

Содержание:

image not found or type unknown



Введение

RAID

(англ. Redundant array of independent disks — избыточный массив независимых жёстких дисков) — массив из нескольких дисков, управляемых контроллером, взаимосвязанных скоростными каналами и воспринимаемых внешней системой как единое целое. В зависимости от типа используемого массива может обеспечивать различные степени отказоустойчивости и быстродействия. Служит для повышения надёжности хранения данных и/или для повышения скорости чтения/записи информации. Изначально, подобные массивы строились в качестве резерва носителям на оперативной (RAM) памяти, которая в то время была дорогой. Со временем, аббревиатура приобрела второе значение – массив уже был из независимых дисков, подразумевая использование нескольких дисков, а не разделов одного диска, а также дороговизну (теперь уже относительно просто нескольких дисков) оборудования, необходимого для построения этого самого массива.

Рассмотрим, какие бывают RAID массивы. Сперва рассмотрим уровни, которые были представлены учёными из Беркли, потом их комбинации и необычные режимы. Стоит заметить, что если используются диски разного размера (что не рекомендуется), то работать они будут по объёму наименьшего. Лишний объём больших дисков просто будет недоступен.

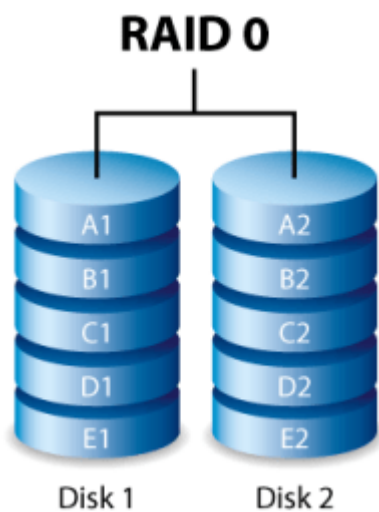
RAID 0. Дисковый массив с чередованием без отказоустойчивости/чётности (Stripe)

Является массивом, где данные разбиваются на блоки (размер блока можно задавать при создании массива) и затем записываются на отдельные диски. В простейшем случае – есть два диска, один блок пишется на первый диск, другой на второй, затем опять на первый и так далее. Также этот режим называется

«чередование», поскольку при записи блоков данных чередуются диски, на которые осуществляется запись. Соответственно, читаются блоки тоже поочерёдно. Таким образом, происходит параллельное выполнение операций ввода/вывода, что приводит к большей производительности. Если раньше за единицу времени мы могли считать один блок, то теперь можем сделать это сразу с нескольких дисков. Основным плюсом данного режима как раз и является высокая скорость передачи данных.

Однако чудес не бывает, а если бывают, то нечасто. Производительность растёт всё же не в N раз (N – число дисков), а меньше. В первую очередь, увеличивается в N раз время доступа к диску, и без того высокое относительно других подсистем компьютера. Качество контроллера оказывает не меньшее влияние. Если он не самый лучший, то скорость может едва заметно отличаться от скорости одного диска. Ну и немалое влияние оказывает интерфейс, которым RAID контроллер соединён с остальной системой. Всё это может привести не только к меньшему, чем N увеличению скорости линейного чтения, но и к пределу количества дисков, установка выше которого прироста давать уже не будет вовсе. Или, наоборот, будет слегка снижать скорость. В реальных задачах, с большим числом запросов шанс столкнуться с этим явлением минимален, ибо скорость весьма сильно упирается в сам жёсткий диск и его возможности.

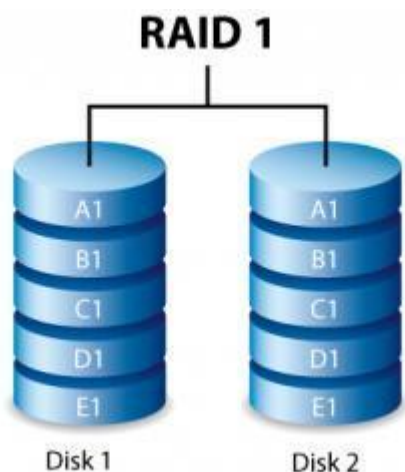
Как видно, в этом режиме избыточности нет как таковой. Используется всё дисковое пространство. Однако, если один из дисков выходит из строя, то, очевидно, теряется вся информация.



