WHITE PAPER

# СОЗДАНИЕ ОТКАЗОУСТОЙЧИВОЙ ВИРТУАЛЬНОЙ ИНФРАСТРУКТУРЫ С ИСПОЛЬЗОВАНИЕМ РЕПЛИКАЦИИ СХД YADRO TATLIN.UNIFIED И СЕРВИСА АППАРАТНОЙ РЕПЛИКАЦИИ ZVIRT

# О WHITE PAPER

Данный документ описывает архитектурные принципы и подходы, а также рекомендации по их применению при построении корпоративных решений на основе продуктов производства компании YADRO.

Документ предназначен для широкой аудитории, но также может включать материалы, требующие специальной профессиональной подготовки, ориентированные на системных архитекторов, администраторов, сервисных инженеров и других ИТ-специалистов, участвующих в планировании, проектировании и внедрении комплексных решений.

! Представленная информация основана на результатах исследований, а также экспертизе специалистов YADRO. Документ носит не предписывающий, а информационный характер, целью которого является информирование получателей документа о предмете произведенного исследования и о его результатах. При этом конкретные результаты и эффективность конечных решений зависят от множества факторов, которые могут различаться в зависимости от особенностей инфраструктуры и условий эксплуатации. Ни при каких обстоятельствах YADRO не несет ответственности за последствия применения получателем документа рекомендаций, инструкций, подходов и решений, указанных в документе.

Запрещается копировать, воспроизводить, изменять, публиковать, передавать, распространять, продавать информацию (часть информации), содержащуюся в документе, или создавать производные работы на основе содержания документа без предварительного письменного разрешения YADRO.

Любая информация или утверждения относительно возможностей, емкости, производительности или других характеристик продукции, указанной в документе, предоставляются исключительно на условиях «как есть» и приведена исключительно в целях описания произведенного исследования. YADRO оставляет за собой право вносить изменения и обновлять информацию в документе без уведомления.

«YADRO» является зарегистрированным товарным знаком ООО «КНС ГРУПП». Другие названия продуктов и компаний, упомянутые здесь, могут быть товарными знаками или торговыми наименованиями их соответствующих владельцев.

Актуальные версии документации публикуются на информационном ресурсе YADRO по адресу service.yadro.com.

# ОГЛАВЛЕНИЕ


1. ВВЕДЕНИЕ ..... 5
   - 1.1. Описание технологии репликации системы хранения YADRO TATLIN.UNIFIED ..... 5
   - 1.2. Поддержка репликации СХД YADRO TATLIN.UNIFIED и платформы виртуализации zVirt в сервисе аппаратной репликации ..... 5
   - 1.3. Основные возможности и преимущества ..... 5
2. АРХИТЕКТУРА И КОМПОНЕНТЫ ..... 6
   - 2.1. Требования к оборудованию и ПО ..... 6
   - 2.2. Архитектура ..... 6
3. НАСТРОЙКА И УСТАНОВКА ..... 6
   - 3.1. Установка оборудования и ПО ..... 6
   - 3.2. Инсталляционный чек-лист ..... 7
   - 3.3. Конфигурирование томов и репликации СХД YADRO TATLIN.UNIFIED ..... 8
     - 3.3.1. Создание дисковых пулов ..... 8
     - 3.3.2. Конфигурация серверных подключений ..... 9
     - 3.3.3. Создание ресурсов (логических томов) ..... 11
     - 3.3.4. Создание репликационной связи между массивами ..... 14
     - 3.3.5. Создание репликационных пар между массивами ..... 15
   - 3.4. Конфигурирование дисковых ресурсов СХД YADRO TATLIN.UNIFIED на серверах zVirt ..... 17
     - 3.4.1. Настройка блочных устройств СХД YADRO TATLIN.UNIFIED на хостах zVirt ..... 17
     - 3.4.2. Конфигурирование политики ввода-вывода (multipath) на уровне гипервизора на хостах zVirt для томов СХД YADRO TATLIN.UNIFIED ..... 19
     - 3.4.3. Подключение домена хранения к ОЦОД ..... 20
   - 3.5. Конфигурирование сервиса аппаратной репликации zVirt ..... 21
     - 3.5.1. Конфигурирование сервиса репликации zVirt с аппаратной репликацией СХД TATLIN.UNIFIED ..... 21
     - 3.5.2. Подключение СХД к сервису аппаратной репликации zVirt ..... 23
   - 3.6. Конфигурирование планов аварийного восстановления ..... 26
     - 3.6.1. Конфигурирование маппинга ресурсов ..... 26
     - 3.6.2. Создание плана восстановления ..... 28
4. УПРАВЛЕНИЕ РЕПЛИКАЦИЕЙ ..... 33
   - 4.1. Мониторинг и управление репликацией ..... 33
   - 4.2. Переключение направления репликации ..... 35
5. АВАРИЙНОЕ И ПЛАНОВОЕ ВОССТАНОВЛЕНИЕ ..... 36
   - 5.1. Плановое переключение сервиса ..... 36
   - 5.2. Обратное перемещение после планового переключения сервиса ..... 38
   - 5.3. Аварийное восстановление ..... 39
   - 5.4. Возврат после аварийного переключения ..... 42
6. УСТРАНЕНИЕ НЕИСПРАВНОСТЕЙ ..... 43
   - 6.1. Распространенные проблемы и решения ..... 43
   - 6.2. Логи и диагностика ..... 45
7. МЕТРИКИ ПРОИЗВОДИТЕЛЬНОСТИ ..... 45
   - 7.1. Описание тестовой инфраструктуры ..... 45
   - 7.2. Результаты тестирования производительности I/O инфраструктуры на 4 ВМ, с одним дисковым устройством и профилем: 70% чтения и 30% записи. ..... 46
   - 7.3. Сравнение метрик производительности I/O инфраструктуры на 4 ВМ и профилем 100% записи ..... 47
8. ЗАКЛЮЧЕНИЕ ..... 47
   - 8.1. Рекомендации по эксплуатации ..... 47
   - 8.2. Дополнительные ресурсы ..... 48
9. ПРИЛОЖЕНИЯ ..... 48
   - 9.1. Сравнение метрик для 4 ВМ (профилем 70% чтения и 30% записи) ..... 48
     - 9.1.1. Сравнение метрик для 4 ВМ без репликации (профилем 70% чтения и 30% записи) ..... 49

9.1.2. Сравнение метрик для 4 ВМ с репликацией (профилем 70% чтения и 30% записи) ..... 57
9.2. Сравнение метрик для 4 ВМ (профилем 100% записи) ..... 64
9.2.1. Сравнение метрик для 4 ВМ с репликацией (100% записи) ..... 65
9.2.2. Сравнение метрик для 4 ВМ без репликации (100% записи) ..... 70

# ВВЕДЕНИЕ

## Описание технологии репликации системы хранения YADRO TATLIN.UNIFIED

Репликация на уровне системы хранения данных YADRO TATLIN.UNIFIED (Array-based Replication) — это технология синхронного копирования данных между физически разнесенными системами хранения (дисковыми массивами), осуществляемая на аппаратном уровне встроенными средствами СХД. Технология репликации позволяет защитить приложения от потери данных и быстро восстановить работу в случае сбоев.

## Поддержка репликации СХД YADRO TATLIN.UNIFIED и платформы виртуализации zVirt в сервисе аппаратной репликации

### Матрица совместимости версий платформы виртуализации zVirt + DR и СХД YADRO TATLIN.UNIFIED:

| Версия zVirt | Поколение массивов | Версия микрокода | Протоколы доступа |
|---|---|---|---|
| zVirt 4.3 | Gen1/Gen2 | 3.1, 3.2 | iSCSI |
| zVirt 4.4 | Gen1/Gen2 | 3.1, 3.2 | iSCSI/FibreChannel |
| zVirt 4.5 | Gen1/Gen2 | 3.1, 3.2 | iSCSI/FibreChannel |
| zVirt 5.0 | Gen1/Gen2 | 3.1, 3.2 | iSCSI/FibreChannel |

## Основные возможности и преимущества

Репликация данных с помощью встроенного функционала СХД позволяет создавать и поддерживать полную синхронную копию продуктивных данных основной СХД на удаленной резервной СХД. Использование синхронной копии позволяет защитить критически важные данные пользователей при аварии на основной СХД или сбоях в инфраструктуре ЦОД, а также обезопасить работу приложений при проведении сервисных работ, требующих перезагрузки основной СХД.

При инициализации репликации запускается процесс фонового копирования продуктивных данных с основной СХД на резервную. Каждая новая операция записи, приходящая от инициатора на реплицируемые тома основной СХД, попадает в кэш основной СХД и отправляется через выделенные порты и каналы репликации на резервную СХД. После получения основной СХД подтверждения записи данных на резервной СХД инициатору отсылается подтверждение о закрытии записи. Данные записываются на диски основной и резервной СХД согласно алгоритмам работы кэша. Описанный процесс называется синхронной репликацией и гарантирует создание точной копии данных на резервной СХД.

Использование механизма синхронной репликации в сочетании с Disaster Recovery функционалом серверных ОС позволяет перемещать приложения, виртуальную инфраструктуру и базы данных между СХД как в локальном, так и в удаленном ЦОД.

# АРХИТЕКТУРА И КОМПОНЕНТЫ

## Требования к оборудованию и ПО

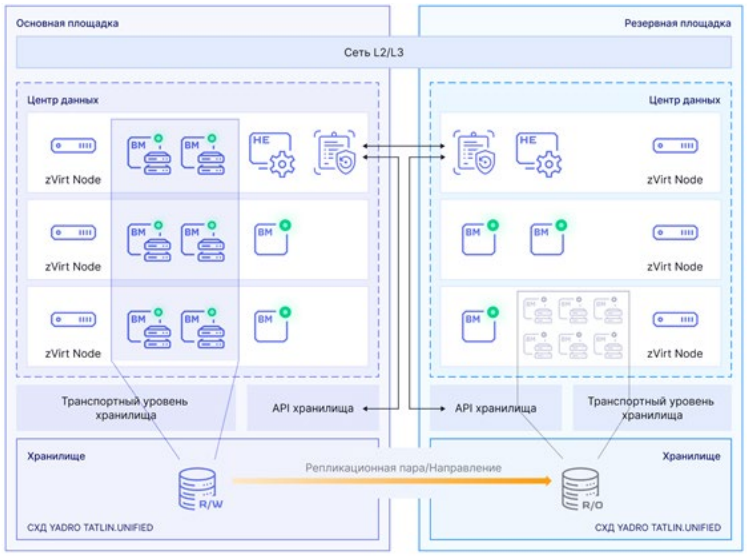
Для построения отказоустойчивой виртуальной инфраструктуры с использованием репликации СХД YADRO TATLIN.UNIFIED и модуля аварийного восстановления zVirt DR требуется:

**Аппаратный уровень**
- 2 СХД YADRO TATLIN.UNIFIED с поддержкой синхронной репликации;
- сетевое оборудование Ethernet 10/25 Гбит/с для варианта iSCSI / FC SAN для подключения ресурсов по FC;
- не менее 2 серверов с Ethernet-HBA и/или FC-HBA.

**Программный уровень**
- Расширенная редакция zVirt для хостов на обеих площадках.

## Архитектура

На схеме ниже представлена архитектура системы:




# НАСТРОЙКА И УСТАНОВКА

## Установка оборудования и ПО

Установка оборудования на двух площадках выполняется в соответствии с рекомендациями вендора.

**YADRO**:
- Руководство владельца TATLIN.UNIFIED;
- Руководство администратора TATLIN.UNIFIED.

**Orion soft**:
- Начало работы. Развертывание zVirt.

Также необходима готовность инфраструктуры:
- Физические подключения FC SAN/Ethernet;
- Настройки FC-зонирования и VLAN;
- Видимость в коммутаторах портов СХД и серверов по WWN/MAC.

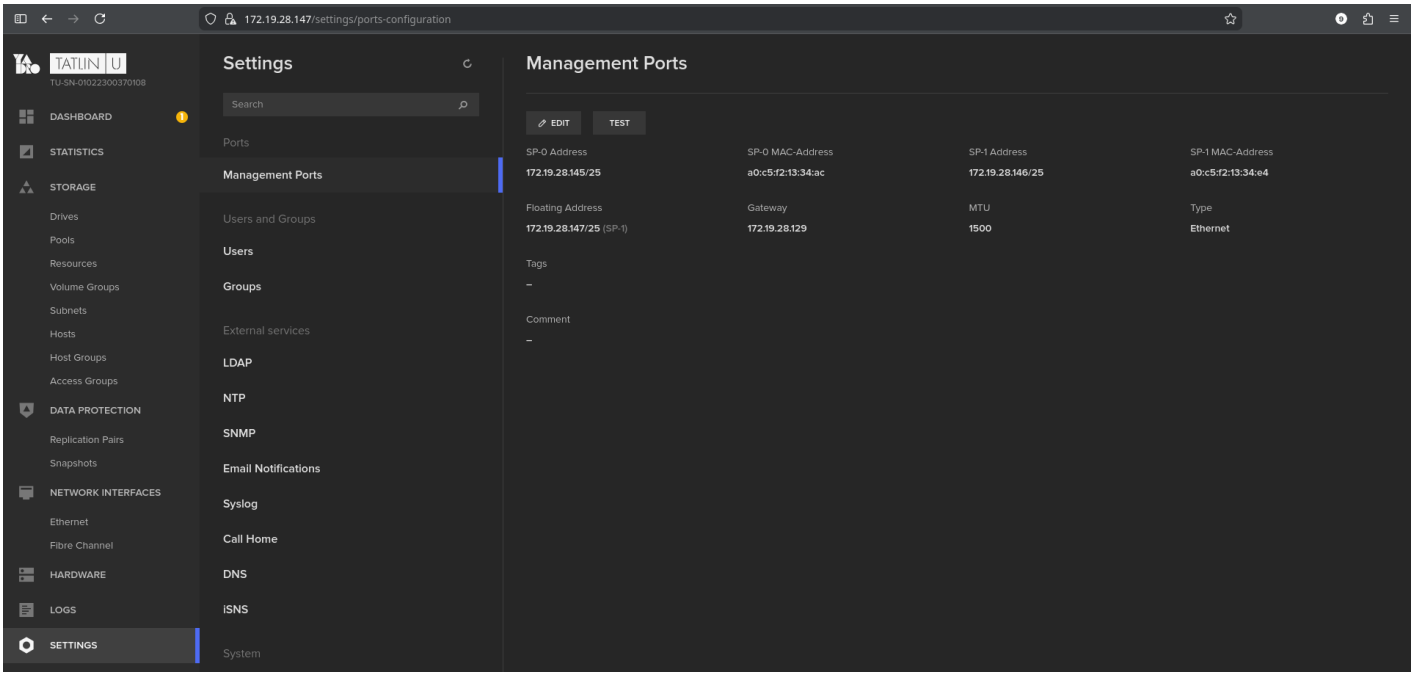
См. WP: Подключение ресурсов к серверной ОС: IP SAN.
См. WP: Подключение ресурсов к серверной ОС: Зонирование FC SAN.

## Инсталляционный чек-лист

### Согласованный список адресов и учетных записей, используемых при тестировании:

| Сайт | Компонент | FQDN | IP |
|---|---|---|---|
| ОЦОД | Менеджер виртуализации | mgmt.zvirt.local | 172.19.28.151 |
| | Хост виртуализации | node1.zvirt.local | 172.19.28.141 |
| | Хост виртуализации | node2.zvirt.local | 172.19.28.140 |
| | TATLIN контроллер SP-0 | | 172.19.28.142 |
| | TATLIN контроллер SP-1 | | 172.19.28.143 |
| | TATLIN контроллер (vIP) | | 172.19.28.144 |
| РЦОД | Менеджер виртуализации | mgmt-dr.zvirt.local | 172.19.28.152 |
| | Хост виртуализации | node3.zvirt.local | 172.19.28.139 |
| | Хост виртуализации | node4.zvirt.local | 172.19.28.138 |
| | TATLIN контроллер SP-0 | | 172.19.28.145 |
| | TATLIN контроллер SP-1 | | 172.19.28.146 |
| | TATLIN контроллер (vIP) | | 172.19.28.147 |

Перед началом конфигурирования СХД убедитесь, что заданные адреса контроллеров доступны в веб-интерфейсе TATLIN на вкладке **Settings** — **Management Ports**.



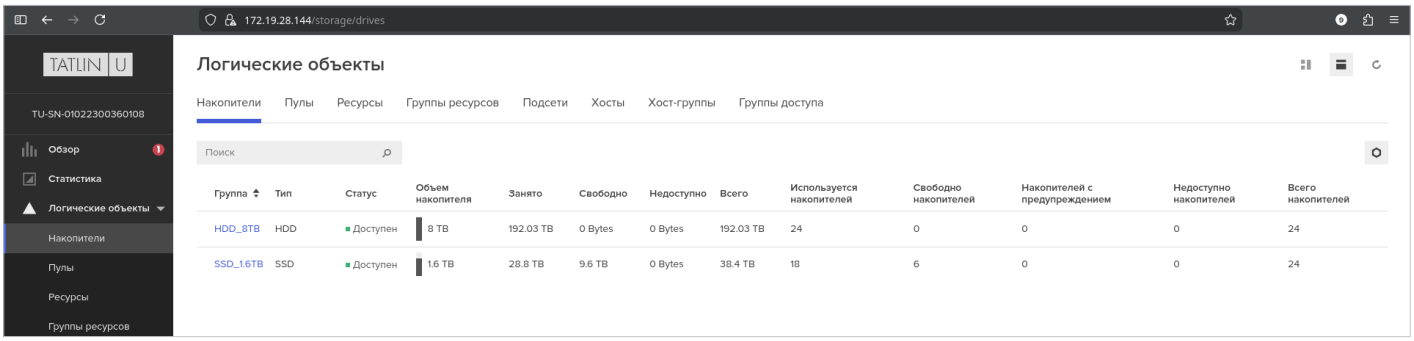
## Конфигурирование томов и репликации СХД YADRO TATLIN.UNIFIED

Для создания тестовой конфигурации ресурсов СХД выполните следующие действия:

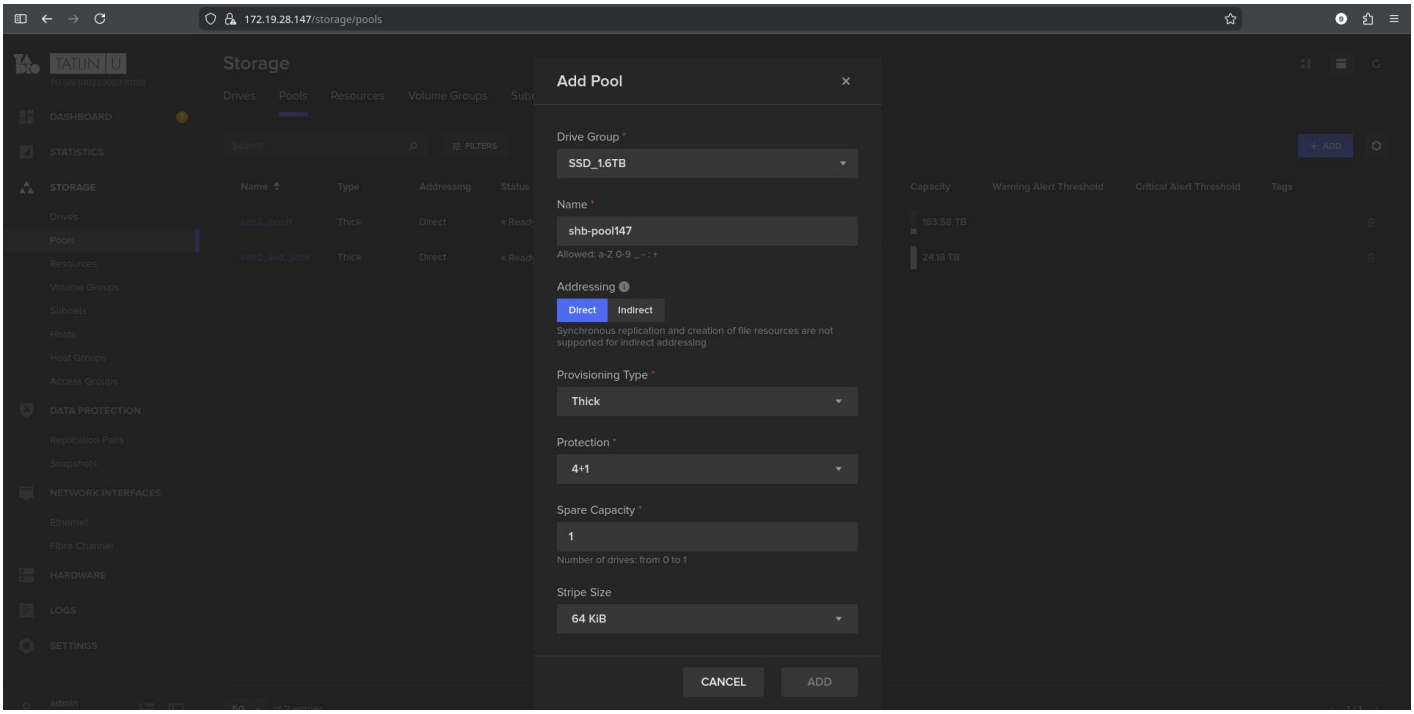
1. Создайте дисковые пулы;
2. Настройте конфигурацию серверных подключений;
3. Создайте ресурсы (логические тома);
4. Подключите хосты;
5. Создайте репликационные связи между массивами;
6. Создайте репликационные пары между массивами.

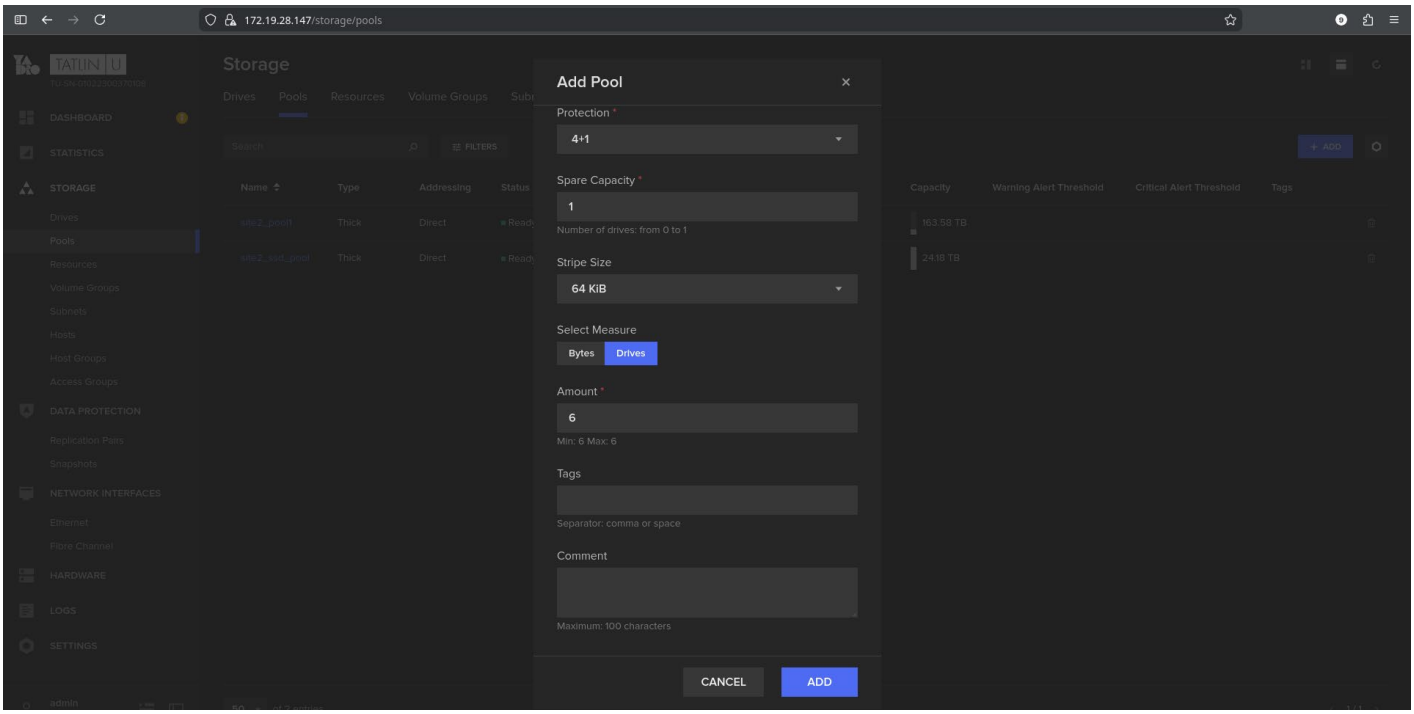
### Создание дисковых пулов

1. Подключитесь к веб-интерфейсу СХД и перейдите на вкладку **Логические объекты** — **Накопители**.
2. Определите, на дисках какой технологии будет создаваться пул.

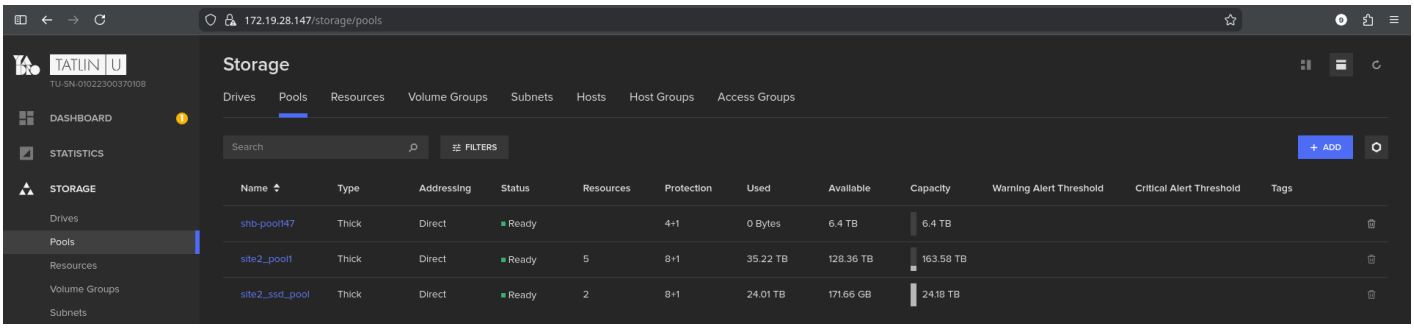



3. Перейдите на вкладку **Pools** и нажмите **ADD**, указав:
   - Тип дисков для пула;
   - Тип адресации пула: Direct;
   - Тип представления Thick: полное использование адресованного пространства;
   - Схему защиты в зависимости от требований к отказоустойчивости и производительности;
   - Объем пула (в ГБ/ТБ или количестве физических дисков);
   - Размер страйпа.






В системе уже были созданы пулы, поэтому пул `shb-pool147` отображается в списке.




Операция по созданию пулов выполняется на обеих СХД, на которых планируется конфигурирование репликации. Конфигурация схемы защиты пулов должна быть идентичной для обеих СХД.

В описываемой конфигурации на второй СХД выполнено создание аналогичного пула с именем `shb-pool144`.

## Конфигурация серверных подключений

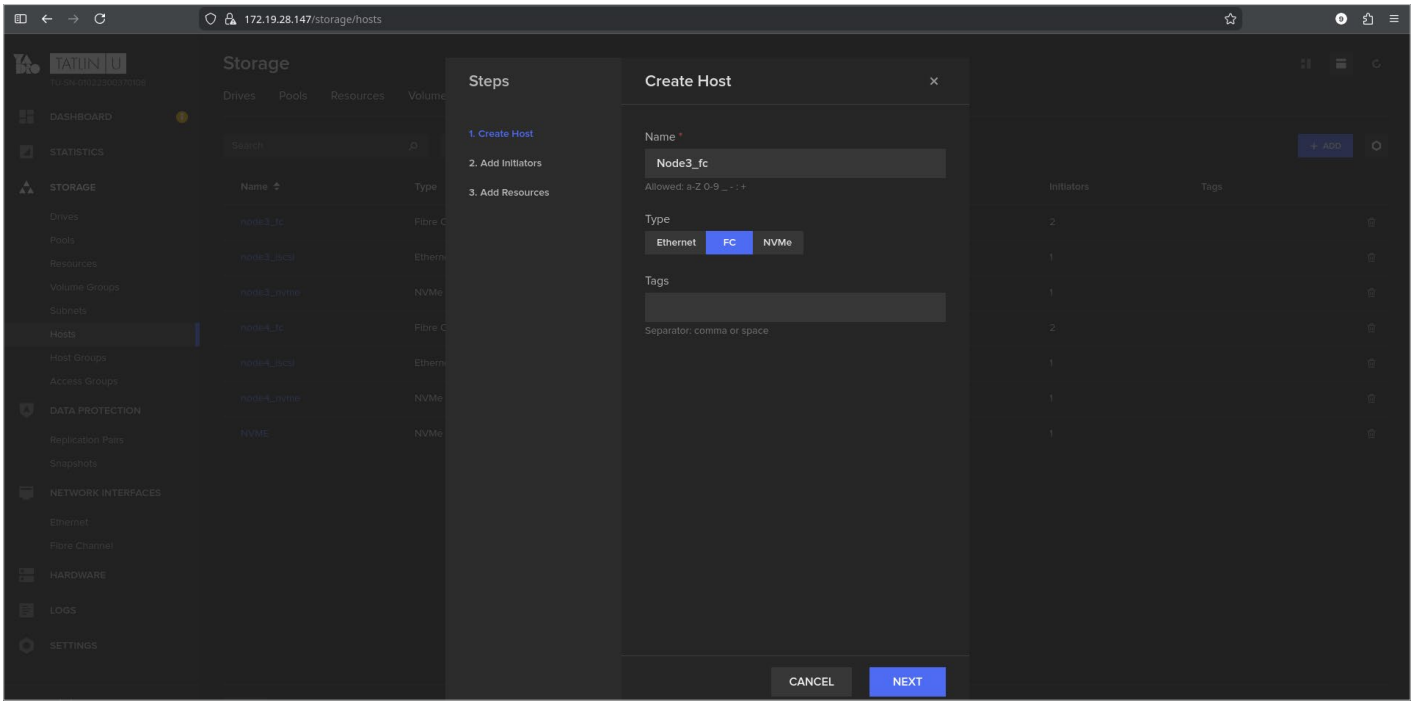
Для предоставления доступа серверов к ресурсам СХД необходимо установить физическое подключение сервера к СХД и зарегистрировать эти подключения на СХД. Сервер регистрируется на СХД в зависимости от выбранной топологии подключения (iSCSI/FC/NVME). При наличии физического подключения сервера к СХД регистрируется по имени IQN (iSCSI Qualified Name) или WWN (World Wide Name). Регистрация IQN/WWN необходима для дальнейших шагов по презентации ресурсов хостам.

1. В разделе **Hosts** нажмите **ADD**.
2. В окне **Create Host** назначьте серверу имя `node3_fc`. Выберите тип подключения FC.

> **!** При регистрации рекомендуется использовать имя хоста, совпадающее с именем хоста zVirt, для точности и удобства идентификации подключений.

При необходимости одновременно выдайте одному физическому серверу тома FC и iSCSI. Для этого создайте два разных хоста, отличающихся по типу подключения (FC/iSCSI).

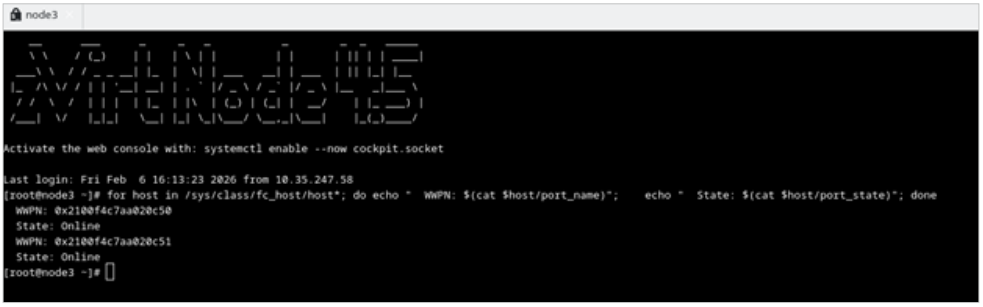
3. Нажмите **NEXT**.



