Игорь Афонин

Надёжность избыточных дисковых массивов

Введение

Проектирование современной ИТ-системы предприятия невозможно без расчёта показателей надёжности. При этом основное внимание уделяется показателям надёжности функционирования системы, доступности сервисов и минимизации последствий простоя. В то же время основную ценность представляют данные.

Несмотря на тенденцию всё большего использования в системах хранения данных твердотельных накопителей (Solid Stade Drive, SSD), основным массовым компонентом хранения продолжают оставаться шпиндельные накопители – жёсткие диски (Hard Disk Drive, HDD). Именно они на текущий момент обеспечивают большие объёмы хранения по минимальной стоимости. Надо отметить, что не уделяется должного внимания проектированию системы хранения данных, для которой ключевым параметром является надёжность хранения и обеспечение доступности данных.

Современные жёсткие диски являются высокотехнологичными устройствами с ёмкостью до 14 Тбайт на устройство. Это практический предел, достижимый при использовании основного на текущий момент перпендикулярного способа записи данных на магнитный диск. Новые технологии, в частности, метод записи HAMR (Heat Assisted Magnetic Recording – тепловая магнитная запись), позволят достичь в ближайшее время ёмкости дисков до 30 Тбайт и более и минимизировать стоимость хранения данных, хотя и текущая ёмкость дисков позволяет строить системы хранения до нескольких петабайт в пределах одного шасси.

Отказы жёсткого диска

Современный жёсткий диск является высокотехнологичным устройством, состоящим из большого количества электронных компонентов, механических деталей и узлов, выполненных с микронными допусками (рис. 1). Несмотря на высокую сложность, конструктивные и технологические решения позволили обеспечить высокую надёжность жёстких дисков со средним временем наработки на отказ достигающим 2–2,5 млн часов. На первой взгляд, это гарантирует бесперебойную работу. Но с постоянно растущими требованиями по объёму хранения данных необходимо увеличивать количество дисков (шпинделей) , что влечёт за собой уменьшение надёжности системы хранения.

Рассмотрим основные причины отказов жёстких дисков. Дерево отказов представлено на рис. 2.

Для жёстких дисков, основная задача которых – хранение данных, отказ – это их потеря, а точнее, невозможность их считывания. Современный подход выделяет два типа отказов дисков: первый – функциональный (явный) отказ, или,

