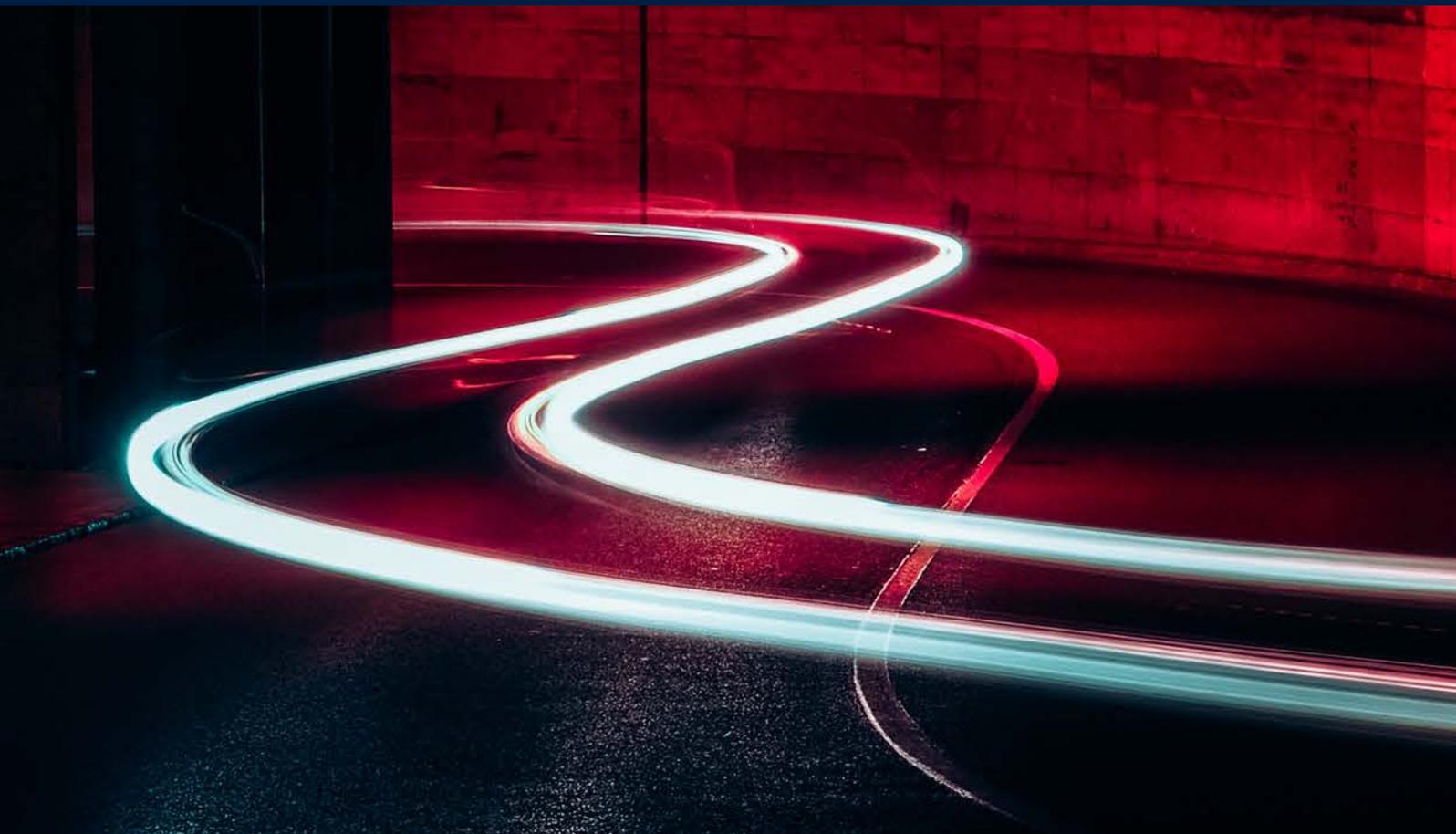


White paper

Как работает SSD-кэширование в системах хранения данных



Содержание

Что такое SSD-кэширование	3
Когда SSD-кэш будет полезен	3
Как работает SSD-кэш в СХД	4
Работа кэша на чтение данных	4
Работа кэша на запись данных	5
Алгоритмы заполнения кэша	5
Особенности SSD-кэширования в RAIDIX	5
1. Категории RRC и RWC	5
Попадание в RRC	6
Попадание в RWC	6
2. Особенности Log-структурированной записи	6
Вытеснение из буфера RRC	6
Вытеснения из буфера RWC	6
Возможности SSD-кэширования в RAIDIX	6
Сокращение износа SSD-накопителей	7
Эффективность SSD-кэширования на различных нагрузках	7
Заключение	8