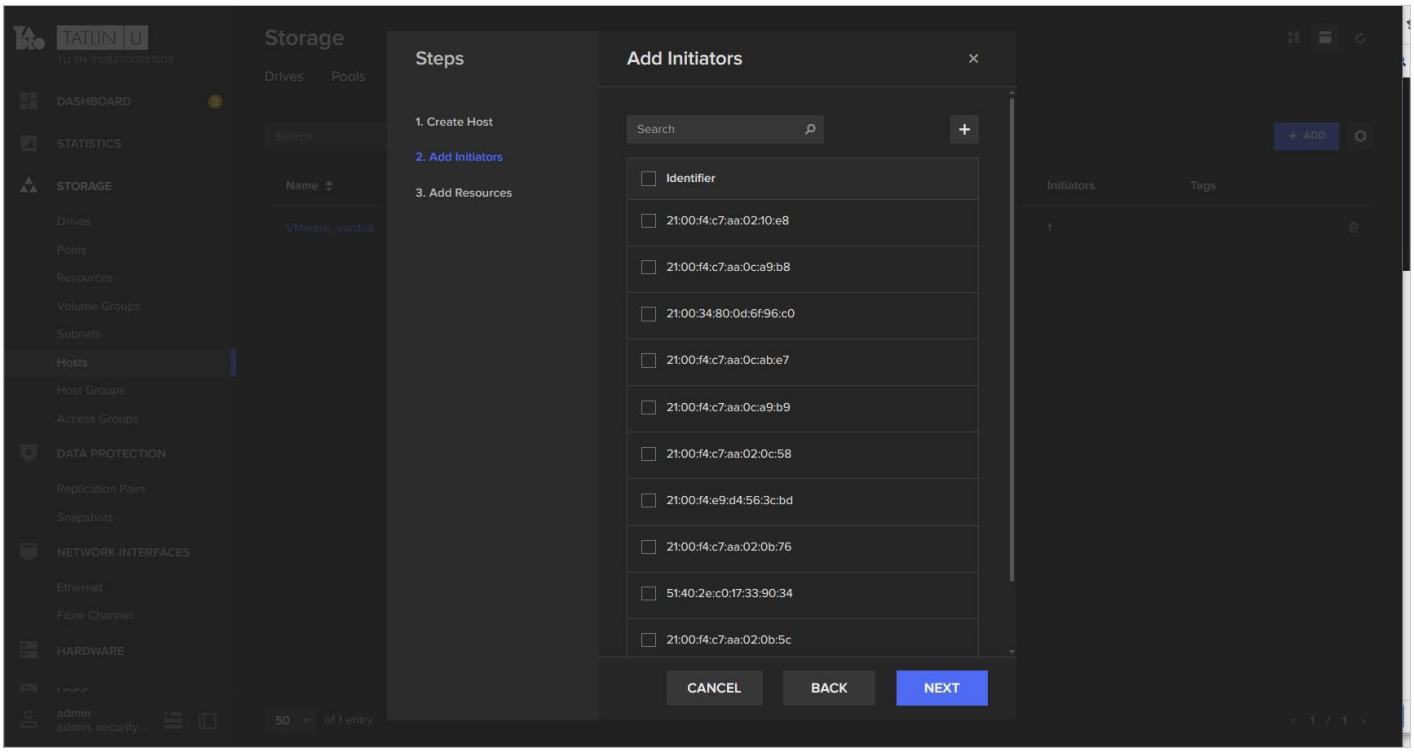

4. Выберите из списка инициаторов, зарегистрированных на СХД, инициаторы, принадлежащие подключаемому серверу.
   Для получения имен инициаторов подключитесь к ним в консоли хостов zVirt и выполните:

```
# for host in /sys/class/fc_host/host*; do echo "  WWPN: $(cat $host/port_name)";    echo "
State: $(cat $host/port_state)"; done
```

**Пример вывода:**




При отсутствии необходимого инициатора в списке добавьте его вручную, используя имя, полученное в выводе команды выше. Для этого нажмите + и введите имя инициатора в появившемся окне.



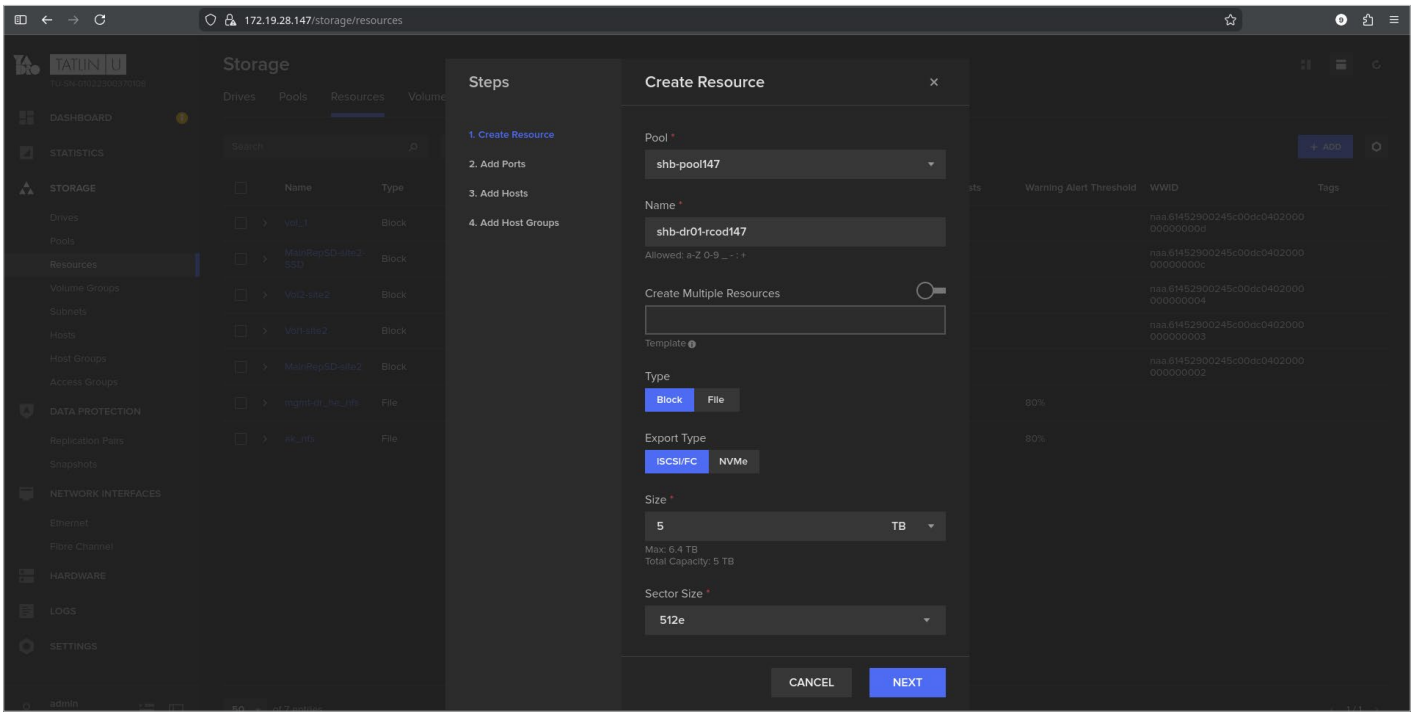

5. Выберите на вкладке **Add Resources** ресурсы, которые планируется презентовать данному хосту. Ресурсы могут быть созданы или выбраны из списка созданных ранее. Для создания нового ресурса нажмите **CREATE**.
   В примере созданы хосты `node1_fc` и `node2_fc`, подключенные к СХД 172.19.28.144, а также `node3_fc` и `node4_fc`, подключенные к СХД 172.19.28.147.

### Создание ресурсов (логических томов)

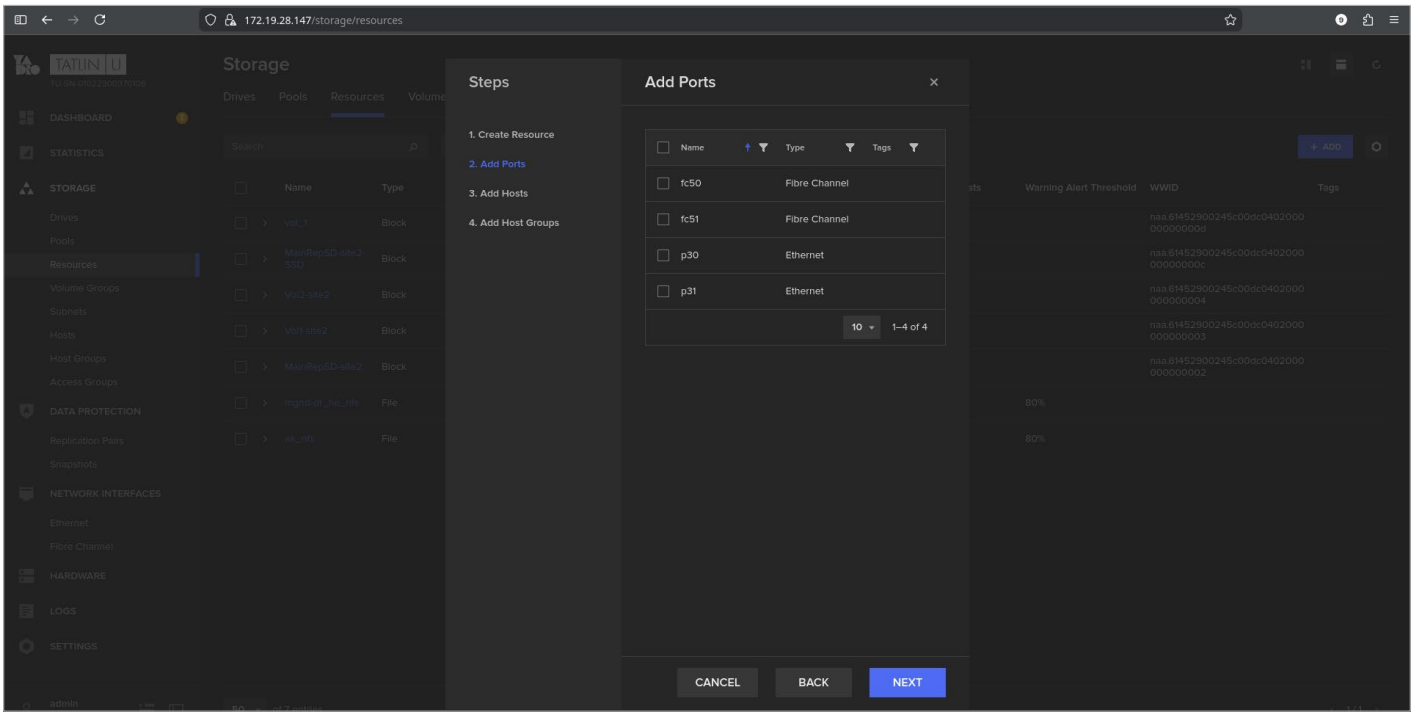
1. Для создания необходимого количества ресурсов перейдите в консоли управления СХД в раздел **Resources** и нажмите **ADD**.
2. В окне **Create Resource** выберите ранее созданный пул и укажите имя создаваемого ресурса `shb-dr01-rcod147`.
3. Выберите объем (например 5 ТБ), укажите размер сектора 512e.

> **!** В текущей версии zVirt не поддерживаются 4-килобайтные блоки.

Назначьте создаваемому ресурсу блочный тип и презентацию по iSCSI/FC.

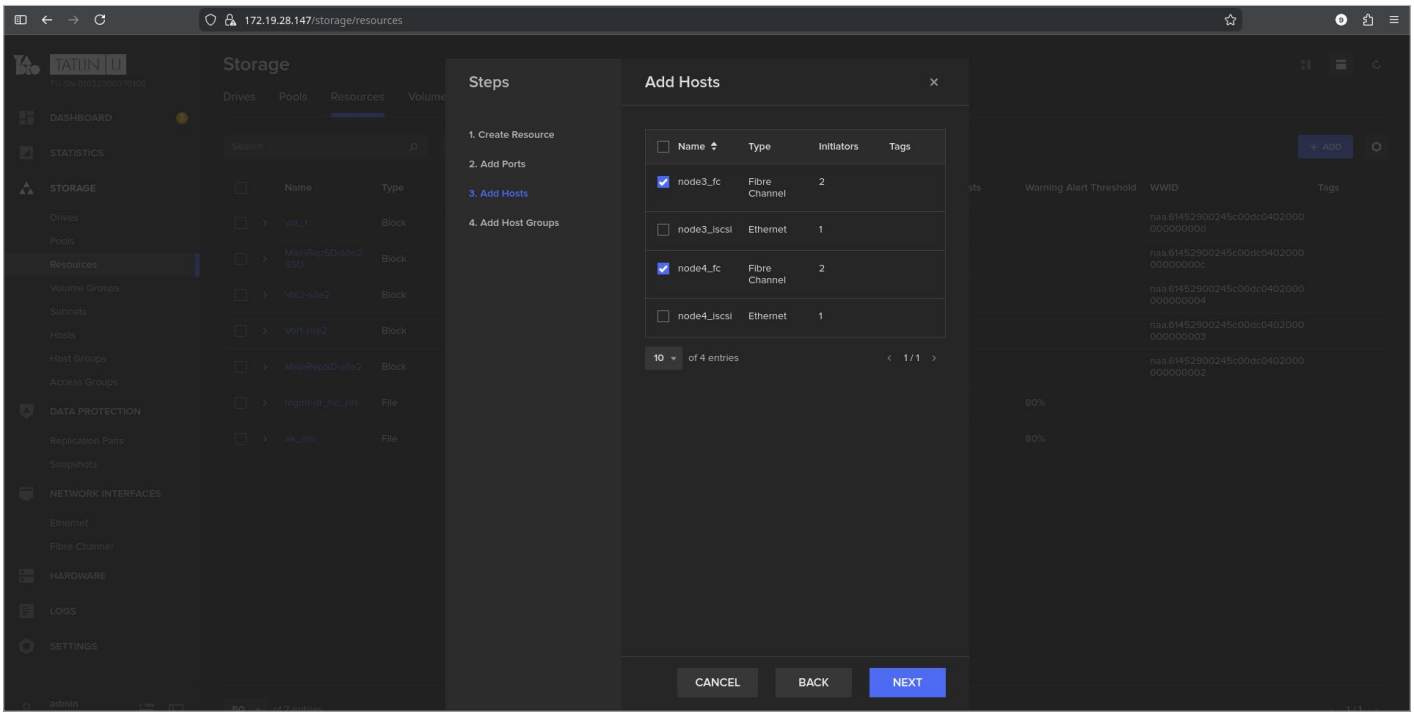



4. На вкладке **Add Ports** выберите порты, через которые ресурс будет презентован хосту.




При FC-подключении установите флажки **fc50** и **fc51**.
При iSCSI-подключении выберите порты типа Ethernet.

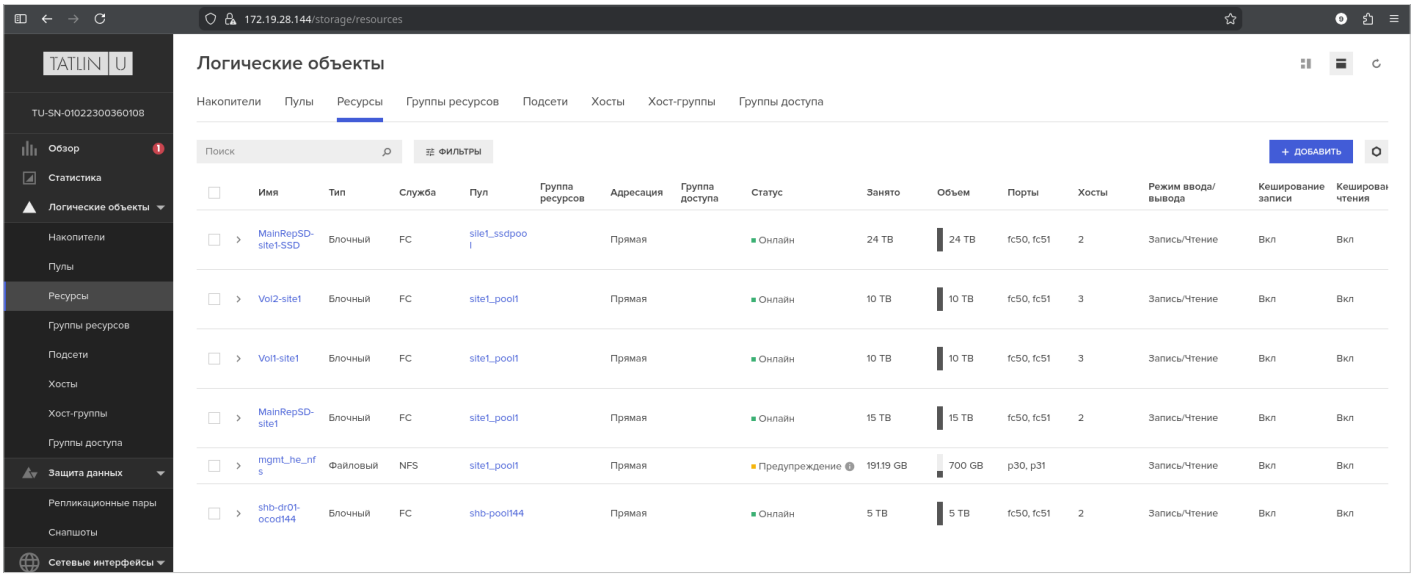
5. Выберите на вкладке **Add Hosts** хосты, созданные ранее на базе зарегистрированных подключений. Это ограничит доступ указанных хостов по выбранным портам выбранными ресурсами (LUN Masking).




6. На вкладке **Add Host Groups**, не добавляя хост-групп, нажмите **CREATE** и завершите создание и презентацию ресурса.
7. Повторите все вышеописанные операции на СХД (https://172.19.28.144/storage/resources). В качестве хостов необходимо выбрать `node1_fc` и `node2_fc`.

> **!** Создаваемые на разных СХД ресурсы должны иметь идентичные объем и схему защиты. Данные параметры являются ключевыми при добавлении ресурсов в пару репликации. Для успешного создания репликационной пары при создании второго ресурса необходимо выбрать аналогичные параметры и объем диска.

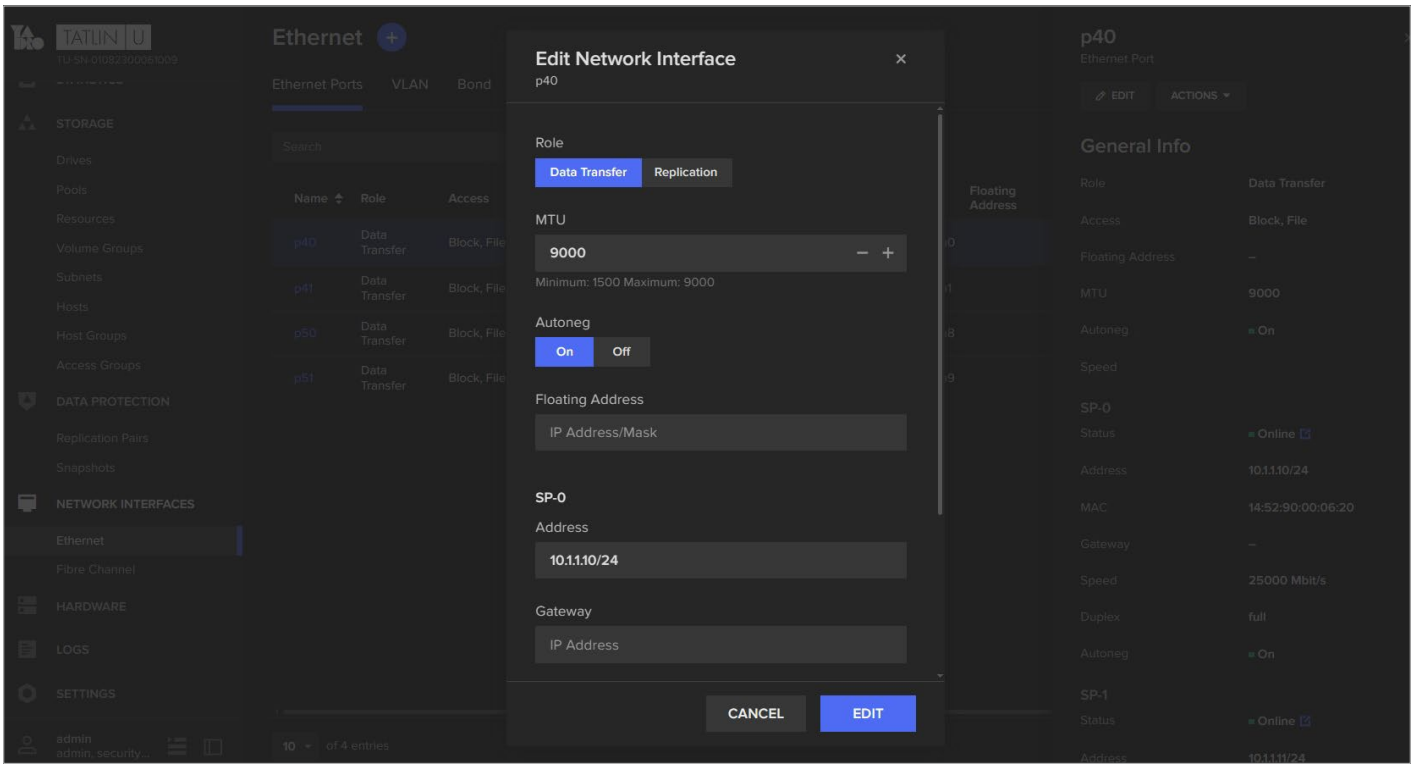
Список созданных ресурсов можно найти в разделе **Resources**.



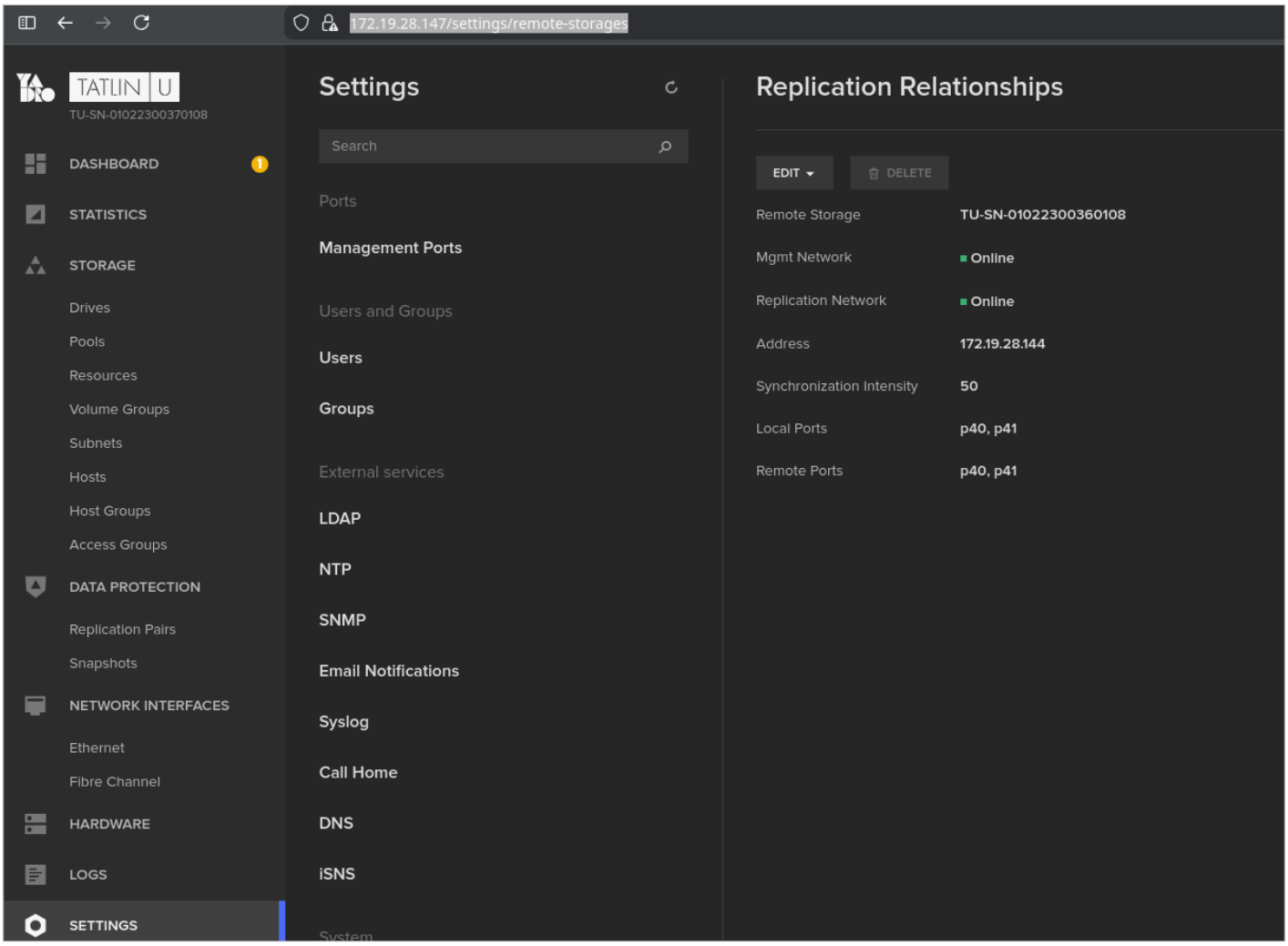
## Создание репликационной связи между массивами

1. Определите репликационные порты. Для этого перейдите в **NETWORK INTERFACES** — **Ethernet** и назначьте выбранным портам роль **Replication**.
2. Выберите номер порта (например, p40) и нажмите **EDIT** в появившемся окне. В окне **Edit Network Interface** укажите роль **Replication**, установите для MTU величину в 9000 и укажите IP-адрес и маску сети репликации для портов на SP-0 и SP-1.
3. Повторите операцию для второй СХД.

> **!** Порты репликации должны быть доступны.
> Команда ping выполняется без потери пакетов. Проверка Jumbo frame показывает планируемое значение: `ping xxx.xxx.xxx.xxx -f -l 8972`.



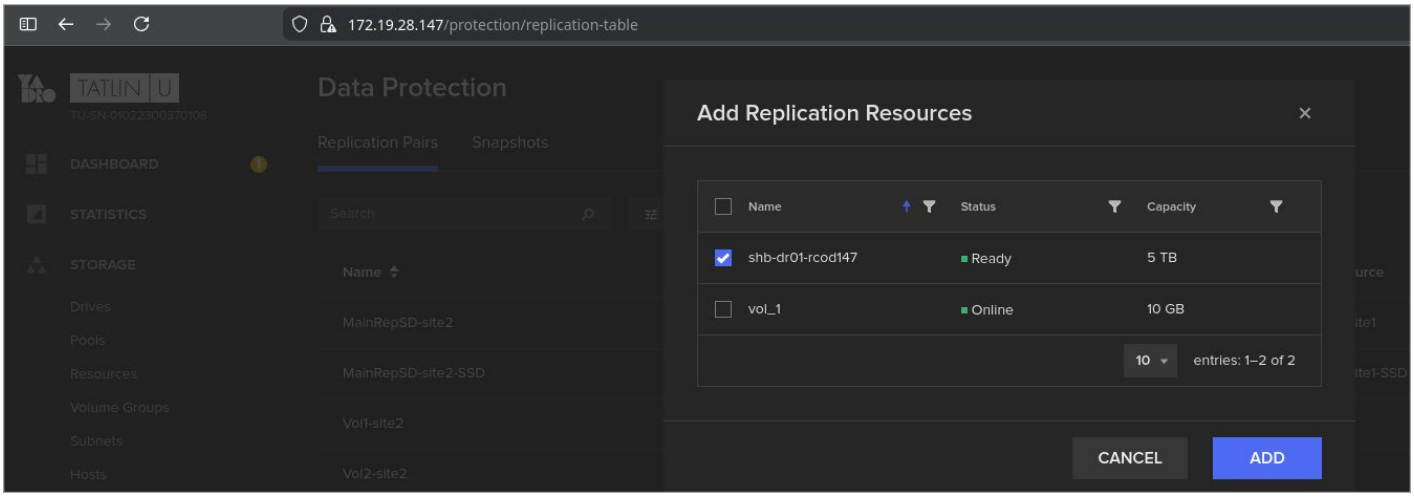

4. После изменения роли портов на **Replication** перейдите в **SETTINGS** — **Replication Relationships** и нажмите **EDIT**. Добавьте вторую СХД, указав ее IP-адрес порта управления и учетные данные УЗ admin. Выберите порты сконфигурированные в предыдущем шаге. По данным портам СХД будут обмениваться данными.



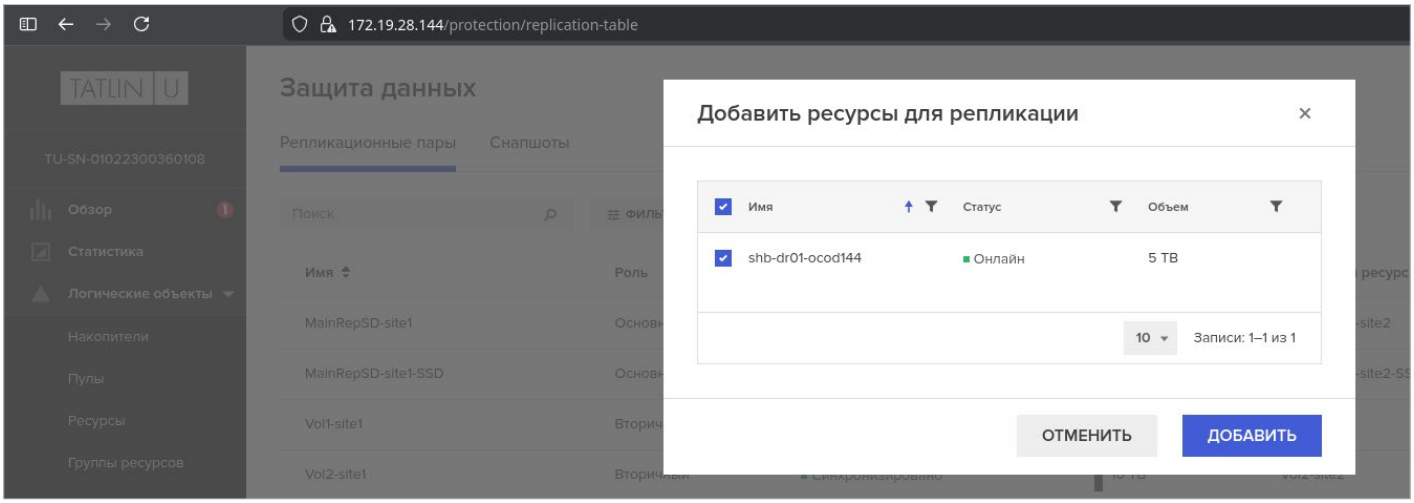

## Создание репликационных пар между массивами

Репликационной парой является объединение двух ресурсов идентичного объема для последующего копирования данных. Ресурсы располагаются на двух разных СХД. Обе СХД соединены между собой репликационными связями через порты репликации.

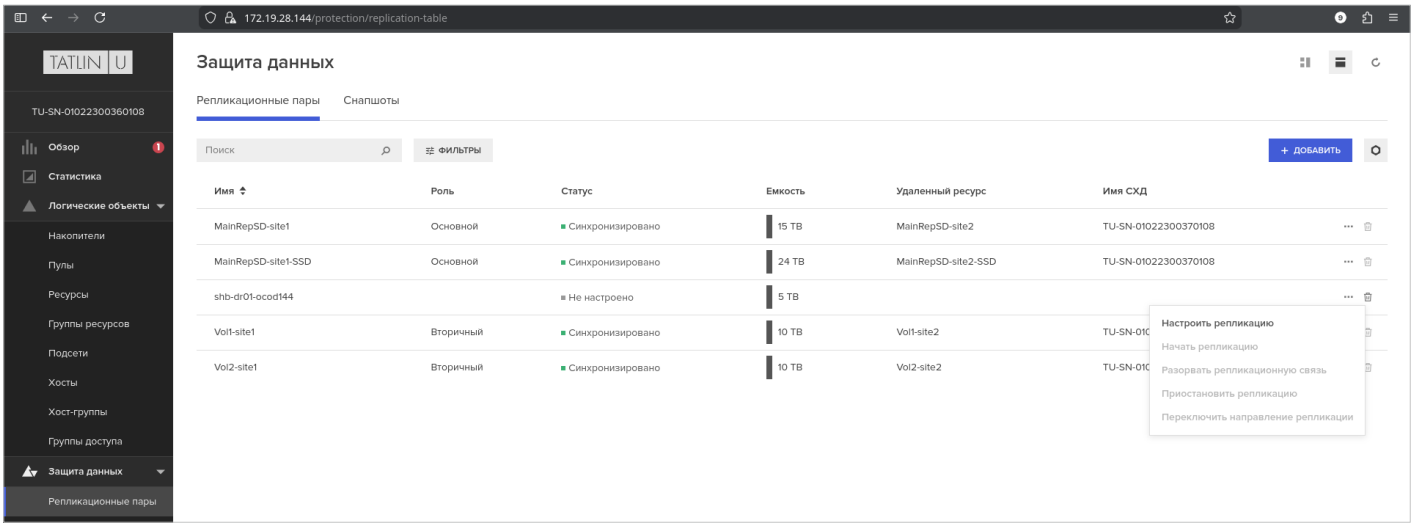
1. Перейдите в **DATA PROTECTION** — **Replication Pairs** и нажмите **ADD**. В окне **Add Replication Resources** выберите запланированные для репликации ресурсы и добавьте их в репликационную пару, нажав **ADD**.



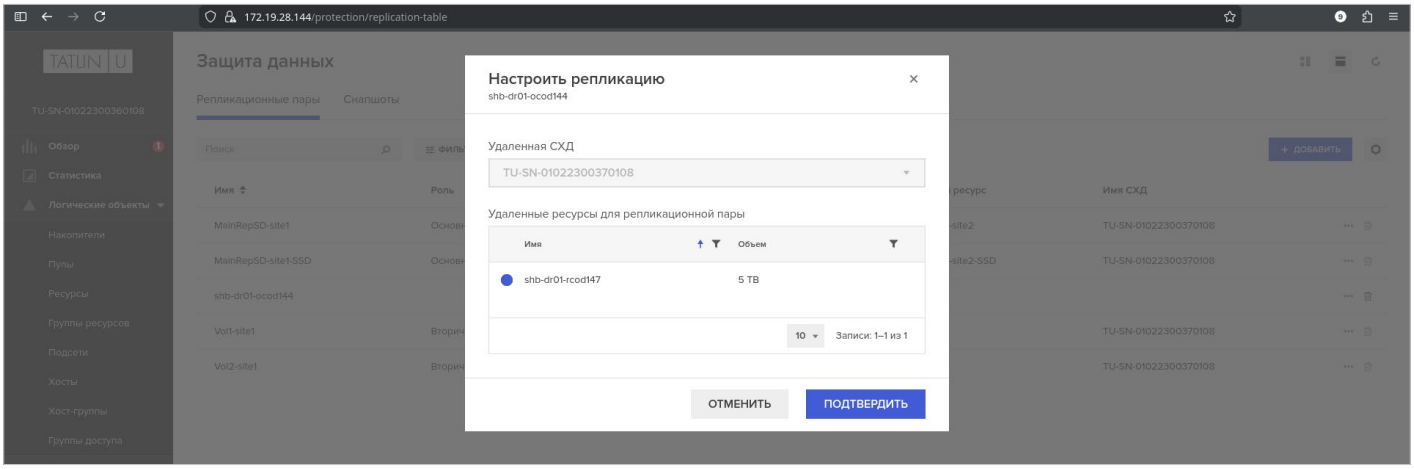
2. Повторите операцию на второй СХД.