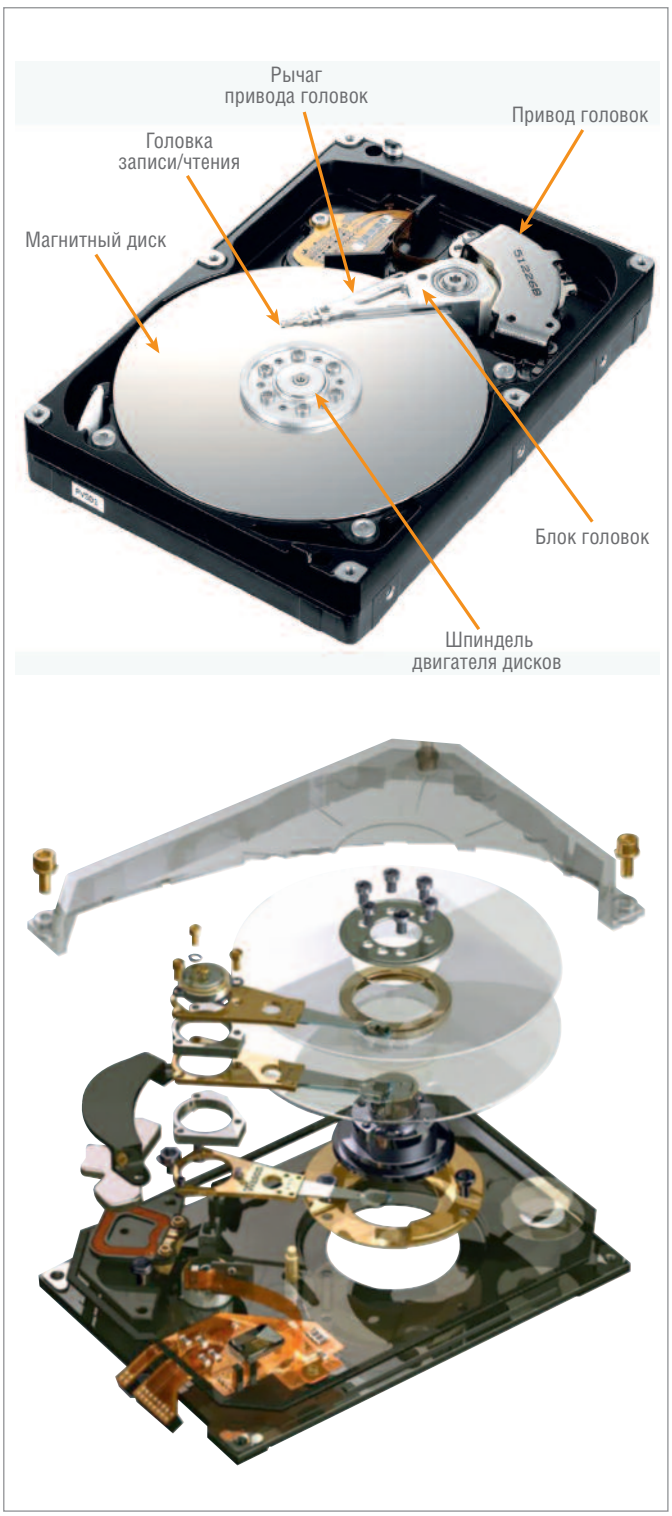

Рис. 1. Устройство жёсткого диска

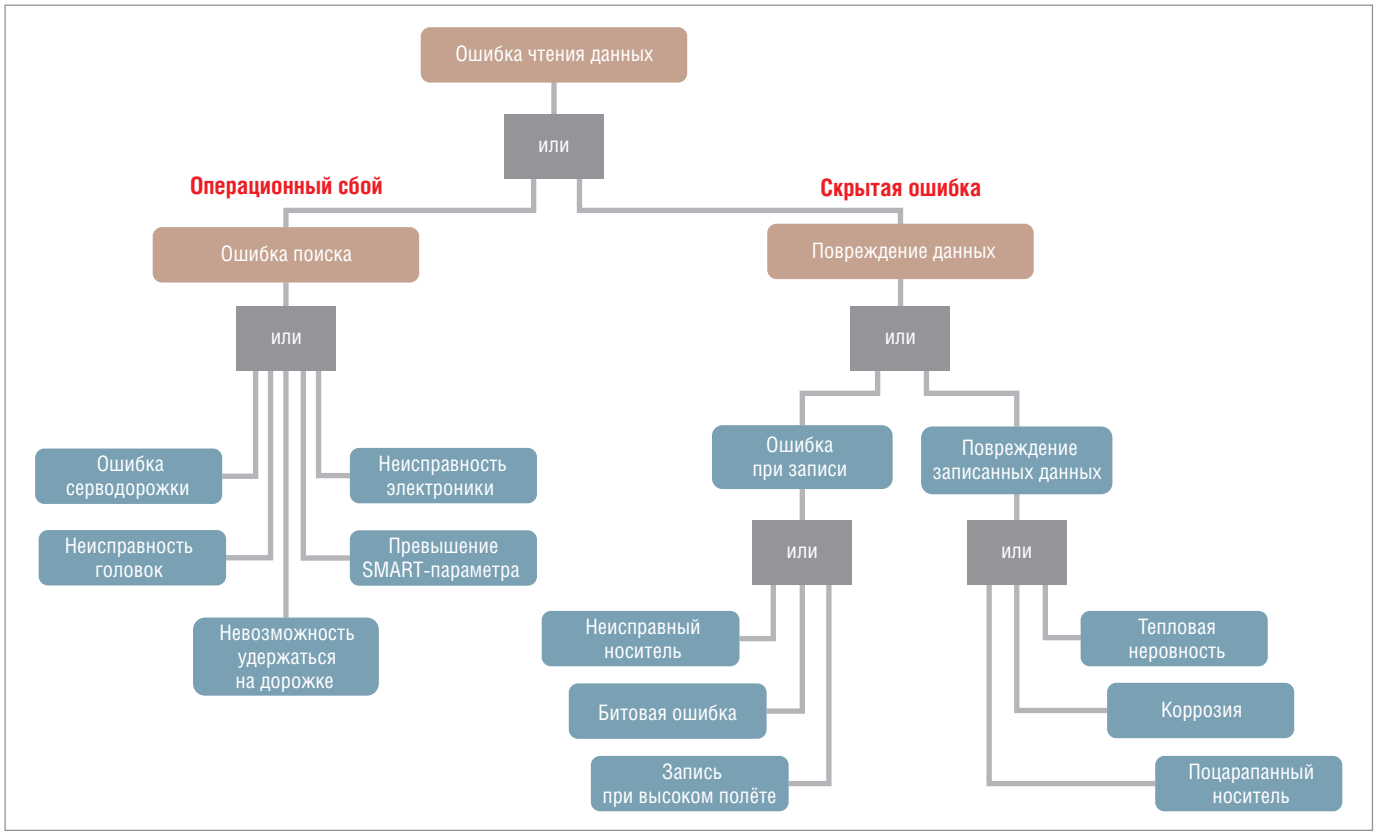

Рис. 2. Дерево отказов жёсткого диска

как принято говорить, операционный сбой, и второй – скрытая ошибка. Функциональный отказ обнаруживается на аппаратном уровне обслуживающим накопитель контроллером и при правильном построении системы не ведёт к потере данных. Скрытые ошибки явно не проявляются, но их последствия могут быть катастрофическими и в итоге способны привести к полной потере данных. Для обнаружения и исправления скрытых ошибок требуется применение специальных алгоритмов и дополнительных вычислительных ресурсов.

Основными причинами функционального отказа являются нарушение серворазметки, сбои системы позиционирования, сбои и отказ в работе электроники накопителя, поломки считывающих головок и превышение лимита критичных SMART-параметров (Self-Monitoring, Analysis and Reporting Technology – технология самоконтроля, анализа и отчётности, или технология оценки состояния жёсткого диска встроенной аппаратурой самодиагностики, а также механизм предсказания времени выхода его из строя).

Функциональный отказ проявляется двумя способами: во-первых, данные не могут быть записаны на жёсткий диск, во-вторых, после того как данные записаны и всё ещё присутствуют на жёстком диске (пластине), электронные или механические неисправности не позволяют считать их. Во втором случае для некоторых отказов возможно восстановление информации с использованием специального оборудования и программного обеспечения. Восстановление данных при отказе блока электроники можно осуществить заменой платы электроники от аналогичного диска.

Среднее время наработки на отказ

Величина функциональных отказов определяет надёжность диска как устройства и выражается через среднее время наработки на отказ – Mean Time Between Failure (*MTBF*).

MTBF является статистическим термином и указывается в технической спецификации устройства. Необходимо понимать, что эта характеристика относится ко всей популяции дисков данной модели, а не к конкретному устройству, и является средней наработкой всех протестированных по специальной методике дисков, отнесённой к количеству отказов.

Значение *MTBF* вычисляется на основании большого (статистически значимого) количества приводов, непрерывно работающих на тестовом сайте, с экстраполяцией данных в соответствии с различными известными статистическими моделями для получения результатов.

Следует отметить, что параметр *MTBF* в целом характеризует надёжность восстанавливаемого устройства и определяется как

$$MTBF = MTTF + MTTR.$$

Здесь *MTTF* (Mean Time To Failure) – средняя наработка до отказа; *MTTR* (Mean Time To Repair) – среднее время до восстановления работоспособности (рис. 3).

Для компонентов системы, в частности, для жёстких дисков производитель, как правило, приводит значение параметра средней наработки на отказ (*MTBF*), которым обычно