RAID 1. Зеркалирование (Mirror)

Суть данного режима RAID сводится к созданию копии (зеркала) диска с целью повышения отказоустойчивости. Если один диск выходит из строя, то работа не прекращается, а продолжается, но уже с одним диском. Для этого режима требуется чётное число дисков. Идея этого метода близка к резервному копированию, но всё происходит «на лету», равно как и восстановление после сбоя (что порой весьма важно) и нет необходимости тратить время на это.

Минусы – высокая избыточность, так как нужно вдвое больше дисков для создания такого массива. Ещё одним минусом является то, что отсутствует какой-либо прирост производительности – ведь на второй диск просто пишется копия данных первого.



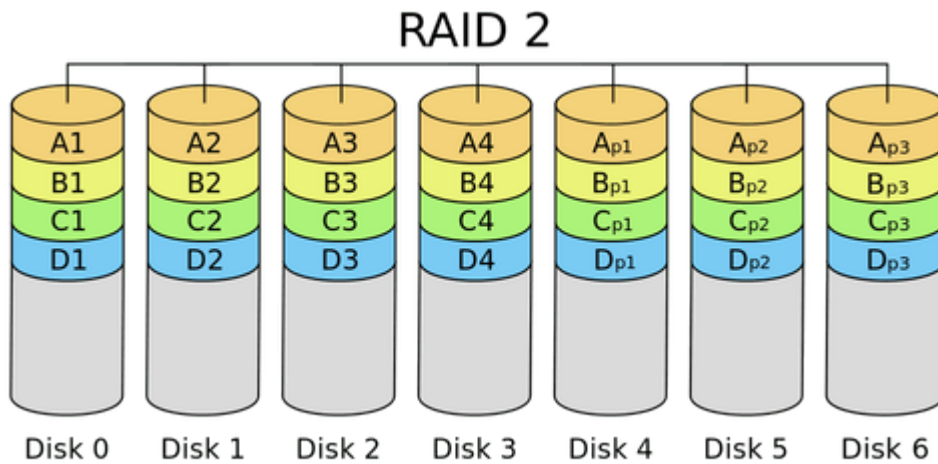
RAID 2 Массив с использованием ошибкоустойчивого кода Хемминга.

Данный код позволяет исправлять и обнаруживать двойные ошибки. Активно используется в памяти с коррекцией ошибок (ECC). В этом режиме диски разбиваются на две группы – одна часть используется для хранения данных и работает аналогично RAID 0, разбивая блоки данных по разным дискам; вторая часть используется для хранения ECC кодов.

Из плюсов можно выделить исправление ошибок «на лету», высокую скорость потоковой передачи данных.

Главным минусом является высокая избыточность (при малом числе дисков она почти двойная, $n-1$). При увеличении числа дисков удельное число дисков хранения ECC кодов становится меньше (снижается удельная избыточность). Вторым

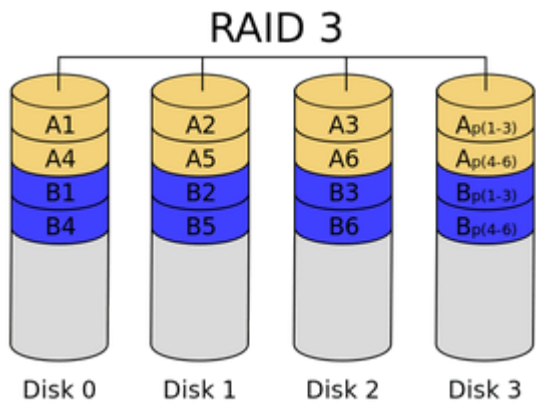
минусом является низкая скорость работы с мелкими файлами. Из-за громоздкости и высокой избыточности с малым числом дисков, данный уровень RAID в данное время не используется, сдав позиции более высоким уровням.



RAID 3. Отказоустойчивый массив с битовым чередованием и чётностью.