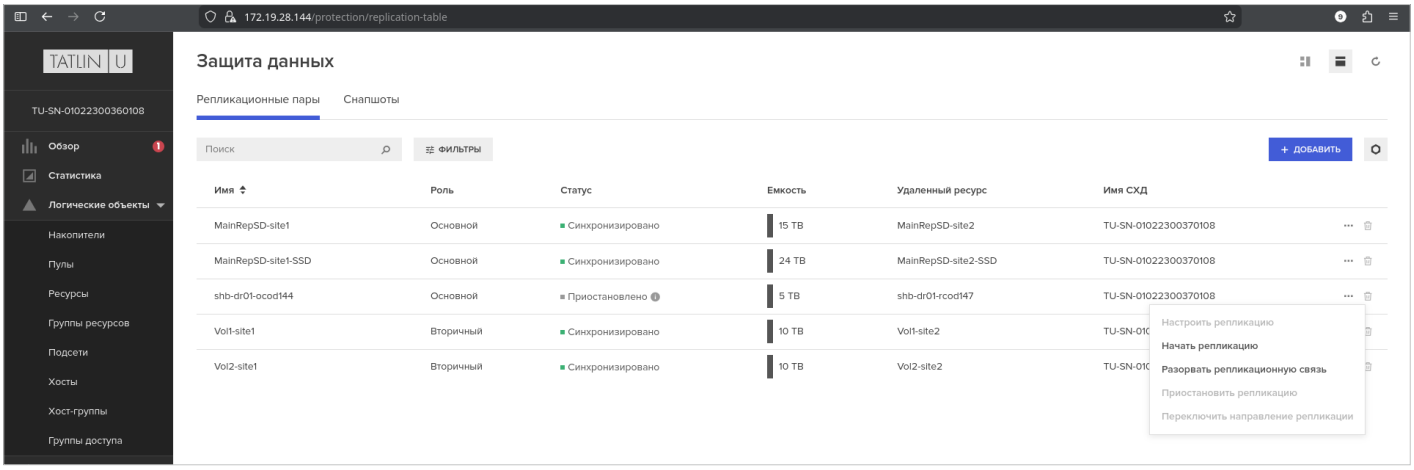

3. После добавления пара появится в списке пар репликации. Пара по умолчанию получит статус **Не настроен**о. Для старта процесса копирования нажмите **Настроить репликацию** в свойствах пары.




4. В окне **Настроить репликацию** отметьте парный ресурс удаленной СХД, на который планируется выполнить копирование данных.



5. Запустите репликацию, нажав **Начать репликацию**.




При конфигурировании репликации исходным ресурсом назначается ресурс на СХД, на которой выполняется конфигурирование и инициация репликации. С этого ресурса данные копируются на парный целевой ресурс удаленной СХД. В примере настройка была выполнена на стороне СХД 144. Соответственно, данная СХД будет автоматически выбрана исходной или основной и с нее будет проводиться репликация в сторону СХД 147, являющейся резервной или целевой.

При данной настройке ресурс доступен на чтение и запись для хостов `node1_fc` и `node2_fc` на стороне ОЦОД. На стороне РЦОД ресурс доступен только на чтение хостам `node3_fc` и `node4_fc`.

Если ресурс, для которого настраивается репликационная пара был подключен к виртуализации и уже содержит данные, настройка репликационной пары для этого ресурса должна проводиться на той стороне, где расположен этот ресурс. Тогда данная сторона будет выбрана основной в репликационной паре и с нее будет проводиться репликация на другую СХД.

## Конфигурирование дисковых ресурсов СХД YADRO TATLIN.UNIFIED на серверах zVirt

В предыдущем разделе был описан процесс создания и презентации ресурсов через порты ввода-вывода. Необходимые для тестов ресурсы созданы и назначены зарегистрированным на СХД хостам через порты FC SAN. После создания и презентации ресурсов на хосте виртуализации необходимо в соответствии с документацией YADRO выполнить следующие действия:

1. Настроить блочные устройства СХД YADRO TATLIN.UNIFIED на хостах zVirt;
2. Сконфигурировать политики ввода-вывода на уровне гипервизора на хостах zVirt для томов СХД YADRO TATLIN.UNIFIED;
3. Подключить домен хранения к ОЦОД.

## Настройка блочных устройств СХД YADRO TATLIN.UNIFIED на хостах zVirt

1. Создайте конфигурационный файл `/etc/udev/rules.d/70-tatlin.rules`:

```
# Set SCSI command timeout for TATLIN devices to 120 seconds

ACTION=="add|change", SUBSYSTEM=="block", ENV{ID_VENDOR}=="YADRO", \

ENV{ID_MODEL}=="TATLIN", ATTR{device/timeout}="120", ATTR{queue/nr_requests}="128"
```

2. Создайте конфигурационный файл `/etc/udev/rules.d/71-tatlin-dm_mpath.rules`:

```
# Set dm-mpath device nr_requests to 256 and scheduler to none

SUBSYSTEM!="block", GOTO="end_rule"

ENV{DEVTYPE}=="partition", GOTO="end_rule"

ACTION!="add|change", GOTO="end_rule"

ATTR{model}!="TATLIN"

KERNEL=="dm-[0-9]*", ATTR{queue/nr_requests}="256", ATTR{queue/scheduler}="none"

LABEL="end_rule"
```

3. Примените новые параметры путем перезагрузки zVirt node или выполните:

```
# lsblk -nd --output NAME,MODEL | grep TATLIN | while read disk_name vendorid; do echo -n
"Disk: $disk_name VendoID:$vendorid CMD:"; cmd="echo 120 > /sys/block/$disk_name/device/
timeout"; echo $cmd; eval $cmd; done

# multipath -ll | grep YADRO | while read mpath dm vendor; do echo -n "Mpath:$mpath DM:$dm
Vendor:$vendor CMD:"; cmd="echo 256 > /sys/devices/virtual/block/$dm/queue/nr_requests && echo
none > /sys/devices/virtual/block/$dm/queue/scheduler"; echo $cmd; eval $cmd; done
```

4. Проверьте применение параметров:

```
# lsblk -nd --output NAME,MODEL | grep TATLIN | while read disk_name vendorid; do printf
"Disk:%s VendoID:%s timeout:%d\n" $disk_name $vendorid $(cat /sys/block/$disk_name/device/
timeout); done
```

**Пример вывода:**




Выданные серверу ресурсы будут сконфигурированы на уровне драйвера дискового хоста виртуализации.

### Конфигурирование политики ввода-вывода (multipath) на уровне гипервизора на хостах zVirt для томов СХД YADRO TATLIN.UNIFIED

Для обеспечения отказоустойчивости путей доступа к ресурсу и балансировки нагрузки необходимо настроить политику ввода-вывода multipath в соответствии с документом YADRO Настройка ОС и платформ виртуализации. Для этого:

1. Создайте конфигурационный файл `/etc/multipath/conf.d/tatlin.conf`:

```
devices {
device {
vendor "YADRO"
product "TATLIN"
path_grouping_policy "multibus"
path_selector "service-time 0"
path_checker tur
no_path_retry 32
max_sectors_kb 1024
}
}
```

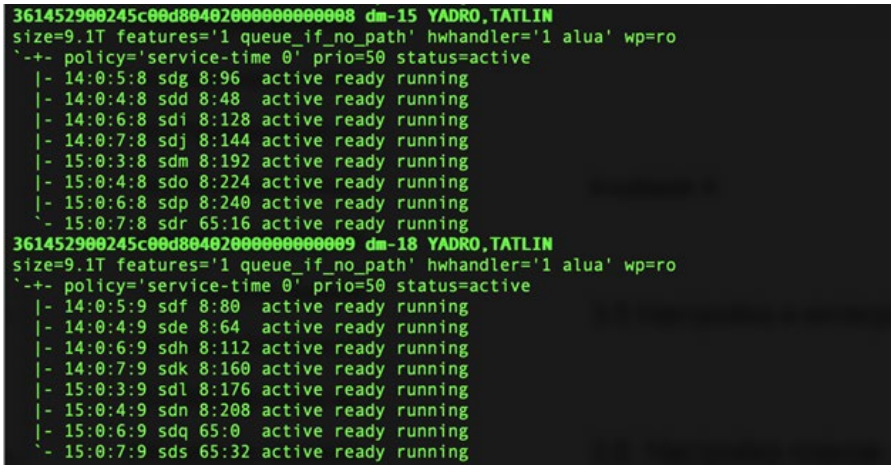
2. Для применения настроек выполните перезагрузку службы multipathd:

```
# systemctl reload multipathd
```

3. Проверьте применение настроек multipath:

```
# multipath -ll
```

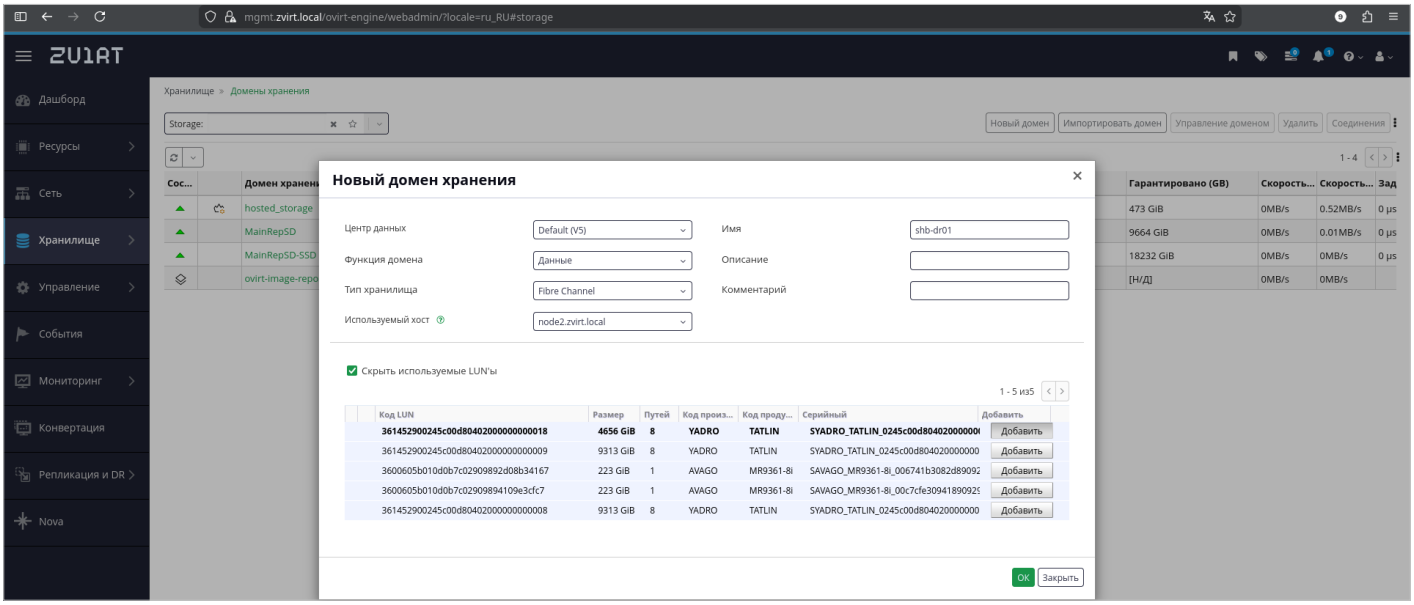
**Пример вывода:**



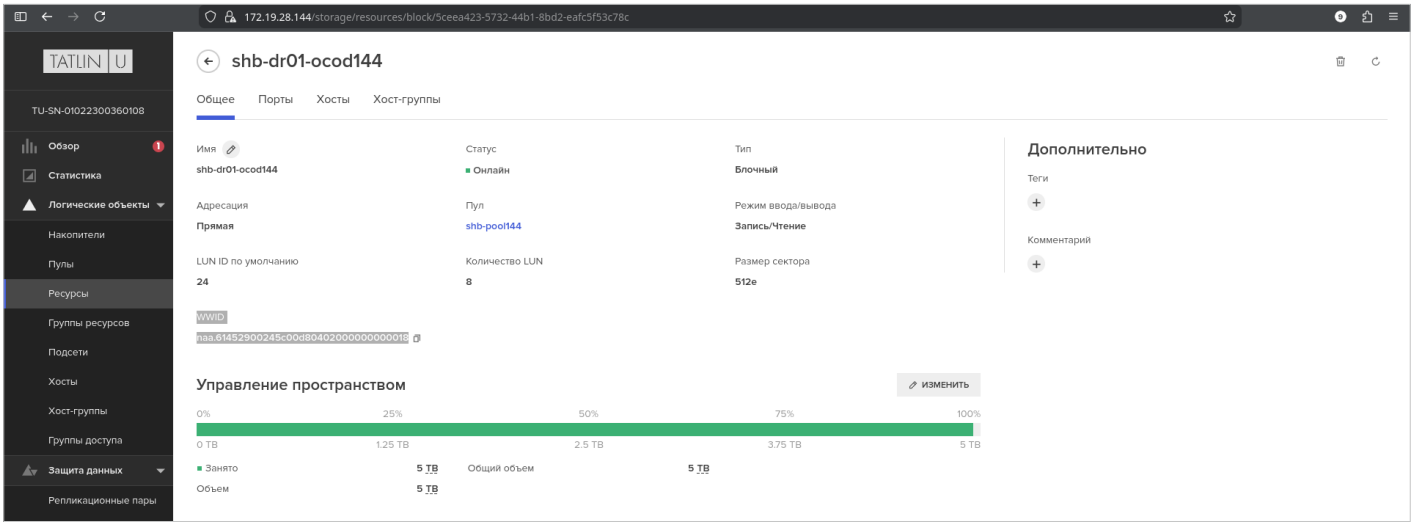
## Подключение домена хранения к ОЦОД

В предыдущих пунктах были созданы ресурсы `shb-dr01-ocod144` на основной СХД в виртуальном основном ЦОД (ОЦОД) и `shb-dr01-rcod147` на резервной СХД в виртуальном резервном ЦОД (РЦОД). Сторона ОЦОД при конфигурировании была назначена исходной в репликационной паре. Ресурс `shb-dr01-ocod144` на сервере ОЦОД подключается как обычный домен хранения, доступный для чтения и записи.

1. Для конфигурирования домена хранения перейдите в интерфейс zVirt (https://mgmt.zvirt.local/ovirt-engine/webadmin/?locale=ru_RU#storage).
2. Перейдите в **Хранилище** — **Домены хранения** и нажмите **Новый домен**. В окне **Новый домен хранения** выберите тип хранилища Fiber Channel. Будет выведен список доступных доменов хранения.




3. Нажмите **Добавить** напротив домена `shb-dr01`. Определить нужный домен можно по размеру или по идентификатору ресурса (LUN ID). LUN ID можно найти в свойствах ресурса в консоли управления СХД.




4. В мастере добавления домена хранения укажите имя `shb-dr01` для домена хранения и нажмите OK.

> **!** В составе имени домена хранения не рекомендуется указывать имя площадки, так как после переключения домена хранения на резервную площадку его имя сохраняется.

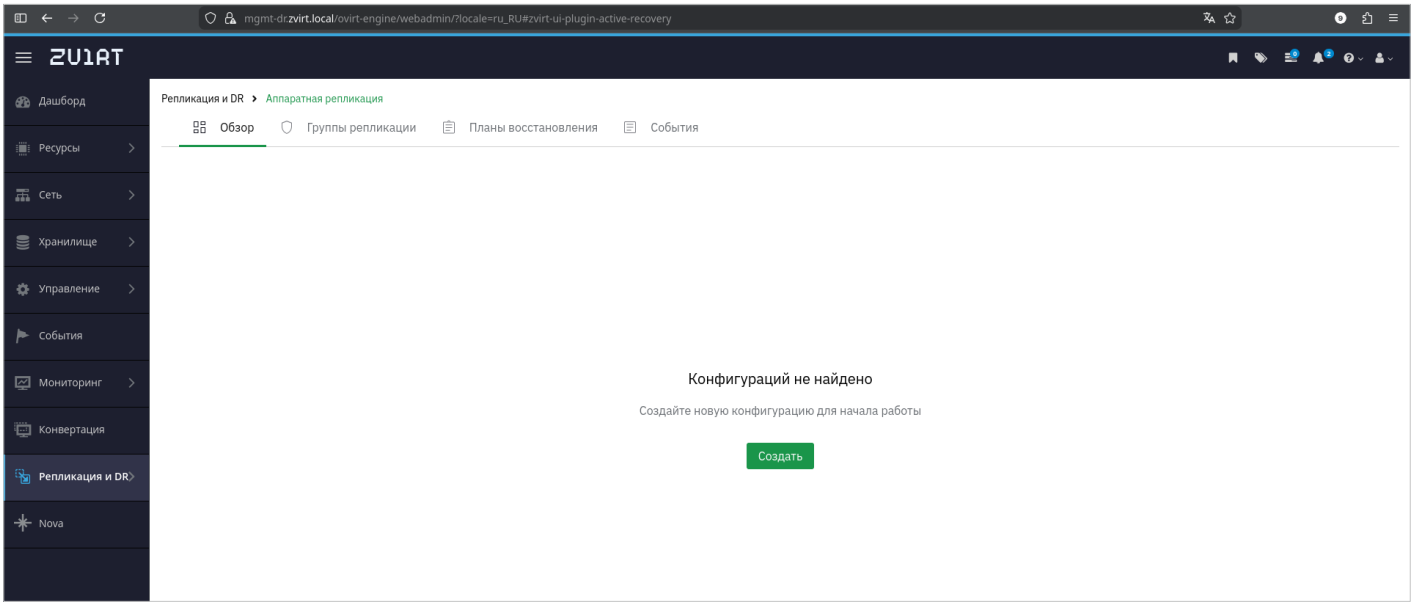
## Конфигурирование сервиса аппаратной репликации zVirt

Для интеграции аппаратной репликации СХД с сервисом репликации zVirt необходимо добавить в сервис информацию о портах управления zVirt и СХД на каждой площадке и указать имена площадок. Для этого необходимо выполнить следующие действия:

1. Сконфигурировать сервис репликации zVirt с аппаратной репликацией СХД TATLIN.UNIFIED;
2. Подключить СХД к сервису аппаратной репликации zVirt.

## Конфигурирование сервиса репликации zVirt с аппаратной репликацией СХД TATLIN.UNIFIED

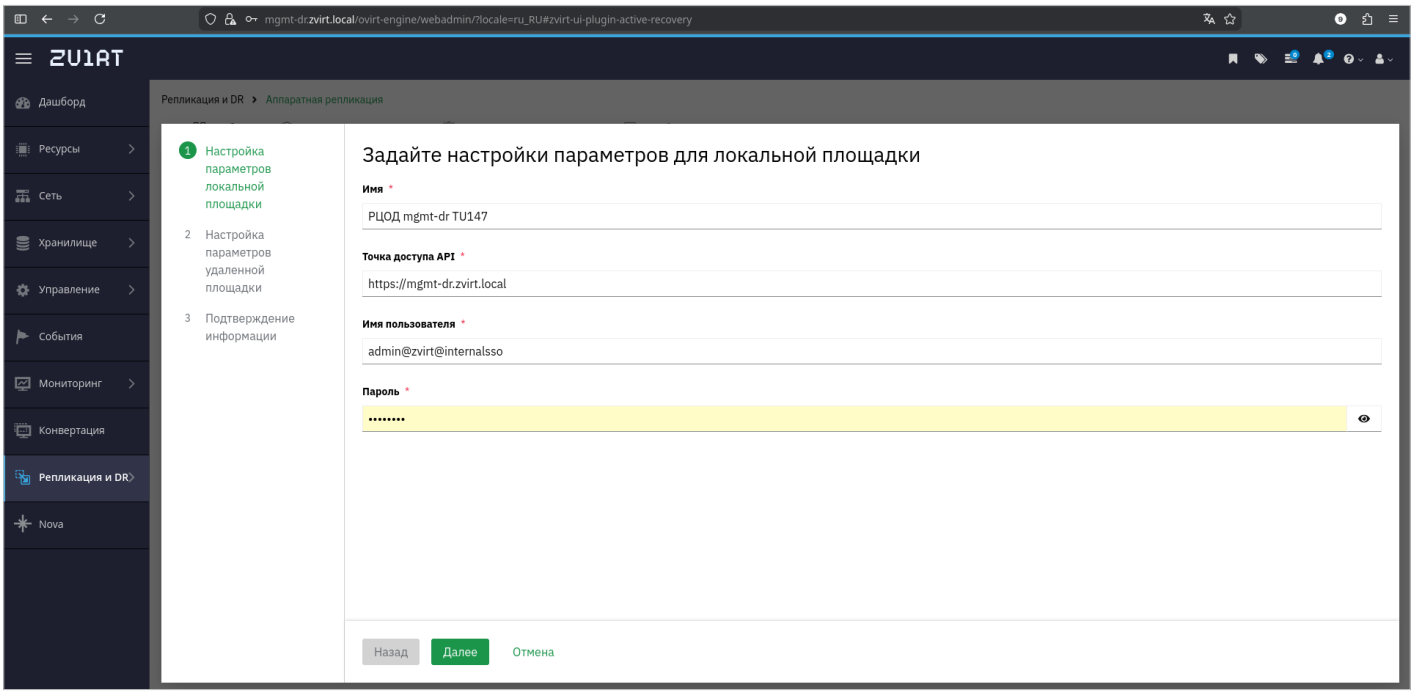
1. Подключитесь к веб-интерфейсу менеджера управления в РЦОД (https://mgmt-dr.zvirt.local/ovirt-engine/).
2. Перейдите в **Репликация и DR** — **Аппаратная репликация** — **Обзор** (https://mgmt-dr.zvirt.local/ovirt-engine/webadmin/plugin/zvirt-ui-plugin-active-recovery/index.html#zvirt-ui-plugin-active-recovery) и нажмите **Создать**.




3. Задайте настройки на вкладке **Настройка параметров локальной площадки**:

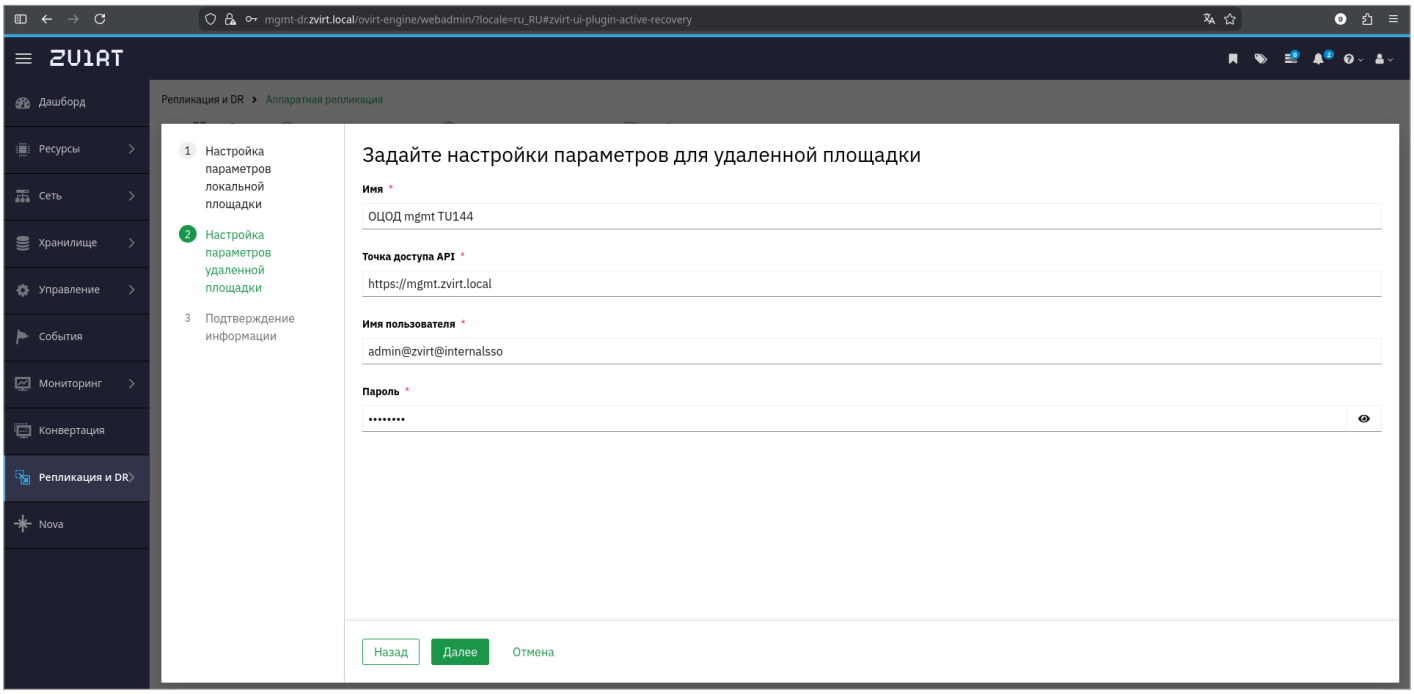
> ! Так как браузер подключен к `mgmt-dr.zvirt.local`, локальной площадкой будет являться `mgmt-dr.zvirt.local`.

| НАСТРОЙКА ПАРАМЕТРОВ ЛОКАЛЬНОЙ ПЛОЩАДКИ | |
|---|---|
| Имя | РЦОД mgmt-dr TU147 |
| Точка доступа API | https://mgmt-dr.zvirt.local |
| Имя пользователя | admin@zvirt@internalsso |
| Пароль | Пароль пользователя admin@zvirt@internalsso |



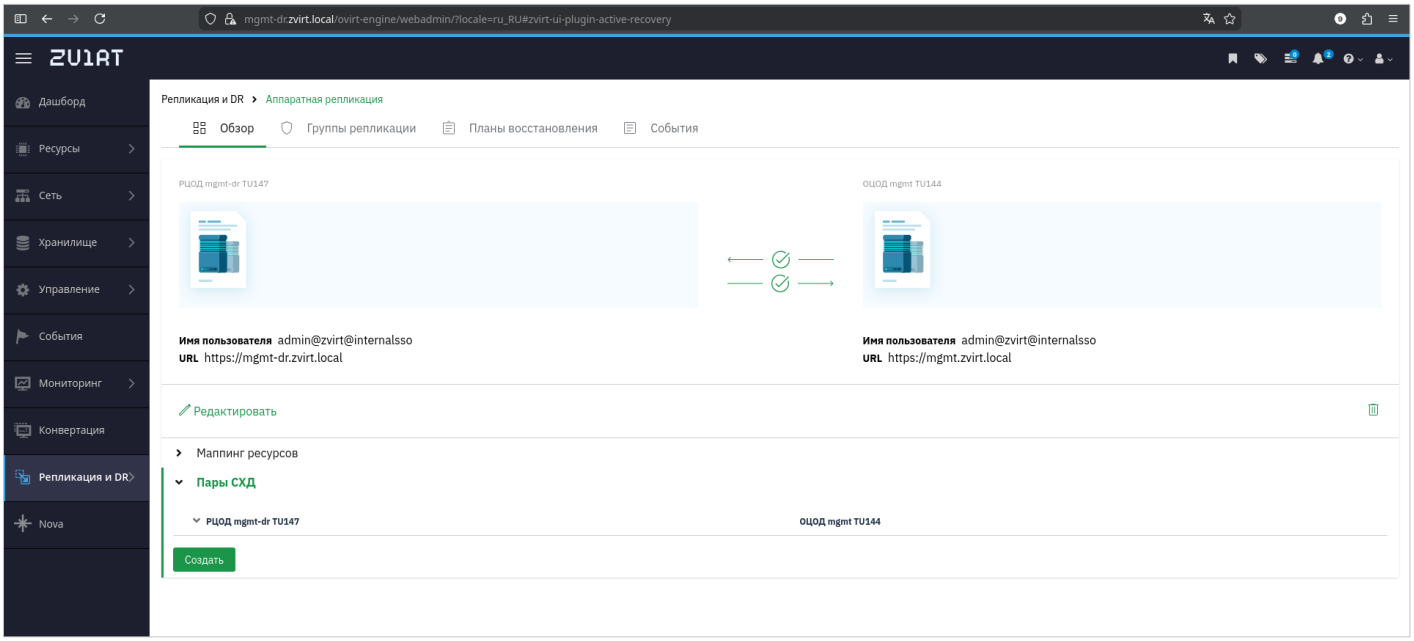

4. На вкладке **Настройка параметров удаленной площадки** укажите настройки для удаленной (основной) площадки:

| НАСТРОЙКА ПАРАМЕТРОВ УДАЛЕННОЙ ПЛОЩАДКИ | |
|---|---|
| Имя | ОЦОД mgmt TU144 |
| Точка доступа API | https://mgmt.zvirt.local |
| Имя пользователя | admin@zvirt@internalsso |
| Пароль | Пароль пользователя admin@zvirt@internalsso |



## Подключение СХД к сервису аппаратной репликации zVirt

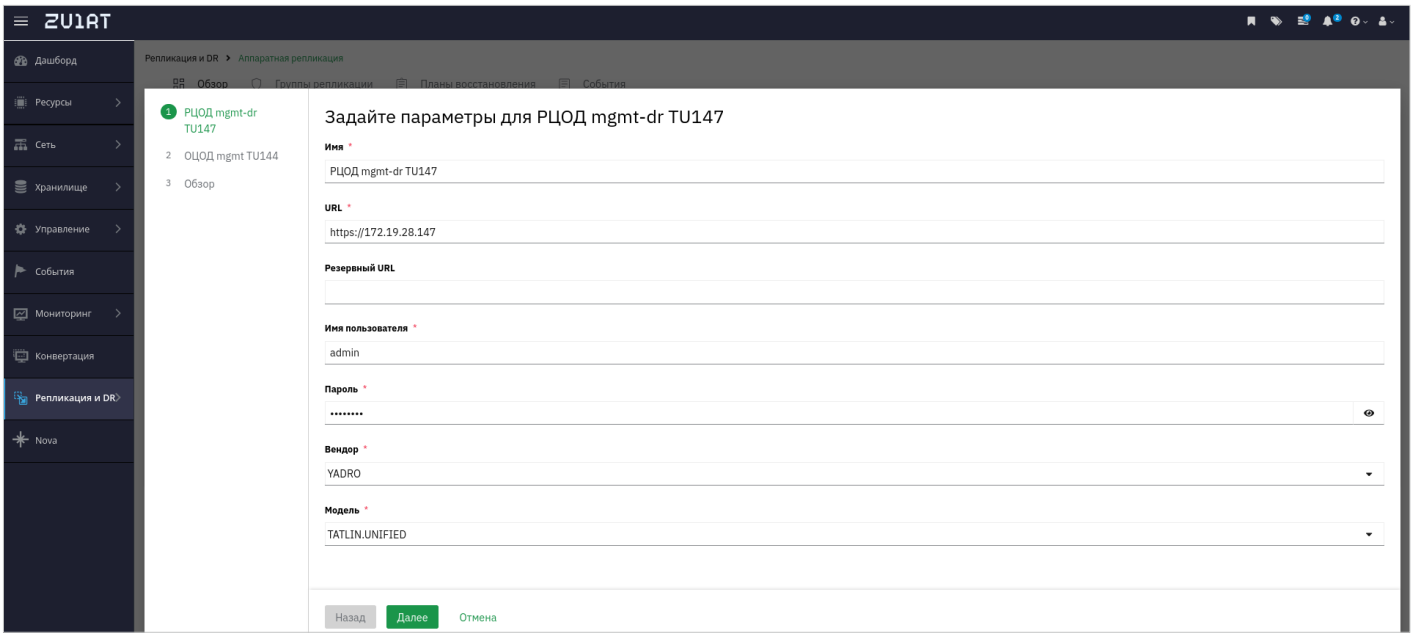
1. Для запуска мастера настройки пар СХД в веб-интерфейсе менеджера управления резервной площадки перейдите в **Репликация и DR** — **Аппаратная репликация** — **Обзор**. Разверните список **Пары СХД** и нажмите **Создать**




2. На вкладке **РЦОД mgmt-dr TU147** укажите параметры:

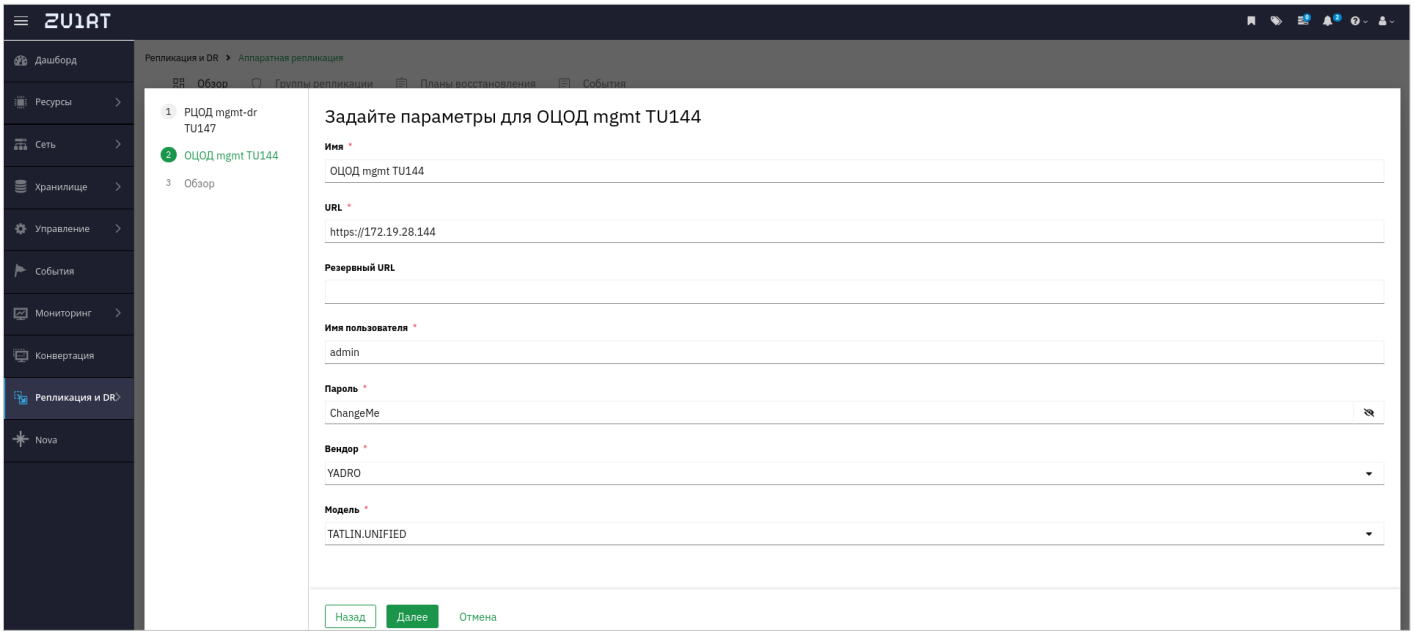
> **!** В поле **Имя** рекомендуется назначить то же имя, которое было дано соответствующей площадке на предыдущем шаге.
> В поле **URL** указывается плавающий адрес для управления СХД РЦОД. Данный адрес является общим для двух контроллеров и используется при входе в консоль управления СХД.
> Поле **Резервный URL** не заполняется.

| РЦОД mgmt-dr TU147 | |
|---|---|
| Имя | РЦОД mgmt-dr TU147 |
| URL | https://172.19.28.147 |
| Резервный URL: | |
| Имя пользователя | admin |
| Пароль | Пароль пользователя admin на СХД |
| Вендор | YADRO |
| Модель | TATLIN.UNIFIED |

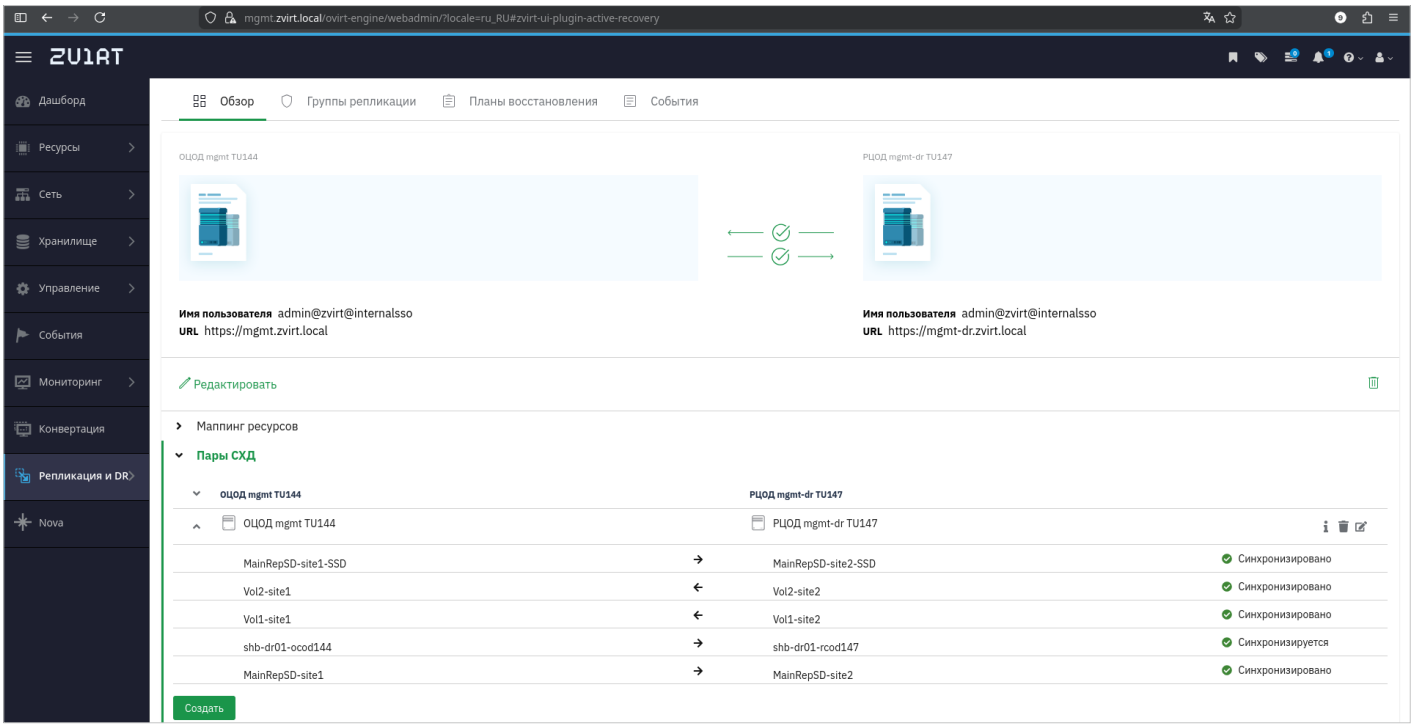



3. На вкладке **ОЦОД mgmt TU144** укажите параметры, относящиеся к ОЦОД:

| ОЦОД mgmt TU144 | |
|---|---|
| Имя | ОЦОД mgmt TU144 |
| URL | https://172.19.28.144 |
| Резервный URL: | |
| Имя пользователя | admin |
| Пароль | Пароль пользователя admin на СХД |
| Вендор | YADRO |
| Модель | TATLIN.UNIFIED |

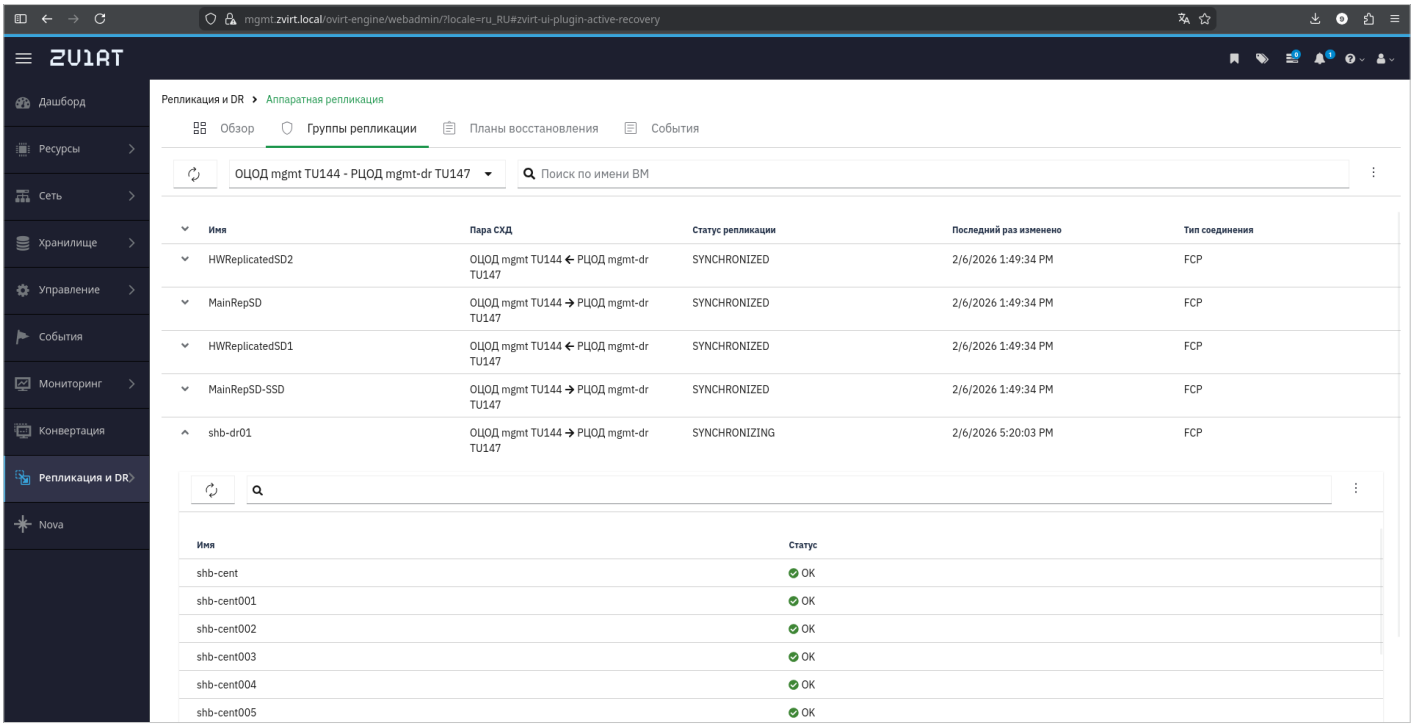


4. Удостоверьтесь, что на вкладке **Обзор** все данные введены корректно и нажмите **Сохранить**. Новая конфигурация должна появиться на вкладке **Репликация и DR** — **Обзор**.




В примере на СХД настроена репликация ресурса `shb-dr01-ocod144` в ОЦОД на ресурс `shb-dr01-rcod147` в РЦОД. **Статус репликации** — **Синхронизируется**. Данный статус указывает на процесс инициирующего копирования данных из ОЦОД в РЦОД.

Данные также можно увидеть на вкладке **Группы репликации**. После создания ВМ на домене хранения данные будут отображаться списком.



## Конфигурирование планов аварийного восстановления

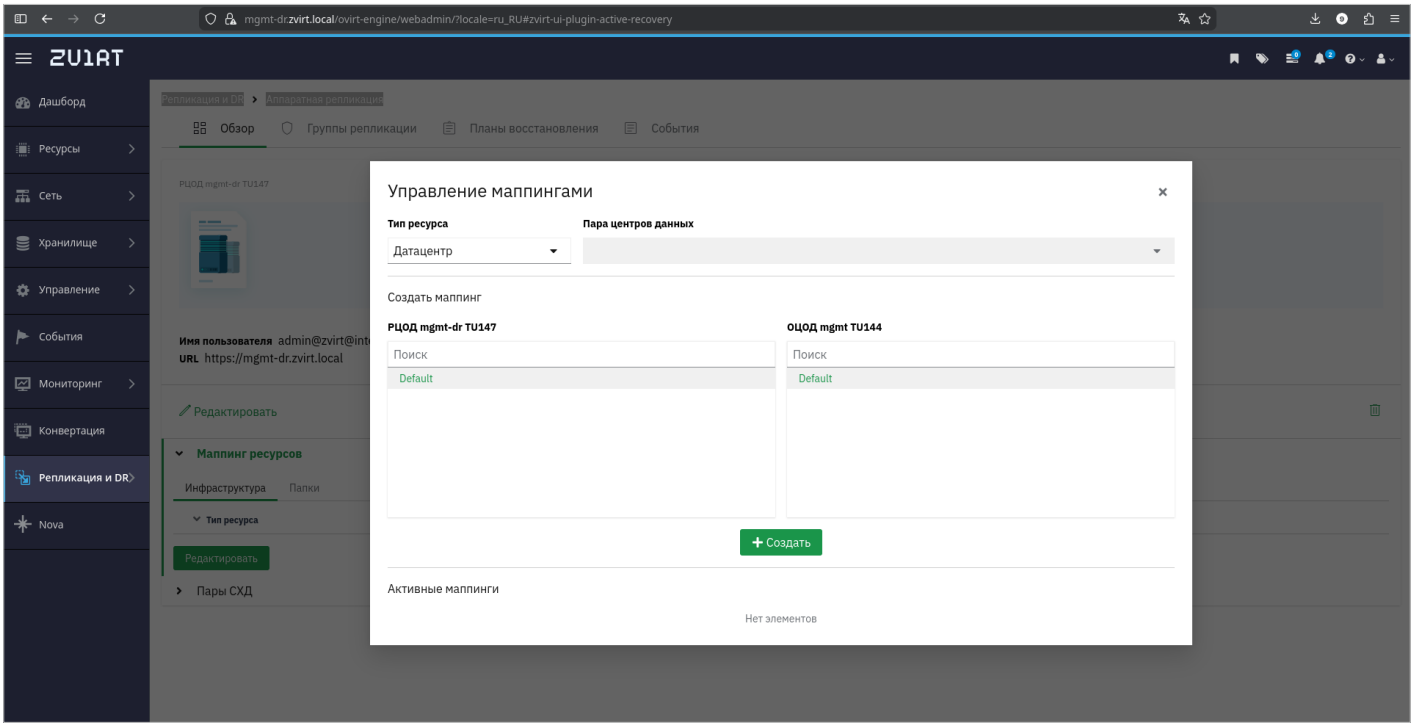
Конфигурирование сервисов репликации и пар репликации может быть выполнено на любой из площадок. Настройки синхронизируются между площадками.

Конфигурирование маппинга ресурсов и планов восстановления выполняется на каждой площадке отдельно. Если планируется защищать только ВМ в ОЦОД, планы восстановления следует создавать только на стороне РЦОД. Созданный план даст возможность производить переключение как из ОЦОД в РЦОД, так и из РЦОД в ОЦОД.

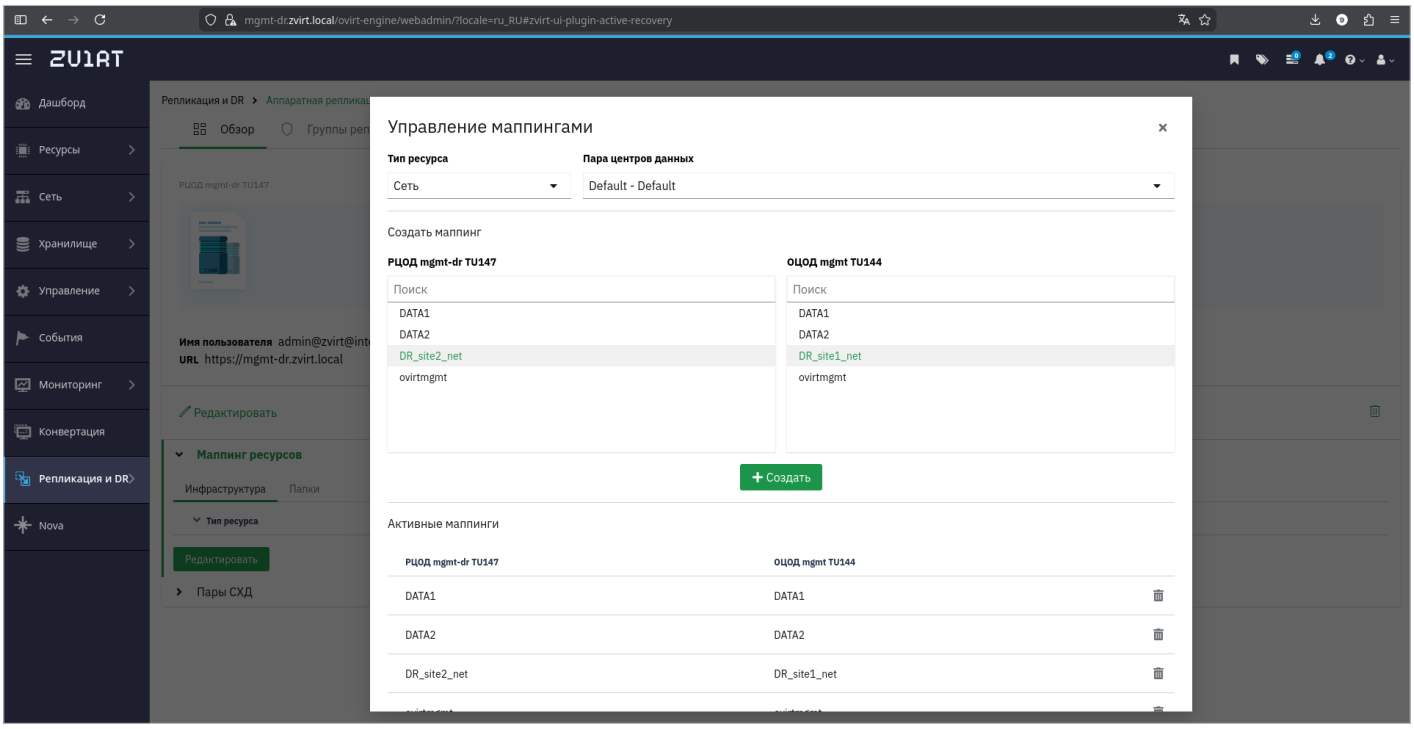
## Конфигурирование маппинга ресурсов

Маппинг ресурсов рекомендуется настраивать на каждой из площадок, участвующих в репликации. В этом случае необходимые параметры для каждой ВМ будут устанавливаться автоматически.

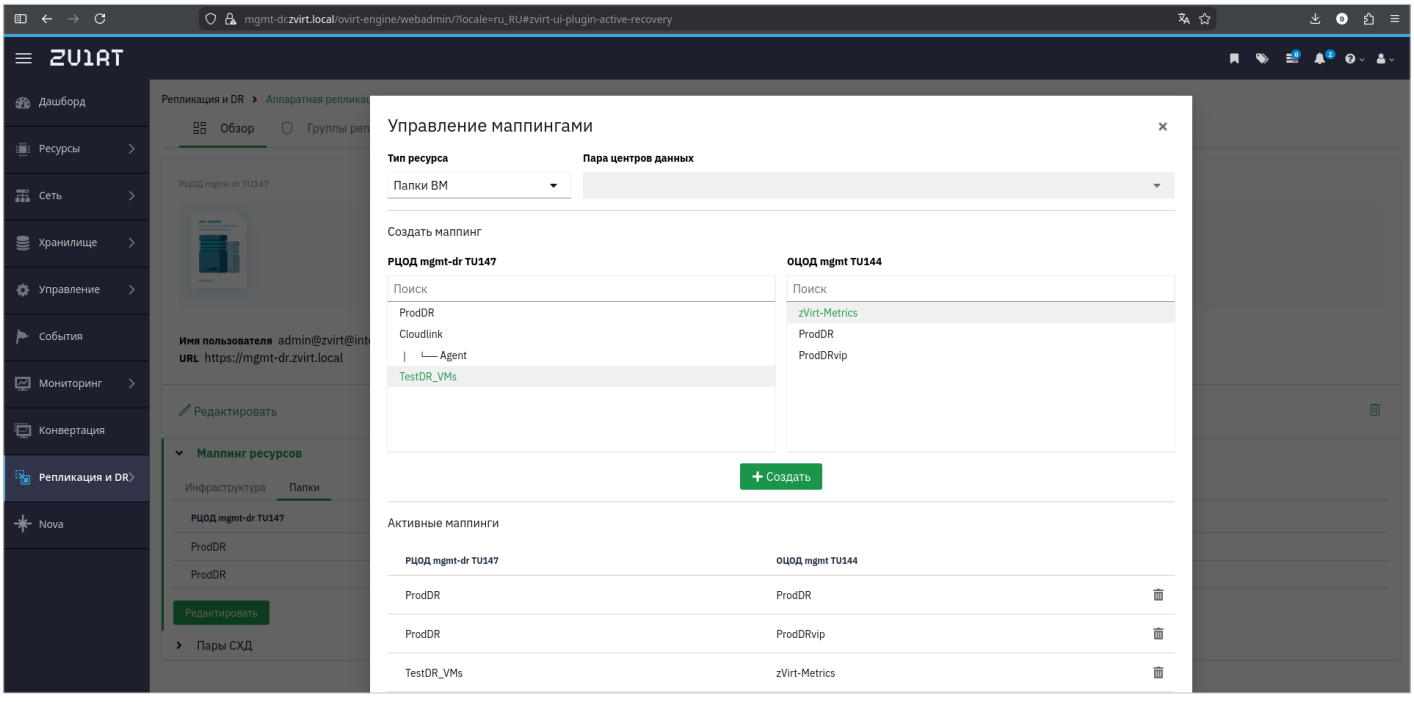
1. Для конфигурирования маппинга в интерфейсе управления платформы виртуализации (https://mgmt-dr.zvirt.local/) перейдите на вкладку **Репликация и DR** — **Аппаратная репликация** — **Обзор** — **Маппинг ресурсов** и нажмите **Редактировать**.
2. В окне **Управление маппингами** укажите в полях названия площадок ЦОД, указав привязку ВМ к этим ЦОД для их автоматической регистрации при переводе ресурсов между площадками.
3. Нажмите **Создать**. В окне **Тип ресурса** по умолчанию выбран параметр **Датацентр**, поэтому будет создан маппинг ресурсов к площадкам (датацентрам).




Маппинг для параметров сетей, профилей сетей, кластеров, папок ВМ выполняется аналогично. При выборе соответствующих пунктов в окне **Тип ресурса** требуемые маппинги будут появляться в списках ОЦОД/РЦОД.




Маппинг папок ВМ имеет древовидное представление.

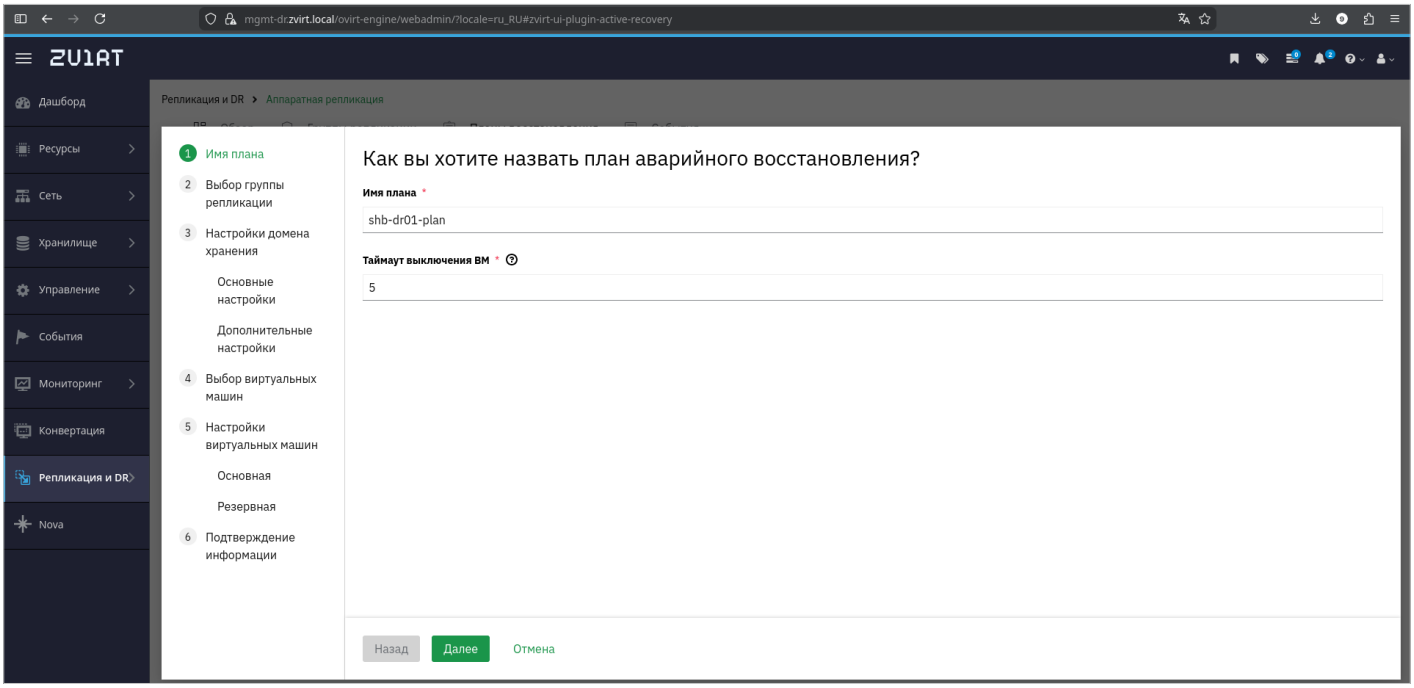


## Создание плана восстановления

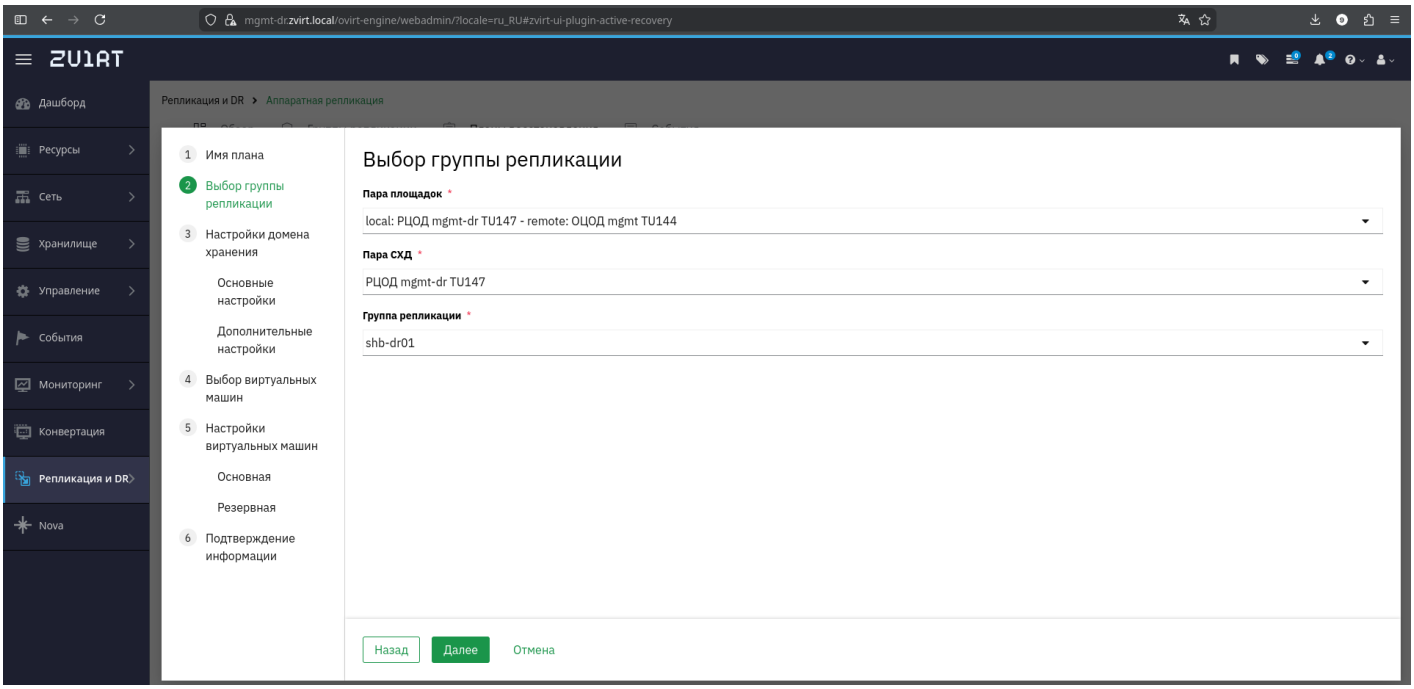
План аварийного восстановления применяется для защиты ВМ, расположенных в ОЦОД, от потенциальных сбоев оборудования в ОЦОД или потери всей площадки. План аварийного восстановления создается на стороне РЦОД, что позволяет использовать его даже при полном отказе площадки ОЦОД.

Для создания плана аварийного восстановления:

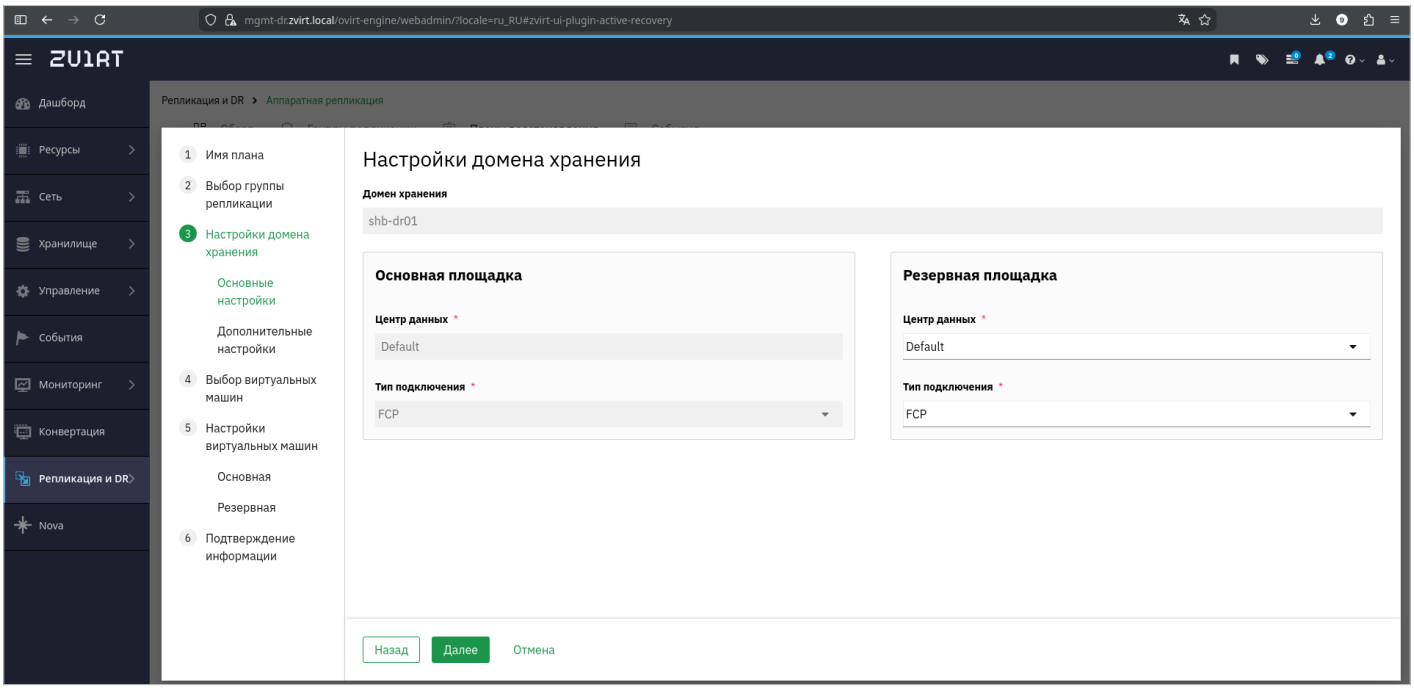
1. В интерфейсе zVirt перейдите на вкладку **Репликация и DR** — **Аппаратная репликация** — Планы восстановления и нажмите **Создать план**.
2. На вкладке **Имя плана** в поле **Имя плана** укажите имя `shb-dr01-plan`.



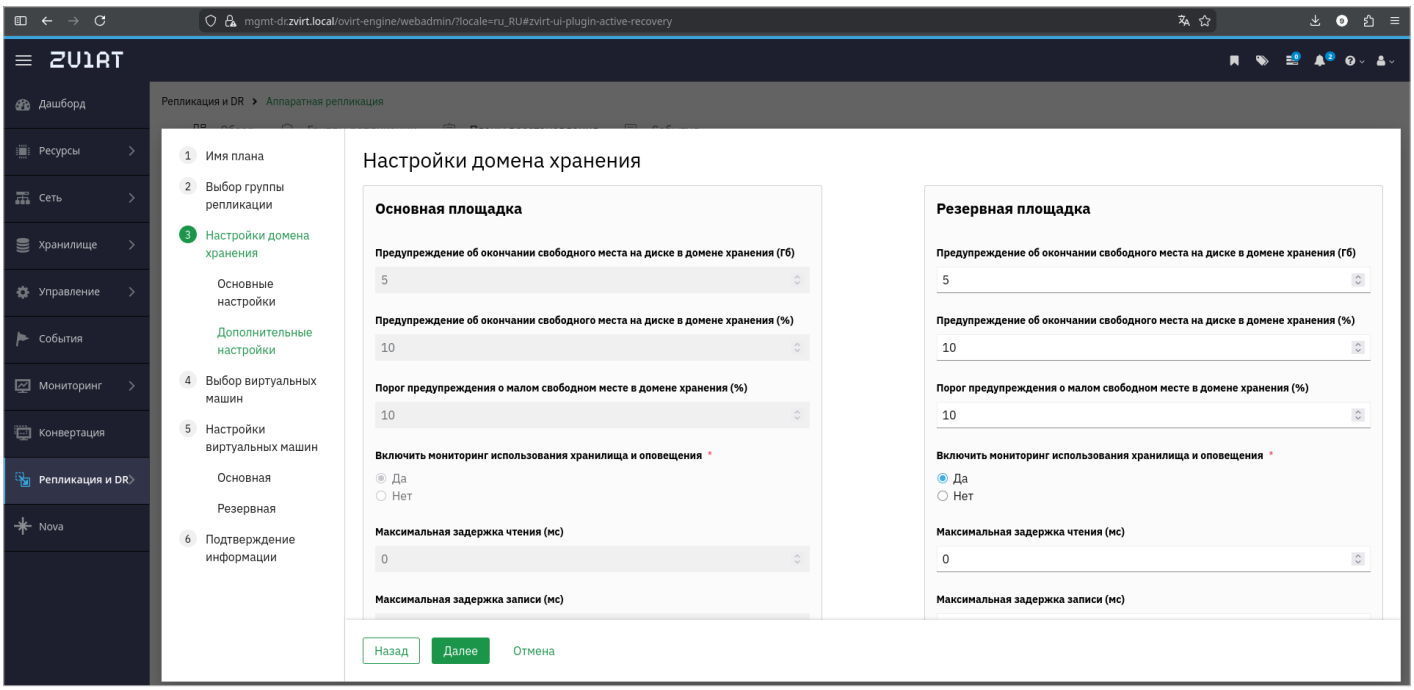

3. На вкладке **Выбор группы репликации** укажите в поле **Пара СХД** укажите `РЦОД mgmt-dr TU147`. В поле **Группа репликации** укажите `shb-dr01`.