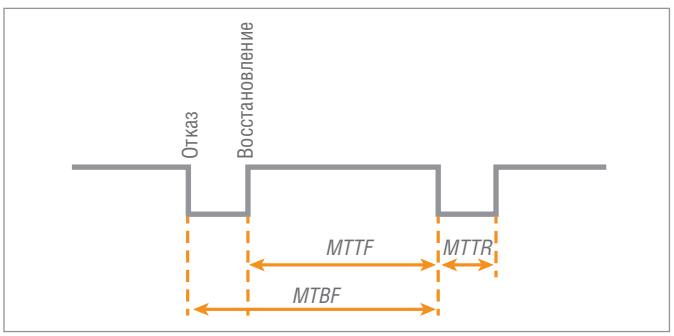

Рис. 3. Среднее время наработки между отказами

оперируют при расчётах надёжности системы. При расчётах показателей надёжности систем с несколькими промежуточными состояниями необходимо оперировать показателем времени наработки до отказа. Обычно в системах хранения данных диск заменяется на исправный из «горячего» резерва автоматически или обслуживающим персоналом из комплекта ЗИП. Поэтому можно считать, что значение *MTTR* равно нулю и таким образом получаем, что *MTBF* = *MTTF*. Даже в случае, когда замена идёт из удалённых запасов и это время исчисляется неделями или месяцами, время замены (*MTTR*) значительно меньше *MTBF* и можно принять *MTBF* равным *MTTF*.

MTBF системы в целом состоит из *MTBF* компонентов и вычисляется по формуле:

$$MTBF_{sys} = \left(\sum_{i=1}^{N} MTBF_i^{-1} \right)^{-1}.$$

Здесь $MTBF_i$ – наработка на отказ *i*-го компонента системы.

Если *MTBF* компонентов системы одинаковы, что характерно для дисковой подсистемы, в которой обычно используются однотипные диски, то для данного случая получаем следующее выражение

$$MTBF_{sys} = \frac{MTBF_{disk}}{N}.$$

Здесь $MTBF_{disk}$ – *MTBF* диска, *N* – количество дисков.

Значения *MTBF* современных жёстких дисков, как было указано ранее, составляют до 2 500 000 часов, что может привести к выводу о том, что спецификация диска обещает более 100 лет его непрерывной работы. Но *MTBF* – это лишь расчётная величина, получаемая на основании параллельного сбора статистики отказов большого числа экземпляров нового исправного оборудования, у которого интенсивность отказов минимальна.

По мере старения и износа вероятность отказов возрастает, также это происходит, если в системе эксплуатируется большое количество дисков.

$$Q_{sys} = \frac{N \times T}{MTBF}.$$

Здесь Q_{sys} – вероятность отказа системы, *N* – количество дисков в системе, *T* – период времени.

Типичный пример для системы, в которой используются 114 однотипных дисков, с *MTBF* = 1 000 000 часов (144 года): за год её работы мы должны ожидать, что один диск выйдет из строя.

$$Q_{114} = \frac{114 \times 1}{114} = 1.$$

Таким образом, для такой системы, как минимум, один диск нужно держать в ЗИП.

Ежегодная вероятность сбоев

Значение *MTBF* имеет смысл только для расчётов параметров надёжности системы, о чём будет рассказано далее. И *MTBF* также не является гарантией относительной надёжности семейства продуктов. Более высокий показатель *MTBF* просто предлагает более надёжную серию (семейство) механизмов (в зависимости от согласованности используемых статистических моделей).

Стоит отметить, что практические значения *MTBF*, которые включают все возвраты дисков производителям независимо от причины, обычно составляют 50–60% от прогнозируемого *MTBF*. На эту величину и следует ориентироваться при эксплуатации изделий.

Вместо *MTBF* гораздо практичнее пользоваться параметром *AFR* — Annual Failure Rate, или ежегодная вероятность сбоев (ещё его называют Reliability – показатель надёжности), выводимым из *MTBF*.

Он вычисляется как

$$AFR = \frac{Failures}{Year} = \frac{Failures}{Hours} \times \frac{Hours}{Years} = \frac{1}{MTBF} \times POH = \frac{POH}{MTBF}.$$

Здесь *POH* – время работы в год (Power-On-Hours per year).

Для режима работы системы 24/7, то есть 24 часа в день × 365 дней в году , типичное значение *POH* составляет 8760, и, соответственно, получаем выражение:

$$AFR = \frac{8760}{MTBF}.$$

Это выражение даёт оценочную долю устройств, которые откажут в течение года. Тогда 1–*AFR* – это доля устройств, которые не откажут в течение года.

Итак, *MTBF* является вероятным средним числом часов работы между отказами, AFR является вероятным процентом отказов в год, исходя из общего количества установленных единиц системы аналогичного типа.

Скрытые ошибки данных

Под скрытыми ошибками данных (Silent Date Corruption) понимают не обнаруживаемые электроникой накопителя в процессе работы ошибки. Причиной этого может быть:

- нарушение данных на соседних дорожках при записи;
- отсутствие модификации оригинальных данных при записи;
- ошибки чтения данных при неправильной интерпретации кодов коррекции ошибок (в случае множественных ошибок);
- считывание неверных данных из-за ошибок позиционирования.