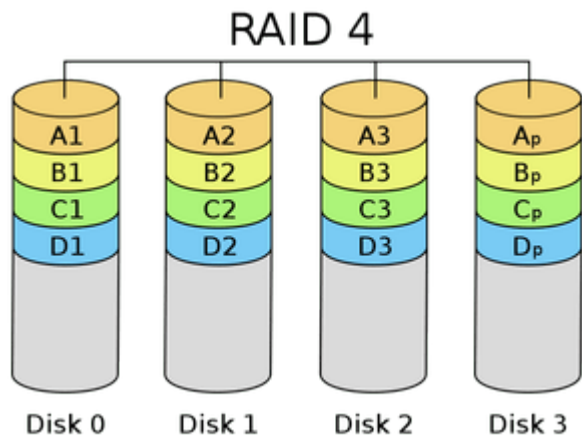
Данный режим записывает данные по блокам на разные диски, как RAID 0, но использует ещё один диск для хранения четности. Таким образом, избыточность намного ниже, чем в RAID 2 и составляет всего один диск. В случае сбоя одного диска, скорость практически не меняется.

Из основных минусов надо отметить низкую скорость при работе с мелкими файлами и множеством запросов. Связано это с тем, что все контрольные коды хранятся на одном диске и при операциях ввода/вывода их необходимо переписывать. Скорость этого диска и ограничивает скорость работы всего массива. Биты чётности пишутся только при записи данных. А при чтении – они проверяются. По причине этого наблюдается дисбаланс в скорости чтения/записи. Одиночное чтение небольших файлов также характеризуется невысокой скоростью, что связано с невозможностью параллельного доступа с независимых дисков, когда разные диски параллельно выполняют запросы.



RAID 4

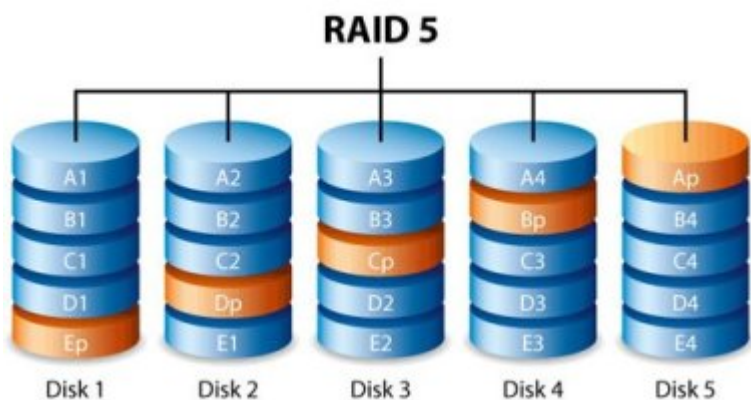
Данные записываются блоками на разные диски, один диск используется для хранения битов чётности. Отличие от RAID 3 заключается в том, что блоки разбиваются не по битам и байтам, а по секторам. Преимущества заключаются в высокой скорости передачи при работе с большими файлами. Также высока скорость работы с большим числом запросов на чтение. Из недостатков можно отметить доставшиеся от RAID 3 – дисбаланс в скорости операций чтения/записи и существование условий, затрудняющих параллельный доступ к данным.



RAID 5. Дисковый массив с чередованием и распределённой чётностью.

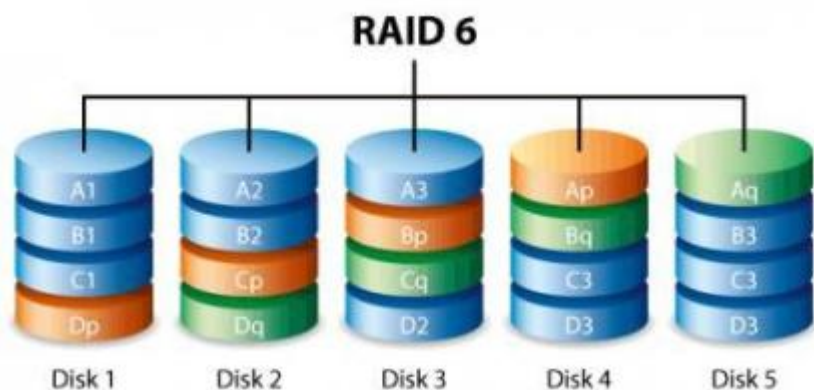
Метод похож на предыдущий, но в нём для битов чётности выделяется не отдельный диск, а эта информация распределяется между всеми дисками. То есть, если используется N дисков, то будет доступен объём N-1 диска. Объём одного