4. При использовании FC-подключения к СХД на вкладке **Настройки домена хранения** используйте параметры по умолчанию. При использовании iSCSI-подключения перестройте подключение к target на СХД.

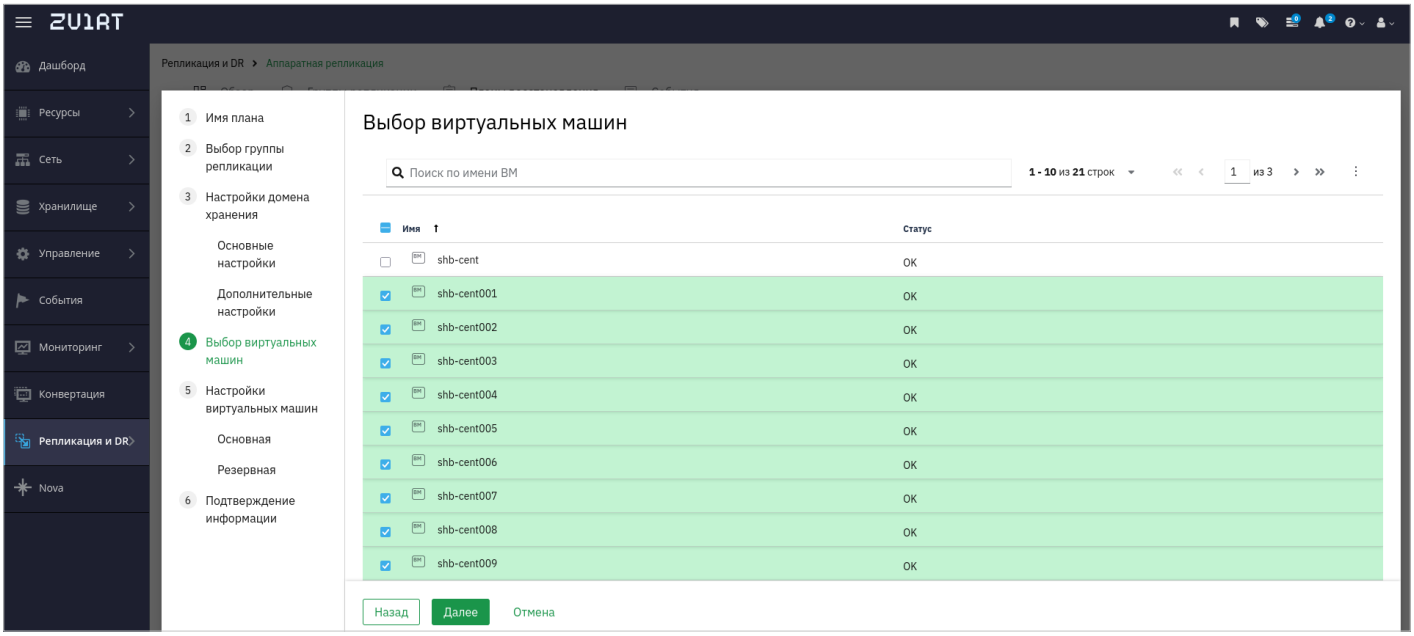



На странице дополнительных настроек указываются параметры домена хранения. Для большинства конфигураций применимы настройки по умолчанию.

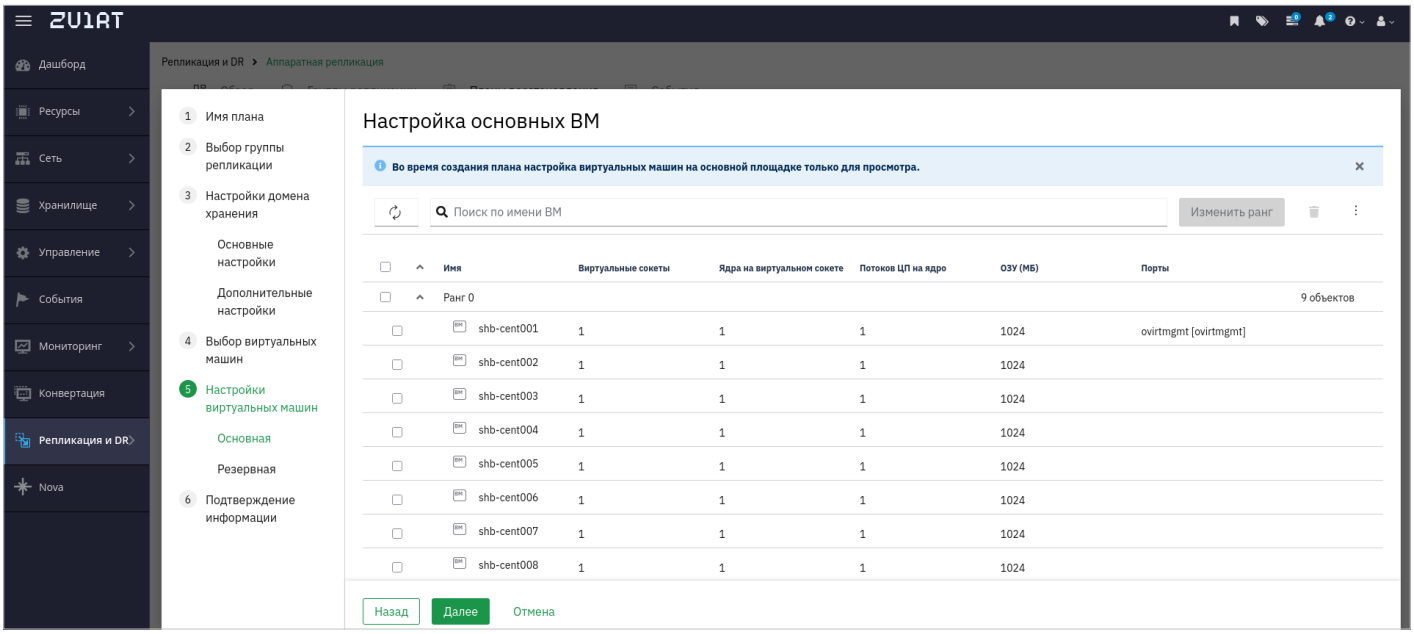


5. На вкладке **Выбор виртуальных машин** выберите те ВМ, которые планируется защитить от сбоев ОЦОД, перезапустив их в РЦОД. В план восстановления могут быть добавлены все имеющиеся ВМ или их часть.

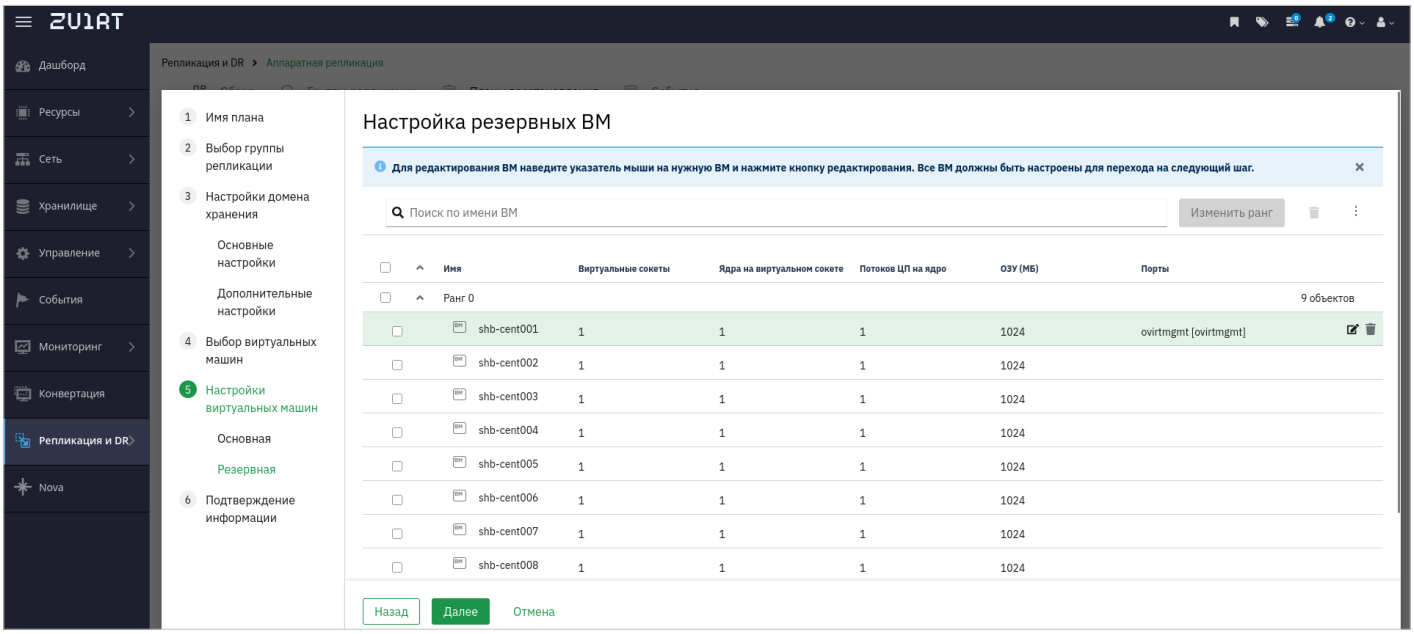
> **!** При выполнении планового переключения все ВМ, расположенные в домене хранения ОЦОД, будут выключены.




На вкладке **Настройки виртуальных машин** — **Основная** отображаются настройки выбранных ВМ на площадке ОЦОД.



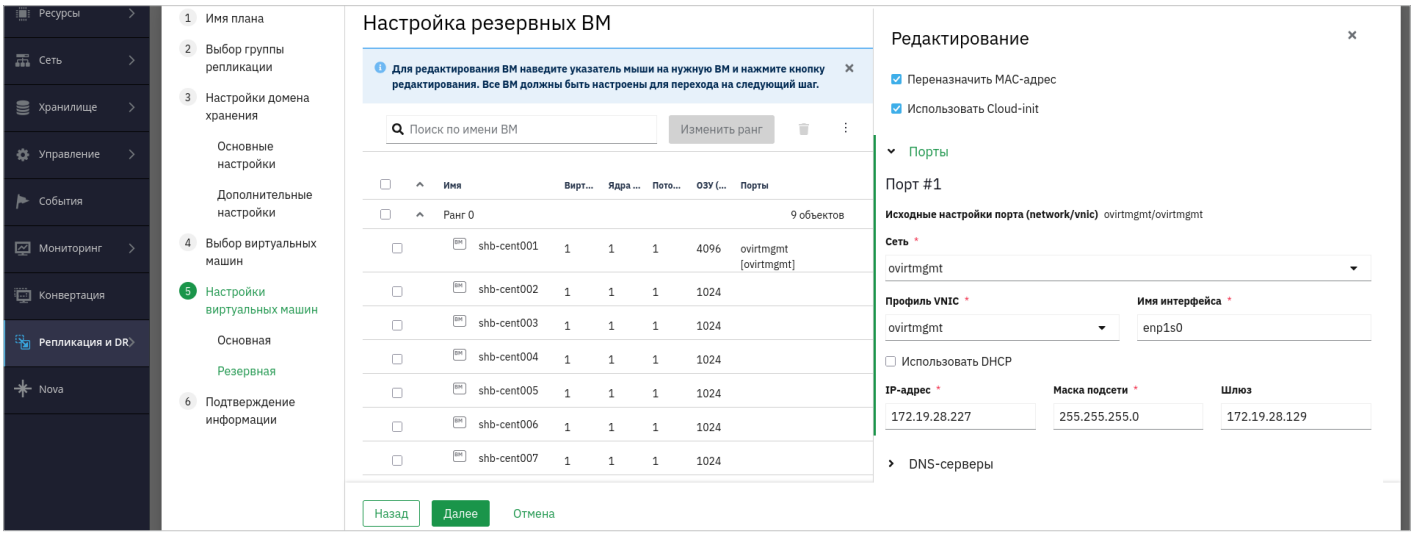

На вкладке **Настройки виртуальных машин** — **Резервная** отображается аналогичный список ВМ. При необходимости для каждой ВМ можно отредактировать параметры в контекстном меню свойств ВМ.



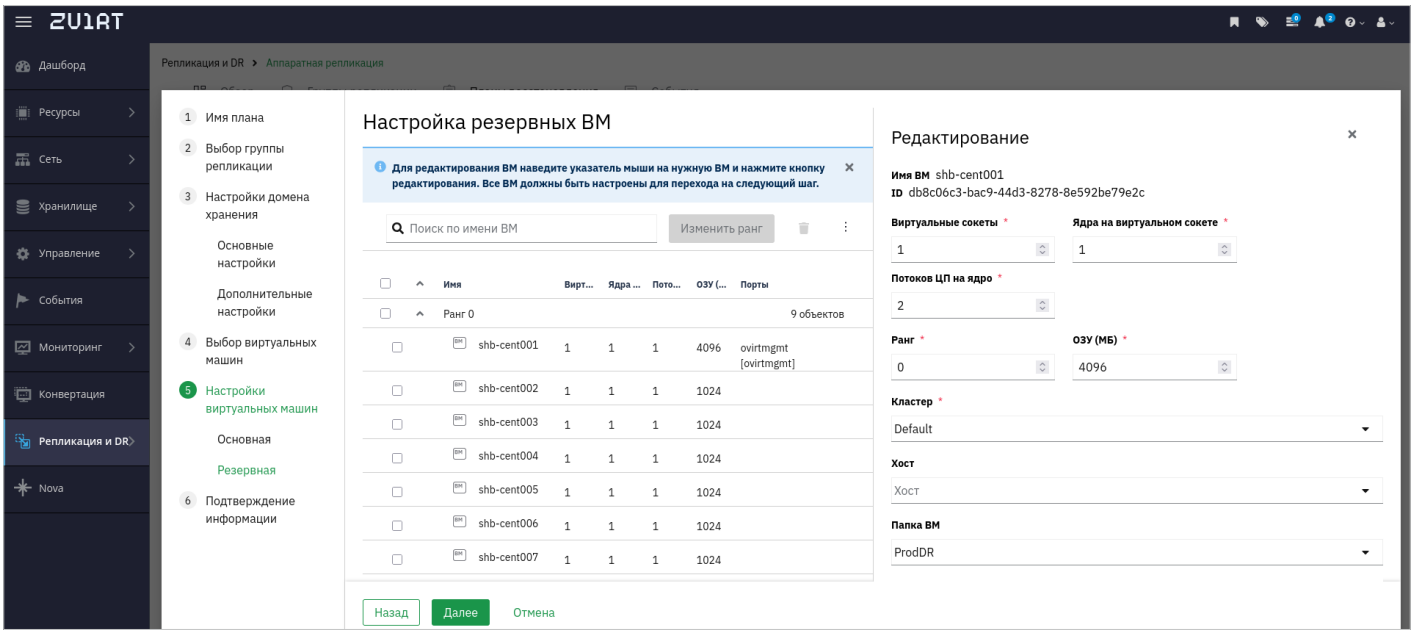

В примере приведены примеры заполненных параметров резервных ВМ, обеспечивающие гарантированный старт ВМ в РЦОД.

Для ВМ, использующих ОС Linux, доступно изменение параметров сети с помощью cloud-init.
Для назначения нового IP-адреса включите редактирование и установите флажок **Использовать Cloud-init**. Введите адрес и сопутствующие настройки в соответствующих окнах и сохраните. В примере новый IP-адрес назначен для ВМ `shb-cent001`.

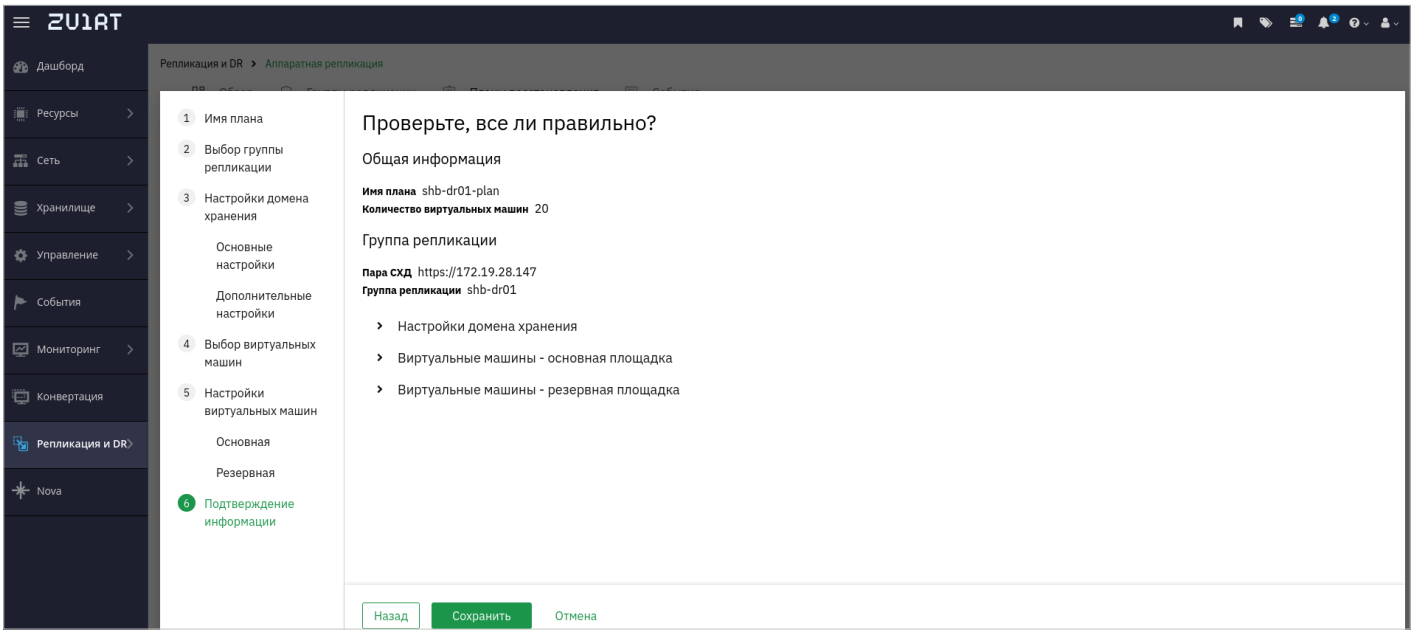
| НАСТРОЙКА РЕЗЕРВНЫХ ВМ | |
|---|---|
| Имя интерфейса | enp1s0<br>(указанное в данном поле имя должно совпадать с именем интерфейса в ОС ВМ) |
| IP-адрес | 172.19.28.227 |
| Маска подсети | 255.255.255.0 |
| Шлюз | 172.19.28.129 |



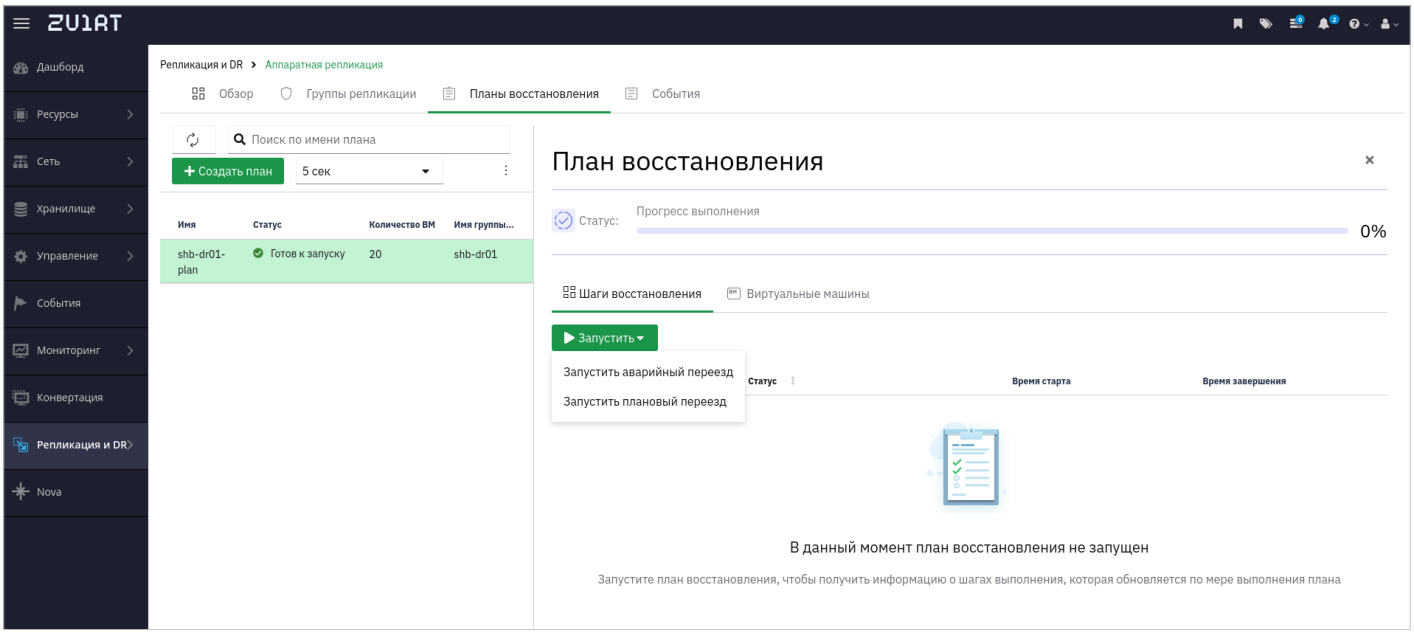
В примере новый IP-адрес назначен для ВМ `shb-cent001`.
Сохраните изменения.




Проверьте правильность настроек на вкладке **Подтверждение информации** и нажмите **Сохранить**.




На вкладке **План восстановления** появится новый план. Для каждого плана восстановления доступны опции аварийного и планового переводов ВМ.



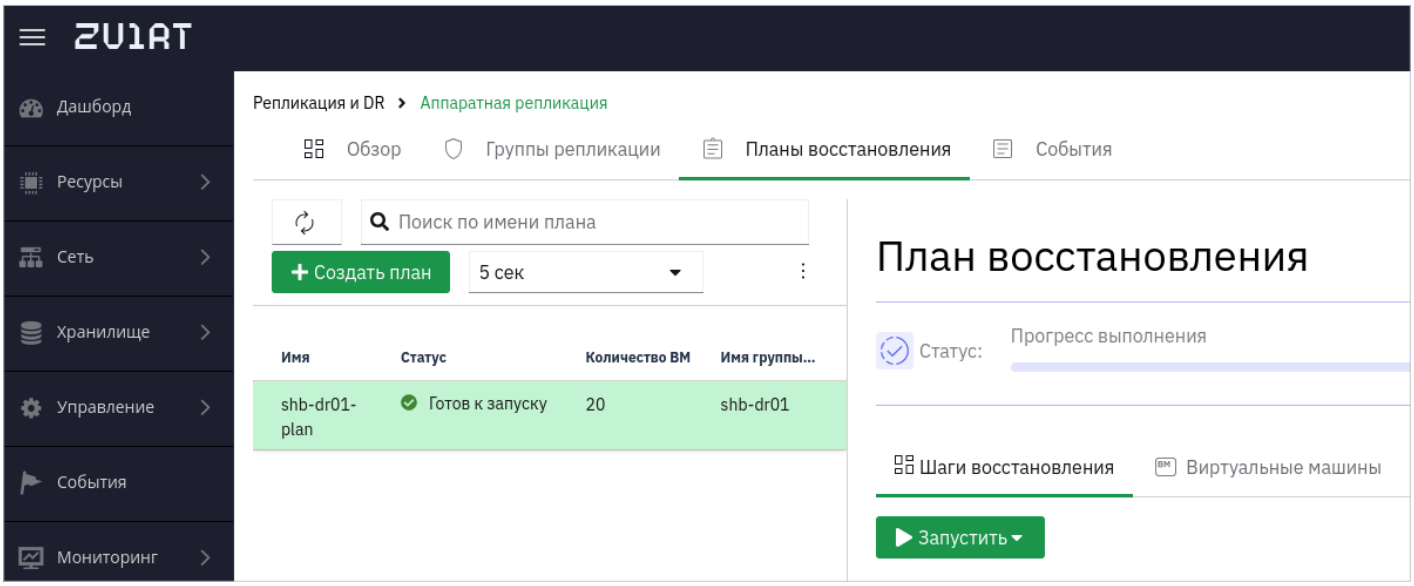

Этим шагом завершается конфигурирование защищенной инфраструктуры. Созданную конфигурацию ВМ можно переводить между ОЦОД/РЦОД запланировано или в случае сбоя одной из площадок.

# УПРАВЛЕНИЕ РЕПЛИКАЦИЕЙ

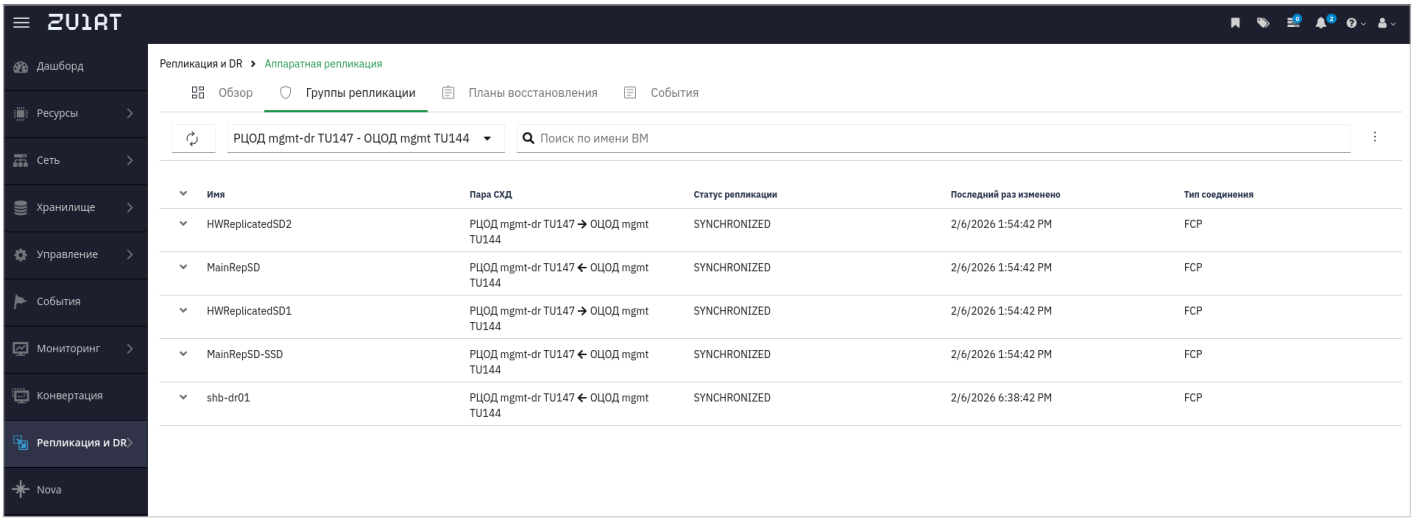
## Мониторинг и управление репликацией

После настройки планов восстановления инфраструктура виртуализации становится подготовлена к эксплуатации. Все компоненты и сервисы по умолчанию функционируют в штатном режиме. Для мониторинга сервиса аппаратной репликации следует обратить внимание на состояние планов восстановления.

1. Проверьте статус плана восстановления на вкладке **Планы восстановления**.

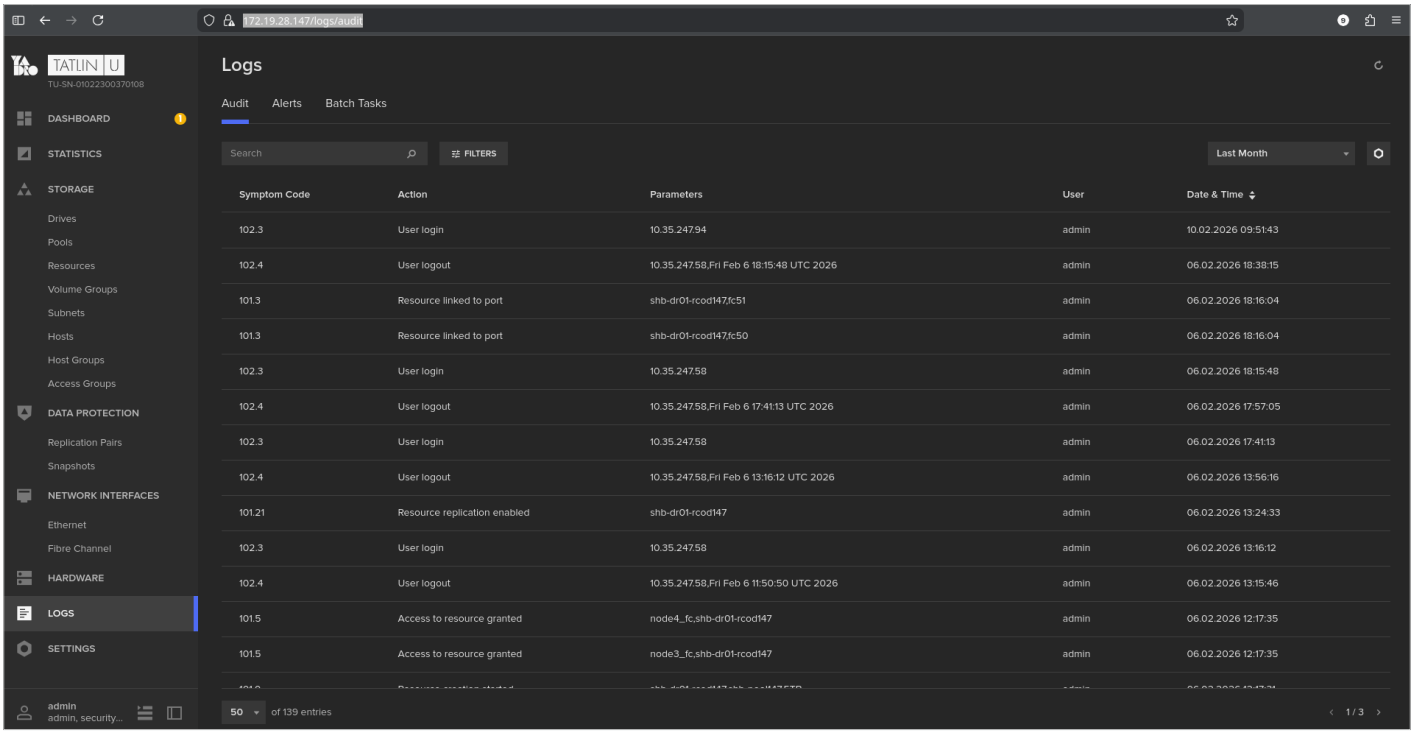


Статус плана хранения определяется статусом репликационных пар томов, для которых сконфигурирован DR. Состояние **Готов к запуску** означает, что план находится в рабочем состоянии и готов к запуску.
Если статус плана восстановления имеет значение **Предупреждение**, проверьте состояние пар репликации на вкладке **Группы репликации**.

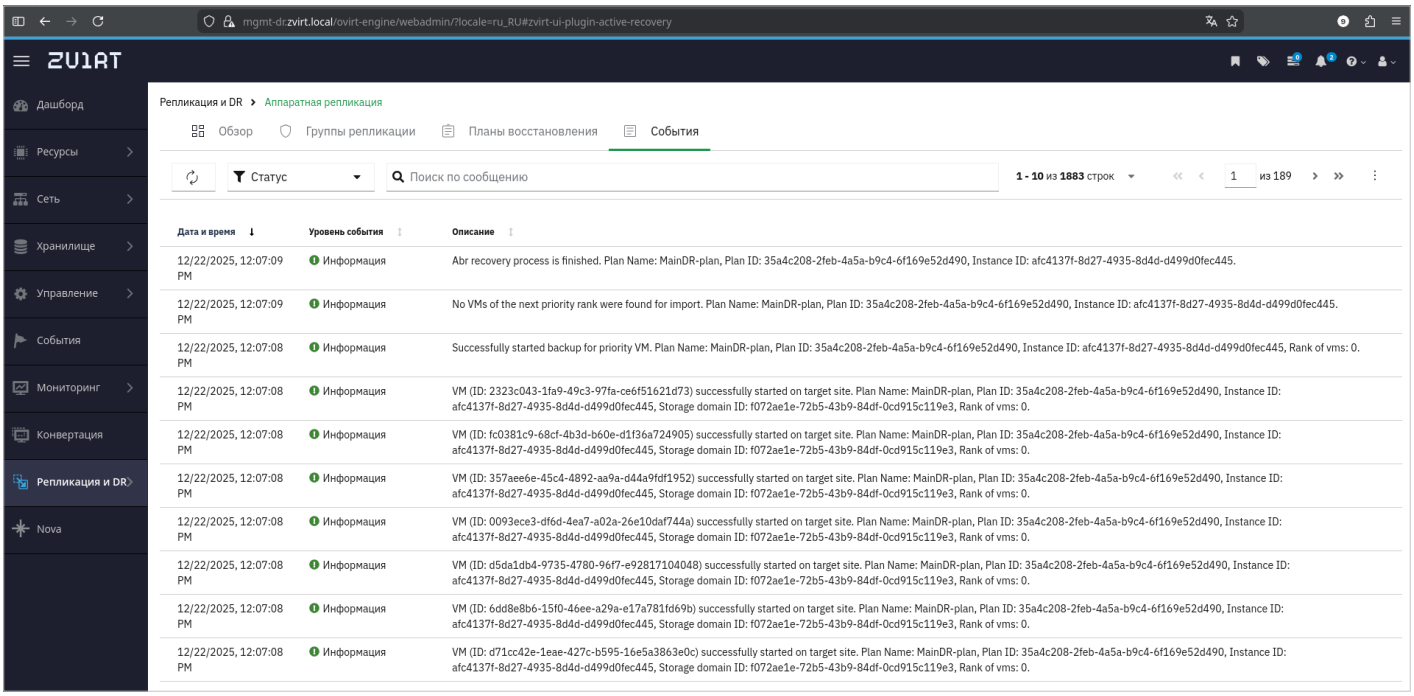



На вкладке **Группы репликации** публикуется информация о состоянии пар репликации.

