Лучшие практики АЭРОДИСК ENGINE AQ LITE

AERODISK

Оглавление

1 Рекомендации по защите данных **2**
Особенности RDG 2

2 Лучшие практики по производительности СХД **15**

Примечание

Документация находится в процессе обновления

Содержание

Рекомендации по защите данных

Настоящий документ содержит описание и практические примеры организации хранения данных в системах AERODISK:

- RAID Distributed Group (RDG);

Особенности RDG

Отличительными особенностями реализации RDG в системах AERODISK являются:

- RDG состоят из виртуальных устройств (VDEV), каждое из которых имеет заданную структуру RAID: 1/10, 5/50, 6/60, 6P/60P (тройная четность);
- в RDG поддерживается и файловый (NFS, CIFS), и блочный доступ (FC, iSCSI);
- виртуальные устройства последовательно объединяются в одну виртуальную группу RDG за счет чего количество дисков в группах (и для данных, и для четности) не ограничено;
- вне зависимости от объема тома или файловой системы все диски в группе участвуют в вводе-выводе для данного тома или файловой системы;
- диски горячей замены являются глобальными;
- любая группа может быть, как стандартной, так и гибридной;
- RAM-кэш включен по умолчанию и работает только на чтение;
- скорость перестроения RAID можно регулировать политикой перестроения;
- RDG поддерживает 2 уровня хранения, SSD+SAS или SSD+NL-SAS;
- RDG поддерживает ускорение операций ввода/вывода за счет использования SSD дисков в качестве SSD кэша и/или SSD Online-tiering;
- RDG наилучшим образом подходит для операций последовательного чтения/записи данных, а так же для операций случайного чтения.

Уровни RDG

В таблице ниже приведены примеры уровней дисковых групп (RDG) с указанием четности для одного виртуального устройства (VDEV) в RDG, процента полезной емкости, а также характеристик производительности. Администратор по своему усмотрению может создавать шаблоны для виртуальных устройств (VDEV).

Таблица 1. Примеры уровней дисковых групп и производительность на различных профилях нагрузки

Уровень	Диски данных	Диски четности	Последовательное чтение	Последовательная запись	Произвольное чтение	Произвольная запись	Полезная емкость
RAID 1/10 (mirror)	1	1	Очень высокая	Очень высокая	Высокая	Средняя	50,0%
RAID 5/50 (2D+1P)	2	1	Высокая	Высокая	Высокая	Низкая	66,6%
RAID 5/50 (4D+1P)	4	1					80,0%
RAID 5/50 (6D+1P)	6	1					85,7%
RAID 6/60 (3D+2P)	3	2	Высокая	Высокая	Высокая	Очень низкая	60,0%
RAID 6/60 (5D+2P)	5	2					71,4%
RAID 6/60 (7D+2P)	7	2					77,8%
RAID 6P/60P (5D+3P)	5	3	Высокая	Высокая	Средняя	Очень низкая	62,5%
RAID 6P/60P (7D+3P)	7	3					70.0%
RAID 6P/60P (9D+3P)	9	3					75.0%

Общее количество дисков в RDG, выделенное для четности (степень избыточности) равно произведению количества дисков четности единичного виртуального устройства и количества виртуальных устройств. Например,

- RDG-группа из 121 диска с уровнем 60P (8+3).
- RDG состоит из 11 виртуальных устройств, т.е. 121=(8+3)x11.
- Каждое виртуальное устройство поддерживает потерю до 3-х дисков.
- Суммарно RDG поддерживает потерю до 33 дисков (т.е. 3*11).

Организация RDG

Примеры организации RDG

RAID-1/10 (mirror)

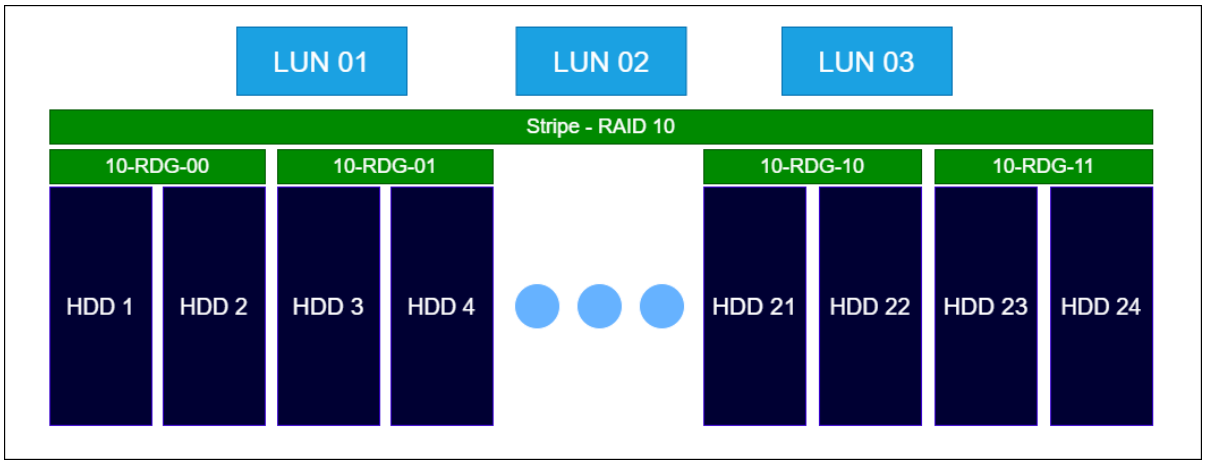
Минимальное количество физических дисков для организации группы типа RDG-1/10 – это 2 физических диска для конфигурации RAID-1 и 4 диска или более для конфигурации RAID-10. RDG-1/10 рекомендуется использовать для хранения данных с высокими требованиями к производительности (и чтение, и запись).

