



СХД AERODISK ENGINE

Организация групп хранения данных

Дата	24.12.2017
Версия документа	1.43
Версия ПО	A-CORE v1.3

Содержание

1	Введение.....	3
2	Описание уровней RDG.....	4
3	Примеры организации RDG	5
3.1	Организация группы RAID-1/10 (mirror).....	5
3.2	Организация группы RAID-5/50.....	9
3.3	Организация группы RDG-6/60	13
3.4	Организация группы RDG-6P/60P.....	17
3.5	Организация группы – SPARE диски.....	21
3.6	Организация группы – SSD диски	22
4	Распределение групп между контроллерами.....	26

1 Введение

Настоящий документ содержит описание и практические примеры организации RAID-групп (RDG) в системах AERODISK.

Отличительными особенностями реализации RAID в системах AERODISK (отличие виртуальных RDG от классических RAID-групп) являются:

- RDG состоят из виртуальных устройств, каждое из которых имеет заданную структуру RAID (0, 1/10, 5/50, 6/60, 6P/60P)
- виртуальные устройства последовательно объединяются в одну виртуальную группу RDG за счет чего количество дисков в группах (и для данных, и для четности) не ограничено;
- производительность групп возрастает пропорционально количеству дисков без деградации;
- вне зависимости от объема тома или файловой системы все диски в группе участвуют в вводе-выводе для данного тома или файловой системы;
- приоритет производительности для различных томов и файловых систем в группе задается с помощью QOS;
- диски горячей замены возможно назначить на одну группу, на несколько или на все;
- любая группа может быть, как гибридной, так и стандартной.

2 Описание уровней RDG

В таблице ниже приведены поддерживаемые уровни дисковых групп (RDG) с указанием четности для одного виртуального устройства в RDG, процента полезной емкости, а также характеристик производительности.

Уровень	Диски для данных в одном виртуальном устройстве	Диски для четности в одном виртуальном устройстве	Производительность (чтение)	Производительность (запись)	Полезная емкость
RAID-1/10 (mirror)	1	1	высокая	высокая	49,82%
RAID-5/50 (2D+1P)	2	1	высокая	средняя	66,43%
RAID-5/50 (4D+1P)	4	1			79,71%
RAID-5/50 (6D+1P)	6	1			83,50%
RAID-6/60 (3D+2P)	3	2	средняя	низкая	59,05%
RAID-6/60 (5D+2P)	5	2			66,10%
RAID-6/60 (7D+2P)	7	2			75,63%
RAID-6P/60P (4D+3P)	4	3	высокая	средняя	56,66%
RAID-6P/60P (6D+3P)	6	3			61,08%
RAID-6P/60P (8D+3P)	8	3			72,69%

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. $121=(8+3) \times 11$
- Каждое виртуальное устройство поддерживает потерю до 3-х дисков
- Суммарно RDG поддерживает потерю до 33 дисков (т.е. 3×11)

3 Примеры организации RDG

3.1 Организация группы 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 низкой критичности, но с высокими требованиями к производительности (и чтение, и запись).

Внутри RDG диски группируются в виртуальные устройства (VDEV).

Физические диски парами объединяются в виртуальные устройства, нумерация V-DEV присваивается автоматически и соответствует типу RDG-группы 10-RDG-00, 10-RDG-01, 10-RDG-02 и т.п.

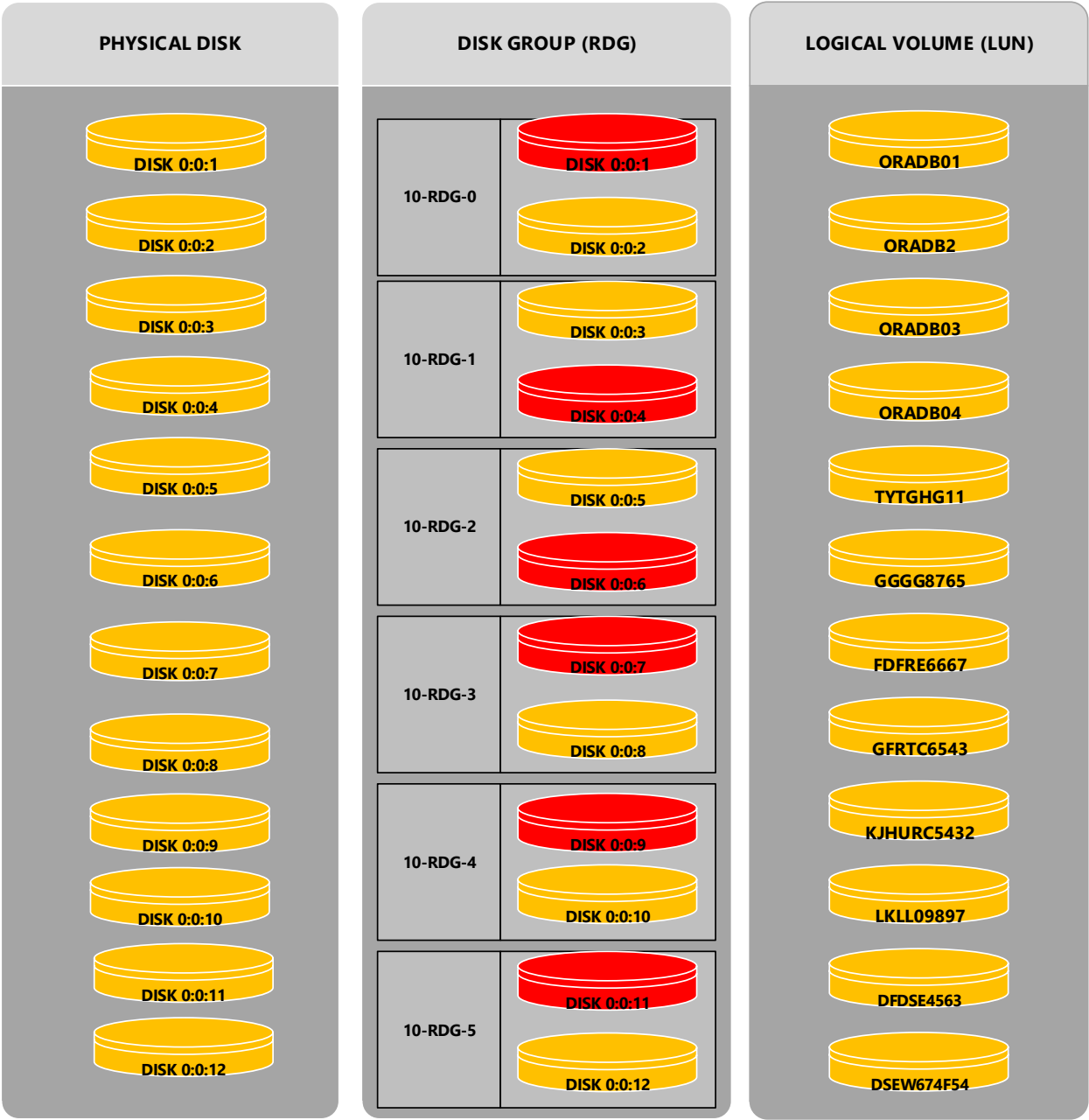
Из V-DEV строится дисковая группа. Дисковая группа с типом RDG-10 допускает выход из строя одного физического диска в одном V-DEV, то есть если у вас используется 6 V-DEV (12 физических дисков), то в группе допускается выход из строя шести дисков (по одному в каждом V-DEV).