будет выделен под биты чётности, как и в RAID 3,4. Но они хранятся не на отдельном диске, а разделены. На каждом диске есть $(N-1)/N$ объёма информации и $1/N$ объёма заполнено битами чётности. Если в массиве выходит из строя один диск, то он остаётся работоспособным (данные, хранившиеся на нём, вычисляются на основе чётности и данных других дисков «на лету»). То есть, сбой проходит прозрачно для пользователя и порой даже с минимальным падением производительности (зависит от вычислительной способности RAID контроллера). Из преимуществ отметим высокие скорости чтения и записи данных, как при больших объёмах, так и при большом числе запросов. Недостатки – сложное восстановление данных и более низкая, чем в RAID 4 скорость чтения.



RAID 6. Дисковый массив с чередованием и двойной распределённой чётностью.

Всё отличие сводится к тому, что используются две схемы чётности. Система устойчива к отказам двух дисков. Основной сложностью является то, что для реализации этого приходится делать больше операций при выполнении записи. Из-за этого скорость записи является чрезвычайно низкой.



Комбинированные (nested) уровни RAID.

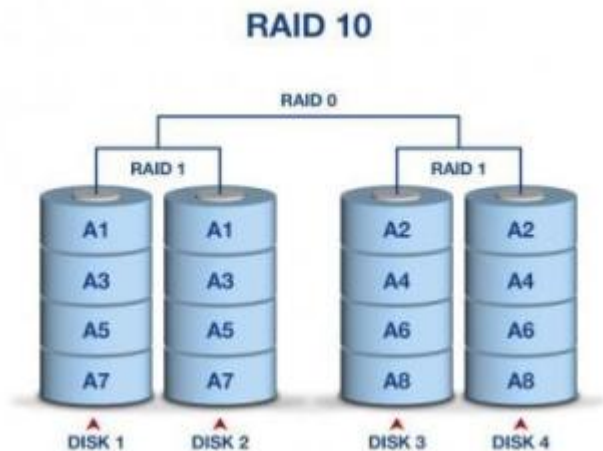
Поскольку массивы RAID являются прозрачными для ОС, то вскоре пришло время и созданию массивов, элементами которых являются не диски, а массивы других уровней. Обычно они пишутся через плюс. Первая цифра означает то, массивы какого уровня входят в качестве элементов, а вторая цифра – то, какую организацию имеет верхний уровень, который объединяет элементы.

RAID 0+1

Комбинация, которая является массивом RAID 1, собранным на базе массивов RAID 0. Как и в массиве RAID 1, доступным будет только половина объёма дисков. Но, как и в RAID 0, скорость будет выше, чем с одним диском. Для реализации такого решения необходимо минимум 4 диска.

RAID 1+0

Также известен, как RAID 10. Является страйпом зеркал, то есть, массивом RAID 0, построенным из RAID 1 массивов. Практически аналогичен предыдущему решению.



RAID 0+3

Массив с выделенной чётностью над чередованием. Является массивом 3-го уровня, в котором данные блоками разбиваются и пишутся на массивы RAID 0. Комбинации, кроме простейших 0+1 и 1+0 требуют специализированных контроллеров, зачастую достаточно дорогих. Надёжность данного вида ниже, чем у следующего варианта.

RAID 3+0

Также известен, как RAID 30. Является страйпом (массивом RAID 0) из массивов RAID 3. Обладает весьма высокой скоростью передачи данных, вкуче с неплохой отказоустойчивостью. Данные сначала разделяются на блоки (как в RAID 0) и попадают на массивы-элементы. Там они опять делятся на блоки, считается их чётность, блоки пишутся на все диски кроме одного, на который пишутся биты чётности. В данном случае, из строя может выйти один из дисков каждого из входящих в состав RAID 3 массива.