Внутри RDG диски группируются в виртуальные устройства (VDEV).

Физические диски парами объединяться в виртуальные устройства (VDEV), нумерация VDEV присваивается автоматически и соответствует типу RDG-группы 10-RDG-00, 10-RDG-01, 10-RDG-02 и т.п.

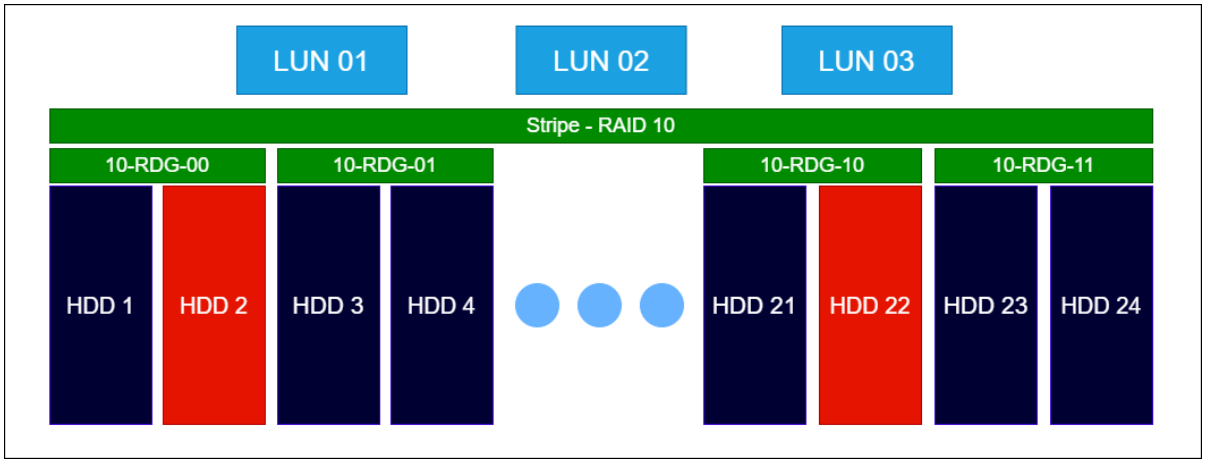
Дисковая группа строится из VDEV. Дисковая группа с типом RDG-10 допускает выход из строя одного физического диска в одном VDEV, то есть если у вас используется 6 VDEV (12 физических дисков), то в группе допускается выход из строя шести дисков (по одному в каждом VDEV).

Пример 1. Организация RDG-1/10 из 24 дисков



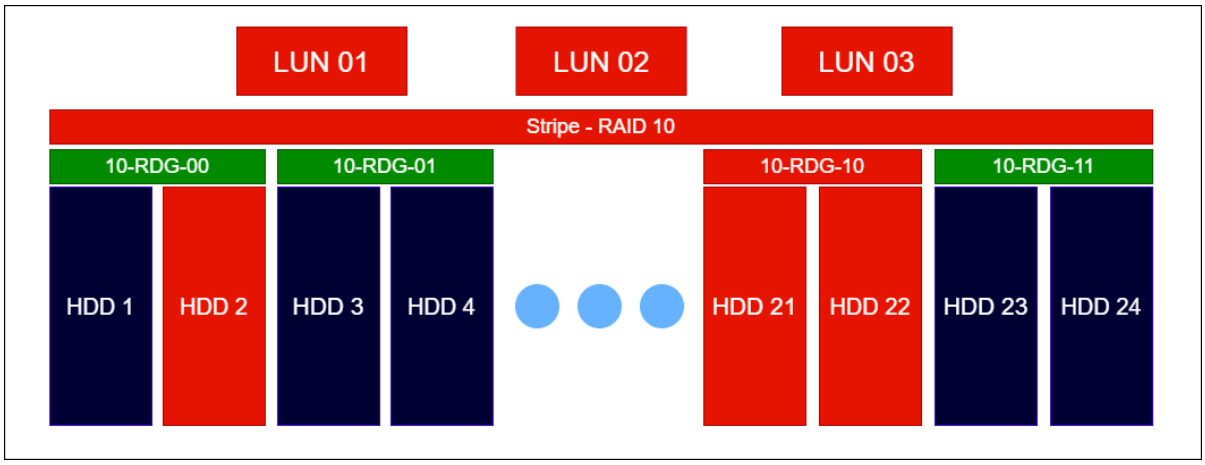
В **примере 1** показана организация RDG группы уровня RDG-1/10 из 24 дисков. Физические диски объединены в 12 VDEV.

Пример 2. RDG-1/10 из 24 дисков с вышедшими из строя несколькими дисками



В **примере 2** отражен выход из строя нескольких физических дисков в разных VDEV с типом 10-RDG. Такая ситуация считается критичной, данные доступны для чтения/записи, но отказоустойчивость отсутствует, выход из строя ещё одного из дисков в том же VDEV грозит потерей всех данных. **Следует незамедлительно заменить сбойные диски**.