Пример 1:



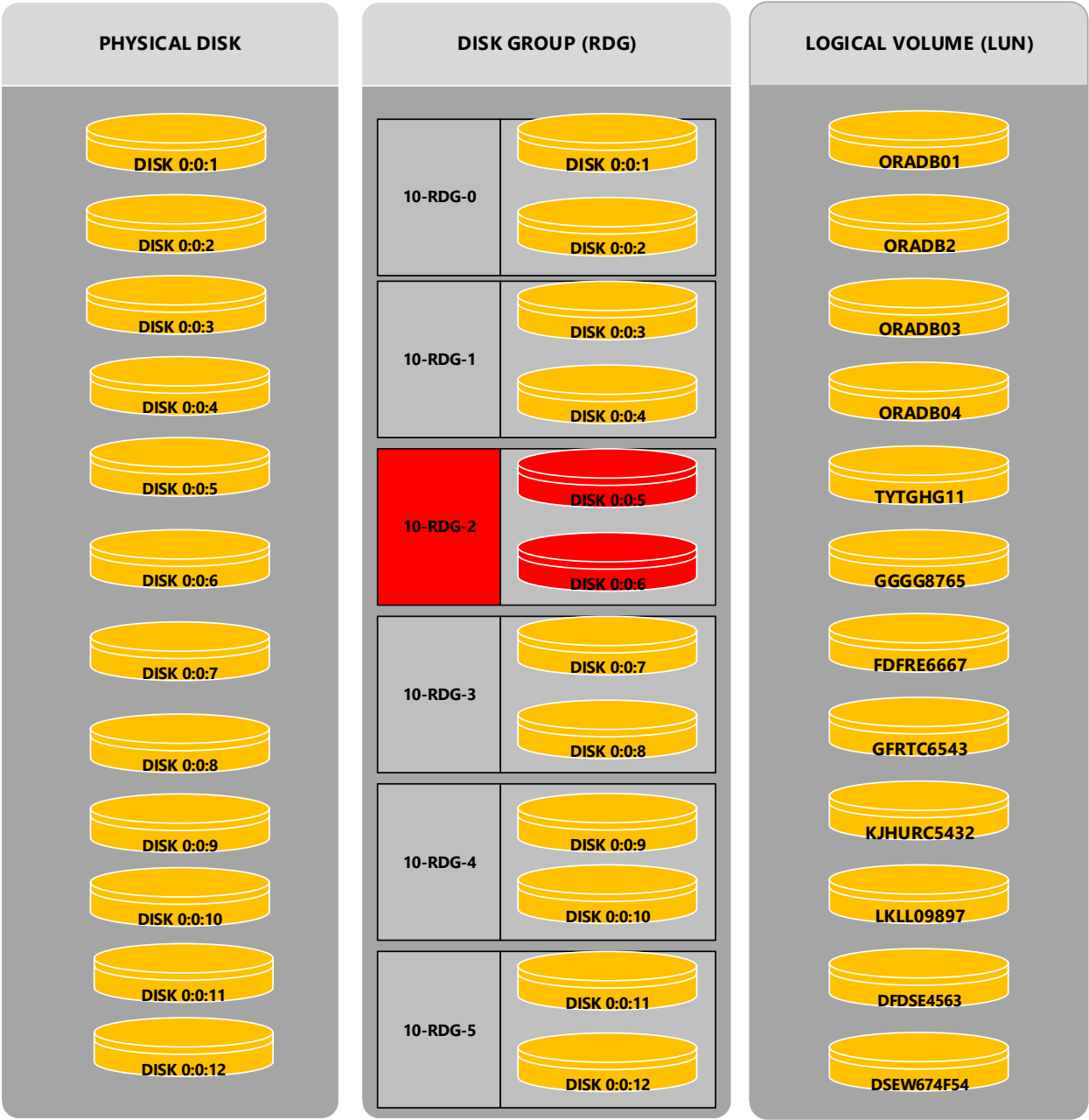
В **Примере-1** показана организация RDG группы уровня RDG-1/10. Физические диски объединены в V-DEV.

Пример 2:



В **Примере-2** отражен выход из строя нескольких физических дисков в разных V-DEV с типом 10-RDG. Такая ситуация считается критичной, данные 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:



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

3.2 Организация группы 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.

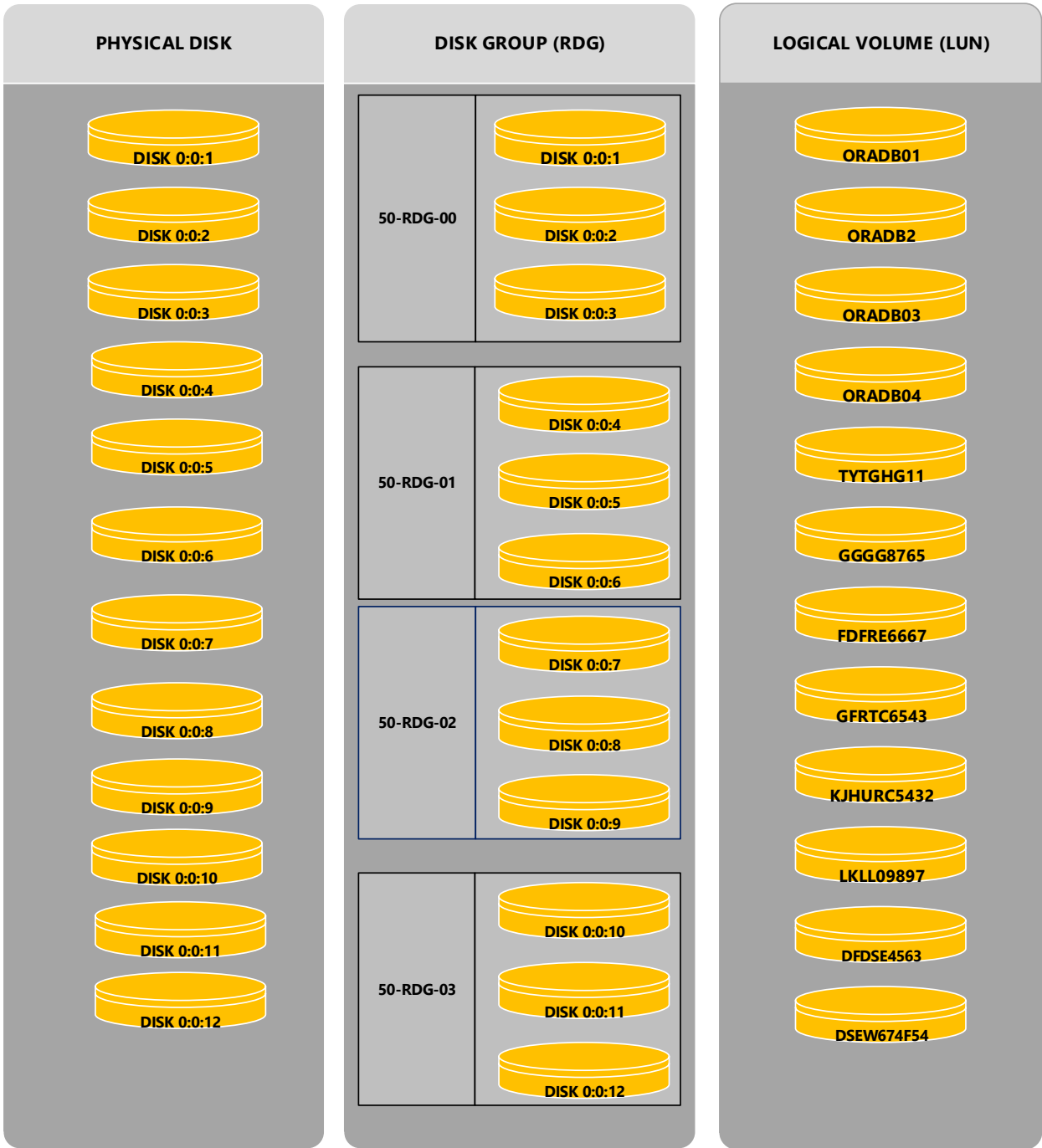
RDG-5/50 имеет следующие варианты конфигурации:

Уровень	Диски для данных в одном виртуальном устройстве	Диски для четности в одном виртуальном устройстве	Производительность (чтение)	Производительность (запись)	Полезная емкость
RAID-5/50 (2D+1P)	2	1	высокая	средняя	66,43%
RAID-5/50 (4D+1P)	4	1			79,71%
RAID-5/50 (6D+1P)	6	1			83,50%

От количества дисков в виртуальном устройстве зависит объем группы, чем больше дисков в виртуальном устройстве, тем больше полезная емкость. При этом следует помнить, что меньшее количество виртуальных устройств будет давать меньше производительности, а также снижать надежность (т.к. меньшее количество дисков во всех V-DEV будут использоваться для четности). Исходя из этого при конфигурации группы следует искать разумный баланс между емкостью и производительностью.