RAID 5+0 (50)

Создаётся путём объединения массивов RAID 5 в массив RAID 0. Обладает высокой скоростью передачи данных и обработки запросов. Обладает средней скоростью восстановления данных и хорошей стойкостью при отказе. Комбинация RAID 0+5 также существует, но больше теоретически, так как даёт слишком мало преимуществ.

RAID 5+1 (51)

Сочетание зеркалирования и чередования с распределённой четностью. Также вариантом является RAID 15 (1+5). Обладает очень высокой отказоустойчивостью. Массив 1+5 способен работать при отказе трех дисков, а 5+1 – пяти из восьми дисков.

RAID 6+0 (60)

Чередование с двойной распределённой четностью. Иными словами – страйп из RAID 6. Как уже говорилось применительно к RAID 0+5, RAID 6 из страйпов не получил распространения (0+6). Подобные приёмы (страйп из массивов с четностью) позволяют повысить скорость работы массива. Ещё одним преимуществом является то, что так можно легко повысить объём, не усложняя ситуации с задержками, необходимыми на вычисление и запись большего числа битов четности.

RAID 100 (10+0)

RAID 100, также пишущийся, как RAID 10+0, является страйпом из RAID 10. По своей сути, он схож с более широким RAID 10 массивом, где используется вдвое больше дисков. Но именно такой «трехэтажной» структуре есть своё объяснение. Чаще всего RAID 10 делают аппаратным, то есть силами контроллера, а уже страйп из них делают программно. К такой уловке прибегают, чтобы избежать проблемы,

о которой говорилось в начале статьи – контроллеры имеют свои ограничения по масштабируемости и если воткнуть в один контроллер двойное число дисков, прироста можно при некоторых условиях вообще не увидеть. Программный же RAID 0 позволяет создать его на базе двух контроллеров, каждый из которых держит на борту RAID 10. Так, мы избегаем «бутылочного горлышка» в лице контроллера. Ещё одним полезным моментом является обход проблемы с максимальным числом разъёмов на одном контроллере – удваивая их число, мы удваиваем и число доступных разъёмов.

Нестандартные режимы RAID

Двойная четность

Распространённым дополнением к перечисленным уровням RAID является двойная четность, порой реализованная и потому называемая «диагональной четностью». Двойная четность уже внедрена в RAID 6. Но, в отличие от нее, четность считается над другими блоками данных. Недавно спецификация RAID 6 была расширена, потому диагональная четность может считаться RAID 6. Если для RAID 6 четность считается как результат сложения по модулю 2 битов, идущих в ряд (то есть сумма первого бита на первом диске, первого бита на втором и т.д.), то в диагональной четности идет смещение. Работа в режиме сбоя дисков не рекомендуется (ввиду сложности вычисления утраченных битов из контрольных сумм).

RAID-DP

Является разработкой NetApp RAID массива с двойной четностью и подпадает под обновленное определение RAID 6. Использует отличную от классической RAID 6 реализации схему записи данных. Запись ведется сначала на кеш NVRAM, снабжённый источником бесперебойного питания, чтобы предотвратить потерю данных при отключении электричества. Программное обеспечение контроллера, по возможности, пишет только цельные блоки на диски. Такая схема предоставляет большую защиту, чем RAID 1 и имеет более высокую скорость работы, нежели обычный RAID 6.

RAID 1,5

Был предложен компанией Highpoint, однако теперь применяется очень часто в контроллерах RAID 1, без каких-либо выделений данной особенности. Суть сводится к простой оптимизации – данные пишутся как на обычный массив RAID 1 (чем 1,5 по сути и является), а читают данные с чередованием с двух дисков (как в RAID 0). В конкретной реализации от Highpoint, применявшейся на платах DFI серии LanParty на чипсете nForce 2, прирост был едва заметным, а порой и нулевым. Связано это, вероятно, с невысокой скоростью контроллеров данного производителя в целом в то время.