Пример 3. RDG-1/10 из 24 дисков с вышедшими из строя несколькими дисками в одном VDEV



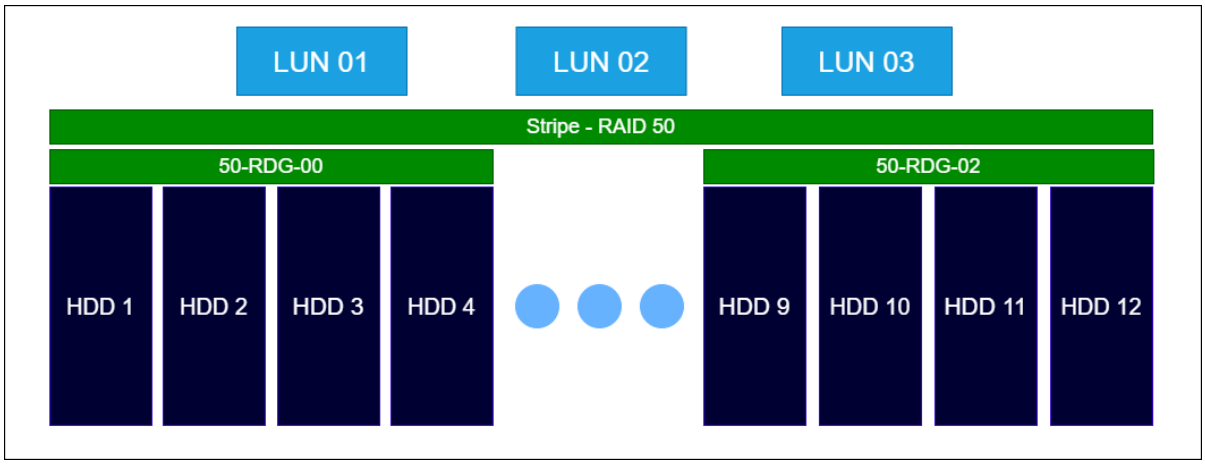
В **примере 3** отражен выход из строя 2-х физических дисков в одной VDEV с типом 10-RDG. Такая ситуация считается фатальной, все данные в RDG недоступны для чтения/записи, следует **заменить диски и восстановить данные из резервных копий**.

RAID-5/50

Минимальное количество физических дисков для организации группы типа RDG-5/50 – это 3 физических диска для конфигурации для RAID-5 и 6 дисков или более для конфигурации RAID-50.

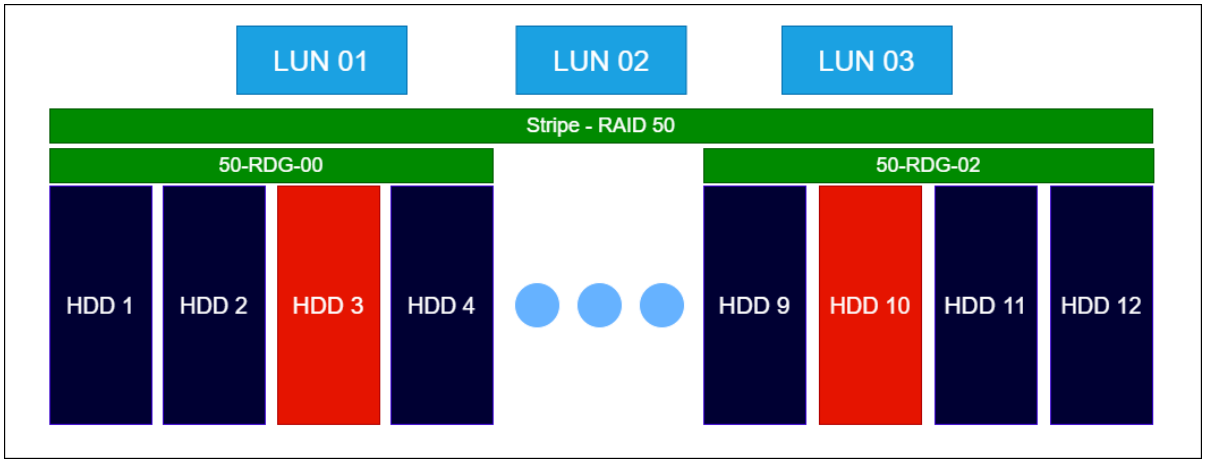
RDG-5/50 рекомендуется использовать для хранения данных невысокой критичности на NL-SAS дисках, а так же данных высокой критичности на SAS, SSD дисках, с высокими требованиями к производительности на чтение и средними требования к производительности на запись. При этом хорошей практикой является компенсация снижения производительности на запись путем использования SSD дисков в качестве кэша на чтение/запись, а также увеличение скорости чтения путем использования SSD дисков в качестве online-tiering устройств.

Пример 1. Организация RDG группы уровня RDG-50 (3D+1P)



В **примере 1** показана организация RDG группы уровня RDG-50 (3D+1P), т.е. в каждом виртуальном устройстве 3 диска используется для хранения данных и один для четности. Физические диски объединены в 3 VDEV-а.