2. Если статус пар отличается от **SYNCHRONIZED**, проверьте состояние пар со стороны СХД. Причинами могут выступать продолжающееся первичное копирование данных, сбой репликационной связи, выполненный средствами CX Failover и др.
Мониторинг СХД осуществляется с помощью консоли управления СХД (https://172.19.28.147/logs/audit) в разделе **Журналы (Logs)**. Сообщения группируются по вкладкам **Аудит (Audit)**, **Системные события (Alerts)** и **Пакетные задачи (Batch Tasks)**.



Для сервиса аппаратной репликации платформы виртуализации zVirt системные события и ошибки отображаются на вкладке **События**. На вкладке также отображается информация о шагах выполнения плана восстановления, времени и статусе выполнения каждого шага.




## Переключение направления репликации

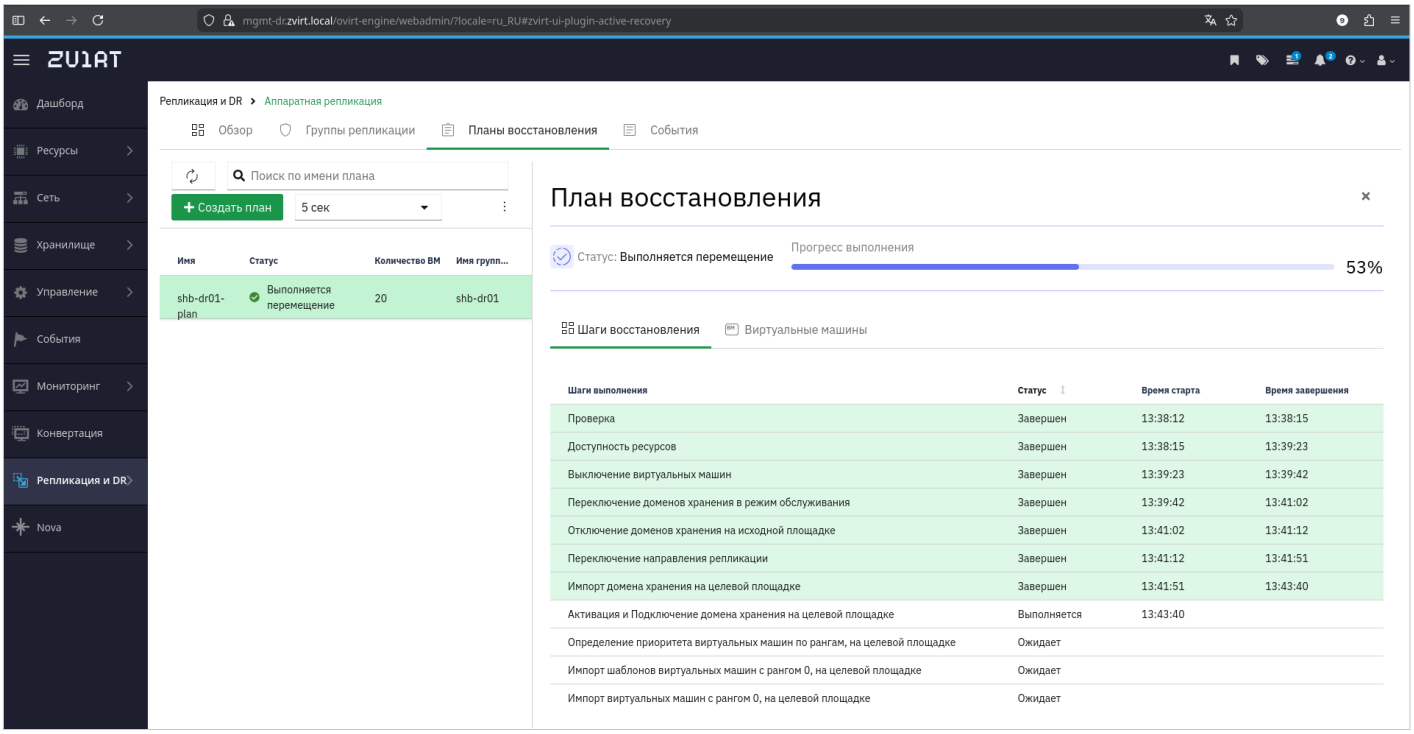
При использовании сервиса аппаратной репликации zVirt все операции по переключению ресурсов между площадками необходимо выполнять в интерфейсе zVirt. Выполнение переключения ресурсов между площадками средствами СХД приведет к рассогласованию статусов конфигурации и недоступности ресурсов. Все операции с парами репликации выполняются в разделе **Репликация и DR**.

# АВАРИЙНОЕ И ПЛАНОВОЕ ВОССТАНОВЛЕНИЕ

## Плановое переключение сервиса

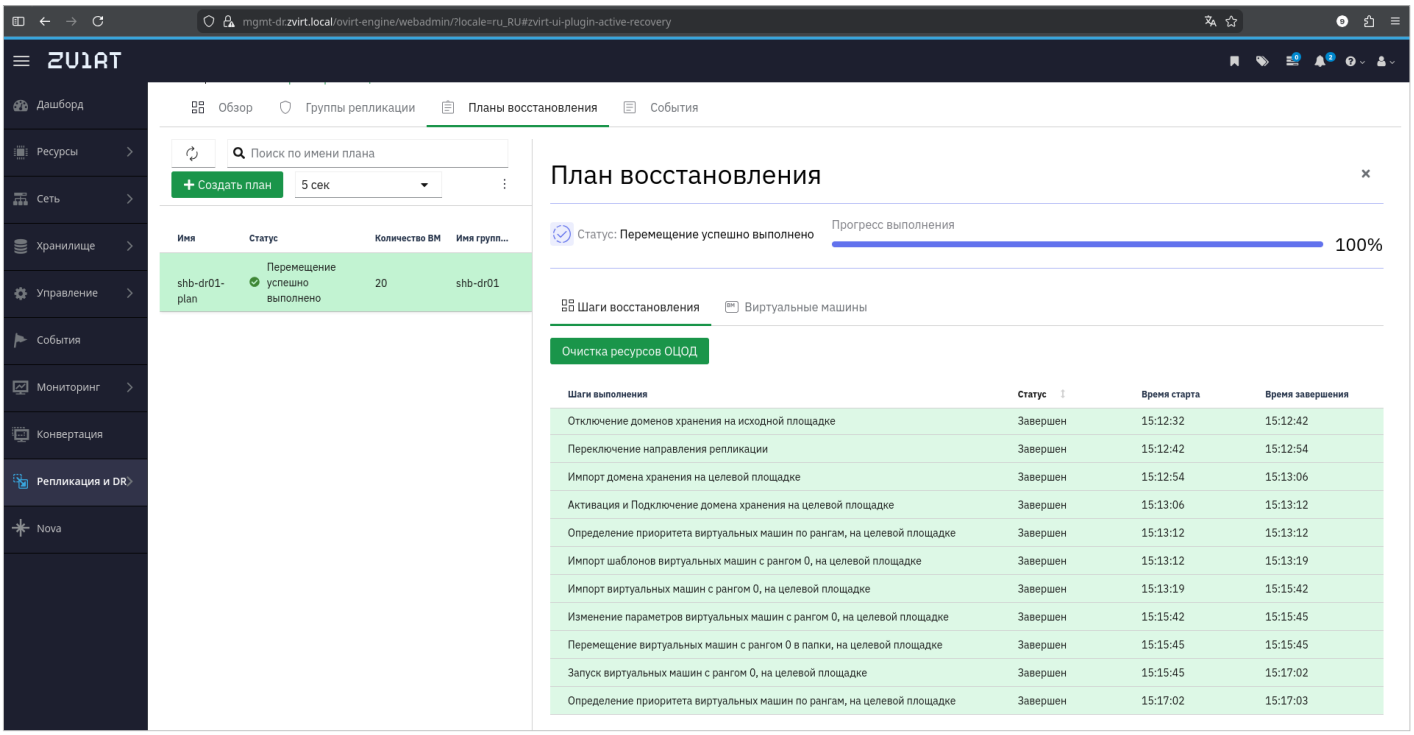
Плановое переключение ресурсов характеризуется отсутствием сбоев в системах и полной доступностью ресурсов. Чтобы выполнить плановое переключение сервиса:

1. В интерфейсе zVirt перейдите на вкладку **Репликация и DR** — **Аппаратная репликация** — **Планы восстановления** и выберите план, для которого планируется переключение.
   Справа появится кнопка **Запустить**.
2. Предусмотрено два варианта запуска плана. Выберите **Запустить плановый переезд** и подтвердите действие, нажав **Да**.
   Шаги и фазы выполнения плана восстановления отображаются в окне мониторинга процесса.

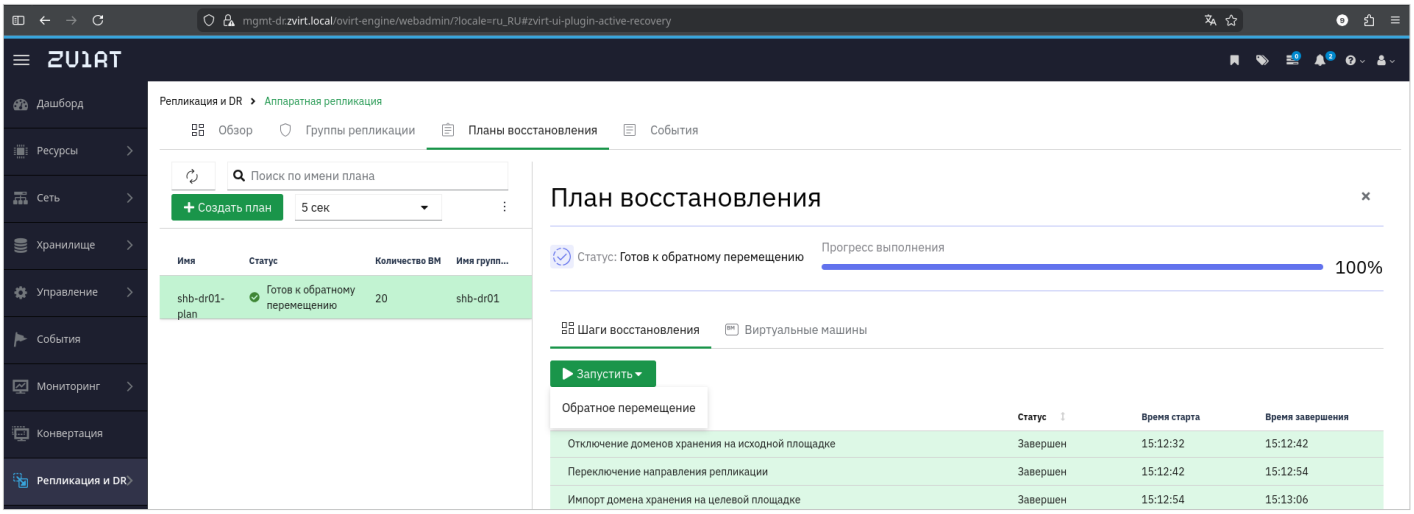



Во время выполнения плана ВМ, затронутые восстановлением, выключаются на площадке ОЦОД. Для этого процедура отправляет запрос на выключение ВМ средствами ОС. После отправки запроса включается таймаут ожидания, выставленный по умолчанию в настройках плана восстановления. По умолчанию таймаут равен пяти минутам. По истечении этого времени ВМ принудительно отключаются по питанию.

В результате работы плана на СХД произойдет переключение направления репликации и активируются площадка РЦОД и ВМ, хранящиеся на СХД РЦОД. При загрузке будут использоваться параметры и настройки, указанные в активированном плане восстановления.




3. После успешного выполнения планового перемещения нажмите на кнопку очистки ресурсов. Это удалит ненужные ресурсы с ОЦОД при наличии. При штатной работе планового перемещения данная кнопка просто выполнит проверки ресурсов в ОЦОД.



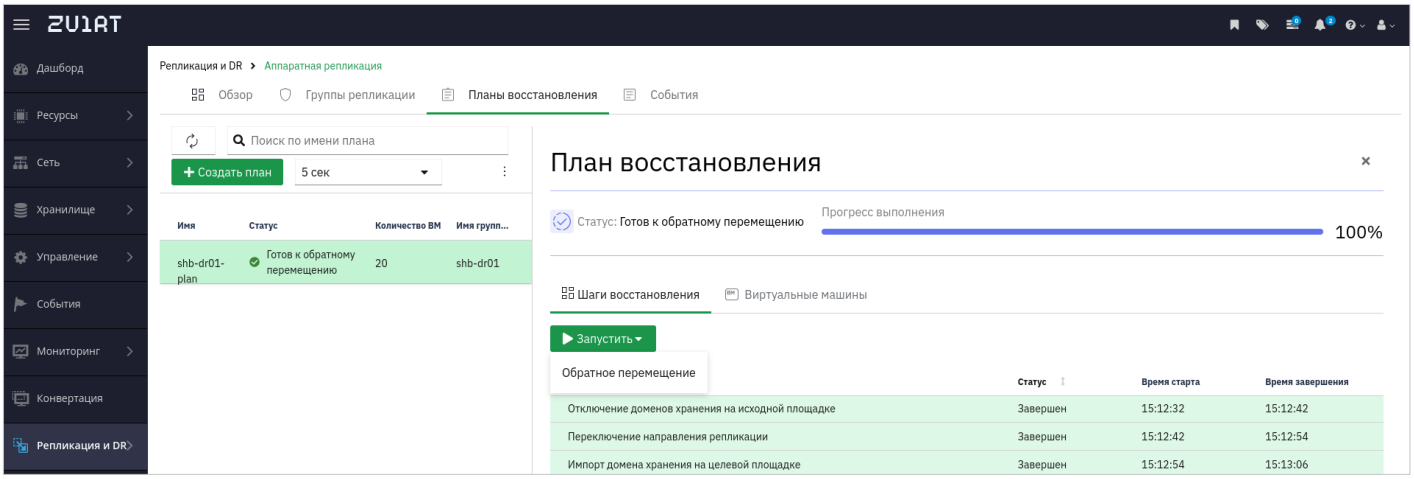

В результате описанных действий конфигурация будет подготовлена к обратному перемещению в рамках того же плана восстановления. Для выполнения обратного перемещения нажмите **Обратное перемещение**.

## Обратное перемещение после планового переключения сервиса

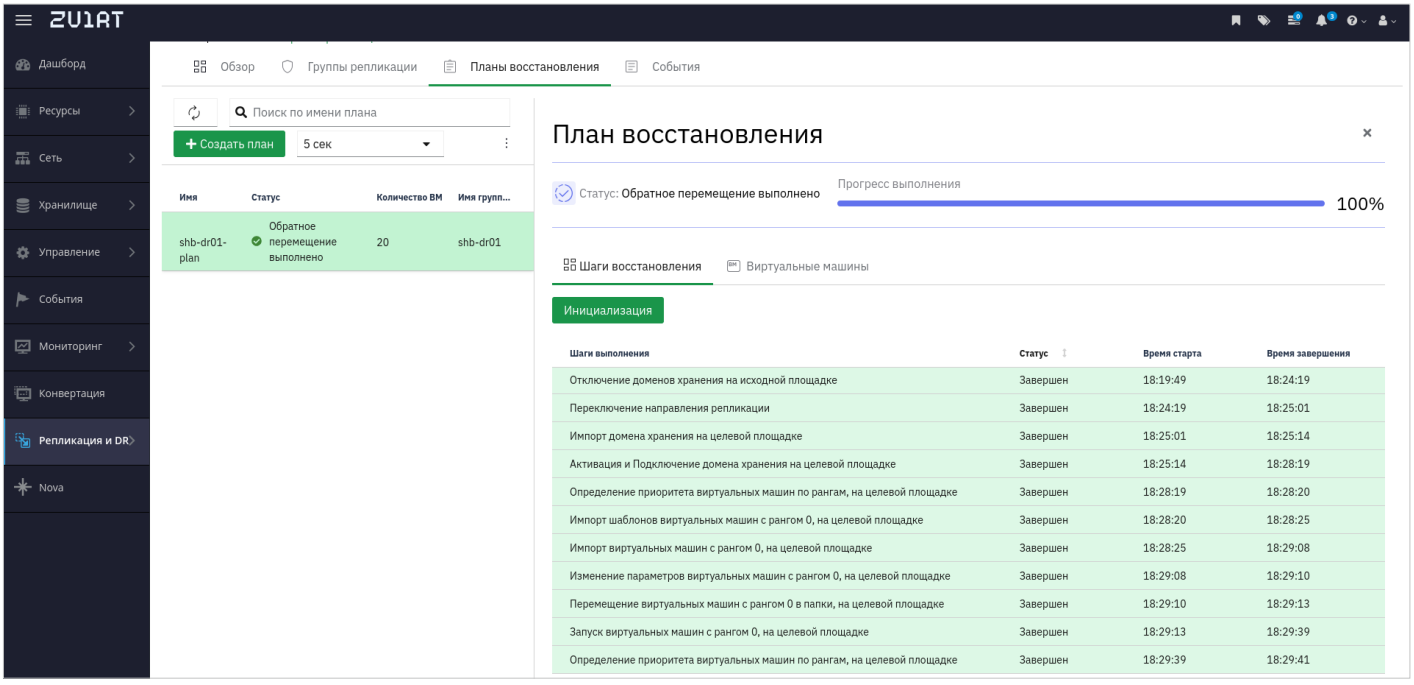
> **!** При обратном перемещении создавать отдельный план восстановления на противоположной площадке настоятельно не рекомендуется. Обратное перемещение должно выполняться в рамках того же плана, по которому выполнялось плановое переключение.

Для возврата ВМ на площадку ОЦОД:

1. Запустите процедуру возврата, нажав **Обратное перемещение**. ВМ на РЦОД будут выключены данной процедурой.
   В ходе процедуры произойдет отключение сервисов хранения на площадке РЦОД и разворот направления репликации (РЦОД → ОЦОД изменится на ОЦОД → РЦОД). Таким образом, площадка ОЦОД станет основной стороной, а РЦОД — резервной. Стартуют сервисы хранения, расположенные на СХД ОЦОД. Будет произведен запуск ВМ на площадке ОЦОД.




Более подробно шаги в процессе выполнения будут показаны в интерфейсе.



После выполнения плана ВМ будут включены и активны на площадке ОЦОД.

2. Для возвращения плана восстановления в исходное состояние нажмите **Инициализация**. Запустится процесс возврата настроек плана восстановления к исходным значениям. После окончания процесса плановое и аварийное восстановление снова станет доступно.

### Аварийное восстановление

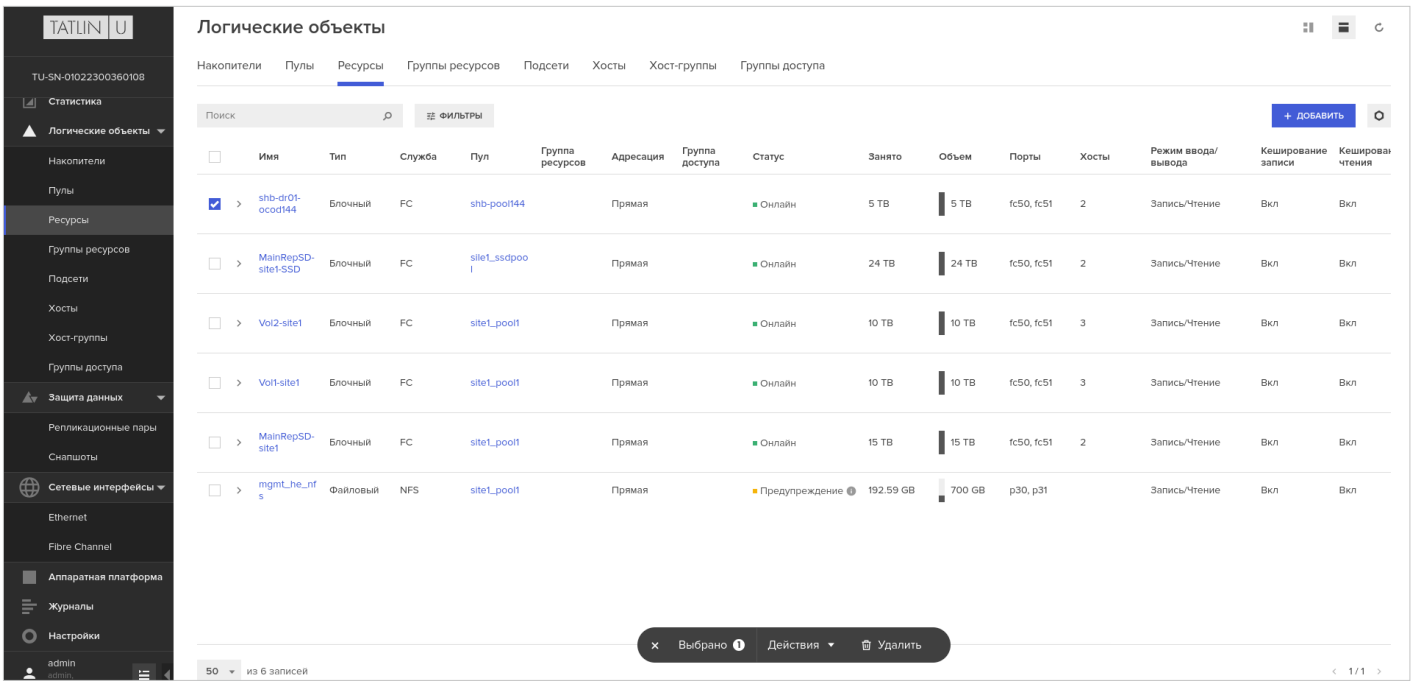
Аварийное восстановление применяется при возникновении сбоя или цепи сбоев на основной площадке. Например, при полной потере электропитания в ОЦОД, отсутствии доступа к сети или другой подобной аварии, требующей длительного устранения.

Для проведения данного теста используется физическое отключение оборудования от электропитания на площадке ОЦОД. В качестве варианта возможна имитация отказа одного из важных узлов, например СХД (отключение ресурсов от портов серверного ввода-вывода как наиболее щадящее действие).

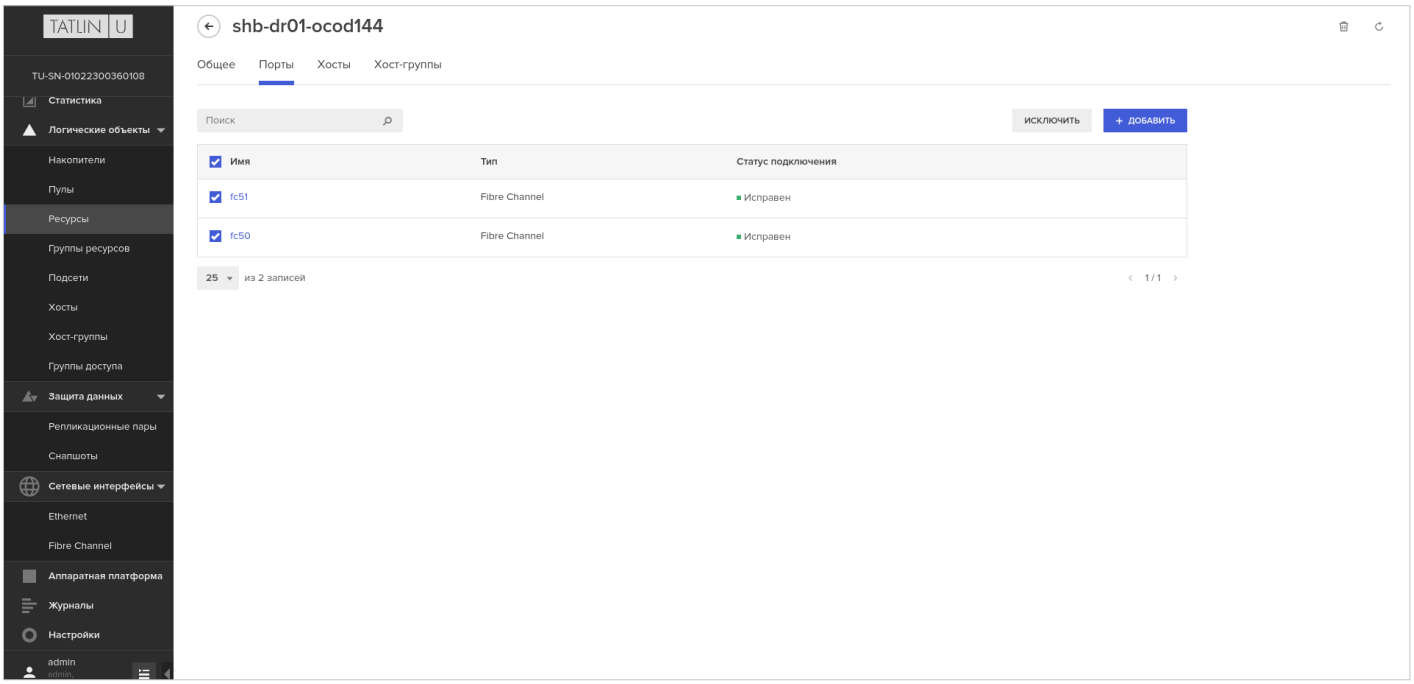
Проведение тестирования допустимо при следующих условиях:

- Репликация направлена с основной площадки на резервную (ОЦОД→ РЦОД);
- Группа репликации, указанная в плане, существует на СХД. Пары репликации находятся в статусе **Synchronized** (допускаются статусы **Interrupted** и **Suspended**);
- Отсутствуют запущенные планы восстановления с группой репликации, указанной в плане восстановления;
- План восстановления не запущен;
- На резервной площадке нет активных ресурсов (виртуальные машины отсутствуют, домен хранения откреплен или удален).

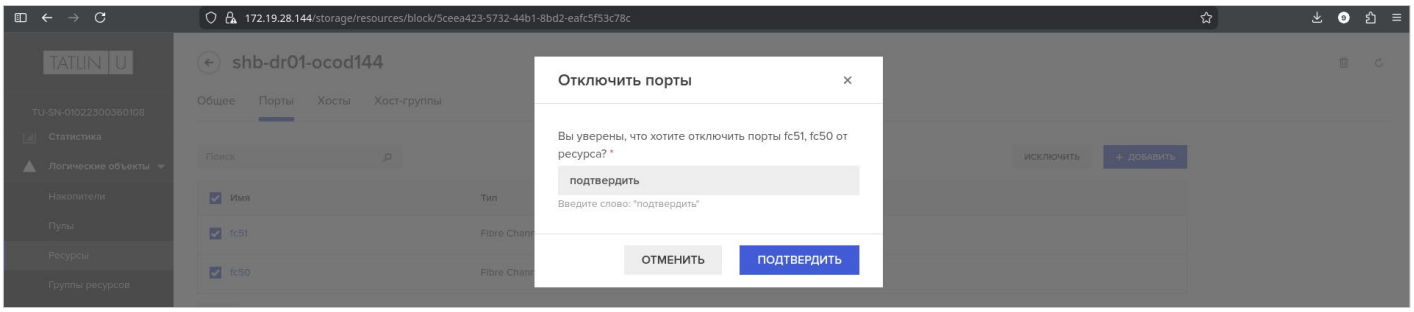
1. Отключите ресурс от портов ввода-вывода. Для этого:
   1.1 Перейдите в веб-интерфейс СХД основной площадки на вкладку **Логические объекты** — **Ресурсы** (https://172.19.28.144/storage/resources).




1.2 Выберите ресурс `shb-dr01-ocod144` и на вкладке **Порты** исключите оба FC-порта.

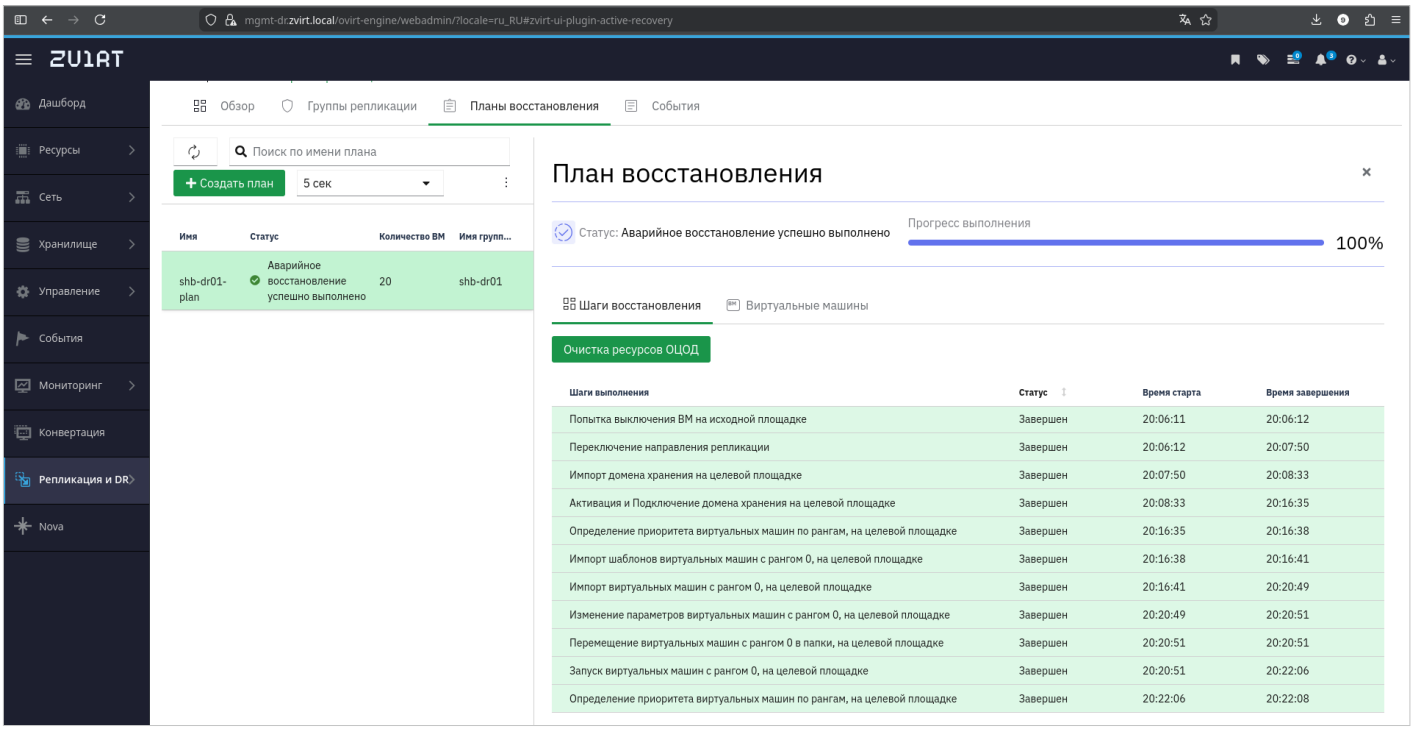



1.3 Подтвердите свой выбор по требованию системы.

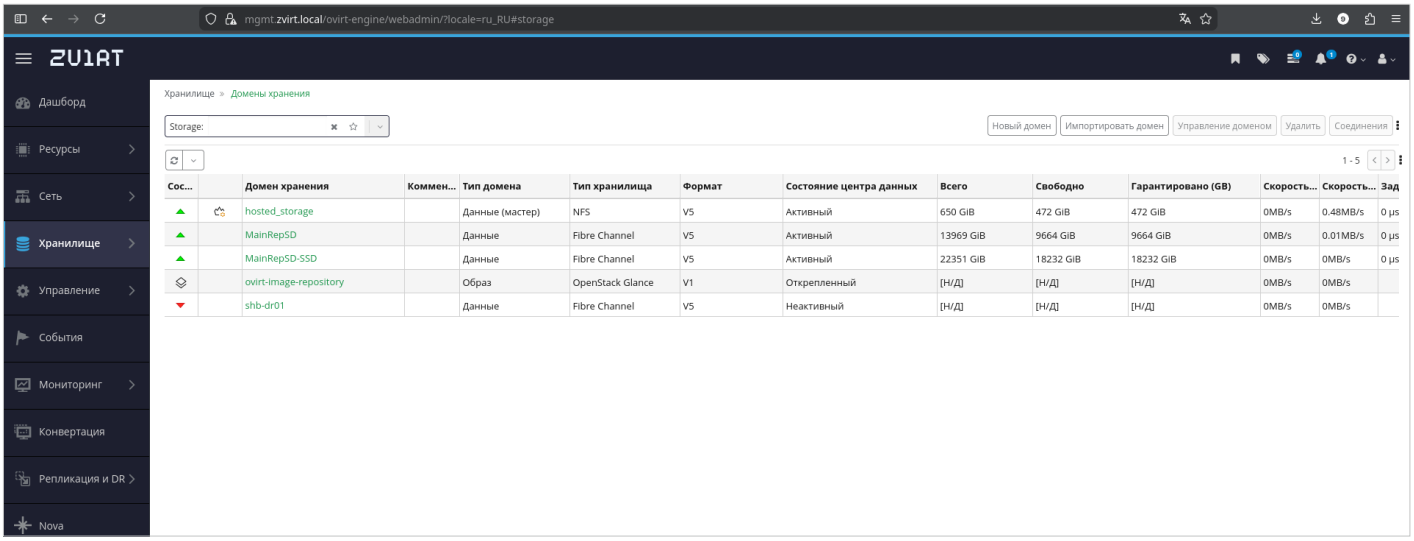



После отключения ресурса на СХД ОЦОД от портов ввода-вывода расположенные на нем ВМ прекратят работу (зависнут).