Что такое SSD-кэширование

Большая часть хранимых данных имеет небольшое количество повторных обращений, такие данные принято называть «холодными». Они составляют значительную часть как в крупных файловых архивах, так и на жестком диске вашего домашнего компьютера. Если к данным обращаются повторно, то они будут называться «теплыми» или «горячими». Последние обычно представляют собой блоки служебной информации, которая считывается при загрузке приложений или выполнении каких-либо стандартных операций.

SSD-кэширование — это технология, при которой твердотельные накопители используются в качестве буфера для часто запрашиваемых данных. Система определяет степень «теплоты» данных и перемещает их на быстрый накопитель. За счет этого чтение и запись этих данных будут выполняться с большей скоростью и с меньшей задержкой.

Про SSD-кэширование часто говорят, когда речь идет о системах хранения данных, где эта технология дополняет HDD-массивы, повышая производительность за счет оптимизации случайных запросов. Устройство HDD-накопителей позволяет им успешно справляться с последовательным паттерном нагрузки, но имеет естественное ограничение для работы со случайными запросами. Объем SSD-кэша при этом обычно составляет около 5–10% от емкости основной дисковой подсистемы.

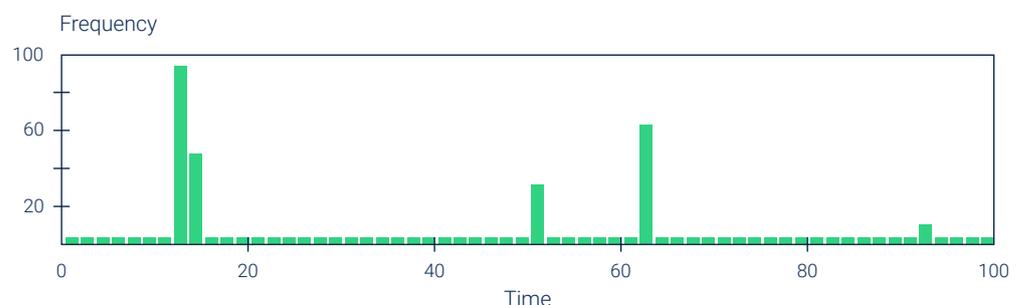
Системы, которые используют SSD-кэш вместе с HDD-дисками, принято называть гибридными. Они популярны на рынке СХД, так как значительно доступнее по цене, чем all-flash конфигурации, но при этом способны эффективно работать с достаточно широким спектром задач и нагрузок.

Когда SSD-кэш будет полезен

SSD-кэш подходит для ситуаций, когда система хранения данных получает не только последовательную нагрузку, но и определенный процент случайных запросов. При этом эффективность SSD-кэширования будет значительно выше в ситуациях, когда случайные запросы характеризуются пространственной локальностью, то есть на определенном адресном пространстве формируется область «горячих» данных.

Технология SSD-кэширования будет особенно полезна, например, при работе с потоками в видеонаблюдении. В таком паттерне преобладает последовательная нагрузка, но могут встречаться также случайные запросы на чтение и запись. Если не обрабатывать такие пики при помощи SSD-кэширования, то система будет пытаться справиться с ними при помощи HDD-массива, следовательно, производительность существенно снизится.

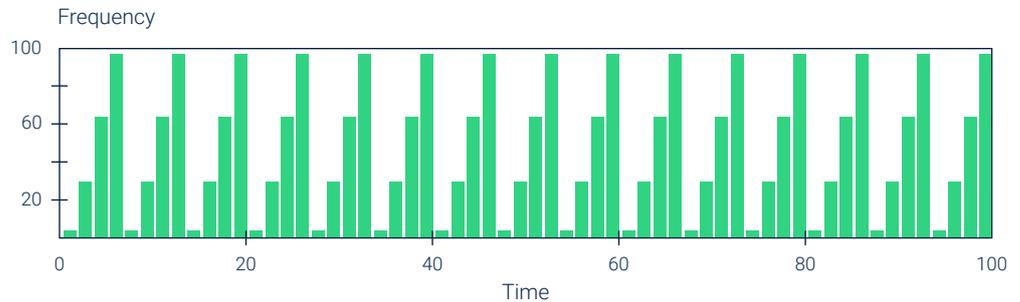
Рисунок 1.
Неравномерный
временной интервал
с непредсказуемой
частотой обращений



На практике появление случайных запросов среди равномерной последовательной нагрузки совсем не редкость. Это может происходить при одновременной работе на сервере нескольких различных приложений. Например, одно имеет установленный приоритет и работает с последовательными запросами, а другие время от времени обращаются к данным (в том числе, повторно) в случайном порядке. Другим примером возникновения случайных запросов может быть так называемый I/O Blender Effect, который перемешивает последовательные запросы.

