



Интеллектуальное восстановление диска: предотвращение ошибок поверхности и отказов с помощью S.M.A.R.T

Краткое содержание

В статье описывается передовая технология Infortrend для интеллектуального восстановления дисков (Intelligent Drive Recovery technology - IDR), обеспечивающая полную целостность данных. В нее входит сканирование поверхности носителей, эффективное предотвращение ошибок и регенерация диска. На примерах применения IDR в системах RAID 5 и RAID 6 иллюстрируются преимущества, которые дает IDR в системах хранения Infortrend, как средство эффективной экономии времени и повышения характеристик одновременно с эффективной защитой данных.

Содержание

| | |
|---|-----------|
| Введение | 3 |
| Применимость настоящего документа..... | 3 |
| Что такое ошибки носителя?..... | 3 |
| Почему возникают ошибки носителя?..... | 3 |
| Влияние..... | 3 |
| Защита RAID 6 | 4 |
| Почему RAID 6?..... | 4 |
| Сравнение RAID 5 и RAID 6..... | 4 |
| Количество дисков и конфигурация RAID..... | 4 |
| Разрешить S.M.A.R.T | 5 |
| Что такое S.M.A.R.T?..... | 5 |
| Режимы S.M.A.R.T..... | 6 |
| Сканирование поверхности | 7 |
| Что такое сканирование поверхности?..... | 7 |
| Как сканирование поверхности улучшает целостность данных..... | 7 |
| Автоматическая проверка ошибок поверхности..... | 9 |
| Заключение | 11 |

Введение

Одно из основных свойств RAID системы хранения заключается в том, что она обеспечивает защиту от отказов дисковых накопителей. Технология интеллектуального восстановления дисков Infortrend (**Intelligent Drive Recovery - IDR**) гарантирует сохранность данных даже при полном выходе дисков из строя. В некоторых случаях дефектные сектора дисков могут не обнаруживаться, так как хост редко обращается к ним. Со временем, когда пользователю потребуются данные с этих дефектных секторов, особенно в критических ситуациях, таких, как восстановление томов, данные могут быть утрачены без возможности их восстановления. Однако, если мы сможем обнаружить и исправить эти дефектные сектора на поверхности носителей заранее, непоправимых ошибок можно избежать до того, как они случатся.

Применимость настоящего документа

Семейство EonStor DS

Семейство EonNAS

Что такое ошибки носителя?

Ошибки носителя - это дефекты, пятна или поврежденные участки поверхности, которые не могут надежно сохранять данные для их последующего считывания. Они могут появляться на любом диске от любого производителя. Это известно и воспринимается как неизбежная данность при производстве жестких дисков, в особенности дисков, предназначенных для применения в системах с высокой плотностью.

Почему возникают ошибки носителя?

Срок службы дисков повысился до уровня, где большинство дисков энтерпрайз-класса имеют среднее время наработки на отказ (MTBF - mean time between failures), превышающее один миллион часов. Несмотря на все улучшения, полностью ошибок носителя избежать не удастся. Большинство дисков в системах хранения подготовлены для использования в системах с высокой плотностью упаковки и хорошо защищены, работают в течение длительного времени и способны хранить тома данных большого объема. Однако, со временем даже диски энтерпрайз-класса достигают своего механического порога, и на поверхности диска появляются дефекты. Для дисков большой емкости (4 TB и более), вероятность появления плохих секторов возрастает просто из-за их большого размера.

Влияние

Дефекты носителя влияют на способность дискового накопителя считывать данные с конкретного сектора поверхности. Их наличие не означает, что диск непригоден к использованию. На любом диске данные с этих дефектных областей могут быть безвозвратно потеряны, однако данные с других секторов можно будет считывать много раз в течение нескольких лет. С помощью RAID защиты мы можем использовать механизмы RAID четности, чтобы восстанавливать данные на этих дефектных секторах через другие

существующие связанные данные. Если на диске остается свободное пространство, мы можем даже игнорировать ошибки поверхности, если мы их обнаружим и восстановим поврежденные данные заранее. Конечно, если ошибки поверхности начинают накапливаться, распространяться на весь диск, и ситуация ухудшается (это означает, что все больше ошибок обнаруживается после очередной проверки), это может означать, что диск близок к полной выработке своего ресурса. Чтобы избежать риска полного отказа диска и потери данных, необходимо запланировать замену диска на новый, как только это будет возможно.