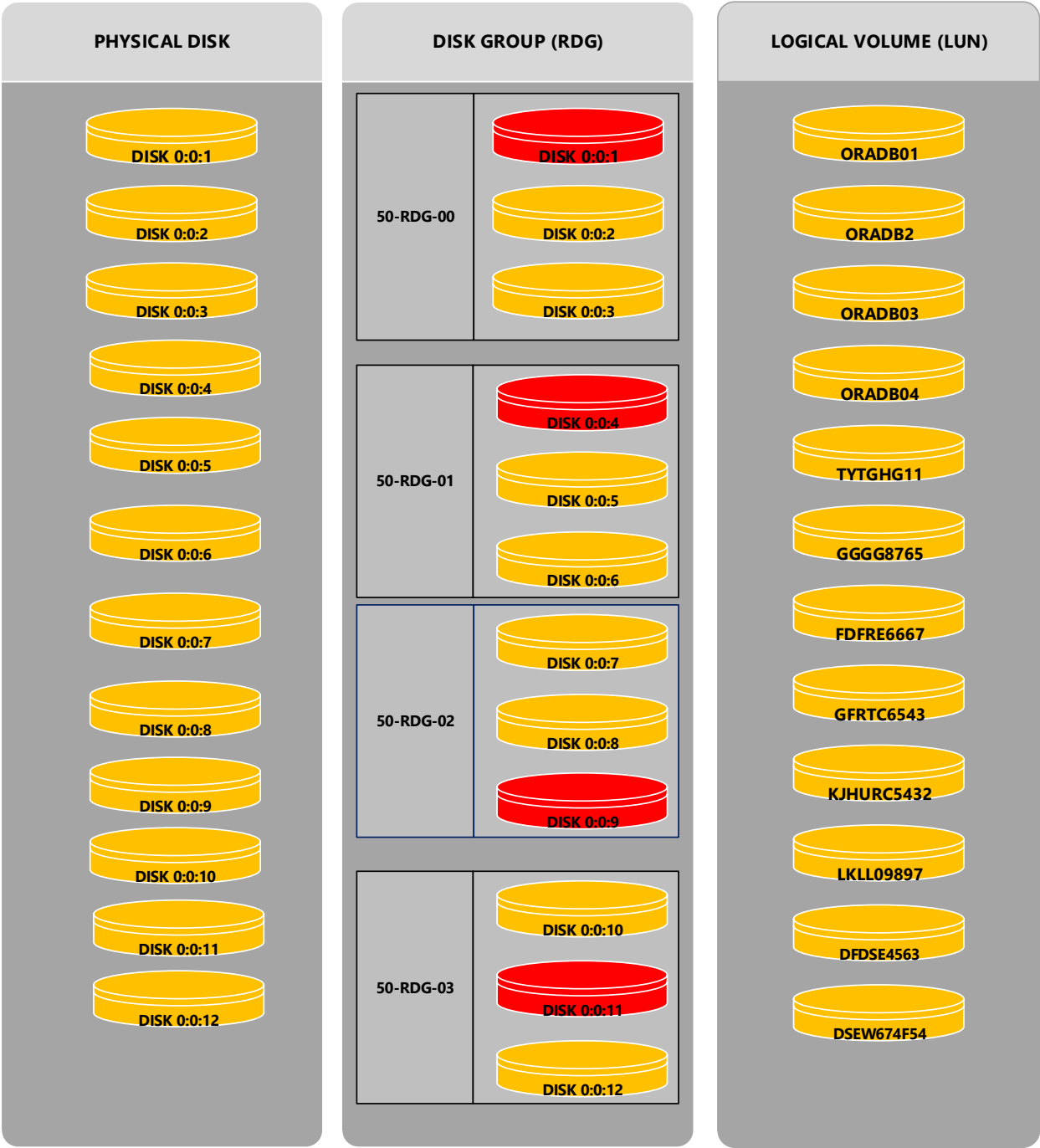
RDG-5/50 рекомендуется использовать для хранения данных низкой критичности, с высокими требованиями к производительности на чтение и средними требованиями к производительности на запись. При этом хорошей практикой является компенсация снижения производительности на запись путем использования SSD дисков в качестве кэша на запись.

Пример 4:



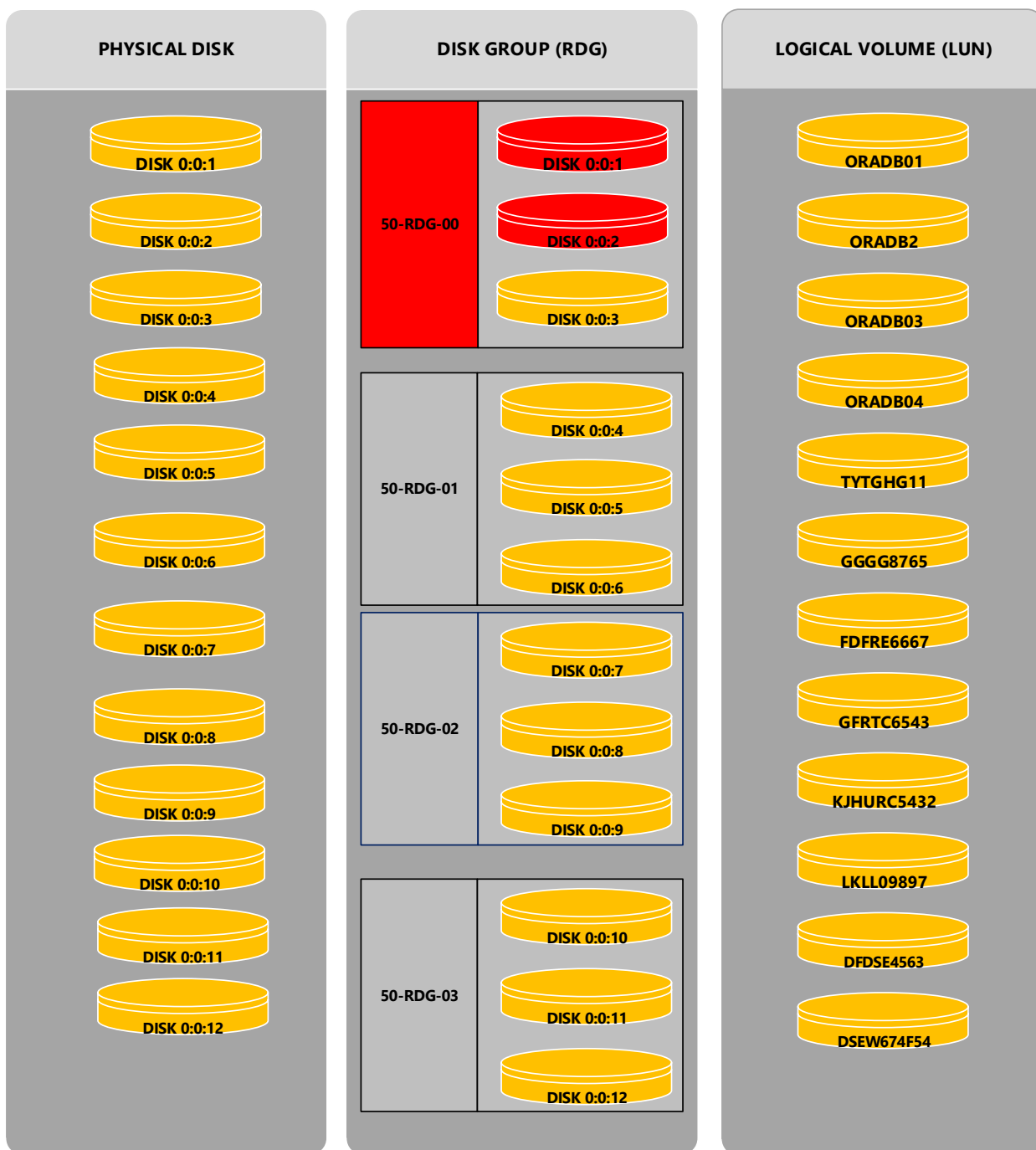
В **Примере-4** показана организация RDG группы уровня RDG-50 (2D+1P), т.е. в каждом виртуальном устройстве 2 диска используется для хранения данных и один для четности. Физические диски объединены в V-DEV.

Пример 5:



В **Примере 5** отражен выход из строя нескольких физических дисков в разных виртуальных V-DEV 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 6:



В **Примере-6** отражен выход из строя 2-х физических дисков в одной V-DEV с типом 50-RDG. В данном примере стало недоступно виртуальное устройство V-DEV – 50-RDG-00. Такая ситуация считается фатальной, данные недоступны для чтения-и записи, следует заменить диски и восстановить данные из резервных копий.

