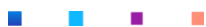




# Лучшие практики по производительности системы хранения данных АЭРОДИСК

Дата: 27.09.2024

Версия: 5.4



## Оглавление

Общие рекомендации.....	3
RDG группы.....	3
Выбор типа RAID.....	3
Соотношение размера блока тома и типа/размера RAID .....	4
SSD-кэширование для групп из шпиндельных дисков .....	5
Количество логических томов .....	5
Рекомендации для типовых задач и приложений .....	6
DDP пулы.....	6
Выбор типа RAID.....	7
Количество логических томов .....	8
Рекомендации для типовых задач и приложений .....	8
Рекомендации к конфигурации аппаратных ресурсов СХД .....	8
Общие рекомендации по типу используемых накопителей.....	9

## Общие рекомендации

Для обеспечения высокого быстродействия рекомендуется равномерно распределять нагрузку между двумя контроллерами. Распределение RDG и DDP групп между двумя контроллерами позволяет максимально задействовать все накопители подсистемы хранения и вычислительные мощности обоих контроллеров.

Для работы со шпиндельными дисками SAS 10k или NL-SAS 7.2k рекомендуется использовать RAID Distributed Group (RDG). Для работы с твердотельными накопителями SSD оптимальным вариантом с точки зрения производительности будет использование Dynamic Disk Pool (DDP).

## RDG группы

RDG группа состоит из одного или более vdev одинакового типа RAID и размера. Для шпиндельных дисков SAS 10k или NL-SAS 7.2k добавляется SSD кэш (SSD кэш+mcache) и опционально SSD Online-Tiering.

При создании логического тома на RDG группе можно выбрать размер блока тома: 4K, 16K, 32K, 64K, 128K. Значение по умолчанию – 128K.

Размер блока тома нужно выбирать максимально близким к размеру того блока, с которым работает приложение, данные которого размещаются на томе. Соответствие размера блока приложения размеру блока тома особенно важно при случайном характере нагрузки. Также соответствие размера блока приложения размеру блока тома особенно важно для групп на шпиндельных дисках.

## Выбор типа RAID

Для приложений, создающих случайную нагрузку высокой интенсивности маленьким блоком (4K, 8K), предпочтительнее использовать RAID10.

Для тех случаев, когда рабочий объем более важен, чем производительность, используются твердотельные диски или шпиндельные диски SAS 10k в RAID5. Не рекомендуется использовать RAID5 со шпиндельными дисками большого объема (от 4ТБ) из-за длительного перестроения RAID и потенциальной опасности потери данных при выходе из строя второго диска в vdev.

Для приложений, создающих последовательную нагрузку большим блоком, рекомендуется использовать RAID5 для твердотельных дисков SSD и шпиндельных дисков SAS 10k; RAID6/60 или RAID6P/60P для шпиндельных дисков большой емкости NL-SAS 7.2k.

## Соотношение размера блока тома и типа/размера RAID

Размер блока тома определяет какими «кусками» происходит взаимодействие (запись/чтение) с vdev, входящих в состав группы. Т.е. если на том происходит запись блоком большего размера, чем блок логического тома, то входящий блок разбивается на несколько блоков (при этом на уровне группы для каждого такого блока формируется контрольная сумма). Если размер входящего блока меньше блока логического тома, то выполняется следующая последовательность действий: чтение блока с vdev, модификация и запись.

Блок логического тома (логический блок) далее на уровне vdev размещается на физических дисках по схеме, определяемой формулой используемого RAID, и vdev взаимодействует с дисками блоками, равными физическому размеру блока диска (PHY\_SIZE).

Для vdev RAID50/60/60P предпочтительнее, чтобы запись на них происходила полным страйпом. Поэтому, при размере физического блока диска 4K, эффективнее писать на vdev блоками размером N\*4K (N – количество дисков данных в RAID).

Рекомендованные формулы vdev для RAID5/6/6P с точки зрения производительности приведены в таблице ниже.

Таблица 1 Рекомендуемые формулы для уровней RAID

Размер блока логического тома	RAID5	RAID6
4K	Не рекомендуется	Не рекомендуется
8K	Не рекомендуется	Не рекомендуется
16K	4+1, 8+1	4+2,8+2
32K	4+1,8+1	4+2,8+2,10+2
64K	4+1,8+1	4+2,8+2,10+2
128K	4+1,8+1	4+2,8+2,10+2

Использование более 12 накопителей в одном vdev не рекомендуется для медленных дисков большого размера из-за длительного перестроения RAID.