Защита RAID 6

Почему RAID 6?

Традиционные конфигурации RAID 3 и RAID 5 были созданы для повышения надежности хранения данных за счет использования одного избыточного диска для хранения информации о четности и коррекции ошибок. Проблема заключается в том, что если один из дисков выходит из строя, или на нем обнаруживаются ошибки поверхности именно в тот момент, когда выполняется операция восстановления данных (RAID rebuilding), система не сможет запросить данные с него, которые требуются для восстановления, и они будут безвозвратно потеряны, так как один из дисков отказал.

Системы уровня RAID 6 хранят два набора информации о четности и о коррекции ошибок, которые организованы таим образом, что если даже один диск отказывает, или обнаруживаются ошибки поверхности во время процесса восстановления данных, то система будет продолжать работу и никакие данные не будут утрачены благодаря более высокому уровню защиты по сравнению с другими RAID конфигурациями. Это особенно полезно при использовании дисков высокой плотности.

Сравнение RAID 5 и RAID 6

Конфигурация RAID 5 обладает лучшими характеристиками, так как системе приходится вычислять только один набор четности. Системы RAID 5 также имеют большую емкость, так как только один диск используется для хранения данных о четности. Однако, чем больше дисков мы установили в систему хранения, тем меньше значения имеет это обстоятельство по сравнению с риском потерять данные. Вот почему мы рекомендуем использовать RAID 6, а не RAID 5, если есть достаточно дисков и емкости и требуется более надежная RAID защита.

Количество дисков и конфигурация RAID

Встроенное программное обеспечение Infortrend предполагает, по умолчанию, что все диски могут быть использованы для создания логических томов. В зависимости от их количества, мы рекомендуем следующие конфигурации:

- Восемь или более дисков: RAID 6
- Менее восьми дисков (от трех до семи): RAID 5

- Менее трех дисков: по умолчанию допускается только персонализированная конфигурация пользователя