3.3 Организация группы 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 V-DEV для дисковой группы типа RDG-60 – это 2 V-DEV (или один V-DEV для RDG-6). Физические диски объединяются в V-DEV по пять жестких дисков в одном V-DEV или более.

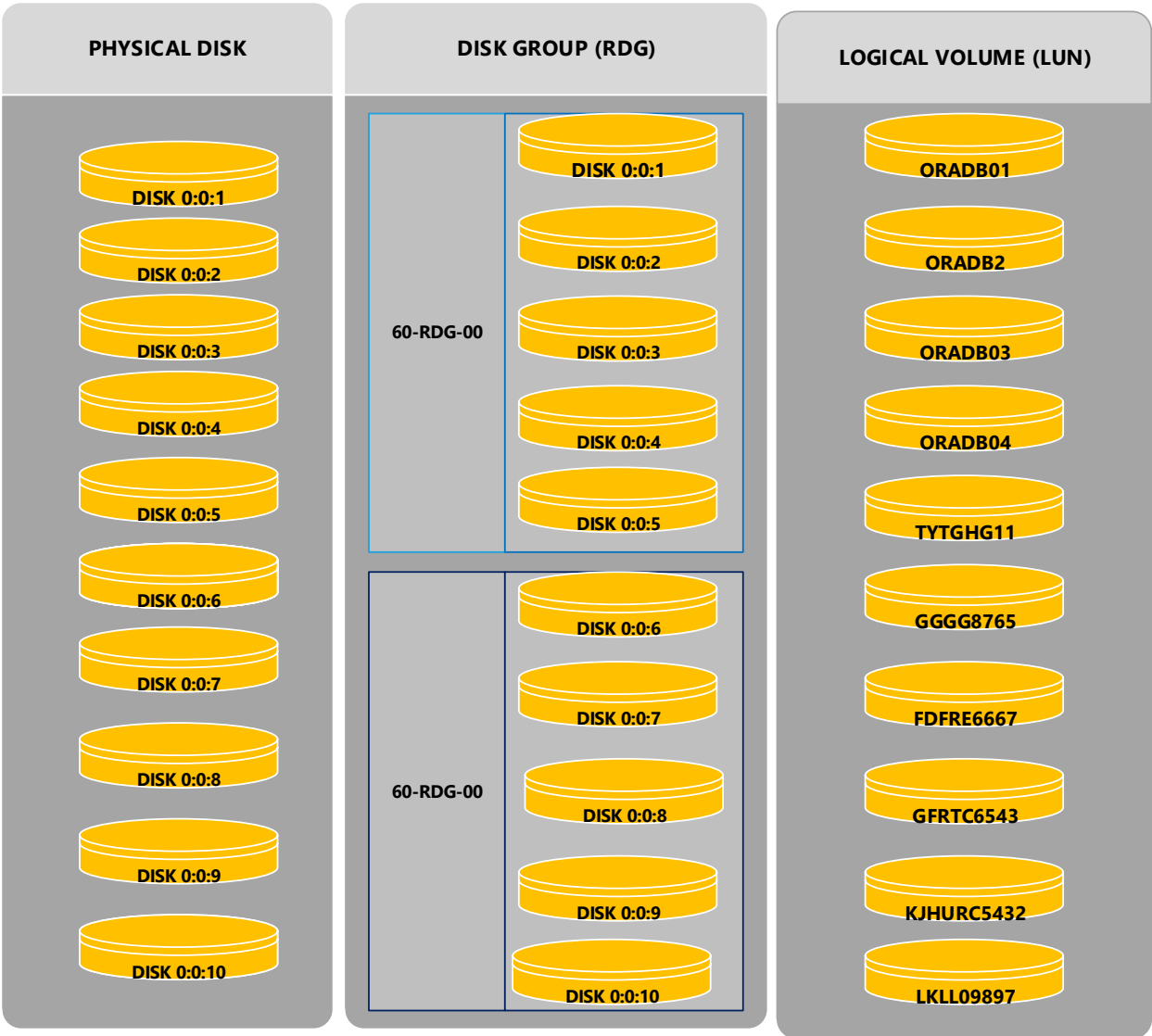
RDG-6/60 имеет следующие варианты конфигурации:

Уровень	Диски для данных в одном виртуальном устройстве	Диски для четности в одном виртуальном устройстве	Производительность (чтение)	Производительность (запись)	Полезная емкость
RAID-6/60 (3D+2P)	3	2	средняя	низкая	59,05%
RAID-6/60 (5D+2P)	5	2			66,10%
RAID-6/60 (7D+2P)	7	2			75,63%

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

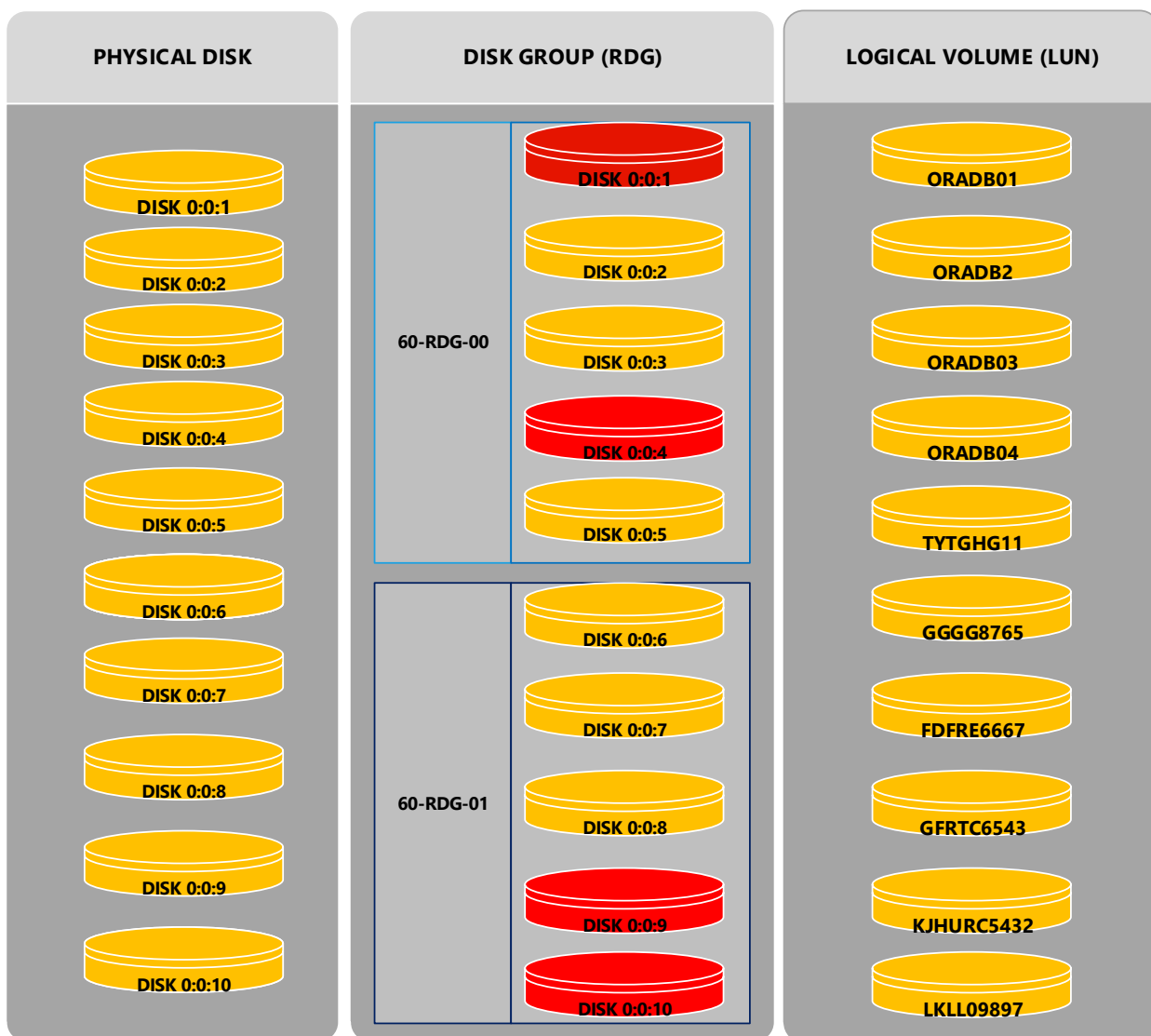
Два диска используются для хранения четности. RDG-60 можно рекомендовать для хранения критичных, но не оперативных данных. Дисковая группа с уровнем RDG-60 допускает выход из строя двух физических дисков в одном V-DEV, выход из строя более двух дисков в одном V-DEV ведет к потере данных.

