллер QCT

В рабочих станциях и малых серверах они позволяют организовать быстрые и надежные массивы для хранения проектов, мультимедиа или локальных архивов. В корпоративной среде RAID‑контроллеры обеспечивают базовый уровень защиты и производительности для критически важных приложений, таких как базы данных, ERP‑системы или виртуальные машины. В дата‑центрах встроенные массивы часто используются в сочетании с внешними СХД: локальный RAID служит для ускорения работы операционной системы и приложений, тогда как масштабируемое хранилище выносится в отдельные NAS или SAN‑решения.

Встроенные серверные решения занимают промежуточное место между программными реализациями и отдельными системами хранения. Они обеспечивают высокую производительность и надежность на уровне конкретного сервера, но при этом не обладают масштабируемостью и функциональной насыщенностью специализированных СХД. Тем не менее, именно такие контроллеры остаются основой для большинства серверных платформ, где RAID используется как базовый инструмент защиты и оптимизации локального хранения данных.

# Заключение

В ходе работы было рассмотрено понятие дисковых массивов, их классификация и основные уровни RAID, а также проанализированы практические области применения данной технологии. Дисковые массивы являются неотъемлемым элементом современной инфраструктуры хранения данных, обеспечивая баланс между производительностью, надежностью и эффективным использованием ресурсов.

Классификация массивов выделяет различные подходы к их построению – от простейших схем JBOD до сложных отказоустойчивых конфигураций. Каждый уровень RAID имеет собственные преимущества и ограничения. Выбор зависит от задач: где‑то важнее гибкость и экономичность программных решений, где‑то – простота NAS, а в корпоративной среде – мощность и устойчивость серверных контроллеров и SAN‑инфраструктуры. Но во всех случаях RAID остается базовым инструментом.

Особое внимание уделено практическому применению RAID. Программные реализации, такие как ZFS или Btrfs, демонстрируют гибкость и функциональную насыщенность, но зависят от ресурсов хоста. Отдельные системы хранения в формате NAS и SAN позволяют разгрузить серверы и предоставить централизованный доступ к данным, что делает их востребованными в домашних условиях, малом и среднем бизнесе, а также в корпоративной среде. Встроенные серверные решения на базе аппаратных RAID‑контроллеров обеспечивают высокую производительность и надежность локального хранения, оставаясь базовым инструментом для большинства серверных платформ.

Дисковые массивы RAID играют ключевую роль в построении информационных систем. Их развитие отражает общие тенденции в области ИТ: рост объемов данных, переход к виртуализации и облачным вычислениям, повышение требований к доступности и безопасности информации. Грамотный выбор уровня RAID и архитектуры хранения позволяет не только повысить эффективность работы конкретной системы, но и обеспечить устойчивость всей инфраструктуры с возможностью дальнейшего расширения системы по необходимости