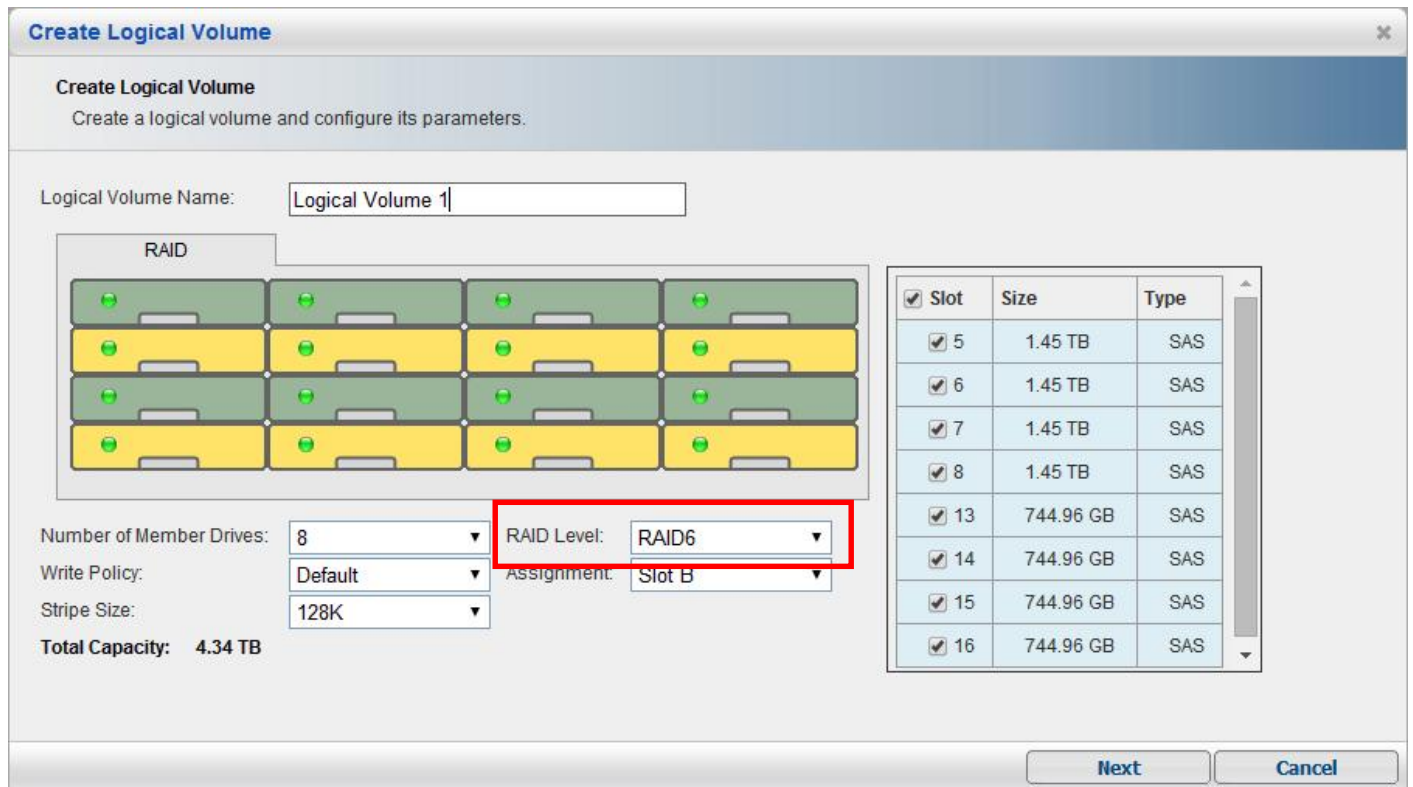


Рис. 1. Выбор RAID 6 как конфигурации для создания логического тома

Мы настоятельно рекомендуем выбирать RAID 6 как более надежную и безопасную конфигурацию. Если пользователь захочет переключиться на другую конфигурацию с выбранной по умолчанию конфигурации RAID 6, появится специальное сообщение, информирующее его об этом действии. В мастере конфигурирования имеются также другие доступные опции: RAID 6 + локальный резерв, RAID 5, и RAID 5 + локальный резерв. Пользователь может выбрать любой поддерживаемый уровень RAID в режиме персонализации.

Разрешить S.M.A.R.T

Системы Infortrend разрешают S.M.A.R.T для всех подключенных дисков.

Что такое S.M.A.R.T?

S.M.A.R.T (Self-Monitoring, Analysis, and Reporting Technology) - это механизм мониторинга дисковых накопителей для проверки различных индикаторов надежности, направленный на предвидение и предотвращение отказов.

Режимы S.M.A.R.T.

RAID (ESVA, EonStor и EonStor DS): система разрешает S.M.A.R.T в режиме “detect-clone-replace”