Причинами возникновения этих ошибок являются производственные дефекты магнитного слоя, коррозийные и физические повреждения магнитного слоя в процессе эксплуатации, временные сбои в позиционировании магнитных головок, например из-за вибраций, ошибки позиционирования из-за термического расширения рабочей поверхности вследствие нарушений температурного режима эксплуатации накопителя.

Фактическим параметром, характеризующим скрытые ошибки данных, является *URE* (Unrecoverable Read Errors – невосстанавливаемые ошибки чтения), определяемый как отношение числа ошибок к объёму (количеству) считанных данных:

$$URE = \frac{N_{error}}{C_{bit}}.$$

Здесь N_{error} – количество ошибок; C_{bit} –объём считанных данных в битах.

Минимальной величиной считывания данных с диска является сектор. Встроенные механизмы позволяют исправить некоторые ошибки чтения сектора, но иногда они не справляются со своей задачей и сектор прочитать не удаётся. Эту ошибку первоначально обозначили как скрытую ошибку сектора – Latent Sector Errors (*LSE*), и она применялась в расчёте на один сектор.

В настоящее время разные производители для различных моделей дисков приводят эти величины как в секторах, так и

в битах, что, в принципе, даёт примерно одинаковые значения при расчётах.

Значение *URE* весьма мало́ и приводится для одной ошибки, то есть $N_{error} = 1$.

Исходя из этого, получаем выражение для объёма считываемых данных, при котором произойдёт ошибка чтения:

$$C_{bit} = \frac{1}{URE}.$$

Типичные значения *URE* и объём считываемых данных, для которых вероятность возникновения ошибки чтения равна 1, приведены в табл. 1. Как видно, значение параметра *URE*, то есть вероятность ошибки считывания 1 бита, весьма мало́. Но следует понимать, что высоконагруженные системы, особенно в режиме восстановления, считывают значительные объёмы данных и возникновение ошибки чтения и, следовательно, потери данных довольно высоко.

Так, при восстановлении RAID-массива происходит считывание данных с работающих дисков и запись информации на новый диск, и общий объём данных будет составлять:

$$C_{read} = C_{bit} \times K$$

или

$$C_{read} = C_{bit} \times (N - R).$$

Здесь C_{read} – объём прочитанных данных (в битах); C_{bit} – объём (размер) одного диска (в битах); K – количество дисков с данными (полезный объём RAID-массива); N – количество дисков в RAID-массиве; R – количество избыточных дисков.

Для типичного дискового массива в 2U-шасси, состоящего из 12 дисков ёмкостью 12 Тбайт каждый, объединённых в RAID 5 (один диск для обеспечения избыточности), получаем объём считываемых данных 132 Тбайт:

$$C_{read} = 12 \text{ Тбайт} \times (12-1) = 132 \text{ Тбайт} = 10^{15} \text{ бит}.$$

При величине *URE*, соответствующей значению 10^{15}, что является типичным значением для современных дисков большой ёмкости, используемых в системах хранения данных, восстановление такого массива может не произойти, то есть информация будет потеряна.

Таким образом, можно определить максимально допустимый объём дискового массива в зависимости от типа используемых дисков, с заданной допустимой вероятностью восстановления данных.

Среднее время восстановления диска в системе хранения данных

Следующим важным параметром при расчёте надёжности дисковой подсистемы, является время до полного восстановления (*MTTR*) неисправного компонента (диска).

В отличие от стандартных систем замена диска в системах хранения данных не означает, что диск становится работоспособным. Для восстановления работоспособности заменённого диска необходимо записать на него данные, то есть перестроить RAID-массив и вернуть его в работоспособное состояние. Таким образом, время восстановления диска в системе хранения данных определяется следующим выражением:

Таблица 1

Типичные значения параметра невосстанавливаемой ошибки чтения (URE)

Тип диска	Значение URE	Объём данных до возникновения сбоя
Desktop	10^{-14}	12,5 Тбайт
Nearline	10^{-15}	125 Тбайт
Entrerprise	10^{-16}	1 250 Тбайт

$$MTTR_{disk} = T_{DIAG} + T_{REP} + T_{RBLD}.$$

Здесь $MTTR_{disk}$ – время ремонта (замены) диска; T_{DIAG} – время обнаружения неисправного диска; T_{REP} – время, необходимое для ремонта (замены) неисправного диска; T_{RBLD} – время восстановления потерянных данных (Rebuild) на новом диске.