Пример 7:



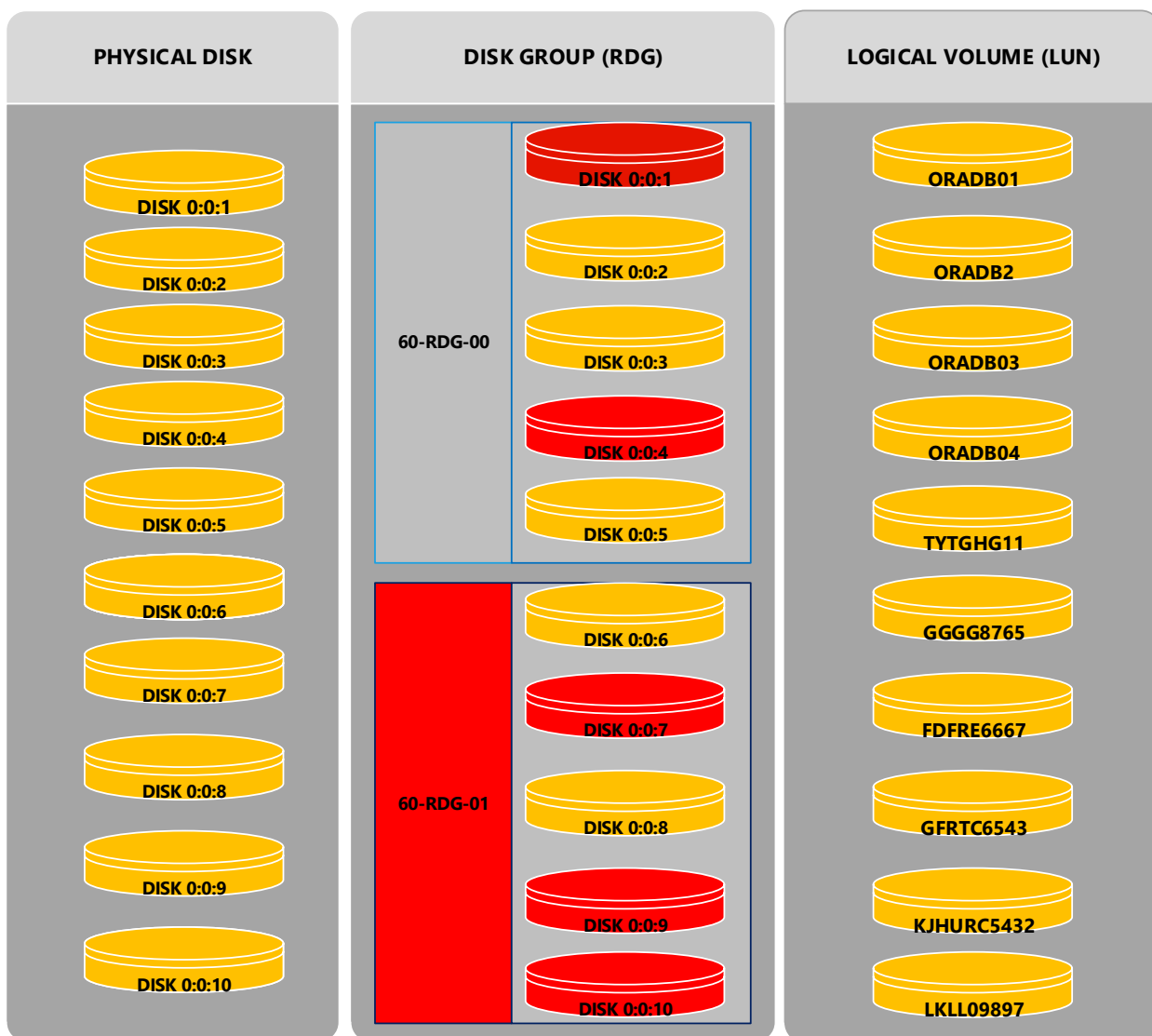
В **Примере-7** показана организация RDG группы уровня 60-RDG. Физические диски объединены в V-DEV по пять дисков в каждом V-DEV.

Пример 8:



В **Примере 8** показан выход из строя нескольких физических дисков – по два в разных V-DEV, 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 грозит потерей всех данных. Следует незамедлительно заменить сбойные диски.

Пример 9:



В **Примере 9** отражен выход из строя нескольких физических дисков в V-DEV 60-RDG-01, 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.

3.4 Организация группы RAID-6P/60P

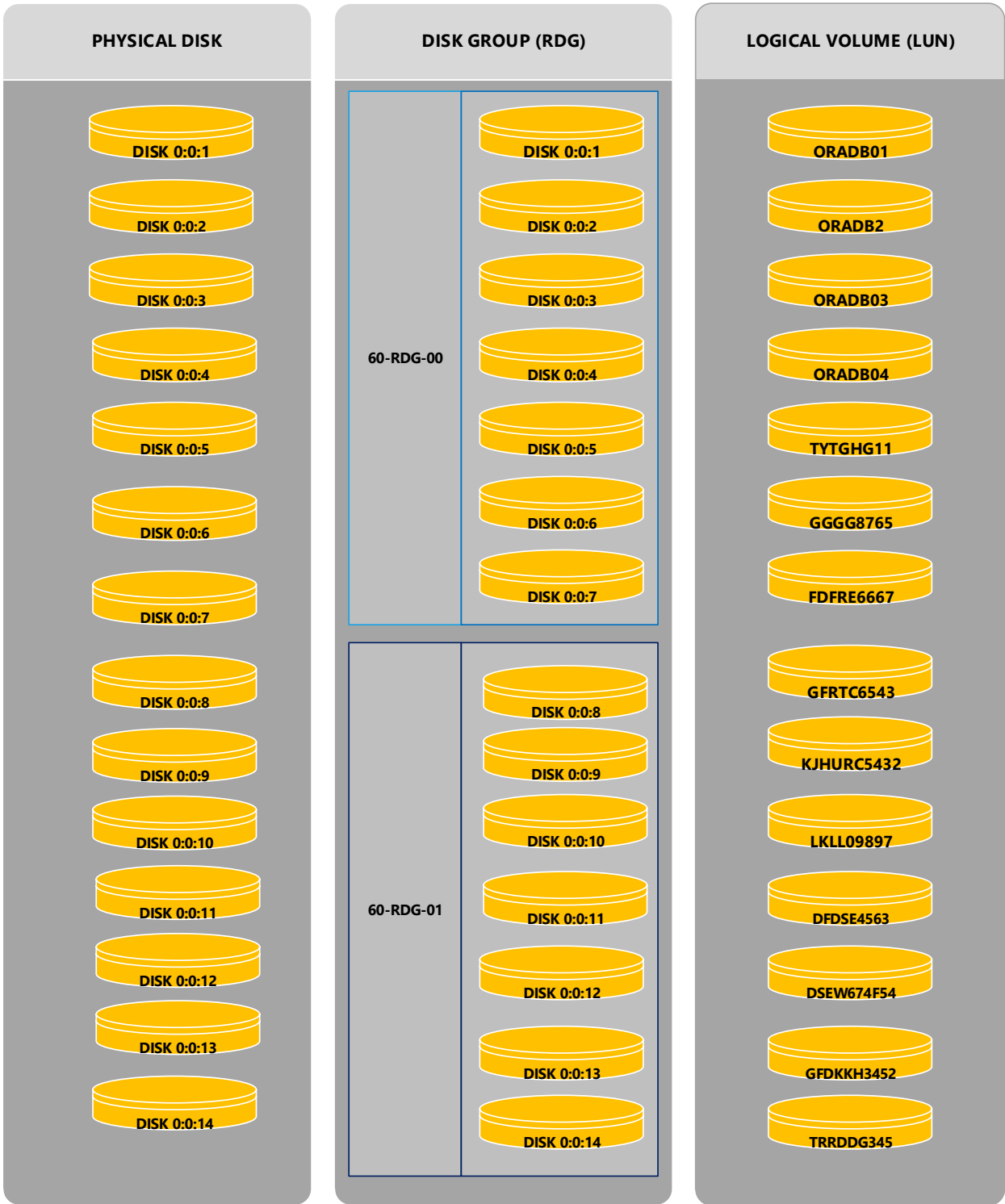
Минимальное количество физических дисков для организации группы типа RDG-6P/60P – это 7 дисков для RDG-6P и 14 физических дисков для конфигурации RDG60P – по семь дисков в каждом V-DEV. Минимальное количество V-DEV для дисковой группы типа RDG-60P – это 2 V-DEV (или один для RDG-6P). Физические диски объединяются в V-DEV по семь или более дисков в одном V-DEV. Три диска используются для хранения четности, четыре или более диска используются для хранения данных. RDG-60P рекомендуется использовать для хранения данных высокой критичности, с высокими требованиями к производительности на чтение и средними требованиями к производительности на запись. Дисковая группа с уровнем RDG-60P допускает выход из строя трех физических дисков в одном V-DEV, выход из строя более трех дисков в одной V-DEV ведет к потере данных.