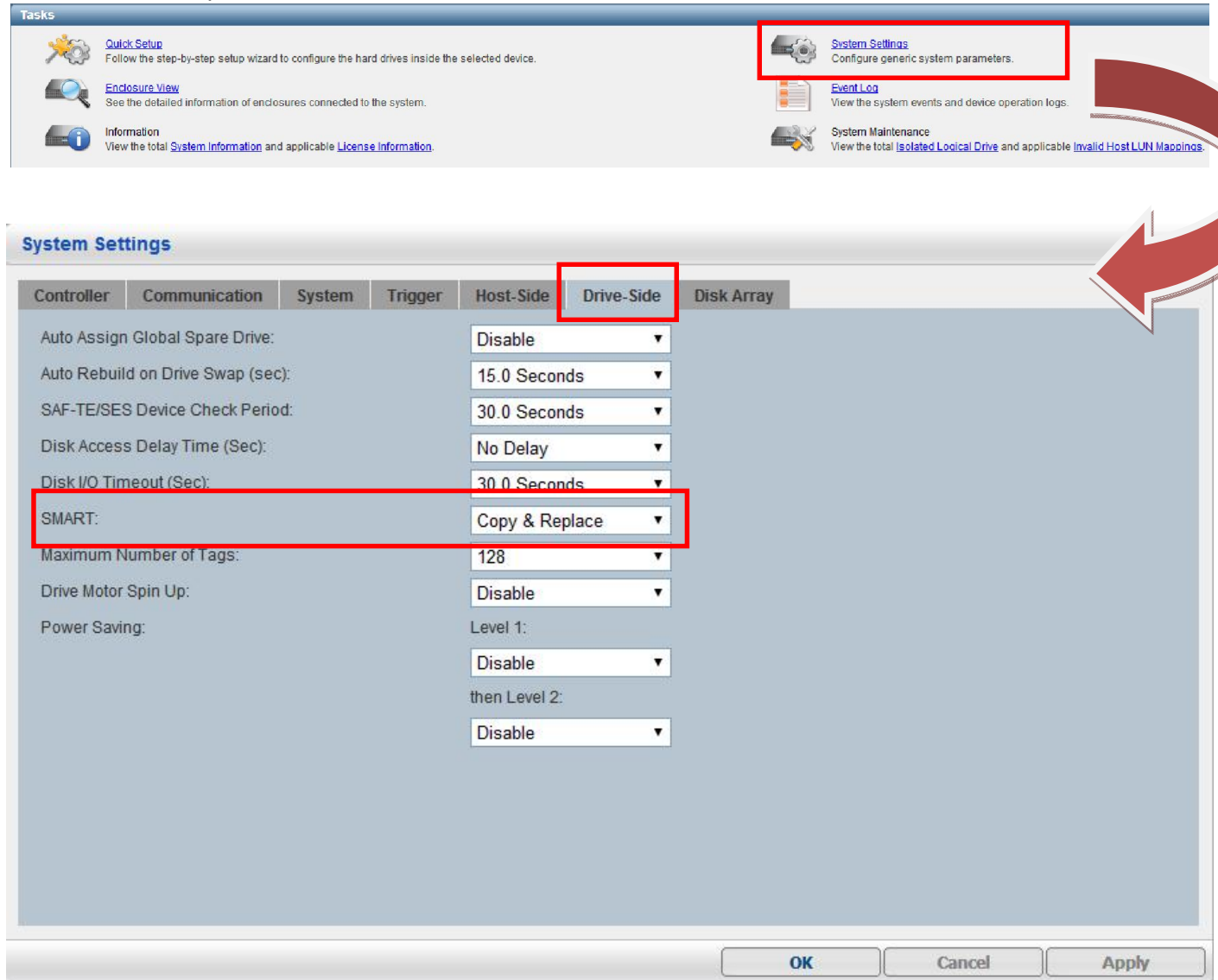


Рис. 2. Проверьте установку разрешения S.M.A.R.T. и связанных опций

Режим “detect-clone-replace”: если S.M.A.R.T обнаруживает ошибку диска, система пытается клонировать весь диск на имеющийся запасной диск и заменить его. Система также информирует пользователя с помощью оповещений о событиях.

EonNAS: система разрешает S.M.A.R.T в режиме “fail drive”. Ошибки диска рассматриваются как аппаратные ошибки, и встроенное ПО EonNAS запускает процесс восстановления, используя запасной диск для спасения данных пользователя и регенерации данных о четности. В отличие от RAID систем блочного уровня, NAS может восстановить определенные области диска через файловую систему, так что нет необходимости клонировать весь диск полностью блок за блоком, как это делает RAID.

Сканирование поверхности

Что такое сканирование поверхности?

Функция сканирования поверхности последовательно проверяет физический диск блок за блоком в поисках дефектных участков. При этой операции обнаруживаются и перемещаются все дефектные сектора. Сканирование выполняется, чтобы снизить вероятность появления “мягких” ошибок (**disk soft error**) поверхности при работе.

Чистка в ZFS для EonNAS

В ZFS имеется ремонтный программный инструмент, именуемый скраб, который обнаруживает искажение данных, непрерывно проверяет их целостность и автоматически устраняет проблемы, вызванные дефектами диска. Он проверяет все логические области с данными способом, похожим на сканирование в блочном режиме, мы применяем эту функцию как механизм проверки данных в EonNAS.

Как сканирование поверхности улучшает целостность данных?

Ошибки поверхности диска неизбежно возникают в течение срока его службы, поэтому производители дисков предусматривают на ней запасные блоки. Предполагается, что эти дефектные блоки будут замещены (переназначены или перемещены) запасными либо автоматически самим диском, либо по команде от встроенного ПО системы хранения. Когда блок выходит из строя, он может быть восстановлен (либо перемещен) в запасную позицию, которая была оставлена именно для этой цели. Не все области диска используются одинаково часто, поэтому сектора с ошибками, которые обычно остаются незамеченными, могут быть обнаружены с помощью регулярного сканирования поверхности.