RAID 1E

Комбинирует в себе RAID 0 и RAID 1. Создаётся минимум на трёх дисках. Данные пишутся с чередованием на три диска, а со сдвигом на 1 диск пишется их копия. Если пишется один блок на три диска, то копия первой части пишется на второй диск, второй части – на третий диск. При использовании четного числа дисков лучше, конечно, использовать RAID 10.

RAID 5E

Обычно при построении RAID 5 один диск оставляют свободным (spare), чтобы в случае сбоя система сразу стала перестраивать (rebuild) массив. При обычной работе этот диск работает вхолостую. Система RAID 5E подразумевает использование этого диска в качестве элемента массива. А объём этого свободного диска распределяется по всему массиву и находится в конце дисков. Минимальное число дисков – 4 штуки. Доступный объём равен $n-2$, объём одного диска используется (будучи распределенным между всеми) для четности, объём еще одного – свободный. При выходе из строя диска происходит сжатие массива до 3-х дисков (на примере минимального числа) заполнением свободного пространства. Получается обычный массив RAID 5, устойчивый к отказу ещё одного диска. При подключении нового диска, массив разжимается и занимает вновь все диски. Стоит отметить, что во время сжатия и распаковки диск не является устойчивым к выходу еще одного диска. Также он недоступен для чтения/записи в это время. Основное преимущество – большая скорость работы, поскольку чередование происходит на большем числе дисков. Минус – что нельзя данный диск назначать сразу к нескольким массивам, что возможно в простом массиве RAID 5.

RAID 5EE

Отличается от предыдущего только тем, что области свободного места на дисках не зарезервированы одним куском в конце диска, а чередуются блоками с битами

четности. Такая технология значительно ускоряет восстановление после сбоя системы. Блоки можно записать прямо на свободное место, без необходимости перемещения по диску.

RAID 6E

Аналогично с RAID 5E использует дополнительный диск для повышения скорости работы и распределения нагрузки. Свободное место разделяется между другими дисками и находится в конце дисков.

RAID 7

Данная технология является зарегистрированной торговой маркой фирмы Storage Computer Corporation. Массив, основывающийся на RAID 3, 4, оптимизированный для повышения производительности. Основное преимущество заключается в использовании кэширования операций чтения/записи. Запросы на передачу данных осуществляются асинхронно. При построении используются диски SCSI. Скорость выше решений RAID 3,4 приблизительно в 1,5-6 раз.

Intel Matrix RAID

Является технологией, представленной Intel в южных мостах, начиная с ICH6R. Суть сводится к возможности комбинации RAID массивов разных уровней на разделах дисков, а не на отдельных дисках. Скажем, на двух дисках можно организовать по два раздела, два из них будут хранить на себе операционную систему на массиве RAID 0, а другие два – работая в режиме RAID 1 – хранить копии документов.

Linux MD RAID 10

Это RAID драйвер ядра Linux, предоставляющий возможность создания более продвинутой версии RAID 10. Так, если для RAID 10 существовало ограничение в виде чётного числа дисков, то этот драйвер может работать и с нечетным. Принцип для трех дисков будет тем же, что в RAID 1E, когда происходит чередование дисков по очереди для создания копии и чередования блоков, как в RAID 0. Для четырех дисков это будет эквивалентно обычному RAID 10. Помимо этого, можно задавать, на какой области диска будет храниться копия. Скажем, оригинал будет в первой половине первого диска, а его копия – во второй половине второго. Со второй половиной данных – наоборот. Данные можно дублировать несколько раз. Хранение копий на разных частях диска позволяет достичь большей скорости доступа в результате разнородности жесткого диска (скорость доступа меняется в

зависимости от расположения данных на пластине, обычно разница составляет два раза).