Время замены может значительно варьироваться: часы, дни или нулевое время при «горячем» резерве (hot spare), когда резервный диск уже установлен в системе и сразу же включается в работу при обнаружении отказа какого-либо диска.

Если принять, что в дисковом массиве операции записи и чтения выполняются одновременно на все диски, а время расчёта контрольных сумм значительно меньше времени чтения/записи на диск, то время восстановления можно оценить как

$$T_{RBLD} = \frac{C_{VOL}}{P_{disk}}.$$

Здесь C_{VOL} – размер тома, P_{DISK} – скорость чтения данных с диска.

Учитывая, что при наличии диска, находящегося в «горячем» резерве, $T_{DIAG} = 0$ и $T_{REPL} = 0$, время восстановления диска будет:

$$MTTR_{disk} = \frac{C_{VOL}}{P_{disk}}.$$

В спецификациях дисков приводится такой параметр, как максимальная постоянная скорость передачи данных по внешнему диаметру (Мбайт/с). Типичная скорость передачи данных современных дисков со скоростью вращения шпинделя 7200 об./мин составляет от 130 до 250 Мбайт/с. Надо понимать, что это максимально возможная потоковая скорость. На внутренних диаметрах она значительно ниже.

На практике восстановление конкурирует ещё и с рабочими запросами ввода-вывода, поэтому скорость передачи данных нужно оценивать как 1/3 от указанной в спецификации или получать экспериментальным путём для заданной конфигурации RAID-массива и рабочей нагрузки.

Ожидаемое время до потери данных

Для систем хранения данных основной метрикой надёжности является среднее время до потери данных – Mean Time To Data Loss (*MTTDL*). Это оценка ожидаемого времени до момента, когда хотя бы один блок данных не сможет быть считан (или восстановлен), то есть время до потери данных.

Для обеспечения надёжного хранения данных современные системы используют механизмы избыточного кодирования, в том числе технологии RAID-массивов (Redundant Array of Independent Disks – избыточный массив независимых дисков) различного уровня. Уровень RAID-массива выбирается из

Рис. 4. Потеря данных для дискового массива RAID 5

критериев обеспечения необходимого уровня надёжности хранения данных, а точнее, возможности их восстановления в случае отказа одного или нескольких входящих в дисковый массив дисков, а также требуемой производительности и ёмкости дисковой подсистемы. Следует отметить, что RAID-массивы обеспечивают возможность восстановления данных только в случае функциональных сбоев дисков, входящих в RAID-массив.

По обеспечению сохранности и доступности информации RAID-массивы можно разделить на следующие типы:

1) без обеспечения сохранности данных при выходе из строя любого накопителя – RAID 0;
2) обеспечение сохранности данных при отказе одного накопителя – RAID 1, RAID 1E, RAID 5, RAID 5EE;
3) обеспечение сохранности данных при отказе двух любых накопителей – уровни RAID 6, RAID DP (Dual Parity);
4) обеспечение сохранности при отказе более двух любых накопителей – RAID TP (Triple Parity), RAID 7.3, RAID M+N (отказ M из N);
5) групповые уровни, обеспечивающие доступность данных при выходе из строя нескольких дисков, но из разных групп – уровни RAID 10, RAID 50, RAID 60 и т.д.

RAID 0, с точки зрения обеспечения надёжности, имеет чисто теоретическое значение и применяется в групповых уровнях для повышения производительности RAID-массива.

RAID 5 и RAID 6 – самые распространённые, причём RAID 6 применяется там, где RAID 5 не обеспечивает требую надёжность хранения, и используется в массивах большой ёмкости.

RAID-уровни четвёртой группы – фирменные (Proprietary) разработки. Применяются в случаях, где стандартные уровни, в частности RAID 6, уже не могут гарантировать заданный уровень надёжности хранения данных.

Групповые уровни RAID, как было сказано ранее, применяются для повышения производительности массива данных.

Рассмотрим расчёт среднего времени до потери данных для основных типов RAID-массивов.