Когда возникают трудности с чтением данных с сектора, (так называемая “мягкая” ошибка, **disk soft error**), то он автоматически пытается восстановить данные собственными средствами. Независимо от того, удалось, или нет считать данные, диск извещает об этом операционную систему. Если данные с этого сектора все же не удастся считать, и обнаруживается ошибка поверхности при сканировании, то система хранения может попытаться исправить ситуацию путем реконструкции данных с помощью четности через RAID защиту, и предпринять попытку записать данные на тот же участок поверхности. Если запись не удастся, то система хранения фиксирует это событие и предлагает диску переместить сектор в запасную физическую позицию с помощью команды SCSI. Если такая позиция имеется, то система хранения напрямую переписывает данные в новую физическую позицию и восстанавливает данные из любого дефектного сектора и области. Вот почему сканирование поверхности улучшает целостность данных даже в том случае, если диск сканируется после процесса восстановления.

В следующем примере было создано одно логическое устройство в конфигурации RAID 5 на дисках (HDD1 - HDD4), каждый диск разбит на данные пользователя и данные о RAID четности для восстановления. Пусть при сканировании обнаружена одна ошибка в одной из полос на первом диске (HDD1), как на рис. 3.

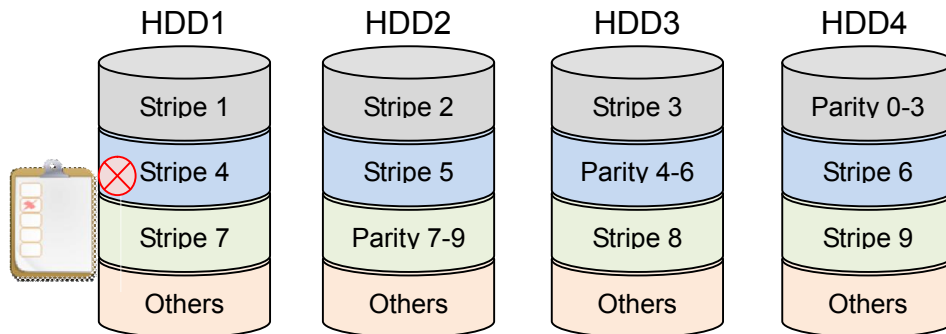


Рис. 3. После запуска сканирования поверхности обнаружена одна ошибка

За счет RAID защиты данные с дефектной поверхности могут быть регенерированы путем XOR вычислений через другие данные пользователя и данные о четности. Регенерированные данные записываются сначала на тот же участок поверхности, если это возможно. Если эта операция окажется успешной, то ошибка поверхности устранена и данные восстановлены.

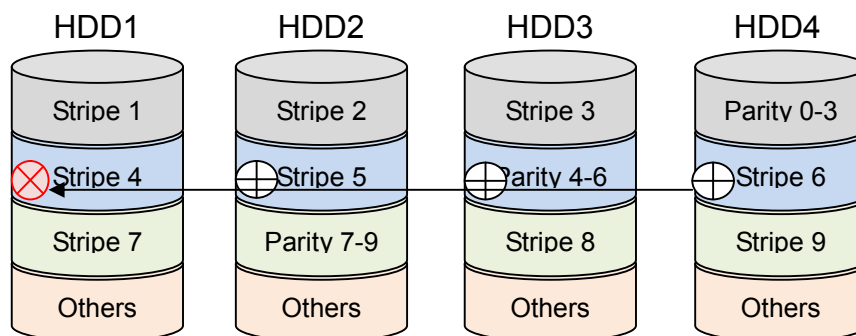


Рис. 4. Использование RAID защиты для регенерации и восстановления данных с дефектной поверхности

Если записать данные в ту же позицию на диске не удастся, то система хранения предлагает диску переназначить для восстановленных данных емкость запасного пространства и записать их еще раз. Данные перенаправляются в новую позицию.