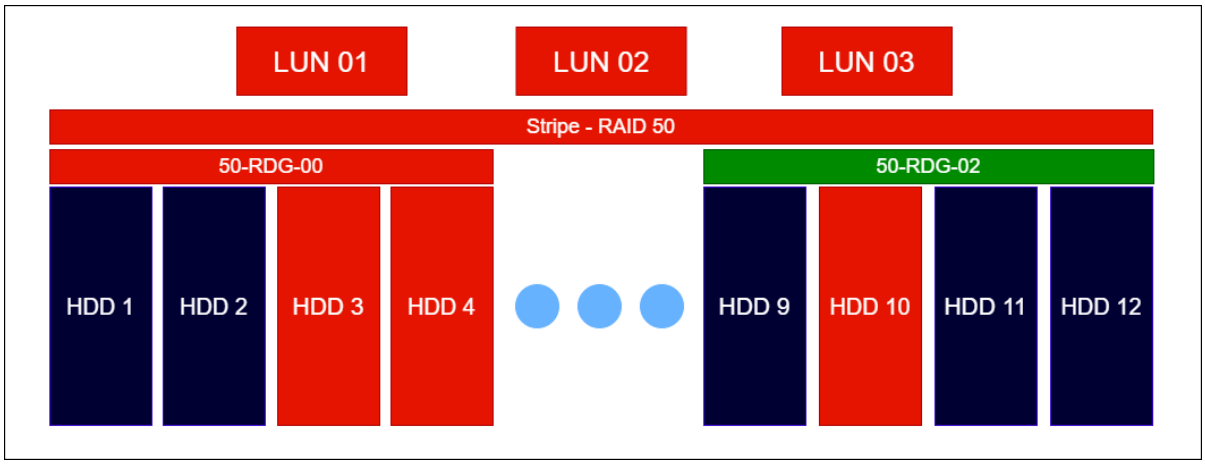
Пример 2. Организация RDG-50 с вышедшими из строя несколькими дисками



В **примере 2** отражен выход из строя нескольких физических дисков в разных виртуальных VDEV с типом 50-RDG. Такая ситуация считается критичной, данные по-прежнему доступны для чтения-записи, но отказоустойчивость отсутствует, выход из строя ещё одного из дисков грозит потерей всех данных. **Следует незамедлительно заменить сбойные диски**.

Пример 3. Организация RDG-50 с вышедшими из строя несколькими дисками в одном VDEV

В **примере 3** отражен выход из строя 2-х физических дисков в одной VDEV с типом 50-RDG. В данном примере стало недоступно виртуальное устройство VDEV – 50-RDG-00. Такая ситуация считается фатальной, данные в дисковой группе RDG недоступны для чтения и записи, следует заменить диски и восстановить данные из резервных копий.



RAID-6/60

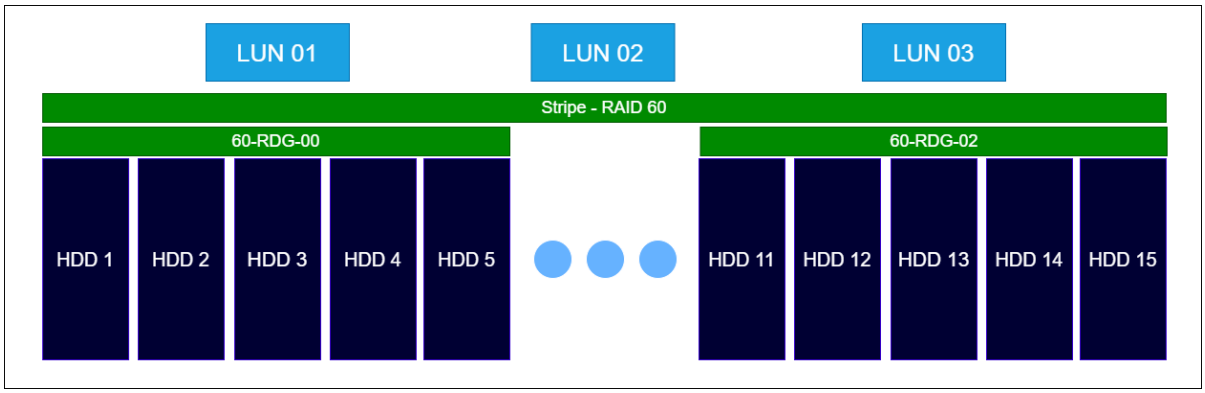
Минимальное количество физических дисков для организации группы типа RDG-6/60 – это 5 физических дисков для RDG-6 и 10 дисков или более для конфигурации RAID-60. Минимальное количество VDEV для дисковой группы типа RDG-60 – это 2 VDEV (или один VDEV для RDG-6).

> **Примечание**
>
> RDG-60 имеет более высокую степень надежности, чем RDG-50, но при этом меньшую производительность, под контрольные суммы выделяется два жестких диска в одном VDEV, что обеспечивает в RDG-60 более высокий уровень защиты данных.

Два диска используются для хранения четности. RDG-60 можно рекомендовать для хранения критичных данных со средними требованиями к производительности на чтение и средними требования к производительности на запись.

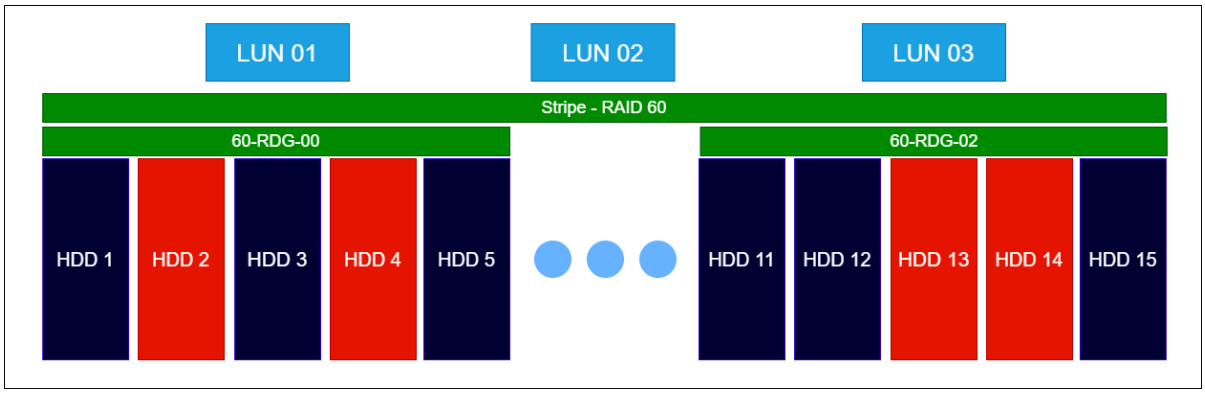
Дисковая группа с уровнем RDG-60 допускает выход из строя двух физических дисков в одном VDEV, выход из строя более двух дисков в одном VDEV ведет к потере данных.