Ожидаемое время до потери данных MTTDL для дискового массива RAID 5

Для обеспечения избыточности RAID 5 необходим один диск, поэтому ёмкость для хранения данных дискового массива RAID 5 будет меньше на ёмкость одного диска, чем ёмкость всех дисков, используемых в нём. При отказе одного диска в массиве он переходит в состояние Degraded (отказавший), в котором отказ ещё одного диска приведёт к потере данных. Поэтому необходимо как можно быстрее установить исправный диск вместо отказавшего и запустить режим восстановления – Rebuild.

Режим восстановления может запускаться как автоматически после замены диска, так и по команде оператора. При наличии в массиве специального уже установленного дополнительного диска «горячей» замены (Hot spare) режим перестройки массива обычно включается автоматически для сокращения времени восстановления.

Для дискового массива RAID 5 возможны два пути потери данных (рис. 4):

- ошибка диска во время восстановления (когда массив находится в Degraded Mode[1]);
- сбой вследствие скрытой ошибки данных (*URE*) во время восстановления.

Время до сбоя первого диска в системе будет составлять:

$$MTTF_{disk\_1} = \frac{MTBF_{disk}}{N}.$$

Здесь $MTBF_{disk}$ – наработка до отказа диска; N – количество дисков в RAID-массиве.

Отказ первого диска не означает потерю данных. Как было сказано ранее, это вызывает переход дискового массива в состояние Degraded и начало восстановления целостности массива (Rebuild).

После первого сбоя в массиве останется (N–1) дисков. Если один из них откажет во время восстановления, то это приведёт к потере данных.

Время до отказа второго диска:

$$MTTF_{disk\_2} = \frac{MTTF_{disk}}{(N-1)}.$$

Отказ второго диска происходит во время восстановления, поэтому вероятность вторичного отказа диска будет

$$Q_{disk\_2} = \frac{MTTR_{disk}}{MTTF_{disk\_2}}.$$

Здесь $MTTR_{disk}$ – время восстановления диска, $MTTF_{disk\_2}$ – время до отказа второго диска.

Таким образом, время до потери данных для RAID 5 при отказе второго диска (функциональный сбой) будет составлять:

$$MTTDL_{RAID5\_MTBF} = \frac{MTTF_{disk\_1}}{Q_{disk\_2}}.$$

[1]Degraded Mode – состояние, когда в дисковом RAID-массиве отказали один или несколько дисков.