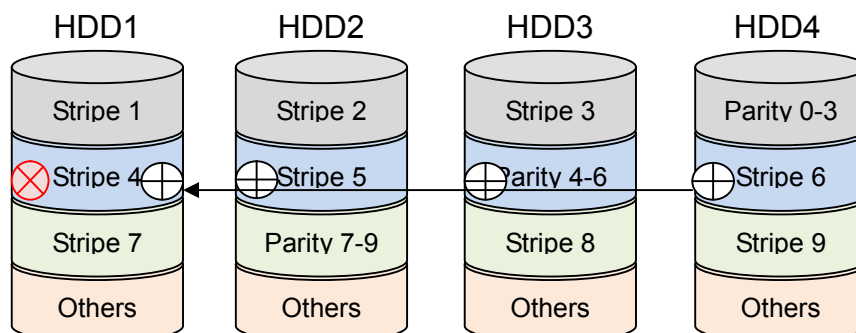


Рис. 5. Если запись в ту же позицию не удастся, диск переназначает пространство и перезаписывает регенерированные данные

Автоматическая проверка ошибок поверхности

Обработка ошибок поверхности хранения является главным приоритетом в RAID устройствах. Сканирование поверхности помогает обнаружить ошибки поверхности в дисковых накопителях, поэтому следует проводить его регулярно, чтобы заблаговременно предотвратить возможную потерю доступности данных вследствие отказа какого-нибудь диска.

Почему проверка по графику предпочтительнее непрерывного фонового сканирования?

Для периодической проверки мы рекомендуем задать график сканирования. Так как в большинстве систем хранения применяются технологии энергосбережения, то они могут переводить диски в спящий режим, чтобы уменьшить потребление энергии. Если система будет постоянно занята непрерывным фоновым сканированием, то все диски никогда не будут в спящем режиме, даже если к ним не будет обращений со стороны хоста. Это означает, что непрерывное сканирование поверхности сводит на нет “зеленые” свойства систем хранения, которые позволяют пользователям экономить деньги.

Периодическая проверка по графику

Добавьте график автоматического сканирования поверхности после создания логического устройства (логический диск для EonStor DS, элемент пула для ESVA и пул для EonNAS).

- Для каждого логического устройства можно задать график сканирования поверхности. Пользователь может удалить или персонализировать этот график вручную.