## SSD-кэширование для групп из шпиндельных дисков

Для групп, создаваемых на шпиндельных дисках SAS 10k или NL-SAS 7.2k, использование SSD-кэша обязательно. SSD-кэш ускоряет запись, группируя записываемые блоки для дальнейшей асинхронной выгрузки на vdev. В случае интенсивной записи в постоянном режиме, например, в системах наблюдения, рекомендуется использование дисков большой емкости с DWPD=3 для предотвращения выхода дисков SSD из строя из-за исчерпания ресурса ячеек флэш-памяти.

Режим SSD кэш+mcache рекомендуется использовать при нагрузке, носящей случайный характер, и при включении функционала дедупликации. В случае использования функционала дедупликации необходимо использовать SSD-диски большой емкости для обеспечения необходимого объема для хранения метаданных.

SSD Online-Tiering рекомендуется использовать при нагрузке, носящей случайный характер и имеющей ярко выраженные «горячие» области, например, для индексов баз данных. Если в нагрузке преобладают операции записи или нагрузка носит последовательный характер, использование Online-Tiering не принесет результат.

Возможно одновременное использование технологий SSD-кэш и Online-Tiering для увеличения производительности RDG-группы и на операциях случайного чтения, и на операциях записи.

Следует отметить, что включение функций кэширования на SSD ведет за собой использование части оперативной памяти RAM-контроллера для хранения связанных с кэшированием метаданных.

## Количество логических томов

Для приложений, создающих интенсивную нагрузку ввода/вывода, требуется разделять ее на множество логических томов. Например, для высоконагруженных баз данных рекомендуется выделять отдельные тома для логов, индексов и размещать файлы базы данных на нескольких логических томах.

Для систем видеонаблюдения с большим количеством камер рекомендуется писать видеопотоки с разных групп камер на разные логические тома.

Для получения максимальной производительности рекомендуется:

- для групп из шпиндельных дисков – не менее 4-х томов на группу;
- для групп из твердотельных дисков SSD – не менее 8-ми томов на группу.

## Рекомендации для типовых задач и приложений

В зависимости от типа нагрузки и требований по отказоустойчивости настройки могут существенно отличаться, в частности по используемым уровням защиты RAID, размеру блока на логических томах (LUN), использованию SSD кэш/мсache и пр.

### Виртуальные среды VMware

- Рекомендуется использование твердотельных дисков SSD. Для низконагруженных сред виртуализации допустимо использование шпиндельных дисков SAS 10k с обязательным использованием SSD кэш + мсache.
- Для лучшей производительности предпочтительней использовать RAID10.
- При использовании RAID5 рекомендуемая формула 4+1 или 8+1.
- Рекомендуется использовать тома для датасторов небольшого размера – 4-16ТБ.
- Требуется создавать не менее 4-х томов на группу.
- Рекомендуемый размер блока тома для виртуальных машин – 32К.
- Для нагруженных баз данных рекомендуется создать отдельные тома с соответствующим размером блока (4К или 8К).

### Системы видеонаблюдения

- Диски NL-SAS.
- Обязательно использовать SSD Cache.
- RAID60 формула 8+2, 10+2 или RAID60P формула 8+3, 10+3.
- Не менее 4-х томов на группу.
- Размер блока тома 128К.
- Не допускается размещать базу каталога «Системы Видеонаблюдения» на группах NL-SAS дисков, рекомендуется выделять под нее отдельные группы и тома на дисках SAS или SSD.

### DDP пулы

Dynamic Disk Pool (DDP) состоит из двух или более дисков одинакового типа и размера. При создании DDP можно выбрать размер блока, которым происходит выделение пространства. По умолчанию размер блока равен 4МБ, опционально можно выбрать 16МБ.

Возможно создание «толстых» томов, «тонких» томов и томов с поддержкой дедупликации и компрессии. Для достижения максимальной производительности рекомендуется использовать толстые тома.

## Выбор типа RAID

При создании логического тома выбирается уровень RAID: 0, 1, 10, 5, 6. В отличие от RDG, в данном случае распределение четности происходит не между физическими дисками, а между блоками, которые выделяются под конкретный логический том во время его создания.