RDG-6P/60P имеет следующие варианты конфигурации:

Уровень	Диски для данных в одном виртуальном устройстве	Диски для четности в одном виртуальном устройстве	Производительность (чтение)	Производительность (запись)	Полезная емкость
RAID-6P/60P (4D+3P)	4	3	высокая	средняя	56,66%
RAID-6P/60P (6D+3P)	6	3			61,08%
RAID-6P/60P (8D+3P)	8	3			72,69%

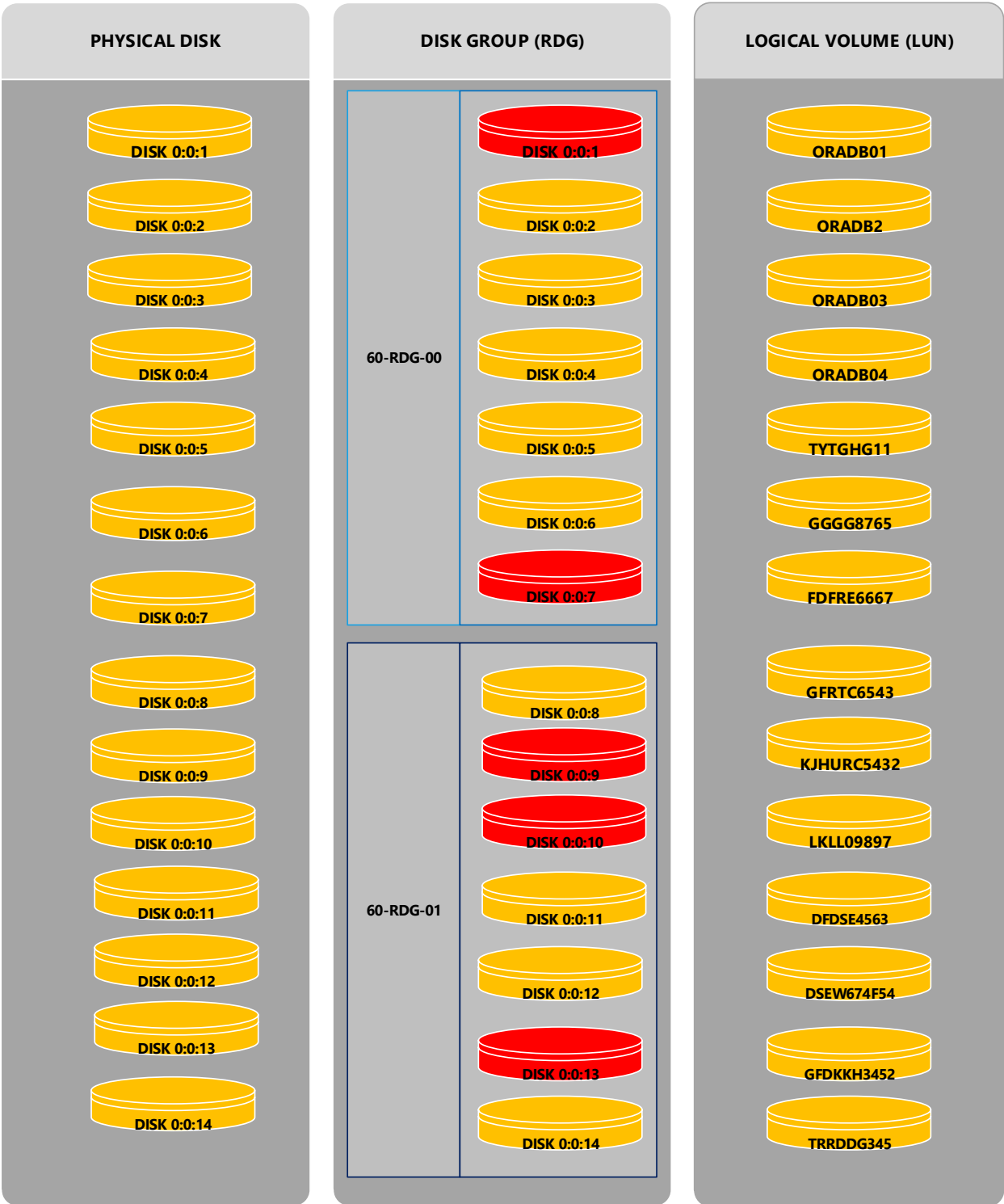
Дисковая группа RDG-60P по организации похожа на гибрид RDG-60 и RDG-50, но имеет более высокую степень надежности чем RDG-50 и лучшую производительность по сравнению с RDG-60 (т.е. по производительности уступает лишь RDG-10). Под контрольные суммы выделяется три жестких диска в одном V-DEV. RDG-60P обеспечивает очень высокий уровень надежности хранения данных и самую высокую степень их консолидации.

Пример 10:



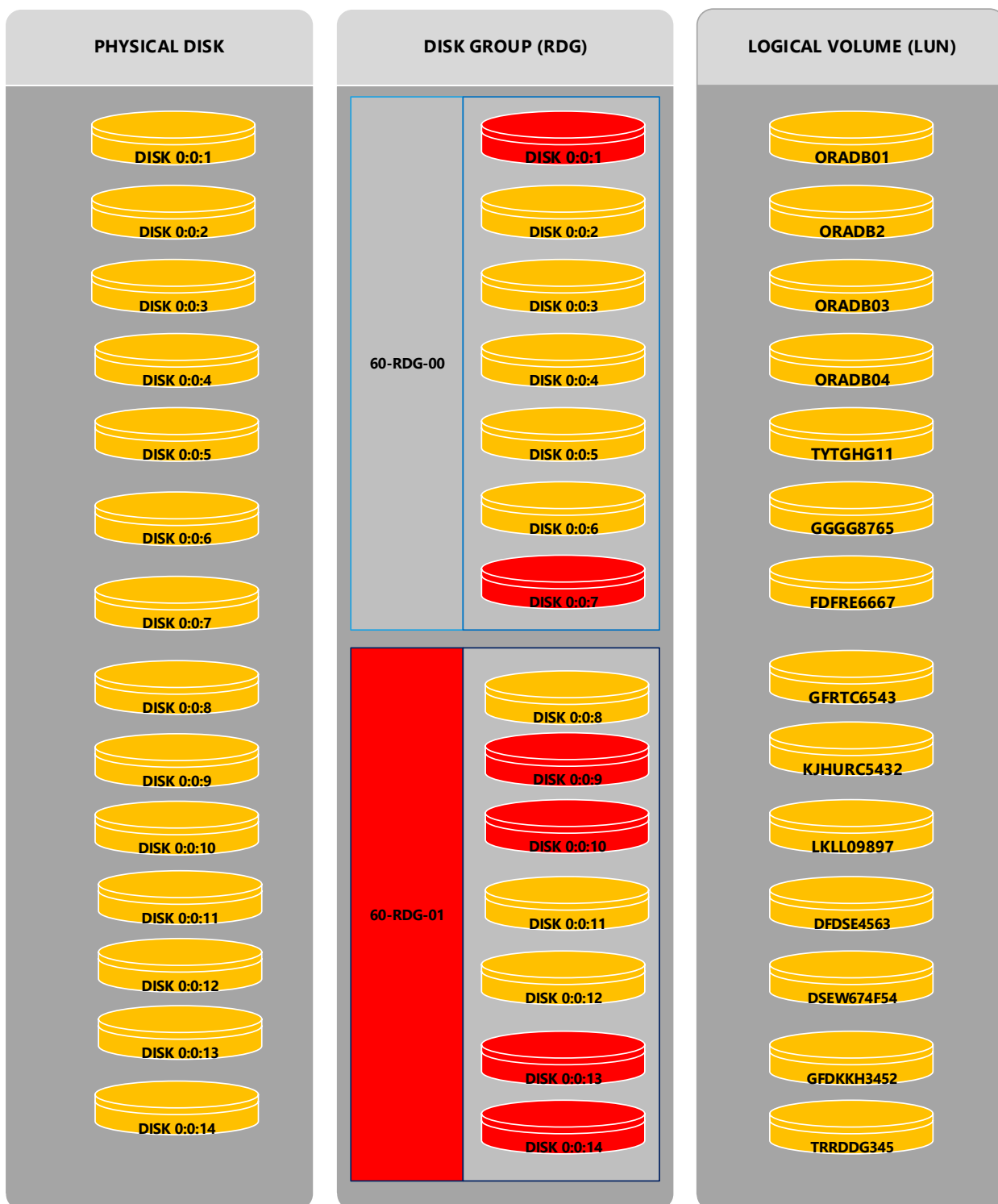
В **Примере 10** показана организация RDG группы уровня 60P-RDG. Физические диски объединены в два V-DEV устройства по 7 физических дисков в каждом для обеспечения максимальной производительности и отказоустойчивости.