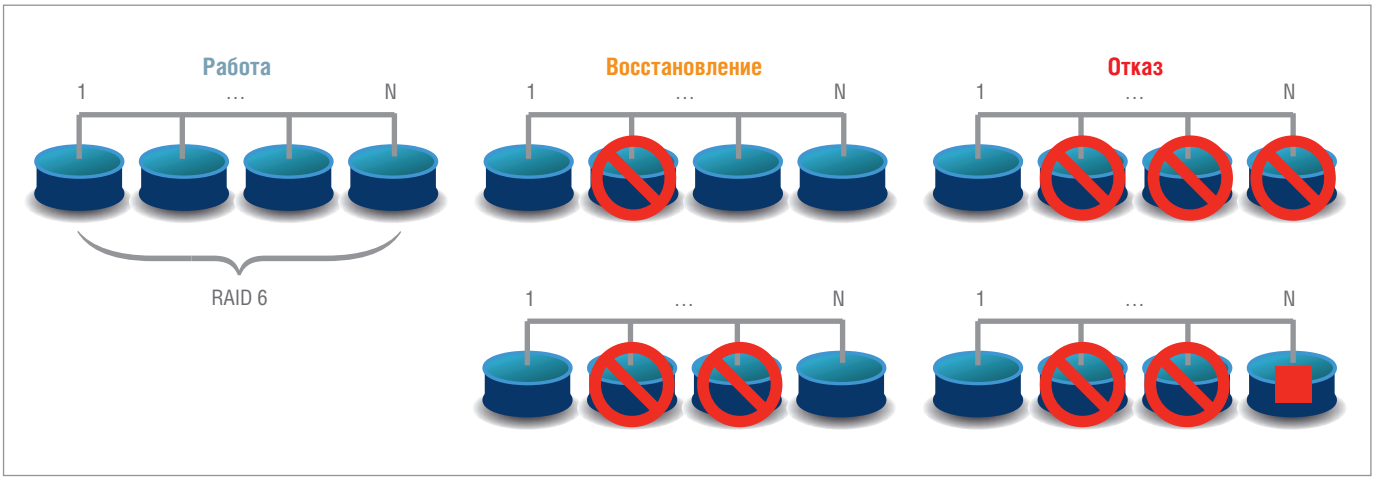

Рис. 5. Потеря данных для дискового массива RAID 6

или

$$MTTDL_{RAID5\_MTBF} = \frac{MTBF_{disk}^{\ 2}}{MTTR_{disk} \times N \times (N-1)}.$$

После отказа первого диска в системе остаётся (N–1) дисков, которые должны быть прочитаны, чтобы восстановить чётность в RAID-массиве.

Вероятность безошибочного чтения всего диска зависит от величины невосстановимых ошибок чтения и ёмкости диска и определяется как

$$P_{disk} = (1-URE)^{C_{bit}}.$$

Здесь URE – количество невосстановимых ошибок чтения; C_{bit} – размер диска в битах.

Тогда вероятность ошибки во время восстановления RAID 5 из-за невосстановимой ошибки чтения:

$$Q_{URE}=1-P_{disk}^{(N-1)}$$

Здесь N–1 – количество дисков, которые нужно считать.

Время до потери данных для массива RAID 5:

$$MTTDL_{RAID5\_URE} = \frac{MTTF_{disk\_1}}{Q_{URE}},$$

или

$$MTTDL_{RAID5\_URE} = \frac{MTBF_{disk}}{Q_{URE} \times N},$$

или

$$MTTDL_{RAID5\_URE} = \frac{MTBF_{disk}}{N \times \left(1 - P_{disk}^{(N-1)}\right)}.$$

Таким образом, время до потери данных в массиве RAID 5 будет

$$MTTDL_{RAID5} = (MTTDL^{-1}_{RAID5\_MTBF} + MTTDL^{-1}_{RAID5\_URE})^{-1}.$$

Ожидаемое время до потери данных MTTDL для дискового массива RAID 6

В RAID 6 используется схема двойной чётности, поэтому для её обеспечения в дисковом массиве необходимы два дополнительных диска. Полезная ёмкость дискового массива будет меньше на два диска, чем общее число дисков в RAID-наборе. Благодаря механизму двойной чётности данные могут быть восстановлены при двух одновременно отказавших дисках.

Таким образом, для дискового массива в RAID 6 потеря данных может произойти в следующих случаях (рис. 5):

1) отказ трёх дисков во время восстановления;
2) отказ двух дисков во время восстановления, плюс произошла ошибка из-за скрытой ошибки данных.

По аналогии с RAID 5 получаем следующие выражения.

Ожидаемое время до потери данных при отказе трёх дисков будет составлять:

$$MTTDL_{RAID6\_MTBF} = \frac{2 \times MTBF_{disk}^{\ 3}}{MTTR_{disk}^{\ 2} \times N \times (N-1) \times (N-2)}.$$

Ожидаемое время до потери данных при отказе двух дисков и возникновения ошибки чтения:

$$MTTDL_{RAID6\_URE} = \frac{MTBF_{disk}^{\ 2}}{N \times (N-1) \times \left(1 - (1 - P_{disk})\right)^{(N-2)} \times MTTR_{disk}}.$$

И время до потери данных в массиве RAID 6 вычисляется как:

$$MTTDL_{RAID6} = (MTTDL^{-1}_{RAID6\_MTBF} + MTTDL^{-1}_{RAID6\_URE})^{-1}.$$

Таким образом, в статье в упрощённом виде рассмотрены модели отказоустойчивых дисковых массивов, организованных в RAID 5 и RAID 6, с учётом показателей надёжности дисков: наработки на отказ (*MTBF*) и возможных невосстановимых ошибок чтения данных (скрытой ошибки данных – *URE*) при восстановлении массива, и предложены формулы расчёта надёжности таких систем.

Данные модели и формулы наиболее точно отражают реальную надёжность избыточных дисковых массивов и могут применяться для расчёта показателей надёжности. ●

Автор – сотрудник фирмы «Адвантикс»
Телефон: (495) 232-1693
E-mail: info@advatix-pc.ru