Пример 1. Организация RDG-60



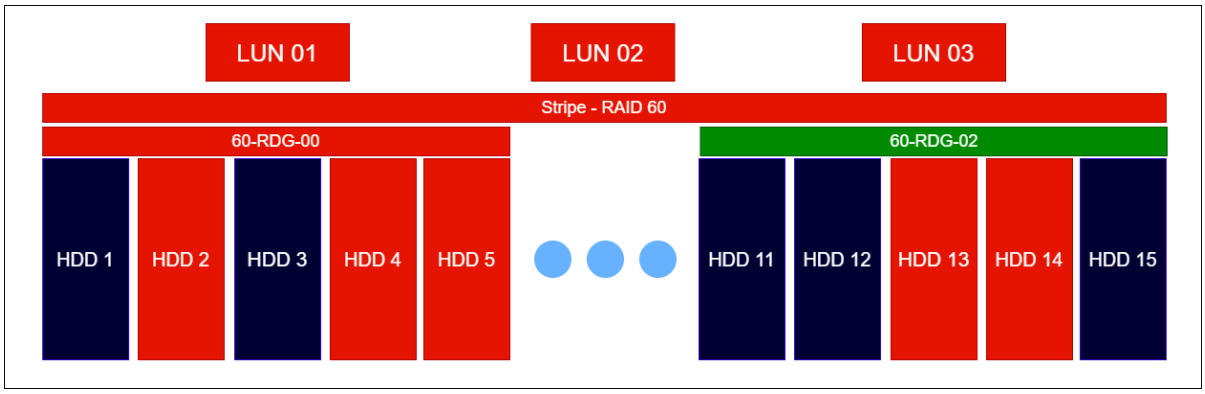
В **примере 1** показана организация RDG группы уровня 60-RDG. Физические диски объединены в 3 VDEVa по 5 дисков в каждом VDEV.

Пример 2. Организация RDG-60 с вышедшими из строя несколькими дисками



В **примере 2** показан выход из строя нескольких физических дисков – по два в разных VDEV, с типом 60-RDG. Такая ситуация считается критичной, но не фатальной, данные по-прежнему доступны для чтения и записи, но отказоустойчивость отсутствует. Выход из строя ещё одного из дисков в затронутых VDEV грозит потерей всех данных. **Следует незамедлительно заменить сбойные диски**.

Пример 3. Организация RDG-60 с вышедшими из строя несколькими дисками в одном VDEV



В **примере 3** отражен выход из строя трех физических дисков в VDEV 60-RDG-00, с типом 60-RDG. Такая ситуация считается фатальной, так как вышло из строя одновременно три жестких диска (при возможных двух), данные недоступны для чтения и записи. **Следует заменить диски и восстановить данные из резервных копий**.

RAID-6/60P

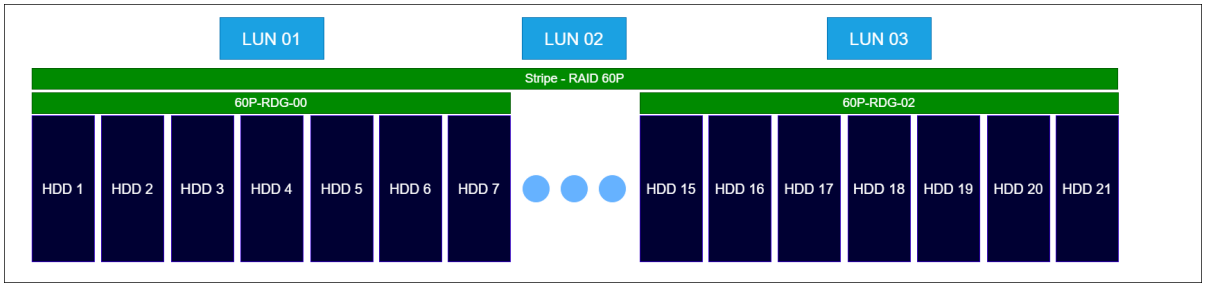
Минимальное количество физических дисков для организации группы типа RDG-6P/60P – это 7 дисков для RDG-6P и 14 физических дисков для конфигурации RDG60P – по семь дисков в каждом VDEV. Минимальное количество VDEV для дисковой группы типа RDG-60P – это 2 VDEV (или один для RDG-6P). Физические диски объединяться в VDEV по семь или более дисков в одном VDEV. Три диска используются для хранения четности, четыре или более диска используются для хранения данных.

RDG-60P рекомендуется использовать для хранения данных высокой критичности, с невысокими требованиями к производительности на чтение и невысокими требования к производительности на запись. Дисковая группа с уровнем RDG-60P допускает выход из строя трех физических дисков в одном VDEV, выход из строя более трех дисков в одной VDEV ведет к потере данных.