RAID-K

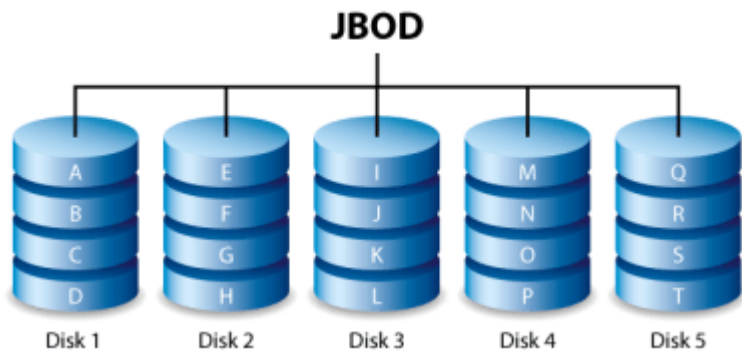
Разработан компанией Kaleidescape для использования в своих медиа устройствах. Схож с RAID 4 с использованием двойной четности, но использует другой метод отказоустойчивости. Пользователь может легко расширять массив, просто добавляя диски, причём в случае, если он содержит данные, данные будут просто добавлены в него, вместо удаления, как это требуется обычно.

RAID-Z

Разработка компании Sun. Самой большой проблемой RAID 5 является потеря информации в результате отключения питания, когда информация из дискового кеша (который является энергозависимой памятью, то есть не хранит данные без электричества) не успела сохраниться на магнитные пластины. Такое несовпадение информации в кеше и на диске называют не когерентностью. Сама организация массива связана с файловой системой Sun Solaris – ZFS. Используется принудительная запись содержимого кеш-памяти дисков, восстанавливать можно не только весь диск, но и блок «на лету», когда контрольная сумма не совпала. Ещё немаловажным аспектом является идеология ZFS – она не меняет данные при необходимости. Вместо этого она пишет обновлённые данные и потом, убедившись, что операция прошла уже удачно, меняет указатель на них. Таким образом, удаётся избежать потери данных при модификации. Мелкие файлы дублируются вместо создания контрольных сумм. Это тоже делается силами файловой системы, поскольку она знакома со структурой данных (массивом RAID) и может выделять место под эти цели. Существует также RAID-Z2, которая, подобно RAID 6 способна выдержать отказ двух дисков с помощью использования двух контрольных сумм.

JBOD

То, что не является RAID в принципе, но часто вместе с ним употребляется. Дословно переводится как «просто набор дисков» (just a bunch of disks) Технология объединяет все диски, установленные в системе в один большой логический диск. То есть, вместо трех дисков будет виден один крупный. Используется весь суммарный объем дисков. Ускорения ни надежности, ни производительности нет.



Drive Extender

Функция, заложенная в Window Home Server. Совмещает в себе JBOD и RAID 1. При необходимости создания копии, она не дублирует сразу файл, а ставит NTFS разделе метку, указывающую на данные. При простое система копирует файл так, чтобы место на дисках было максимальным (использовать можно диски разного объема). Позволяет достичь многих преимуществ RAID – отказоустойчивости и возможности простой замены вышедшего из строя диска и его восстановления в фоновом режиме, прозрачности местонахождения файла (вне зависимости от того, на каком диске он находится). Также можно проводить параллельный доступ с разных дисков с помощью вышеуказанных меток, получая сходную с RAID 0 производительность.

UNRAID

Разработана компанией Lime technology LLC. Эта схема отличается от обычных RAID массивов тем, что позволяет смешивать диски SATA и PATA в одном массиве и диски разных объема и скорости. Для контрольной суммы (четности) используется выделенный диск. Данные не чередуются между дисками. В случае отказа одного диска, теряются только файлы, на нём хранящиеся. Однако, с помощью четности они могут быть восстановлены. UNRAID внедрен как добавление к Linux MD (multidisk).

Большинство видов RAID массивов не получило распространения, часть используется в узких сферах применения. Наиболее массовыми, от простых пользователей до серверов начального уровня стали RAID 0, 1, 0+1/10, 5 и 6. Нужен ли вам рейд-массив для ваших задач – решать вам. Теперь вы знаете, в чём их отличия друг от друга.

Список используемых источников

- <https://web-profi.by/raid-massivy-klassifikaciya-osobennosti-primeneniye/>

(Дата обращения 02.04.19)

- <https://compress.ru/article.aspx?id=21065>

(Дата обращения 02.04.19)

- <https://studfiles.net/preview/1175523/page:22/>

(Дата обращения 02.04.19)