1. Для восстановления работоспособности ВМ стартуйте их на площадке РЦОД. Для этого:
   1.1 В интерфейсе zVirt перейти на вкладку **Репликация и DR** — **Аппаратная репликация** — **Планы восстановления**.
   1.2 Выберите план, для которого планируется выполнить переключение и выделите его. Справа появится кнопка **Запустить**.
   1.3 Выберите **Аварийный режим** для запуска плана восстановления.




После запуска плана будут выполнены все ранее перечисленные шаги процедуры перевода ресурсов между площадками: выключение ВМ, переключение направления репликации, активация сервисов хранения, применение параметров плана восстановления и старт ВМ. После успешного выполнения всех шагов процедуры ВМ успешно стартуют на площадке РЦОД. Сервисы хранения определят хранилище в ОЦОД неактивным.



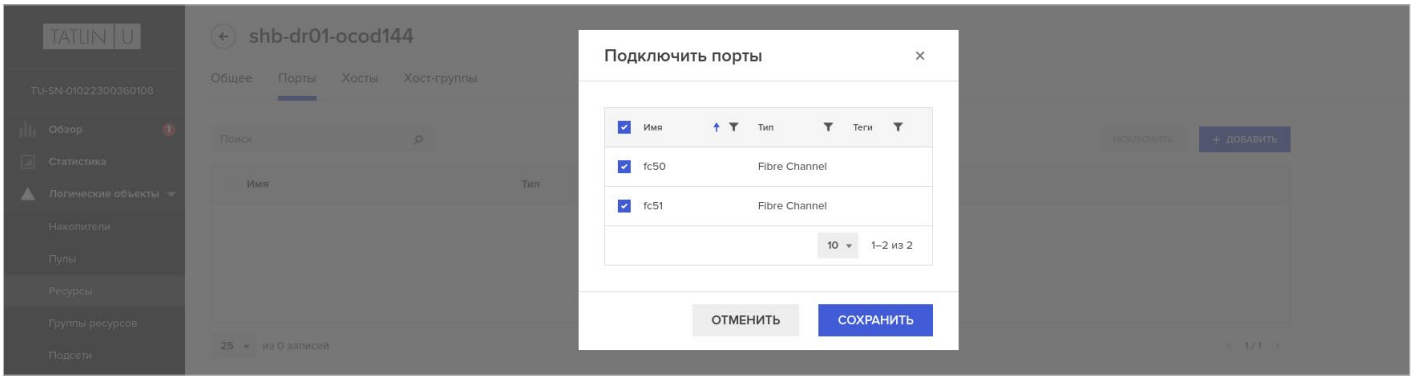

После запуска процедуры очистки ресурсов ОЦОД хранилище `shb-dr01` будет отключено от системы виртуализации на стороне ОЦОД, чтобы при запуске обратного перемещения его можно было успешно подключить обратно.

## Возврат после аварийного переключения

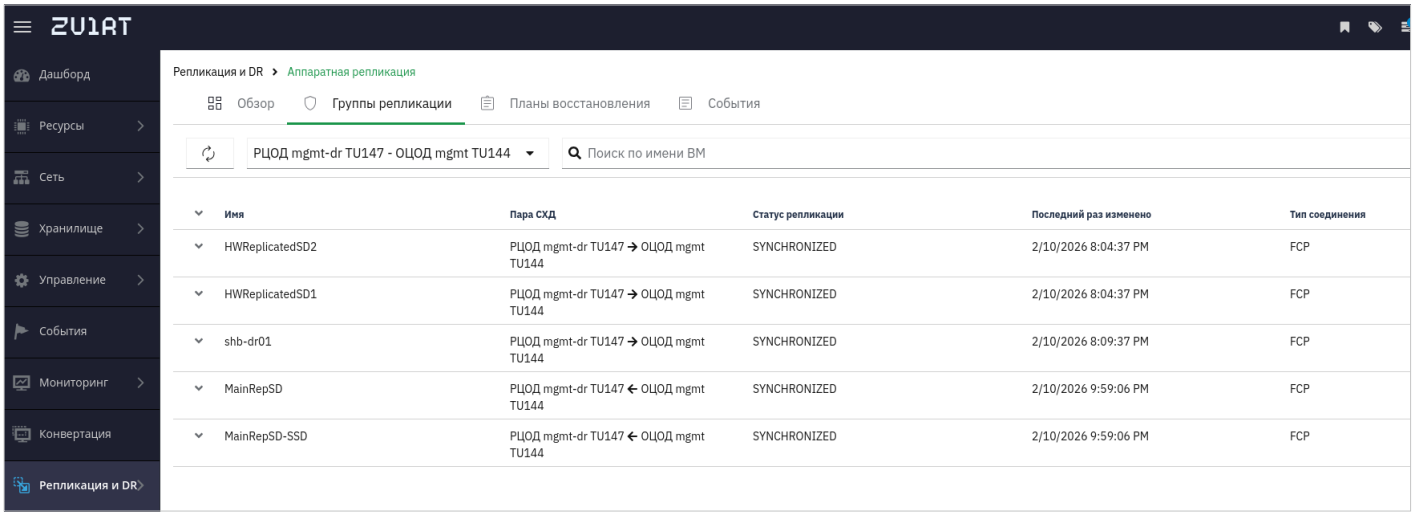
> **!** В тестировании не производилось отключений СХД, поэтому статус пар репликации не изменялся.
> В случае более сложного отказа (потеря питания ОЦОД, сбой контроллера СХД, отключение сетевых коммутаторов) перед началом процедуры возврата ВМ в ОЦОД требуется восстановление пар репликации. Эти работы выполняются на стороне СХД.
>
> При разрыве пары репликации требуется ее пересоздание и полное копирование данных. В примере копирование пошло бы в направлении РЦОД → ОЦОД. После завершения полного копирования и получения статуса пары **Синхронизировано** можно начинать процесс возврата ВМ по описанной ниже процедуре.

Для возврата ресурсов в ОЦОД после ликвидации последствий аварии, как и в случае планового перевода ресурсов между площадками, выполняется процедура обратного перемещения.

1. Для запуска процедуры нажмите **Обратное перемещение**.
   В примере на стороне ОЦОД отключались только порты ввода-вывода для ресурса `shb-dr01-ocod144`.
   Для ликвидации последствий симуляции аварии необходимо вернуть доступ к указанному ресурсу через FC-подключение.
2. В консоли управления СХД (https://172.19.28.144/storage/resources) для ресурса `shb-dr01-ocod144` на вкладке **Ресурсы** — **Порты** нажмите **ДОБАВИТЬ** и установите в окне **Подключить порты** флажки **fc50** и **fc51**.
   Нажмите **СОХРАНИТЬ**.




Поскольку физических сбоев на СХД ОЦОД не имитировалось, статус ресурса в паре репликации не изменился.
На вкладке **Группы репликации** должно отображаться корректное направление репликации (РЦОД → ОЦОД) и статус **SYNCHRONIZED**.



3. Для возврата ВМ на площадку ОЦОД на вкладке **Планы восстановления** нажмите **Запустить** — **Обратное перемещение** напротив плана `shb-dr01-plan`.
   После успешного завершения выполнения плана ВМ будут запущены в ОЦОД. Станет возможным запуск процесса инициализации параметров и установки их исходных значений.
4. При необходимости запуска инициализации нажмите **Инициализация**.
   Более подробно конфигурирование и управление репликацией описаны в документации YADRO и Orion soft.

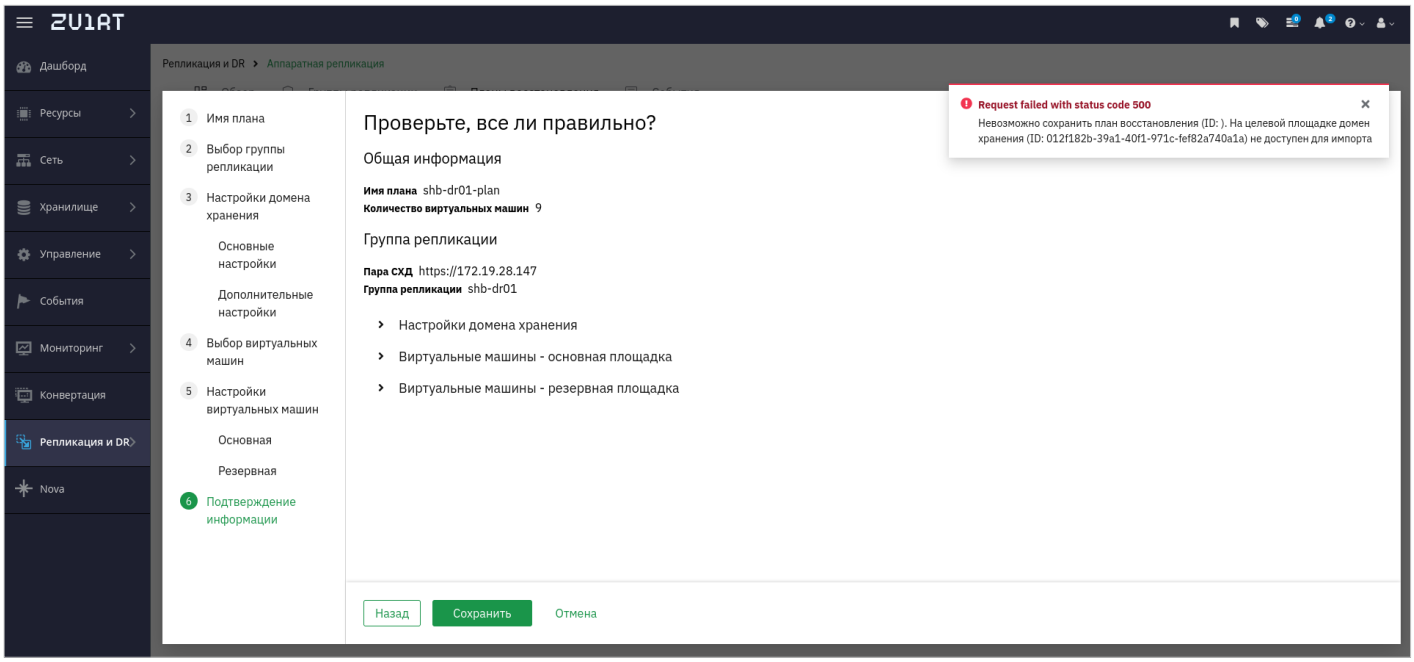
   **YADRO**:
   - Описание функциональности: репликация

   **Orion soft**:
   - Репликация и DR. Программная репликация

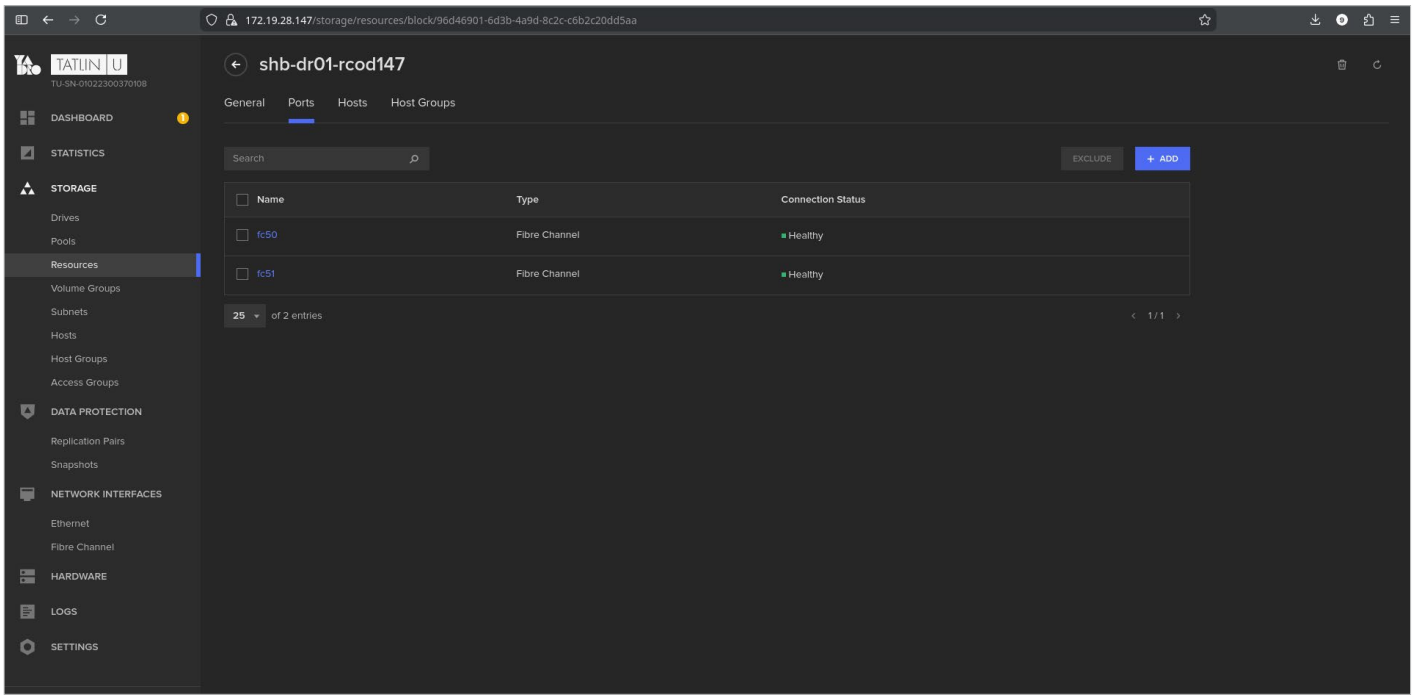
# УСТРАНЕНИЕ НЕИСПРАВНОСТЕЙ

## Распространенные проблемы и решения

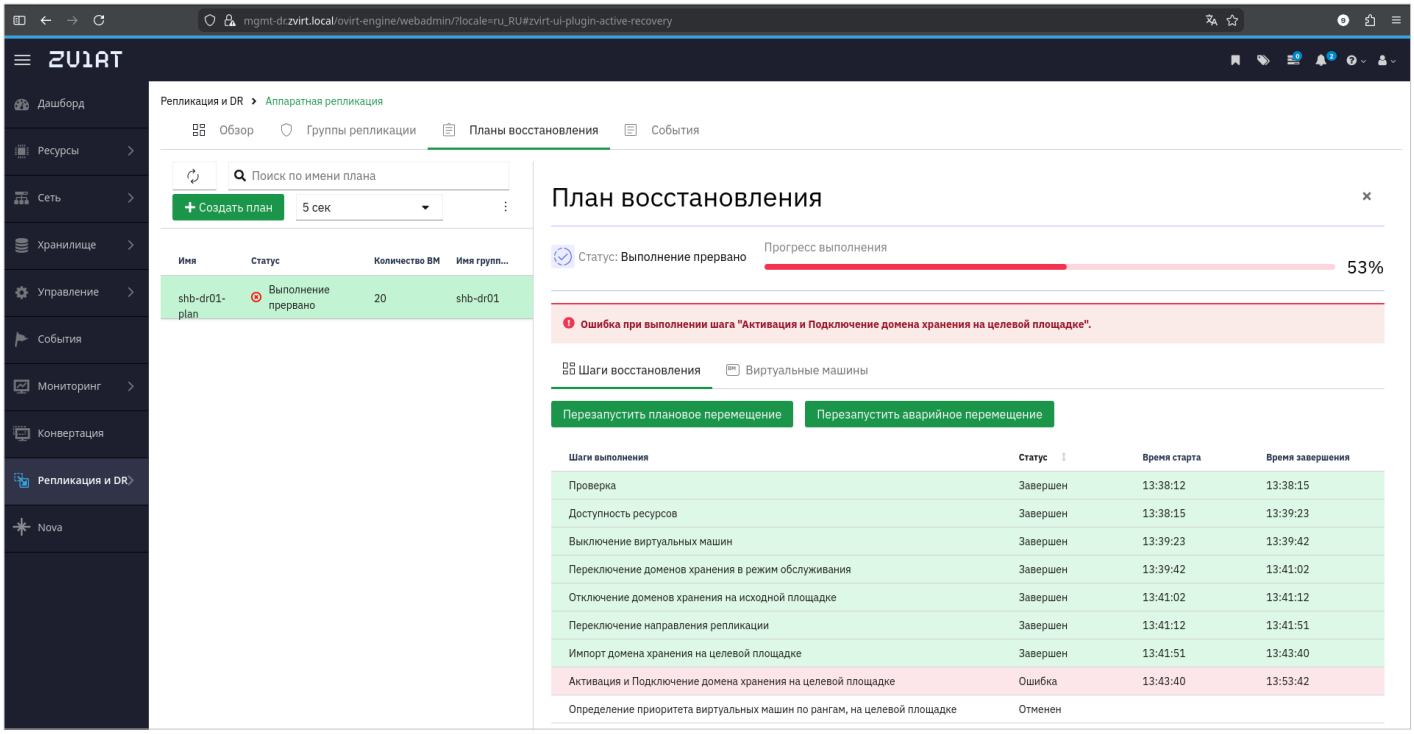
Если при сохранении плана возникает ошибка недоступности домена хранения для импорта, проверьте на СХД, презентован ли ресурс хостам в РЦОД.



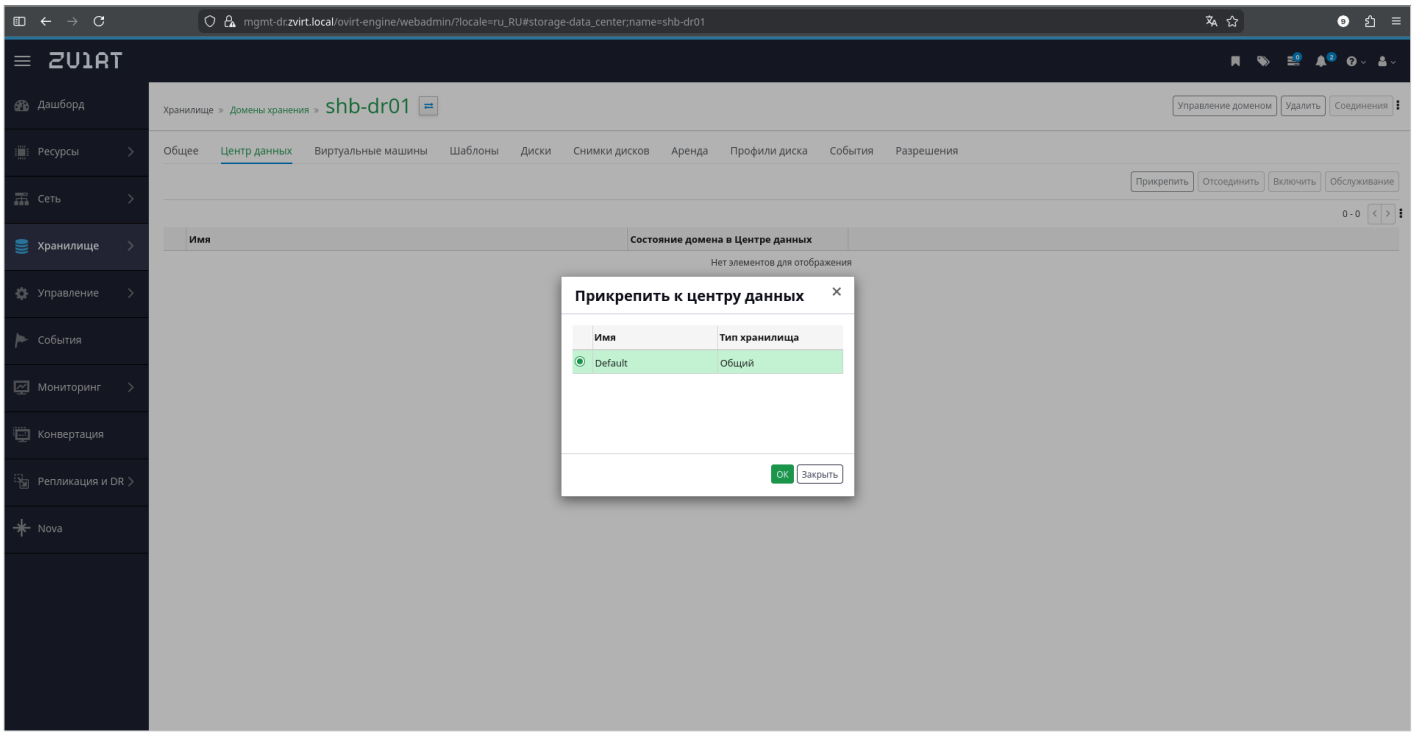

В примере обнаружено, что для ресурса не были выбраны порты, по которым должен предоставляться доступ.




План при сохранении проверяет, доступен ли ресурс на резервной стороне для подключения к виртуализации. Если при выполнении планового перемещения возникла ошибка, то процесс может быть перезапущен. Выполнение плана начнется с первого шага, но те шаги, которые были пройдены успешно, будут проверены и пропущены.
Перед повторным запуском плана рекомендуется определить причину предыдущего сбоя и устранить ее.




В примере показана ошибка при активации домена хранения. Причину подобных ошибок следует искать на стороне хостов виртуализации. В данном случае причина в некорректно настроенном multipath.




В примере после устранения причины повторный запуск плана успешно прошел этап подключения домена хранения.

### Логи и диагностика

Логи сервиса аппаратной репликации расположены на менеджерах управления `mgmt.zvirt.local` и `mgmt-dr.zvirt.local`:
Для версии 4.4 и более ранних версий логи хранятся в `/var/log/backend/backend.log`.
Начиная с версии 4.5 хранятся в `/var/log/zvirt-engine-backend/backend.log`.
Лог следует просматривать на том менеджере управления, с которого осуществлялся запуск плана.

## МЕТРИКИ ПРОИЗВОДИТЕЛЬНОСТИ

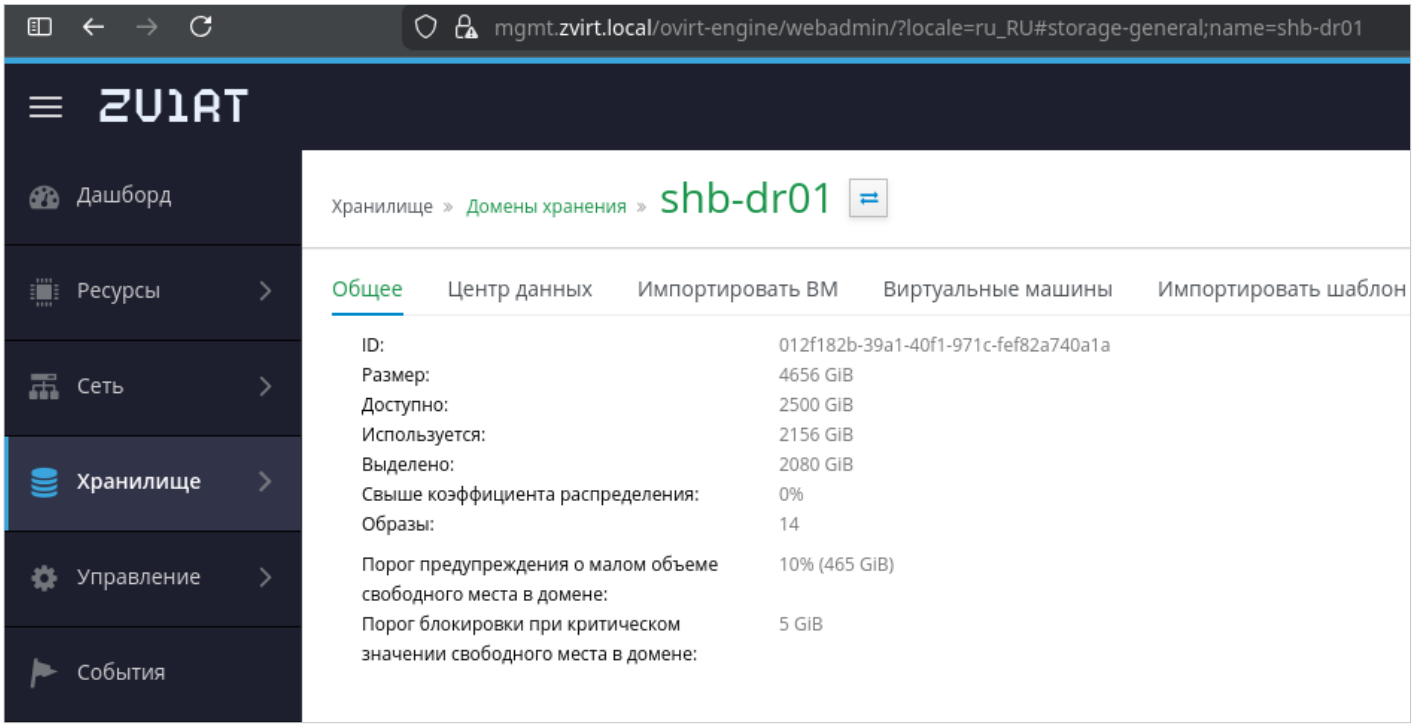
### Описание тестовой инфраструктуры

Тестирование производительности подсистемы ввода-вывода осуществляется при помощи утилиты `fio` в гостевой ОС CentOS.

Целью тестирования является анализ влияния синхронной репликации на производительность дисковой подсистемы.

Тестирование выполняется на домене хранения FC. Тестовую нагрузку создают 4 ВМ. ВМ выполняют запись и чтение только с одной СХД. Данные в фоне синхронизируются на вторую СХД для теста с включенной репликацией. В ином случае синхронизация приостанавливается для получения результатов без синхронной репликации. Профиль тестирования одинаковый для обоих тестов.

С обеих сторон пулы состоят из 4 дисков SSD для данных, 1 диска SSD для четности и 1 диска SSD Hot-swap. Диски предоставлены хостам по FC. На стороне СХД применяются карты 32 Гбит/с. На стороне хоста подключено 2 порта по 32 Гбит/с. Линк репликации работает на Ethernet и состоит из 2 линков по 25 Гбит/с.




Тестирование выполнялось на ВМ:

- `shb-app01`;
- `shb-app02`;
- `shb-app03`;
- `shb-app04`.

Для каждой ВМ использовался диск для тестирования объемом 500 ГБ, диски формата raw, preallocated, интерфейс VirtIO.

### Результаты тестирования производительности I/O инфраструктуры на 4 ВМ, с одним дисковым устройством и профилем: 70% чтения и 30% записи

Было проведено 2 типа тестов с `fio` в конфигурациях: без репликации и с синхронной репликацией. Подробный отчет `fio` находится в Приложении. Ниже представлены полученные результаты при профиле нагрузки 70 % чтения и 30 % записи.

Репликация привела к незначительному снижению суммарной пропускной способности домена хранения. Агрегированная скорость чтения снизилась с 988 МиБ/с до 948 МиБ/с (упала на ~4 %). Скорость записи снизилась с 423 МиБ/с до 406 МиБ/с (упала на ~4 %).

Данное снижение обусловлено потреблением части ресурсов СХД и сетевой инфраструктуры на операции синхронного дублирования данных на удаленный узел.

Важным параметром являются задержки:

- **Операции записи**:
  Средняя задержка (clat): 665 мкс (без репликации — 522 мкс); Увеличение задержек записи соответствует архитектурным особенностям синхронной репликации, при которой каждая операция записи подтверждается только после завершения записи на пассивном узле;
- **Операции чтения**:
  Средняя задержка: 782 мкс (без репликации — 764 мкс).

### Сравнительная таблица метрик профилем 70% чтения и 30% записи

| Параметр | Без репликации | С репликацией | Изменение |
|---|---|---|---|
| Пропускная способность чтения (MiB/s) | 988 | 948 | ▼ 4.0 % |
| Пропускная способность записи (MiB/s) | 423 | 406 | ▼ 4.0 % |
| IOPS чтения | 126000 | 121000 | ▼ 4.0 % |
| IOPS записи | 54200 | 52000 | ▼ 4.0 % |
| Средняя задержка чтения (мкс) | 764 | 782 | ▲ 2.3 % |
| Средняя задержка записи (мкс) | 522 | 665 | ▲ 27.4 % |
| Утилизация домена хранения (%) | 100 | 100 | 0 % |

### Сравнение метрик производительности I/O инфраструктуры на 4 ВМ и профилем 100% записи

### Сравнительная таблица метрик профилем 100% записи

| Параметр | Без репликации | С репликацией | Изменение |
|---|---|---|---|
| Пропускная способность записи (MiB/s) | 7150 | 5679 | ▼ 20.6 % |
| IOPS записи | 114500 | 90900 | ▼ 20.6 % |
| Средняя задержка записи (мкс) | 1173 | 1408 | ▲ 19.5 % |
| Утилизация домена хранения (%) | 100 | 100 | 0 % |

Включение синхронной репликации приводит к предсказуемому снижению общей пропускной способности и количества операций ввода-вывода (IOPS) на 20,6 %. Суммарная производительность снизилась с 7150 МиБ/с до 5679 МиБ/с. Это связано с тем, что каждая операция записи теперь должна быть подтверждена не только локальным пулом, но и удаленной системой хранения, что увеличивает время отклика.
Репликация оказывает влияние на задержки. Средняя задержка увеличилась на 0,3 мс.

## ЗАКЛЮЧЕНИЕ

### Рекомендации по эксплуатации

Проведенное тестирование подтвердило работоспособность решения распределенной кластеризации виртуальной среды zVirt на базе репликации СХД TATLIN.UNIFIED. Были протестированы и подтверждены возможности как планового, так и аварийного перевода ресурсов с основной площадки на резервную. Также были протестированы и подтверждены возможности возврата ресурсов на исходную площадку после ее восстановления от сбоя.

Приведены пошаговые рекомендации по настройке как ресурсов и репликации СХД, так и инфраструктуры виртуализации. Основными рекомендациями по эксплуатации решения являются:

- Использование синхронной репликации ресурсов СХД;
- Контроль состояния пар репликации перед выполнением процедуры перевода ресурсов;
- Выполнение всех операций по переключению ресурсов только на стороне zVirt.

Для перевода ресурсов используется план восстановления, включающий в себя набор ВМ и список параметров для их запуска. План создается и используется на стороне РЦОД.

! При выполнении плана переключаются между площадками только указанные в нем ВМ. ВМ, не включенные в план восстановления, но входящие в тот же домен хранения, будут выключены.

Замеры производительности при нагрузке с синхронной репликацией показывают снижение на 20 %, при максимальном использовании портов репликации и репликационных связей на запись как для ввода-вывода, так и для пропускной способности комплекса. Снижение обусловлено дополнительными операциями передачи данных по репликационным связям и ожиданием подтверждения приема от удаленной СХД, являющимися условиями для закрытия операции на локальной СХД. Продолжительность данной задержки прямо определяется временем отклика сетевой среды, использующейся для репликации данных. Снижение производительности для профиля 70/30 чтение/запись составляет 4 %, что демонстрирует несущественное влияние репликации на продуктивную нагрузку. Синхронная репликация начинает заметно влиять на продуктивную нагрузку при смещении баланса профиля нагрузки в сторону записи. При 100% записи рост задержек составил 19%, что является существенным ростом общего времени отклика ресурсов.

Данные показатели являются базовыми и получены в лабораторной среде. При эксплуатации решения в инфраструктурной среде заказчиков возможно получение значений, отличных от лабораторных, ввиду возможных различий в инфраструктуре репликации данных: использования каналов иной пропускной способности, иного количества связей и портов, иного профиля нагрузки по записи и т.п.