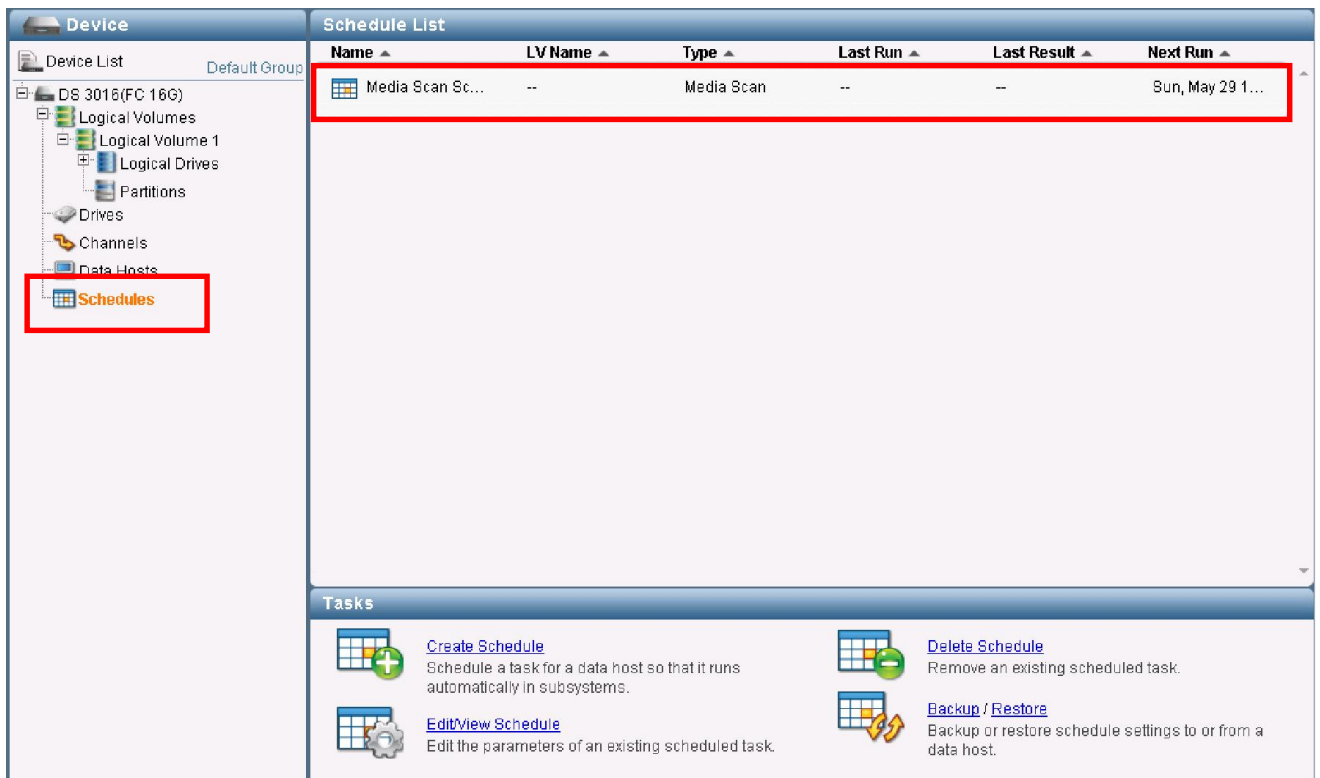


Рис. 6. Пользователь может удалять или персонализировать графики вручную

Интеллектуальное восстановление диска: предотвращение ошибок поверхности и отказов

- Если восстановить заводские настройки по умолчанию, то существующие графики сканирования будут снова добавлены автоматически после того, как заводские настройки восстановлены.
- После обновления встроенного ПО все логические устройства сохраняют заданные для них графики сканирования. Для архивных устройств без графиков обновленное ПО позволяет задать соответствующие настройки.
- Фоновое сканирование поверхности не влияет на характеристики, так как оно приостанавливается на время обработки любого I/O запроса и возобновляется только тогда, когда система бездействует.

Политика графика автоматического сканирования поверхности

Приоритет: низкий

Время начала : заданный период после создания (или восстановление заводских настроек)

Периодичность : каждые две недели для конфигураций RAID 1, RAID 3, RAID 5 и RAID 6

Просроченный график: продолжается после завершения предыдущей задачи

Если предыдущая задача не закончена, а следующая задача приближается, это означает, что система хранения постоянно занята и не может завершить сканирование поверхности в заданное время. Следующая задача начнется после завершения предыдущей, и точное время начала сканирования может быть пропущено. Пользователь будет оповещен сообщением, если система занята и приоритет сканирования или период проверки не согласуется с ними. Пользователь может изменить политику сканирования, повысив приоритет до нормального или высокого, чтобы ускорить процесс, однако это может повлиять на характеристики во время операции сканирования. Можно также увеличить период сканирования, однако интервал проверки также увеличится, что может сказаться на защите - все зависит от предпочтений пользователя.

Внимание:

1. Старые модели могут не поддерживать IDR и автоматического сканирования поверхности без обновления ПО до последней версии. Они, тем не менее, позволяют создавать графики сканирования вручную.
2. Сканирование может быть начато только тогда, когда логическое устройство находится в состоянии "good". Если логические диски находятся в состоянии "rebuilding" или "drive adding", сканирование будет остановлено и возобновится, когда появятся условия для него.
3. Если контроллер выходит из строя, текущее сканирование возобновится, когда второй контроллер возьмет управление на себя (только для моделей с двойным избыточным контроллером).
4. Нормальный сброс системы также может приостановить текущую задачу сканирования, она возобновится, когда система снова будет полностью онлайн.

Заключение

Системы хранения Infortrend представляют собой тщательно разработанные завершённые продукты, в которых сканирование поверхности носителей и технология предотвращения ошибок усилены интеллектуальным восстановлением диска (IDR). В дополнение к стандартной защите данных, обеспеченной RAID технологией, незначительные и трудно обнаруживаемые дефекты дисковых накопителей локализуются и сглаживаются за счёт того, что Infortrend прилагает дополнительные усилия для развития связанных технологий - в данном случае IDR. Эти механизмы предотвращают возможные потери данных и устраняют не восстанавливаемые другими способами ошибки поверхности, чтобы минимизировать их влияние и получить в итоге диски без деградации. Дело в том, что незначительные дефекты, которые легко не заметить, обрабатываются так, как будто отказал полностью весь диск, так что недостающие данные могут быть регенерированы без негативного влияния на все диски RAID массива.