Если на СХД поступает нагрузка с большой частотой случайных и мало повторяющихся запросов, то эффективность SSD-кэша будет снижаться.

Рисунок 2.
Равномерный
временной интервал
с предсказуемой
частотой обращений



При большом количестве таких обращений пространство SSD-накопителей будет быстро заполняться, и производительность системы будет стремиться к скорости работы на HDD-накопителях.

Следует помнить, что SSD-кэш является довольно ситуативным инструментом, который будет показывать свою продуктивность далеко не во всех случаях. В общих чертах его использование будет полезным при следующих характеристиках нагрузки:

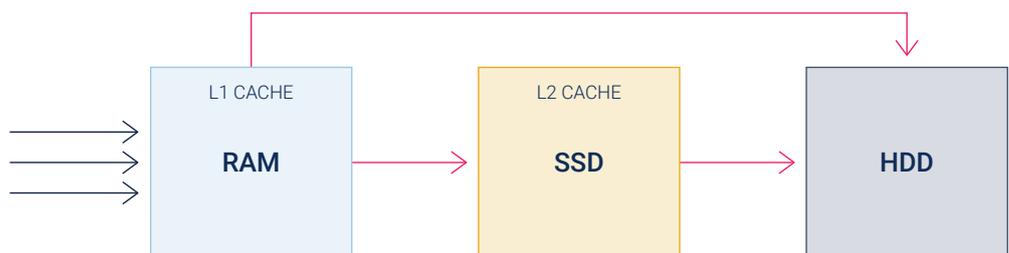
- случайные запросы на чтение или на запись имеют низкую интенсивность и неравномерный временной интервал;
- количество операций ввода-вывода на чтение значительно больше, чем на запись;
- количество «горячих» данных будет предполагаться меньше размеров рабочего пространства SSD.

Как работает SSD-кэш в СХД

Функция кэша — ускорять выполнение операции за счет размещения часто запрашиваемых данных на быстрых носителях. Для кэширования самых «горячих» данных используется оперативная память (RAM), в СХД это кэш первого уровня (L1 Cache).

Кэш первого уровня может быть дополнен менее быстрыми твердотельными накопителями. В этом случае у нас появляется кэш второго уровня (L2 Cache). Такой подход используется для реализации SSD-кэширования в большинстве существующих СХД.

Рисунок 3.
Традиционный SSD-кэш
второго уровня.



Традиционный SSD-кэш второго уровня работает следующим образом: все запросы после RAM попадают в буфер SSD (рисунок 1).

Работа кэша на чтение

Система получает запрос на чтение данных, находит необходимые блоки на основном хранилище (HDD) и производит их чтение. При повторных обращениях система создает копии этих данных на SSD-накопителях. Последующие операции чтения будут выполняться уже с быстрых носителей, что увеличит скорость работы.

Работа кэша на запись

Система получает запрос на запись и размещает необходимые блоки данных на SSD-накопителях. Благодаря быстрым носителям, операция записи и оповещение инициатора происходят с минимальными по времени задержками. По мере заполнения кэша система начинает постепенно передавать на основное хранилище наиболее «холодные» данные.

Алгоритмы заполнения кэша

Один из главных вопросов в работе SSD-кэша — выбор данных, которые будут помещаться в буферное пространство. Так как объем хранения тут ощутимо ограничен, то при его заполнении нужно принимать решение о том, какие блоки данных вытеснить и по какому принципу производить замещение.

Для этого применяются алгоритмы заполнения кэша. Коротко рассмотрим наиболее распространенных в сегменте СХД.

FIFO (First In, First Out) — из кэша последовательно вытесняются наиболее старые блоки, замещаясь наиболее свежими.

LRU (Least Recently Used) — из кэша первыми вытесняются блоки данных с наиболее давней датой обращения.