Примечание

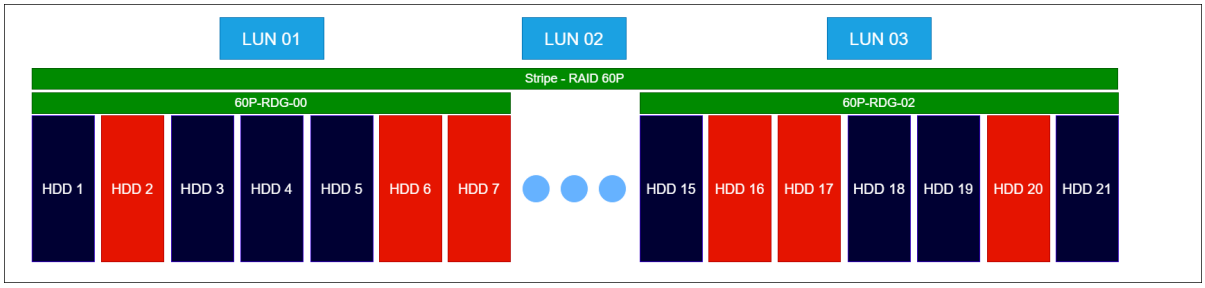
Дисковая группа RDG-60P по организации похожа на RDG-60 и RDG-50, но имеет более высокую степень надежности чем RDG-60. Под контрольные суммы выделяется три жестких диска в одном VDEV. RDG-60P обеспечивает очень высокий уровень надежности хранения данных.

Пример 1. Организация RDG-60P



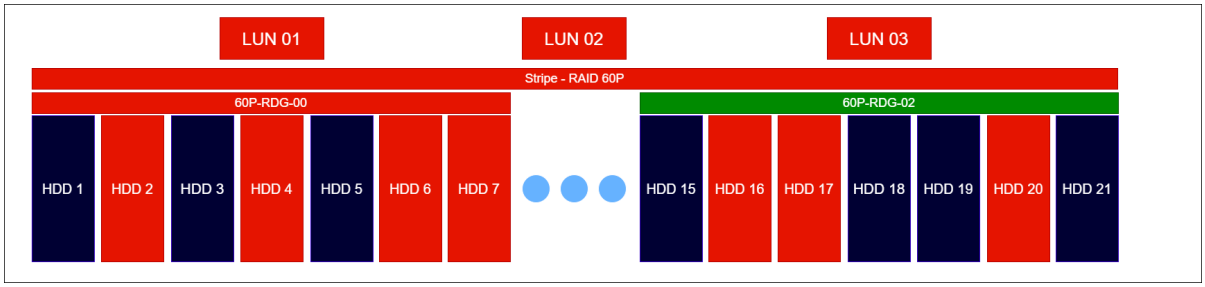
В **примере 1** показана организация RDG группы уровня 60P-RDG. Физические диски объединены в 3 VDEVa устройства по 7 физических дисков в каждом для обеспечения максимальной производительности и отказоустойчивости.

Пример 2. Организация RDG-60P с вышедшими из строя несколькими дисками



В **примере 2** отражен выход из строя 3-х физических дисков в разных VDEV с типом 60P-RDG. Такая ситуация считается критичной, но не фатальной, данные по-прежнему доступны для чтения и записи, но отказоустойчивость отсутствует, выход из строя ещё одного из дисков в затронутых VDEV грозит потерей всех данных. **Следует незамедлительно заменить сбойные диски**.

Пример 3. Организация RDG-60P с вышедшими из строя несколькими дисками в одном VDEV



В **примере 3** отражен выход из строя 4-х физических дисков в VDEV 60P-RDG-00 с типом 60P-RDG. Такая ситуация считается фатальной, так как вышло из строя одновременно четыре жестких диска (при возможных трех) в одном VDEV, данные недоступны для чтения и записи. **Следует заменить диски и восстановить данные из резервных копий.**

SSD кэш и online-tiering

Твердотельные диски обладают гораздо более высокой скоростью доступа к данным и меньшими задержками при работе по сравнению с жёсткими дисками.

СХД AERODISK поддерживает RDG группы как созданные только из SSD дисков, так гибридные RDG группы, когда в рамках одной RDG совмещают 2 типа дисков SAS 10k или NL-SAS для хранения данных и SSD для адаптации ввода-вывода (SSD-кэш и/или Online-tiering). Использование гибридных RDG групп до сих пор является наиболее эффективным по соотношению цена/качество, поскольку SSD меньше по объему и значительно выше по стоимости, если сравнивать их с SAS/NLSAS и SATA дисками. Учитывая этот факт, использование SSD как быстрого слоя для доступа к данным позволит совместить и максимально утилизировать высокую производительность SSD с большими объемами и низкой стоимостью других типов дисков.

В рамках одной RDG группы у администратора есть выбор, под какой тип операций использовать SSD диски (можно использовать оба уровня одновременно):

- Кэш на чтение и запись – тип READ-WRITE-SSD (RW-SSD)
- Дополнительный уровень – тип Online-tiering

При добавлении к RDG SSD диски всегда объединяются в RAID1.