Выбор уровня RAID зависит от требований по отказоустойчивости и от профиля нагрузки. Чем выше процент операций записи, тем более оптимально с точки зрения производительности использование RAID10.

Таблица 2 Рекомендуемые уровни RAID для различных типов нагрузки

Тип нагрузки	Оптимальный уровень защиты	Тип нагрузки
Потоковое чтение, запись	RAID1, RAID10, RAID5, RAID6	Потоковое чтение, запись
100% случайное чтение	RAID1, RAID10, RAID5, RAID6	100% случайное чтение
100% случайная запись	RAID1, RAID10	100% случайная запись
Смешанная случайная нагрузка, 70% чтения и 30% записи	RAID1, RAID10, RAID5, RAID6	Смешанная случайная нагрузка, 70% чтения и 30% записи

Рекомендуется использование следующих уровней защиты: RAID10 с числом дисков 2 до 12, RAID5 4+1, RAID5 8+1, RAID6 4+2, RAID 8+2. В случае размещения логических томов на большем числе дисков быстродействие снижается незначительно, но существенно снижается производительность и возрастает время восстановления отказоустойчивости после выхода диска из строя.

При создании логических томов автоматически проверяется свободное место на дисках DDP пула, том создается на наименее загруженных дисках.

## Количество логических томов

При работе с томами по блочным протоколам Fibre Channel и iSCSI для получения максимальной производительности СХД желательно создавать не менее 4-8 томов на DDP-пул и не менее 8-12 томов на контроллер. Точное число томов для достижения максимальной производительности может варьироваться в зависимости от профиля нагрузки и используемого уровня RAID.

## Рекомендации для типовых задач и приложений

### Виртуальные среды VMware

- Рекомендуется использовать SSD.
- Если нужно обеспечить минимальные задержки, рекомендуется использовать RAID10 для защиты данных.
- Если низкие задержки некритичны или большой процент операций чтения, рекомендуется использовать RAID5 или RAID6.
- Для RAID5 формула 4+1 или 8+1.
- Для RAID6 рекомендуемая формула 4+2 или 8+2.
- Не менее 4-х томов на группу.

### Высокопроизводительные базы данных

- Рекомендуется использовать SSD.
- Если нужно обеспечить минимальные задержки, рекомендуется использовать RAID10 для защиты данных.
- Если низкие задержки некритичны или большой процент операций чтения, рекомендуется использовать RAID5 или RAID6.
- Для RAID5 формула 4+1 или 8+1.
- Для RAID6 рекомендуемая формула 4+2 или 8+2.
- Возможно использование RAID 5 или RAID6 для файлов баз данных, RAID10 для логов.

## Рекомендации к конфигурации аппаратных ресурсов СХД

Каждый из компонентов системы хранения данных имеет свой лимит по числу операций ввода/вывода в секунду и по пропускной способности в МБ/с.

Контроллеры массива;



- SSD/HDD накопители;
- Front-End адаптеры (Fibre Channel, Ethernet);
- Back-end адаптеры (SAS 12Gb).

Необходимо предусмотреть нужное количество дисков, адаптеров ввода-вывода и контроллеров с достаточным числом вычислительных ядер для обеспечения требуемой приложениями производительности.

На этапе планирования системы можно выбрать число процессорных ядер контроллеров и объем оперативной памяти. В зависимости от модели в СХД может быть от 16 до 96 процессорных ядер и от 512 ГБ до 3 ТБ оперативной памяти. При использовании дедупликации рекомендуется добавлять в систему для каждого контроллера не менее 1 ГБ памяти на каждый ТБ полезного объема. Аналогично, при использовании удаленной репликации или функционала метрокластера необходимо предусмотреть 1 ГБ оперативной памяти на каждый ТБ реплицируемого между двумя системами хранения объема.

## Общие рекомендации по типу используемых накопителей

- Хранение резервных копий, архивы, видеонаблюдение – NL\_SAS 7200 rpm диски.
- Малонагруженная виртуализация, инфраструктурные задачи – SAS 10k.
- Виртуализация, базы данных – SSD.

Необходимо, чтобы число установленных накопителей не было узким местом. Для достижения максимального быстродействия All Flash СХД рекомендуется использовать:

- Dynamic Disk пулы, как минимум по 1 на контроллер;
- RAID 10: не менее 24 на систему;
- RAID 5,6: не менее 48 на систему.