LARC (Lazy Adaptive Replacement Cache) — блоки данных попадают в кэш, если они были запрошены как минимум дважды за определенный промежуток времени, а кандидаты на замещение отслеживаются в дополнительной LRU-очереди в оперативной памяти.

SNLRU (Segmented LRU) — данные из кэша вытесняются по принципу LRU, но при этом они проходят несколько категорий (сегментов), обычно это: «холодные», «теплые», «горячие». Степень «теплоты» здесь определяется частотой обращений.

LFU (Least Frequently Used) — из кэша первыми вытесняются те блоки данных, к которым было меньше всего обращений.

LRFU (Least Recently/Frequently Used) — алгоритм комбинирует работу LRU и LFU, вытесняя сначала те блоки, которые попадают под рассчитываемый параметр из даты и количества обращений.

В зависимости от типа алгоритма и качества его реализации будет определяться итоговая эффективность SSD-кэширования.

Особенности SSD-кэширования в RAIDIX

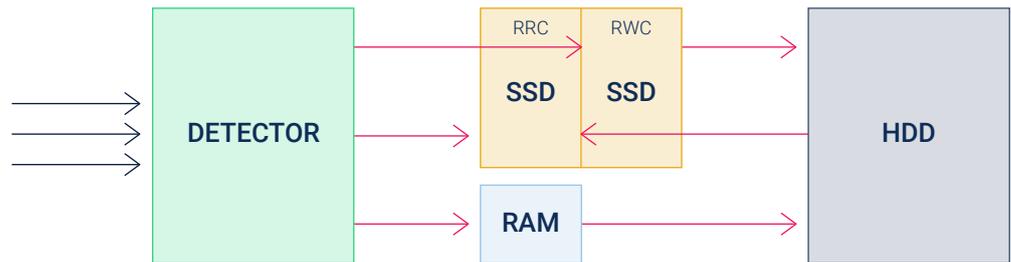
В RAIDIX реализован параллельный SSD-кэш, который имеет две уникальные особенности: разделение входящих запросов на категории RRC (Random Read Cache) и RWC (Random Write Cache) и использование Log-структурированной записи для собственных алгоритмов вытеснения.

1. Категории RRC и RWC

Пространство кэша разделено на две функциональные категории: для случайных запросов на чтение — RRC, для случайных запросов на запись — RWC. Для каждой из этих категорий есть свои правила попадания и вытеснения.

За попадание отвечает специальный детектор, который квалифицирует поступающие запросы.

Рисунок 4. Схема работы SSD-кэша в RAIDIX



Попадание в RRC

В область RRC попадают только случайные запросы с частотой обращения больше 2-х (ghost-очередь).

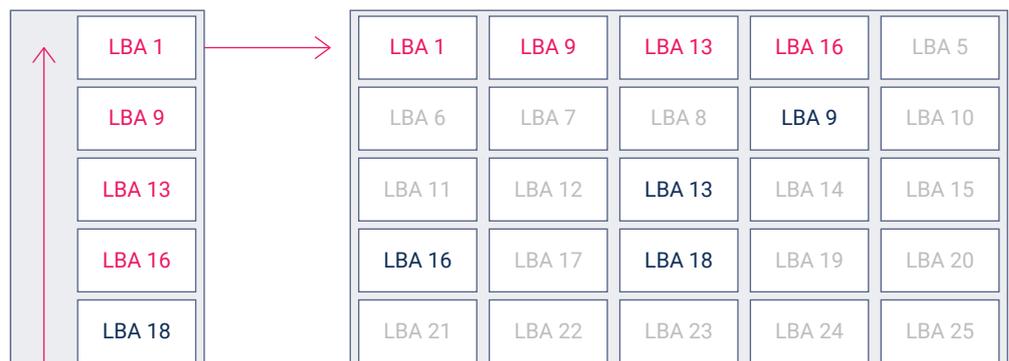
Попадание в RWC

В область RWC попадают все случайные запросы на запись, у которых размер блока меньше устанавливаемого параметра (по умолчанию 32KB).

2. Особенности Log-структурированной записи

Log-структурированная запись — это способ последовательной записи блоков данных без учета их логической адресации (LBA, Logical Block Addressing).

Рисунок 5. Визуализация принципа Log-структурированной записи



В RAIDIX Log-структурированная запись используется для заполнения выделенных областей (с установленным размером в 1 Гб) внутри RRC и RWC. Эти области применяются для дополнительного ранжирования при перезаписи пространства кэша.