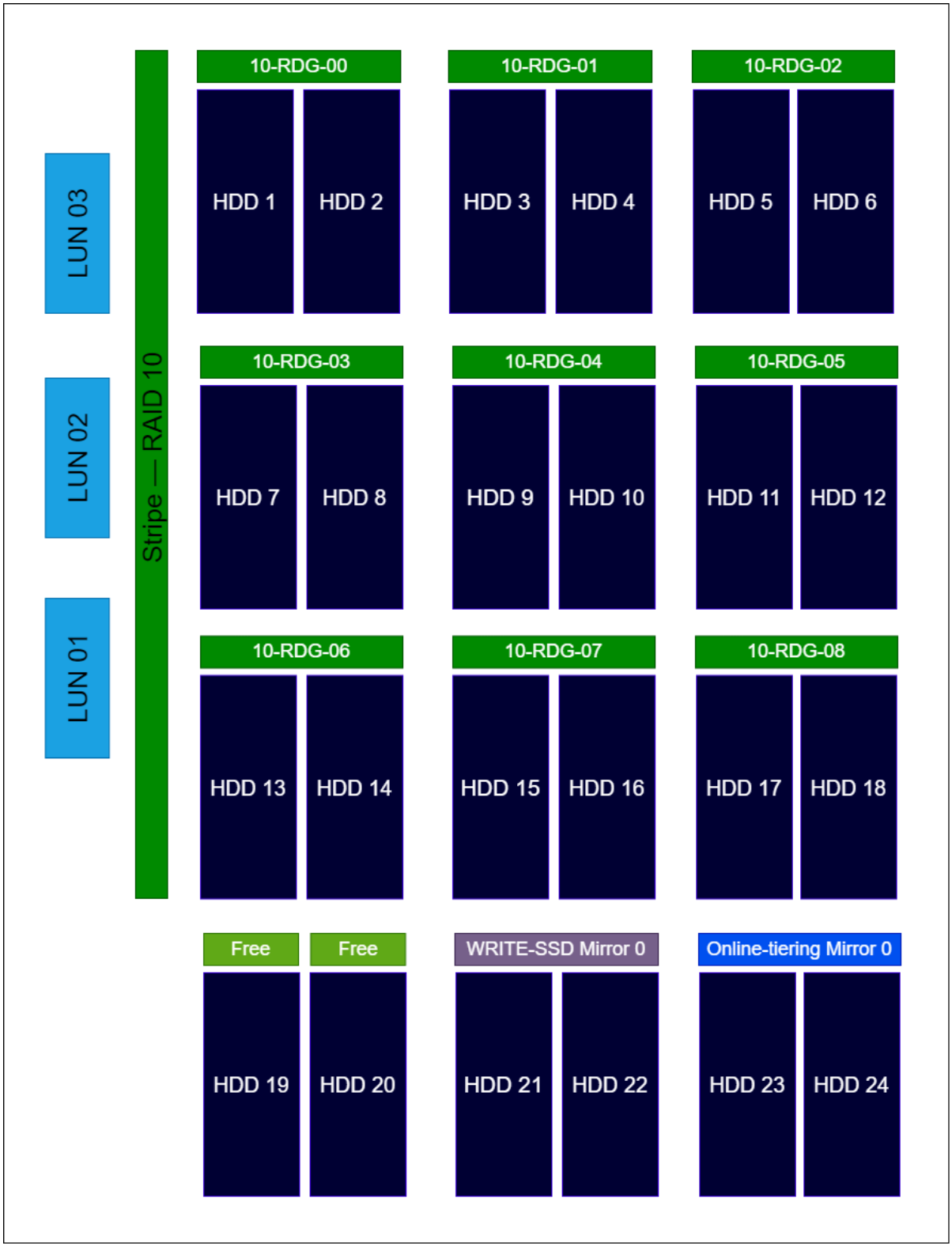
Пример 1. Таблица. Рекомендуемая разметка

Тип диска	Имя	VDEV	Тип	Слот
SAS 10k 2400Gb	R00	10-RDG-00	10-RDG	0:2:0
SAS 10k 2400Gb	R00	10-RDG-00	10-RDG	0:2:1
SAS 10k 2400Gb	R00	10-RDG-01	10-RDG	0:2:2
SAS 10k 2400Gb	R00	10-RDG-01	10-RDG	0:2:3
SAS 10k 2400Gb	R00	10-RDG-02	10-RDG	0:2:4
SAS 10k 2400Gb	R00	10-RDG-02	10-RDG	0:2:5
SAS 10k 2400Gb	R00	10-RDG-03	10-RDG	0:2:6
SAS 10k 2400Gb	R00	10-RDG-03	10-RDG	0:2:7
SAS 10k 2400Gb	R00	10-RDG-04	10-RDG	0:2:8
SAS 10k 2400Gb	R00	10-RDG-04	10-RDG	0:2:9
SAS 10k 2400Gb	R00	10-RDG-05	10-RDG	0:2:10
SAS 10k 2400Gb	R00	10-RDG-05	10-RDG	0:2:11
SAS 10k 2400Gb	R00	10-RDG-06	10-RDG	0:2:12
SAS 10k 2400Gb	R00	10-RDG-06	10-RDG	0:2:13
SAS 10k 2400Gb	R00	10-RDG-07	10-RDG	0:2:14
SAS 10k 2400Gb	R00	10-RDG-07	10-RDG	0:2:15
SAS 10k 2400Gb	R00	10-RDG-08	10-RDG	0:2:16
SAS 10k 2400Gb	R00	10-RDG-08	10-RDG	0:2:17
SAS 10k 2400Gb	FREE	FREE-DISK	SPARE Disk	0:2:18
SAS 10k 2400Gb	FREE	FREE-DISK	SPARE Disk	0:2:19
SAS DWPD3 800GB SSD	R00	WRITE SSD Mirror-0	WRITE SSD	0:2:20
SAS DWPD3 800GB SSD	R00	WRITE SSD Mirror-0	WRITE SSD	0:2:21
SAS DWPD1 1920GB SSD	R00	Online tier Mirror-0	Online tier SSD	0:2:22
SAS DWPD1 1920GB SSD	R00	Online tier Mirror-0	Online tier SSD	0:2:23

В этой таблице показан один из вариантов рекомендуемой разметки СХД AERODISK ENGINE N2 с установленными 20 дисками SAS и четырьмя дисками SSD для конфигурации с RDG-10.