Пример 11:



В **Примере 11** отражен выход из строя нескольких физических дисков в разных V-DEV 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 грозит потерей всех данных. Следует незамедлительно заменить сбойные диски.

Пример 12:



В **Примере 12** отражен выход из строя нескольких физических дисков в V-DEV 60P-RDG-01 с типом 60P-RDG. Такая ситуация считается фатальной, так как вышло из строя одновременно четыре жестких диска (при возможных трех) в одном V-DEV, данные недоступны для чтения и записи, следует заменить диски и восстановить данные из резервных копий.

3.5 Организация группы – SPARE диски

SPARE диски или диски горячей замены - это тип дисков, обеспечивающих максимальную доступность RDG группы.

В обычном режиме жесткий диск SPARE не участвует в операциях ввода-вывода. При выходе из строя жесткого диска в RDG группе, жесткий диск с типом SPARE включается в работу вместо вышедшего из строя жесткого диска. AERODISK рекомендует при возможности использовать SPARE диски, вне зависимости от используемого типа RDG.

Особенности использования дисков SPARE:

- Один диск может быть использован любой RDG группой, состоящей из того же типа дисков (Global hot spare);
- Чем выше уровень RDG, тем меньше необходимость в SPARE дисках;
- Чем больше физических дисков используется в дисковой группе (RDG), тем больше дисков SPARE рекомендуется использовать.

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. После замены вышедшего из строя диска, диск автоматически вернется в изначальный статус SPARE.

3.6 Организация группы – SSD диски

Твердотельные диски обладают гораздо более высокой скоростью доступа к данным и меньшими задержками при работе по сравнению с жёсткими дисками. Для того что бы СХД AERODISK использовала весь потенциал, заложенный в дисках SSD, следует продумать их корректное использование.

СХД AERODISK поддерживает RDG группы, созданные только из SSD дисков, но лучшей практикой является использование гибридных RDG групп, когда в рамках одной RDG совмещают 2 типа дисков SAS (NL-SAS), SATA для хранения данных и SSD для адаптации ввода-вывода (SSD-кэш и/или Online-tiering). Такой шаг является наиболее корректным, поскольку SSD меньше по объему и значительно выше по стоимости, если сравнивать их с SAS\NLSAS и SATA дисками. Учитывая этот факт, использование в данном качестве SSD в поможет 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:

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

Во избежание потери данных диски, которые имеют тип READ-WRITE-SSD и Online-tiering, необходимо объединять в зеркальную пару. Для дисков READ-SSD такая необходимость опциональна.

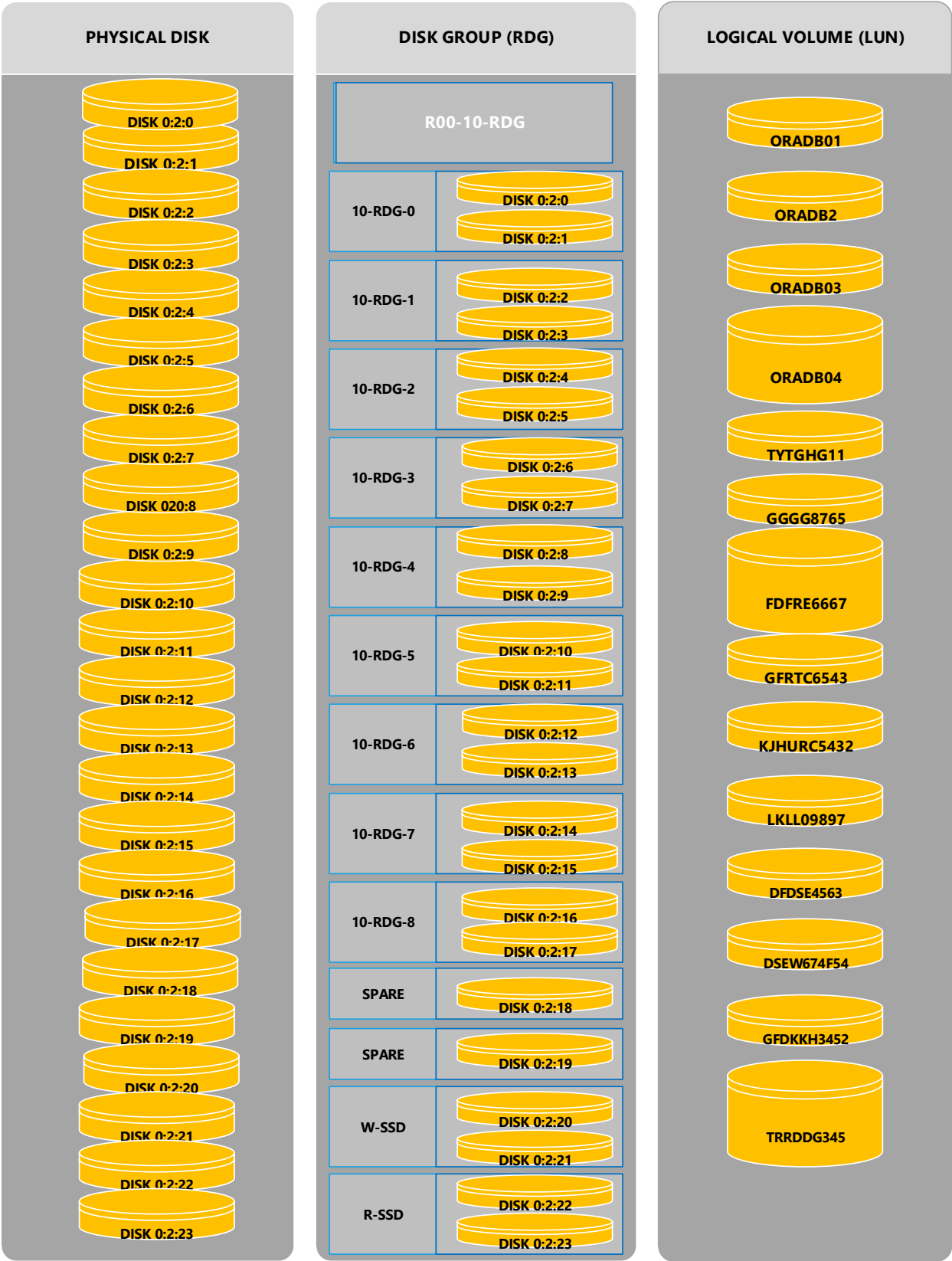
Пример 14:



Тип диска	RDG Name	V-DEV	Тип RDG	Slot
SAS 10k 900GB	R00	10-RDG-00	10-RDG	0:2:0
SAS 10k 900GB	R00	10-RDG-00	10-RDG	0:2:1
SAS 10k 900GB	R00	10-RDG-01	10-RDG	0:2:2
SAS 10k 900GB	R00	10-RDG-01	10-RDG	0:2:3
SAS 10k 900GB	R00	10-RDG-02	10-RDG	0:2:4
SAS 10k 900GB	R00	10-RDG-02	10-RDG	0:2:5
SAS 10k 900GB	R00	10-RDG-03	10-RDG	0:2:6
SAS 10k 900GB	R00	10-RDG-03	10-RDG	0:2:7
SAS 10k 900GB	R00	10-RDG-04	10-RDG	0:2:8
SAS 10k 900GB	R00	10-RDG-04	10-RDG	0:2:9
SAS 10k 900GB	R00	10-RDG-05	10-RDG	0:2:10
SAS 10k 900GB	R00	10-RDG-05	10-RDG	0:2:11
SAS 10k 900GB	R00	10-RDG-06	10-RDG	0:2:12
SAS 10k 900GB	R00	10-RDG-06	10-RDG	0:2:13
SAS 10k 900GB	R00	10-RDG-07	10-RDG	0:2:14
SAS 10k 900GB	R00	10-RDG-07	10-RDG	0:2:15
SAS 10k 900GB	R00	10-RDG-08	10-RDG	0:2:16
SAS 10k 900GB	R00	10-RDG-08	10-RDG	0:2:17
SAS 10k 900GB	R00	SPARE-00	SPARE Disk	0:2:18
SAS 10k 900GB	R00	SPARE-01	SPARE Disk	0:2:19
SAS SSD 400GB	R00	WRITE-SSD Mirror-0	WRITE SSD	0:2:20
SAS SSD 400GB	R00	WRITE-SSD Mirror-0	WRITE SSD	0:2:21
SAS SSD 400GB	R00	READ-SSD Mirror-0	READ SSD	0:2:22
SAS SSD 400GB	R00	READ-SSD Mirror-0	READ SSD	0:2:23

В **Примере 14** – показан один из вариантов рекомендуемой разметки СХД AERODISK ENGINE 2100 с установленными 20 дисками SAS и четырьмя дисками SSD. Как видно из примера 14, СХД AERODISK Engine сконфигурирована оптимальным образом.

Пример 15:

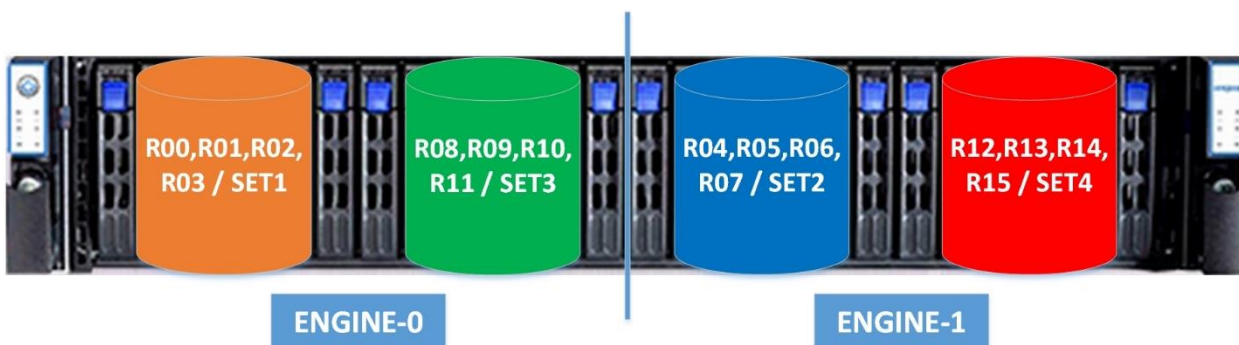


В **Примере 15** показана логическая схема разметки RDG-10 + SPARE + RW SSD. Из схемы видно, что диски с 0 по 17 слот задействованы под организацию RDG группы с типом RDG-10, диски в слотах 18,19

используются как SPARE устройства, SSD диски в слотах с 20 по 23 используются для организации адаптированного ввода-вывода.

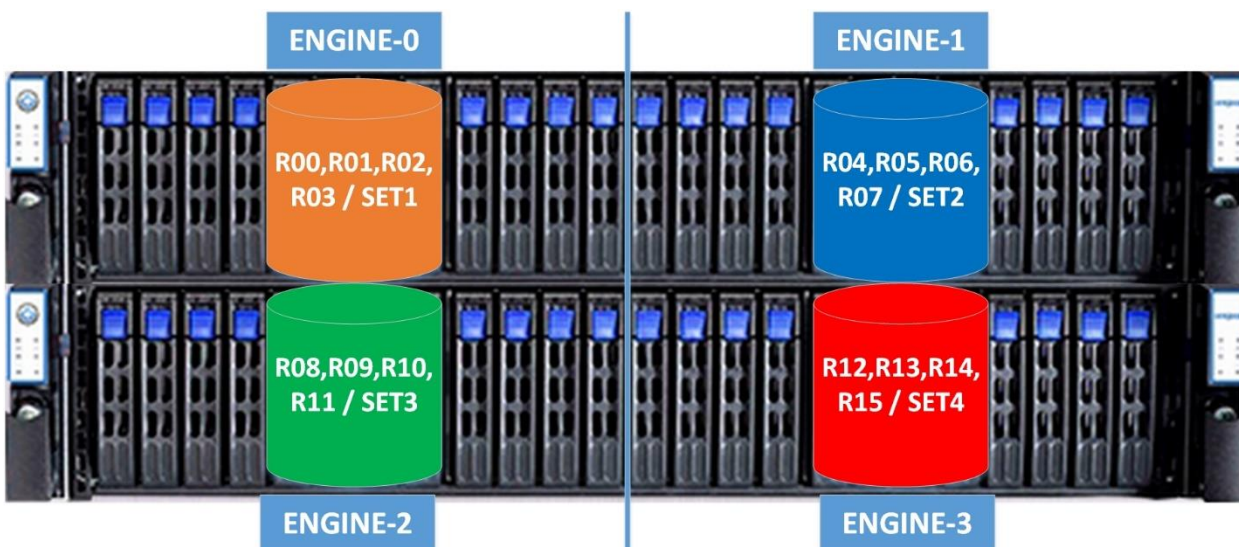
4 Распределение групп между контроллерами

Пример корректной конфигурации для СХД AERODISK Engine N2



Как видно из примера, RDG-группы равномерно распределены по двум системным контроллерам СХД. Системные ресурсы СХД используются сбалансированно, точки отказа отсутствуют.

Пример корректной конфигурации для СХД AERODISK Engine N4.



Как видно из примера, RDG-группы равномерно распределены по четырем системным контроллерам СХД. Системные ресурсы СХД используются сбалансированно, точки отказа отсутствуют.

Распределение групп по системным контроллерам и типам используемых дисков для СХД AERODISK ENGINE (N2-N4).

Приоритет контроллера	PLATINUM (SSD)	GOLD (SSD+SAS)	SILVER (SAS)	BRONZE (NL-SAS)
CL0,1,2,3	R00	R01	R02	R03
CL1,2,3,0	R04	R05	R06	R07
CL2,3,0,1	R08	R09	R10	R11
CL3,0,1,2	R12	R13	R14	R15

Распределение групп по системным контроллерам и типам используемых дисков для СХД AERODISK Hi-Engine (N8).

Приоритет контроллера	PLATINUM (SSD)	GOLD (SSD+SAS)	SILVER (SAS)	BRONZE (NL-SAS)
CL0,1,2,3,4,5,6,7	R00	R01	R02	R03
CL1,2,3,4,5,6,7,0	R04	R05	R06	R07
CL2,3,4,5,6,7,0,1	R08	R09	R10	R11
CL3,4,5,6,7,0,1,2	R12	R13	R14	R15
CL4,5,6,7,0,1,2,3	R16	R17	R18	R19
CL5,6,7,0,1,2,3,4	R20	R21	R22	R23
CL6,7,0,1,2,3,4,5	R24	R25	R26	R27
CL7,0,1,2,3,4,5,6	R28	R29	R30	R31