Вытеснение из буфера RRC

Выбирается самая холодная область RRC, и в нее перезаписываются новые данные из ghost-очереди (данные с частотой обращений больше 2-х).

Вытеснения из буфера RWC

Область выбирается по принципу FIFO, а затем из нее последовательно, в соответствии с LBA (Logical Block Address), вытесняются блоки данных.

Возможности SSD-кэширования в RAIDIX

Параллельная архитектура SSD-кэша в RAIDIX позволяет ему быть не просто буфером для накопления случайных запросов — он начинает выполнять роль «умного распределителя» нагрузки на дисковую подсистему. Благодаря сортировке запросов и особым алгоритмам вытеснения, сглаживание пиков случайной нагрузки происходит быстрее и с меньшим влиянием на общую производительность системы.

Алгоритмы вытеснения используют log-структурированную запись для более эффективного замещения данных в кэше. Благодаря этому снижается количество обращений к flash-накопителям и существенно сокращается их износ.

Сокращение износа SSD-накопителей

Благодаря детектору нагрузки и алгоритмам перезаписи суммарное количество write hits на массив SSD накопителей в RAIDIX составляет 1.8. В аналогичных условиях работа кэша второго уровня с алгоритмом LRU имеет значение 10.8. Это означает, что количество требуемых перезаписей на флеш-накопители в реализованном подходе будет в 6 раз меньше, чем во многих традиционных СХД. Соответственно, SSD-кэш в RAIDIX будет использовать ресурс твердотельных накопителей значительно эффективнее, увеличивая срок их жизни примерно в 6 раз.

Эффективность SSD-кэширования на различных нагрузках

Смешанную нагрузку можно рассматривать как хронологический перечень состояний с последовательным или случайным типом запроса. Системе хранения данных приходится справляться с каждым из этих состояний, даже если оно не является для нее предпочтительным и удобным.

Мы провели тестирование SSD-кэша, эмулируя различные рабочие ситуации с разными типами нагрузок. Сравнив полученные результаты со значениями системы без SSD-кэша, можно наглядно оценить прирост производительности при разных типах запросов.

Конфигурация системы:

SSD кэш: RAID 10, 4 SAS SSD, объем 372 GB

Основное хранилище: RAID 6i, 13 HDD, объем 3072 GB

Тип паттерна	Тип запроса	Значение с SSD-кэшированием	Значение без SSD-кэширования	Увеличение производительности
Случайное чтение (100% попадание в кэш)	random read 100%	85.5K IOPS	2.5K IOPS	В 34 раза
Случайная запись (100% попадание в кэш)	random write 100%	23K IOPS	500 IOPS	В 46 раз
Случайное чтение (80% попадание в кэш, 20% попадание на HDD)	random read 100%	16.5K IOPS	2.5K IOPS	В 6.5 раз
Случайное чтение из кэша, запись на HDD	random read 50% sequential write 50%	40K IOPS 870 Mbps	180 IOPS 411 Mbps	В 222 раза В 2 раза
Случайное чтение и запись (100% попадание в кэш)	random read 50% random write 50%	30K IOPS 19K IOPS	224 IOPS 800 IOPS	В 133 раза В 23 раза
Последовательные запросы с большим блоком, случайная нагрузка 100% попадает в SSD-кэш	random read 25%	2438 IOPS	56 IOPS	В 43 раза
	random write 25%	1918 IOPS	82 IOPS	В 24 раза
	sequential read 25%	668 Mbps	120 Mbps	В 5.5 раз
	sequential write 25%	382 Mbps	76.7 Mbps	В 5 раз

У каждой реальной ситуации будет свой неповторимый «рисунок» нагрузки, и такое фрагментарное представление не дает однозначного ответа об эффективности SSD-кэширования на практике. Но оно помогает сориентироваться в том, где данная технология может быть наиболее полезна.

Заключение

Технология SSD-кэширования позволяет повысить производительность СХД при работе со смешанным типом нагрузки. Это доступный и простой способ получить эффективно работающую систему в случаях, когда HDD накопители не имеют физической возможности обеспечить желаемый результат.

При существующем разнообразии серверных задач и приложений, применение SSD-кэша в гибридных СХД становится все более привлекательным. Но следует помнить, что эта технология требовательна к условиям использования, и она не является универсальным решением всех проблем с производительностью.

SSD-кэш, реализованный в СХД RAIDIX, обладает особым набором свойств, который позволяет ему не только ускорять работу системы, но и продлевать срок используемых SSD-накопителей.