Для SSD кэш рекомендуется использовать диски с DWPD 3 или более для предотвращения преждевременного выхода SSD диска из строя в связи с исчерпанием ресурса ячеек флэш-памяти. Желательно предусмотреть не менее 2 дисков SSD на каждые 20-30 HDD. В ряде случае, например в системах видеонаблюдения, рекомендуется использовать диски объемом 1600GB и более.

Пример 2. Логическая схема разметки RDG-10 + SPARE + RW SSD + ONLINE TIER SSD



В **примере 2** показана логическая схема разметки RDG-10 + SPARE + RW SSD + ONLINE TIER SSD. Из схемы видно, что диски с 1 по 18 задействованы под организацию RDG группы с типом RDG-10, диски 19 и 20 используются как глобальные SPARE устройства, SSD диски с 21 по 24 используются для организации адаптированного ввода-вывода.

SPARE-диски

SPARE диски или диски **горячей замены** – это тип дисков, обеспечивающих максимальную доступность дисковой группы.

Все диски в системе по умолчанию являются дисками глобальной горячей замены (global hot spare).

Особенности использования дисков SPARE:

- один диск может быть использован любой дисковой группой RDG, состоящей из того же типа дисков (Global hot spare);
- чем больше физических дисков используется в дисковой группе, тем больше дисков SPARE рекомендуется резервировать. На 30 физических дисков рекомендуется иметь не менее одного spare диска.

Рекомендуется использовать диски SPARE для всех типов дисков. В случае выхода из строя одного из дисков с данными, данные перестроятся на диск SPARE в ONLINE-режиме, без потери доступа к СХД. В случае, если замененный диск будет возвращен в работу, обратного копирования данных со SPARE-диска на замененный диск автоматически проводиться не будет.

Распределение групп

Группы RDG в SAN режиме рекомендуется распределять равномерно по двум системным контроллерам СХД. При таком распределении системные ресурсы СХД используются сбалансировано, точки отказа отсутствуют.

Режимы ускорения ввода/вывода

Для ускорения ввода/вывода доступны следующие функции:

- SSD-кэш чтение и запись (SSD RW);
- SSD-кэш чтение и запись и хранение метаданных (SSD RW +MCACHE);
- Online-tiering (SSD Online-tiering).

SSD Cache, MCACHE

SSD-кэш или SSD-кэш + MCACHE логически разделяет RDG на 2 плана производительности:

- **стандартный** – где используется один тип дисков и адаптация соответственно выполняется только на уровне оперативной памяти и только для операций чтения;
- **быстрый** – где используются SSD диски для кэширования и/или Online-tiering.

План производительности назначается автоматически на уровне RDG при добавлении SSD-дисков в группу и применяется ко всем LUN-ам и файловым системам, работающим в данной RDG, сразу после добавления.

При создании гибридного хранилища SSD-диски добавляются в кэш пул на запись/чтение (2 диска) в RAID1.

SSD-кэш работает во фронтальном режиме и по умолчанию применяется для всех транзакций. При этом, чтобы исключить переполнение кэша, применяется механизм циклической адаптации (выталкивания) записей из кэша.

> **Примечание**
>
> Система Aerodisk Engine не имеет логического ограничения по объему SSD и RAM кэша, за исключением физического ограничения используемого оборудования.

SSD Online-tiering

SSD Online-tiering – это режим хранения данных, который позволяет перемещать блоки данных между различными уровнями в зависимости от нагрузки на них, позволяя тем самым размещать более «горячие» данные (т.е. часто используемые) на быстрых дисках, а более «холодные» данные (т.е. редко используемые) на медленных.

Перемещение блоков данных между уровнями происходит в онлайн-режиме. На SSD-слое всегда хранится копия данных с HDD-дисков, то есть добавление SSD-дисков в Online-tiering не добавляет общей емкости RDG-группе.

Диски для Online-tiering также добавляются на уровне RDG-группы, после добавления дисков в Online-tiering группа меняет статус на «Быстрая».

Минимальное и рекомендуемое количество дисков на уровень Online-tiering – 2.

В отличие от механизма кэширования, данный функционал хранит данные, пока к ним есть обращения и для режима Online-tiering рекомендуется использовать более емкие SSD-диски по сравнению с режимом SSD-кэш.

Лучшие практики по производительности СХД

Для обеспечения высокого быстродействия СХД рекомендуется равномерно распределять нагрузку между контроллерами. Для этого требуется учитывать не только возможности программного обеспечения и аппаратную конфигурацию СХД, но и тип данных, с которыми необходимо работать.

Актуальную версию документа **Лучшие практики по производительности СХД** можно увидеть на нашем сайте (https://aerodisk.ru) в разделе Документация, Лучшие практики.

В документе даны рекомендации по выбору архитектуры СХД в зависимости от ее назначения и типа данных, с которыми предстоит работать.