### Дополнительные ресурсы

**Orion soft**:
- Wiki: Репликация и DR. Аппаратная репликация

**YADRO**:
- Руководство администратора TATLIN.UNIFIED
- WP: Подключение ресурсов к серверной ОС: IP SAN
- WP: Подключение ресурсов к серверной ОС: Зонирование FC SAN
- Описание функциональности: репликация

## ПРИЛОЖЕНИЯ

### Сравнение метрик для 4 ВМ (профилем 70% чтения и 30% записи)

Тестирование проводилось с одинаковым профилем нагрузки для всех 4 ВМ. Нагрузка запускалась с отдельной ВМ по команде:

```
fio --output=/root/fileio_4VM$(date +"%Y.%d.%m_%H.%M.%S").log \
    --client=172.19.28.196 --remote-config /root/fio_w.fio \
    --client=172.19.28.197 --remote-config /root/fio_w.fio \
    --client=172.19.28.193 --remote-config /root/fio_w.fio \
    --client=172.19.28.194 --remote-config /root/fio_w.fio
```

Содержимое профиля:

```
[global]
direct=1
ioengine=libaio
continue_on_error=none
group_reporting=1
randrepeat=0
rw=randrw
blocksize=8k
ramp_time=10
time_based=1
runtime=900

[LOAD]
filename=/dev/vda
rwmixread=70
cpus_allowed=1-8
numjobs=2
iodepth=16
```

**Сравнение метрик для 4 ВМ без репликации (профилем 70% чтения и 30% записи)**

```
hostname=shb-app04, be=0, 64-bit, os=Linux, arch=x86-64, fio=fio-3.35, flags=1
hostname=shb-app03, be=0, 64-bit, os=Linux, arch=x86-64, fio=fio-3.35, flags=1
hostname=shb-app02, be=0, 64-bit, os=Linux, arch=x86-64, fio=fio-3.35, flags=1
hostname=shb-app01, be=0, 64-bit, os=Linux, arch=x86-64, fio=fio-3.35, flags=1
<shb-app02> LOAD: (g=0): rw=randrw, bs=(R) 8192B-8192B, (W) 8192B-8192B, (T) 8192B-8192B,
ioengine=libaio, iodepth=16
<shb-app01> LOAD: (g=0): rw=randrw, bs=(R) 8192B-8192B, (W) 8192B-8192B, (T) 8192B-8192B,
ioengine=libaio, iodepth=16
<shb-app02> ...
<shb-app01> ...
<shb-app03> LOAD: (g=0): rw=randrw, bs=(R) 8192B-8192B, (W) 8192B-8192B, (T) 8192B-8192B,
```

```
ioengine=libaio, iodepth=16

<shb-app04> LOAD: (g=0): rw=randrw, bs=(R) 8192B-8192B, (W) 8192B-8192B, (T) 8192B-8192B,
ioengine=libaio, iodepth=16

<shb-app03> ...

<shb-app04> ...

<shb-app04> Starting 2 processes

<shb-app03> Starting 2 processes

<shb-app02> Starting 2 processes

<shb-app01> Starting 2 processes

<shb-app01>

LOAD: (groupid=0, jobs=2): err= 0: pid=5828: Tue Mar 31 22:25:04 2026

  read: IOPS=31.6k, BW=247MiB/s (259MB/s)(217GiB/900003msec)

    slat (nsec): min=1121, max=310202, avg=3377.89, stdev=2998.10

    clat (usec): min=60, max=28040, avg=781.31, stdev=601.40

     lat (usec): min=63, max=28044, avg=784.69, stdev=601.47

    clat percentiles (usec):

     |  1.00th=[  293],  5.00th=[  351], 10.00th=[  388], 20.00th=[  420],

     | 30.00th=[  465], 40.00th=[  523], 50.00th=[  603], 60.00th=[  693],

     | 70.00th=[  791], 80.00th=[  963], 90.00th=[ 1434], 95.00th=[ 2057],

     | 99.00th=[ 2802], 99.50th=[ 2999], 99.90th=[ 3884], 99.95th=[ 5407],

     | 99.99th=[19268]

   bw (  KiB/s): min=225440, max=266554, per=100.00%, avg=253084.02, stdev=2426.95, samples=3600

   iops        : min=28180, max=33319, avg=31635.28, stdev=303.37, samples=3600

  write: IOPS=13.6k, BW=106MiB/s (111MB/s)(93.0GiB/900003msec); 0 zone resets

    slat (nsec): min=1095, max=251444, avg=3974.65, stdev=3078.76

    clat (usec): min=114, max=18110, avg=523.66, stdev=332.73

     lat (usec): min=117, max=18112, avg=527.63, stdev=332.90

    clat percentiles (usec):

     |  1.00th=[  243],  5.00th=[  314], 10.00th=[  351], 20.00th=[  388],

     | 30.00th=[  408], 40.00th=[  437], 50.00th=[  474], 60.00th=[  519],

     | 70.00th=[  578], 80.00th=[  652], 90.00th=[  750], 95.00th=[  807],

     | 99.00th=[  947], 99.50th=[ 1037], 99.90th=[ 5211], 99.95th=[ 7832],
```

```
     | 99.99th=[13173]
    bw (  KiB/s): min=98400, max=116499, per=100.00%, avg=108463.56, stdev=1279.44, samples=3600
    iops        : min=12300, max=14562, avg=13557.72, stdev=159.91, samples=3600
  lat (usec)    : 100=0.01%, 250=0.49%, 500=42.08%, 750=30.44%, 1000=13.64%
  lat (msec)    : 2=9.51%, 4=3.73%, 10=0.08%, 20=0.02%, 50=0.01%
  cpu           : usr=4.45%, sys=9.57%, ctx=16327087, majf=0, minf=73
  IO depths     : 1=0.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=100.0%, 32=0.0%, >=64=0.0%
     submit     : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     complete   : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.1%, 32=0.0%, 64=0.0%, >=64=0.0%
     issued rwts: total=28458154,12196245,0,0 short=0,0,0,0 dropped=0,0,0,0
     latency    : target=0, window=0, percentile=100.00%, depth=16
```

Run status group 0 (all jobs):

```
   READ: bw=247MiB/s (259MB/s), 247MiB/s-247MiB/s (259MB/s-259MB/s), io=217GiB (233GB),
run=900003-900003msec

   WRITE: bw=106MiB/s (111MB/s), 106MiB/s-106MiB/s (111MB/s-111MB/s), io=93.0GiB (99.9GB),
run=900003-900003msec
```

Disk stats (read/write):

```
  vdb: ios=28782598/12335504, merge=0/0, ticks=22129210/6309825, in_queue=28439036, util=100.00%

<shb-app03>

LOAD: (groupid=0, jobs=2): err= 0: pid=1970: Tue Mar 31 22:25:04 2026

  read: IOPS=31.4k, BW=245MiB/s (257MB/s)(215GiB/900002msec)

    slat (nsec): min=1113, max=2843.1k, avg=3150.08, stdev=3503.05

    clat (usec): min=51, max=42557, avg=790.16, stdev=637.83

     lat (usec): min=61, max=42560, avg=793.31, stdev=637.98

    clat percentiles (usec):

     |  1.00th=[  293],  5.00th=[  355], 10.00th=[  388], 20.00th=[  408],

     | 30.00th=[  453], 40.00th=[  510], 50.00th=[  594], 60.00th=[  709],

     | 70.00th=[  799], 80.00th=[  988], 90.00th=[ 1483], 95.00th=[ 2089],

     | 99.00th=[ 2835], 99.50th=[ 3064], 99.90th=[ 4490], 99.95th=[ 6390],

     | 99.99th=[20841]
```

```
   bw (  KiB/s): min=223870, max=276928, per=100.00%, avg=251017.24, stdev=3964.30, samples=3600
   iops        : min=27983, max=34616, avg=31376.86, stdev=495.54, samples=3600
  write: IOPS=13.4k, BW=105MiB/s (110MB/s)(92.3GiB/900002msec); 0 zone resets
    slat (nsec): min=1107, max=2803.0k, avg=3193.14, stdev=3594.27
    clat (usec): min=33, max=19411, avg=524.04, stdev=335.58
     lat (usec): min=119, max=19414, avg=527.23, stdev=335.90
    clat percentiles (usec):
     |  1.00th=[  237],  5.00th=[  314], 10.00th=[  355], 20.00th=[  388],
     | 30.00th=[  404], 40.00th=[  424], 50.00th=[  461], 60.00th=[  506],
     | 70.00th=[  578], 80.00th=[  668], 90.00th=[  775], 95.00th=[  816],
     | 99.00th=[  996], 99.50th=[ 1106], 99.90th=[ 5211], 99.95th=[ 7570],
     | 99.99th=[13042]
   bw (  KiB/s): min=95390, max=120608, per=100.00%, avg=107618.92, stdev=1871.80, samples=3600
   iops        : min=11922, max=15076, avg=13452.08, stdev=233.98, samples=3600
  lat (usec)   : 50=0.01%, 100=0.01%, 250=0.53%, 500=43.85%, 750=26.54%
  lat (usec)   : 1000=15.28%
  lat (msec)   : 2=9.79%, 4=3.87%, 10=0.11%, 20=0.02%, 50=0.01%
  cpu          : usr=4.96%, sys=11.02%, ctx=15324369, majf=0, minf=73
  IO depths    : 1=0.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=100.0%, 32=0.0%, >=64=0.0%
     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     complete  : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.1%, 32=0.0%, 64=0.0%, >=64=0.0%
     issued rwts: total=28220423,12099002,0,0 short=0,0,0,0 dropped=0,0,0,0
     latency   : target=0, window=0, percentile=100.00%, depth=16
```

Run status group 0 (all jobs):

```
  READ: bw=245MiB/s (257MB/s), 245MiB/s-245MiB/s (257MB/s-257MB/s), io=215GiB (231GB),
run=900002-900002msec

  WRITE: bw=105MiB/s (110MB/s), 105MiB/s-105MiB/s (110MB/s-110MB/s), io=92.3GiB (99.1GB),
run=900002-900002msec
```

Disk stats (read/write):

```
  vda: ios=28550561/12240580, merge=0/0, ticks=22174412/6254038, in_queue=28428451, util=100.00%
<shb-app04>
LOAD: (groupid=0, jobs=2): err= 0: pid=2633: Tue Mar 31 22:25:04 2026
  read: IOPS=31.9k, BW=249MiB/s (261MB/s)(219GiB/900003msec)
    slat (nsec): min=1200, max=824983, avg=3115.95, stdev=3269.74
    clat (usec): min=56, max=35384, avg=776.28, stdev=601.52
     lat (usec): min=59, max=35388, avg=779.39, stdev=601.65
    clat percentiles (usec):
     |  1.00th=[  293],  5.00th=[  351], 10.00th=[  383], 20.00th=[  408],
     | 30.00th=[  445], 40.00th=[  502], 50.00th=[  594], 60.00th=[  701],
     | 70.00th=[  791], 80.00th=[  963], 90.00th=[ 1434], 95.00th=[ 2057],
     | 99.00th=[ 2802], 99.50th=[ 2999], 99.90th=[ 3851], 99.95th=[ 5342],
     | 99.99th=[18744]
   bw (  KiB/s): min=229954, max=275328, per=100.00%, avg=255168.62, stdev=3453.79, samples=3598
   iops        : min=28744, max=34416, avg=31895.86, stdev=431.74, samples=3598
  write: IOPS=13.7k, BW=107MiB/s (112MB/s)(93.8GiB/900003msec); 0 zone resets
    slat (nsec): min=1167, max=703122, avg=3527.52, stdev=3308.65
    clat (usec): min=30, max=18235, avg=517.68, stdev=334.60
     lat (usec): min=121, max=18242, avg=521.20, stdev=334.87
    clat percentiles (usec):
     |  1.00th=[  235],  5.00th=[  314], 10.00th=[  351], 20.00th=[  388],
     | 30.00th=[  400], 40.00th=[  416], 50.00th=[  453], 60.00th=[  498],
     | 70.00th=[  570], 80.00th=[  652], 90.00th=[  766], 95.00th=[  807],
     | 99.00th=[  979], 99.50th=[ 1090], 99.90th=[ 5211], 99.95th=[ 7635],
     | 99.99th=[13173]
   bw (  KiB/s): min=97025, max=118919, per=100.00%, avg=109351.47, stdev=1643.27, samples=3598
   iops        : min=12128, max=14864, avg=13668.71, stdev=205.41, samples=3598
  lat (usec)   : 50=0.01%, 100=0.01%, 250=0.54%, 500=45.36%, 750=25.76%
  lat (usec)   : 1000=14.98%
  lat (msec)   : 2=9.52%, 4=3.73%, 10=0.08%, 20=0.02%, 50=0.01%
```

```
  cpu          : usr=5.13%, sys=10.70%, ctx=15499823, majf=0, minf=73
  IO depths    : 1=0.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=100.0%, 32=0.0%, >=64=0.0%
     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     complete  : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.1%, 32=0.0%, 64=0.0%, >=64=0.0%
     issued rwts: total=28687891,12294083,0,0 short=0,0,0,0 dropped=0,0,0,0
     latency   : target=0, window=0, percentile=100.00%, depth=16
```

Run status group 0 (all jobs):

```
  READ: bw=249MiB/s (261MB/s), 249MiB/s-249MiB/s (261MB/s-261MB/s), io=219GiB (235GB),
run=900003-900003msec
  WRITE: bw=107MiB/s (112MB/s), 107MiB/s-107MiB/s (112MB/s-112MB/s), io=93.8GiB (101GB),
run=900003-900003msec
```

Disk stats (read/write):

```
  vda: ios=29017005/12433930, merge=0/0, ticks=22154438/6280868, in_queue=28435305, util=100.00%
<shb-app02>
LOAD: (groupid=0, jobs=2): err= 0: pid=2308: Tue Mar 31 22:25:04 2026
  read: IOPS=31.6k, BW=247MiB/s (259MB/s)(217GiB/900003msec)
    slat (nsec): min=1079, max=1617.5k, avg=3466.69, stdev=3220.18
    clat (usec): min=49, max=32171, avg=781.94, stdev=601.70
     lat (usec): min=61, max=32179, avg=785.41, stdev=601.78
    clat percentiles (usec):
     |  1.00th=[  293],  5.00th=[  351], 10.00th=[  388], 20.00th=[  420],
     | 30.00th=[  465], 40.00th=[  523], 50.00th=[  603], 60.00th=[  693],
     | 70.00th=[  791], 80.00th=[  963], 90.00th=[ 1450], 95.00th=[ 2057],
     | 99.00th=[ 2802], 99.50th=[ 2999], 99.90th=[ 3916], 99.95th=[ 5538],
     | 99.99th=[19006]
   bw (  KiB/s): min=226000, max=267952, per=100.00%, avg=253015.55, stdev=2395.44, samples=3600
   iops        : min=28250, max=33494, avg=31626.71, stdev=299.43, samples=3600
  write: IOPS=13.5k, BW=106MiB/s (111MB/s)(93.0GiB/900003msec); 0 zone resets
    slat (nsec): min=1061, max=618620, avg=3533.47, stdev=3191.90
    clat (usec): min=56, max=20517, avg=523.09, stdev=332.53
```

```
     lat (usec): min=119, max=20522, avg=526.62, stdev=332.71
    clat percentiles (usec):
     |  1.00th=[  243],  5.00th=[  314], 10.00th=[  347], 20.00th=[  388],
     | 30.00th=[  408], 40.00th=[  437], 50.00th=[  469], 60.00th=[  515],
     | 70.00th=[  578], 80.00th=[  652], 90.00th=[  750], 95.00th=[  807],
     | 99.00th=[  947], 99.50th=[ 1037], 99.90th=[ 5211], 99.95th=[ 7767],
     | 99.99th=[13042]
   bw (  KiB/s): min=98113, max=116832, per=100.00%, avg=108394.64, stdev=1296.93, samples=3600
   iops        : min=12264, max=14604, avg=13549.09, stdev=162.10, samples=3600
  lat (usec)   : 50=0.01%, 100=0.01%, 250=0.50%, 500=42.18%, 750=30.27%
  lat (usec)   : 1000=13.66%
  lat (msec)   : 2=9.53%, 4=3.75%, 10=0.08%, 20=0.02%, 50=0.01%
  cpu          : usr=4.43%, sys=9.65%, ctx=16260886, majf=0, minf=73
  IO depths    : 1=0.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=100.0%, 32=0.0%, >=64=0.0%
     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     complete  : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.1%, 32=0.0%, 64=0.0%, >=64=0.0%
     issued rwts: total=28449389,12188047,0,0 short=0,0,0,0 dropped=0,0,0,0
     latency   : target=0, window=0, percentile=100.00%, depth=16
```

Run status group 0 (all jobs):

```
  READ: bw=247MiB/s (259MB/s), 247MiB/s-247MiB/s (259MB/s-259MB/s), io=217GiB (233GB),
run=900003-900003msec

  WRITE: bw=106MiB/s (111MB/s), 106MiB/s-106MiB/s (111MB/s-111MB/s), io=93.0GiB (99.8GB),
run=900003-900003msec
```

Disk stats (read/write):

```
  vda: ios=28775047/12326331, merge=0/0, ticks=22134367/6294693, in_queue=28429059, util=100.00%

All clients: (groupid=0, jobs=2): err= 0: pid=0: Tue Mar 31 22:25:04 2026
  read: IOPS=126k, BW=988Mi (1036M)(868GiB/900003msec)
    slat (nsec): min=1079, max=2843.1k, avg=3277.58, stdev=3255.60
    clat (usec): min=49, max=42557, avg=782.39, stdev=610.76
     lat (usec): min=59, max=42560, avg=785.67, stdev=610.87
```

```
   bw (  KiB/s): min=905264, max=1086762, per=100.00%, avg=1012285.43, stdev=3133.02,
samples=14398
   iops        : min=113157, max=135845, avg=126534.70, stdev=391.63, samples=14398
  write: IOPS=54.2k, BW=423Mi (444M)(372GiB/900003msec); 0 zone resets
    slat (nsec): min=1061, max=2803.0k, avg=3557.87, stdev=3310.02
    clat (usec): min=30, max=20517, avg=522.10, stdev=333.87
     lat (usec): min=117, max=20522, avg=525.66, stdev=334.11
   bw (  KiB/s): min=388928, max=472858, per=100.00%, avg=433828.59, stdev=1542.79, samples=14398
   iops        : min=48614, max=59106, avg=54227.59, stdev=192.85, samples=14398
  lat (usec)   : 50=0.01%, 100=0.01%, 250=0.52%, 500=43.37%, 750=28.25%
  lat (usec)   : 1000=14.39%
  lat (msec)   : 2=9.59%, 4=3.77%, 10=0.08%, 20=0.02%, 50=0.01%
  cpu          : usr=4.74%, sys=10.24%, ctx=63412165, majf=0, minf=292
  IO depths    : 1=0.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=100.0%, 32=0.0%, >=64=0.0%
     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     complete  : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.1%, 32=0.0%, 64=0.0%, >=64=0.0%
     issued rwts: total=113815857,48777377,0,0 short=0,0,0,0 dropped=0,0,0,0
```

**Сравнение метрик для 4 ВМ с репликацией (профилем 70% чтения и 30% записи)**

```
hostname=shb-app04, be=0, 64-bit, os=Linux, arch=x86-64, fio=fio-3.35, flags=1

hostname=shb-app03, be=0, 64-bit, os=Linux, arch=x86-64, fio=fio-3.35, flags=1

hostname=shb-app02, be=0, 64-bit, os=Linux, arch=x86-64, fio=fio-3.35, flags=1

hostname=shb-app01, be=0, 64-bit, os=Linux, arch=x86-64, fio=fio-3.35, flags=1

<shb-app03> LOAD: (g=0): rw=randrw, bs=(R) 8192B-8192B, (W) 8192B-8192B, (T) 8192B-8192B,
ioengine=libaio, iodepth=16

<shb-app02> LOAD: (g=0): rw=randrw, bs=(R) 8192B-8192B, (W) 8192B-8192B, (T) 8192B-8192B,
ioengine=libaio, iodepth=16

<shb-app01> LOAD: (g=0): rw=randrw, bs=(R) 8192B-8192B, (W) 8192B-8192B, (T) 8192B-8192B,
ioengine=libaio, iodepth=16

<shb-app03> ...

<shb-app02> ...

<shb-app01> ...

<shb-app04> LOAD: (g=0): rw=randrw, bs=(R) 8192B-8192B, (W) 8192B-8192B, (T) 8192B-8192B,
ioengine=libaio, iodepth=16

<shb-app04> ...

<shb-app04> Starting 2 processes

<shb-app02> Starting 2 processes

<shb-app01> Starting 2 processes

<shb-app03> Starting 2 processes

<shb-app02>

LOAD: (groupid=0, jobs=2): err= 0: pid=2408: Tue Mar 31 23:31:01 2026

  read: IOPS=30.4k, BW=238MiB/s (249MB/s)(209GiB/900003msec)

    slat (nsec): min=1084, max=932852, avg=3473.30, stdev=3248.73

    clat (usec): min=21, max=30954, avg=760.95, stdev=593.46

     lat (usec): min=62, max=30958, avg=764.42, stdev=593.53

    clat percentiles (usec):

     |  1.00th=[  289],  5.00th=[  343], 10.00th=[  379], 20.00th=[  412],

     | 30.00th=[  449], 40.00th=[  502], 50.00th=[  578], 60.00th=[  660],

     | 70.00th=[  766], 80.00th=[  922], 90.00th=[ 1418], 95.00th=[ 2024],

     | 99.00th=[ 2769], 99.50th=[ 2999], 99.90th=[ 3818], 99.95th=[ 5735],

     | 99.99th=[17695]
```

```
   bw (  KiB/s): min=221500, max=259312, per=100.00%, avg=243665.08, stdev=2338.72, samples=3600
   iops        : min=27687, max=32414, avg=30457.88, stdev=292.34, samples=3600
  write: IOPS=13.0k, BW=102MiB/s (107MB/s)(89.6GiB/900003msec); 0 zone resets
    slat (nsec): min=1066, max=1004.6k, avg=3585.09, stdev=3267.96
    clat (usec): min=34, max=19856, avg=662.35, stdev=357.42
     lat (usec): min=197, max=19865, avg=665.94, stdev=357.58
    clat percentiles (usec):
     |  1.00th=[  338],  5.00th=[  396], 10.00th=[  420], 20.00th=[  474],
     | 30.00th=[  529], 40.00th=[  594], 50.00th=[  644], 60.00th=[  701],
     | 70.00th=[  750], 80.00th=[  799], 90.00th=[  865], 95.00th=[  938],
     | 99.00th=[ 1123], 99.50th=[ 1270], 99.90th=[ 6063], 99.95th=[ 7242],
     | 99.99th=[12911]
   bw (  KiB/s): min=94974, max=112448, per=100.00%, avg=104456.60, stdev=1222.56, samples=3600
   iops        : min=11871, max=14056, avg=13056.85, stdev=152.85, samples=3600
  lat (usec)   : 50=0.01%, 100=0.01%, 250=0.18%, 500=34.90%, 750=33.66%
  lat (usec)   : 1000=17.97%
  lat (msec)   : 2=9.54%, 4=3.62%, 10=0.09%, 20=0.02%, 50=0.01%
  cpu          : usr=4.35%, sys=9.48%, ctx=16101819, majf=0, minf=73
  IO depths    : 1=0.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=100.0%, 32=0.0%, >=64=0.0%
     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     complete  : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.1%, 32=0.0%, 64=0.0%, >=64=0.0%
     issued rwts: total=27396964,11744806,0,0 short=0,0,0,0 dropped=0,0,0,0
     latency   : target=0, window=0, percentile=100.00%, depth=16
```

Run status group 0 (all jobs):

```
  READ: bw=238MiB/s (249MB/s), 238MiB/s-238MiB/s (249MB/s-249MB/s), io=209GiB (224GB),
run=900003-900003msec

  WRITE: bw=102MiB/s (107MB/s), 102MiB/s-102MiB/s (107MB/s-107MB/s), io=89.6GiB (96.2GB),
run=900003-900003msec
```

Disk stats (read/write):

```
  vda: ios=27711688/11879068, merge=0/0, ticks=20738309/7724319, in_queue=28462629, util=100.00%
<shb-app01>
LOAD: (groupid=0, jobs=2): err= 0: pid=5921: Tue Mar 31 23:31:01 2026
  read: IOPS=30.4k, BW=238MiB/s (249MB/s)(209GiB/900002msec)
    slat (nsec): min=1125, max=929501, avg=3397.21, stdev=3014.70
    clat (usec): min=60, max=30930, avg=760.86, stdev=594.53
     lat (usec): min=63, max=30931, avg=764.26, stdev=594.61
    clat percentiles (usec):
     |  1.00th=[  289],  5.00th=[  343], 10.00th=[  379], 20.00th=[  408],
     | 30.00th=[  449], 40.00th=[  502], 50.00th=[  578], 60.00th=[  660],
     | 70.00th=[  766], 80.00th=[  922], 90.00th=[ 1418], 95.00th=[ 2024],
     | 99.00th=[ 2769], 99.50th=[ 2999], 99.90th=[ 3818], 99.95th=[ 5669],
     | 99.99th=[17957]
   bw (  KiB/s): min=221232, max=256993, per=100.00%, avg=243699.54, stdev=2414.98, samples=3599
   iops        : min=27654, max=32124, avg=30462.21, stdev=301.87, samples=3599
  write: IOPS=13.0k, BW=102MiB/s (107MB/s)(89.5GiB/900002msec); 0 zone resets
    slat (nsec): min=1082, max=1072.2k, avg=3941.27, stdev=3095.40
    clat (usec): min=178, max=18643, avg=662.96, stdev=356.69
     lat (usec): min=197, max=18645, avg=666.90, stdev=356.86
    clat percentiles (usec):
     |  1.00th=[  338],  5.00th=[  396], 10.00th=[  420], 20.00th=[  474],
     | 30.00th=[  537], 40.00th=[  594], 50.00th=[  644], 60.00th=[  701],
     | 70.00th=[  750], 80.00th=[  799], 90.00th=[  865], 95.00th=[  938],
     | 99.00th=[ 1123], 99.50th=[ 1270], 99.90th=[ 6063], 99.95th=[ 7242],
     | 99.99th=[12911]
   bw (  KiB/s): min=93616, max=112865, per=100.00%, avg=104361.79, stdev=1218.22, samples=3599
   iops        : min=11702, max=14108, avg=13045.00, stdev=152.30, samples=3599
  lat (usec)   : 100=0.01%, 250=0.18%, 500=34.91%, 750=33.57%, 1000=18.04%
  lat (msec)   : 2=9.56%, 4=3.62%, 10=0.09%, 20=0.02%, 50=0.01%
  cpu          : usr=4.24%, sys=9.19%, ctx=16226135, majf=0, minf=73
```

```
  IO depths    : 1=0.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=100.0%, 32=0.0%, >=64=0.0%
     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     complete  : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.1%, 32=0.0%, 64=0.0%, >=64=0.0%
     issued rwts: total=27401369,11734453,0,0 short=0,0,0,0 dropped=0,0,0,0
     latency   : target=0, window=0, percentile=100.00%, depth=16
```

Run status group 0 (all jobs):

```
  READ: bw=238MiB/s (249MB/s), 238MiB/s-238MiB/s (249MB/s-249MB/s), io=209GiB (224GB),
run=900002-900002msec

  WRITE: bw=102MiB/s (107MB/s), 102MiB/s-102MiB/s (107MB/s-107MB/s), io=89.5GiB (96.1GB),
run=900002-900002msec
```

Disk stats (read/write):

```
  vdb: ios=27718923/11871149, merge=0/0, ticks=20753562/7733159, in_queue=28486720, util=100.00%

<shb-app04>

LOAD: (groupid=0, jobs=2): err= 0: pid=2677: Tue Mar 31 23:31:01 2026

  read: IOPS=30.4k, BW=237MiB/s (249MB/s)(209GiB/900002msec)

    slat (nsec): min=1203, max=1452.7k, avg=3317.34, stdev=3637.72

    clat (usec): min=56, max=30892, avg=761.91, stdev=595.43

     lat (usec): min=64, max=30893, avg=765.22, stdev=595.60

    clat percentiles (usec):

     |  1.00th=[  289],  5.00th=[  347], 10.00th=[  379], 20.00th=[  404],

     | 30.00th=[  437], 40.00th=[  494], 50.00th=[  578], 60.00th=[  676],

     | 70.00th=[  783], 80.00th=[  930], 90.00th=[ 1418], 95.00th=[ 2024],

     | 99.00th=[ 2802], 99.50th=[ 2999], 99.90th=[ 3785], 99.95th=[ 5604],

     | 99.99th=[17695]

   bw (  KiB/s): min=219728, max=263662, per=100.00%, avg=243219.66, stdev=3232.32, samples=3600

   iops        : min=27466, max=32956, avg=30402.16, stdev=404.04, samples=3600

  write: IOPS=13.0k, BW=102MiB/s (107MB/s)(89.4GiB/900002msec); 0 zone resets

    slat (nsec): min=1191, max=2305.6k, avg=3707.07, stdev=3701.53

    clat (usec): min=36, max=20412, avg=665.31, stdev=364.75

     lat (usec): min=205, max=20417, avg=669.02, stdev=365.05
```

```
    clat percentiles (usec):
     |  1.00th=[  338],  5.00th=[  392], 10.00th=[  412], 20.00th=[  461],
     | 30.00th=[  519], 40.00th=[  586], 50.00th=[  652], 60.00th=[  709],
     | 70.00th=[  775], 80.00th=[  799], 90.00th=[  881], 95.00th=[  963],
     | 99.00th=[ 1188], 99.50th=[ 1336], 99.90th=[ 6063], 99.95th=[ 7308],
     | 99.99th=[13042]
   bw (  KiB/s): min=93568, max=113840, per=100.00%, avg=104282.33, stdev=1530.24, samples=3600
   iops        : min=11696, max=14230, avg=13034.99, stdev=191.30, samples=3600
  lat (usec)   : 50=0.01%, 100=0.01%, 250=0.18%, 500=36.79%, 750=29.58%
  lat (usec)   : 1000=19.83%
  lat (msec)   : 2=9.85%, 4=3.63%, 10=0.09%, 20=0.02%, 50=0.01%
  cpu          : usr=5.09%, sys=10.93%, ctx=15272771, majf=0, minf=73
  IO depths    : 1=0.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=100.0%, 32=0.0%, >=64=0.0%
     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     complete  : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.1%, 32=0.0%, 64=0.0%, >=64=0.0%
     issued rwts: total=27343303,11723665,0,0 short=0,0,0,0 dropped=0,0,0,0
     latency   : target=0, window=0, percentile=100.00%, depth=16
```

Run status group 0 (all jobs):

```
READ: bw=237MiB/s (249MB/s), 237MiB/s-237MiB/s (249MB/s-249MB/s), io=209GiB (224GB), run=900002-900002msec

WRITE: bw=102MiB/s (107MB/s), 102MiB/s-102MiB/s (107MB/s-107MB/s), io=89.4GiB (96.0GB), run=900002-900002msec
```

Disk stats (read/write):

```
  vda: ios=27660131/11858924, merge=0/0, ticks=20702831/7736920, in_queue=28439751, util=100.00%

<shb-app03>

LOAD: (groupid=0, jobs=2): err= 0: pid=2018: Tue Mar 31 23:31:01 2026

  read: IOPS=30.0k, BW=235MiB/s (246MB/s)(206GiB/900002msec)

    slat (nsec): min=1121, max=1103.1k, avg=3289.61, stdev=3631.72

    clat (usec): min=40, max=38529, avg=771.96, stdev=623.69

     lat (usec): min=63, max=38531, avg=775.25, stdev=623.87
```

```
   clat percentiles (usec):
    |  1.00th=[  289],  5.00th=[  347], 10.00th=[  383], 20.00th=[  404],
    | 30.00th=[  445], 40.00th=[  498], 50.00th=[  586], 60.00th=[  685],
    | 70.00th=[  791], 80.00th=[  947], 90.00th=[ 1450], 95.00th=[ 2057],
    | 99.00th=[ 2802], 99.50th=[ 3032], 99.90th=[ 4178], 99.95th=[ 6325],
    | 99.99th=[19530]
  bw (  KiB/s): min=212184, max=260662, per=100.00%, avg=240466.02, stdev=3321.30, samples=3600
  iops        : min=26522, max=32582, avg=30057.95, stdev=415.17, samples=3600
write: IOPS=12.9k, BW=101MiB/s (105MB/s)(88.4GiB/900002msec); 0 zone resets
   slat (nsec): min=1113, max=839693, avg=3321.59, stdev=3631.44
   clat (usec): min=44, max=22496, avg=670.88, stdev=363.78
    lat (usec): min=195, max=22498, avg=674.21, stdev=364.12
   clat percentiles (usec):
    |  1.00th=[  338],  5.00th=[  392], 10.00th=[  412], 20.00th=[  465],
    | 30.00th=[  523], 40.00th=[  594], 50.00th=[  660], 60.00th=[  717],
    | 70.00th=[  775], 80.00th=[  807], 90.00th=[  889], 95.00th=[  971],
    | 99.00th=[ 1188], 99.50th=[ 1352], 99.90th=[ 6063], 99.95th=[ 7242],
    | 99.99th=[12911]
  bw (  KiB/s): min=91852, max=113538, per=100.00%, avg=103037.22, stdev=1554.97, samples=3600
  iops        : min=11481, max=14192, avg=12879.31, stdev=194.39, samples=3600
lat (usec)    : 50=0.01%, 100=0.01%, 250=0.17%, 500=35.83%, 750=29.73%
lat (usec)    : 1000=20.26%
lat (msec)    : 2=10.13%, 4=3.73%, 10=0.11%, 20=0.02%, 50=0.01%
cpu           : usr=5.01%, sys=11.09%, ctx=15163869, majf=0, minf=73
IO depths     : 1=0.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=100.0%, 32=0.0%, >=64=0.0%
    submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
    complete  : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.1%, 32=0.0%, 64=0.0%, >=64=0.0%
    issued rwts: total=27032876,11583329,0,0 short=0,0,0,0 dropped=0,0,0,0
    latency   : target=0, window=0, percentile=100.00%, depth=16
```

Run status group 0 (all jobs):

```
  READ: bw=235MiB/s (246MB/s), 235MiB/s-235MiB/s (246MB/s-246MB/s), io=206GiB (221GB),
run=900002-900002msec

  WRITE: bw=101MiB/s (105MB/s), 101MiB/s-101MiB/s (105MB/s-105MB/s), io=88.4GiB (94.9GB),
run=900002-900002msec
```

Disk stats (read/write):

```
vda: ios=27343578/11716990, merge=0/0, ticks=20737071/7708029, in_queue=28445099, util=100.00%

All clients: (groupid=0, jobs=2): err= 0: pid=0: Tue Mar 31 23:31:01 2026

  read: IOPS=121k, BW=948Mi (994M)(833GiB/900003msec)

    slat (nsec): min=1084, max=1452.7k, avg=3369.66, stdev=3393.36

    clat (usec): min=21, max=38529, avg=763.89, stdev=601.86

     lat (usec): min=62, max=38531, avg=767.26, stdev=601.98

   bw (  KiB/s): min=874644, max=1040629, per=100.00%, avg=971050.30, stdev=2862.45,
samples=14399

   iops        : min=109329, max=130076, avg=121380.20, stdev=357.81, samples=14399

  write: IOPS=52.0k, BW=406Mi (426M)(357GiB/900003msec); 0 zone resets

    slat (nsec): min=1066, max=2305.6k, avg=3639.75, stdev=3439.76

    clat (usec): min=34, max=22496, avg=665.36, stdev=360.68

     lat (usec): min=195, max=22498, avg=669.00, stdev=360.92

   bw (  KiB/s): min=374010, max=452691, per=100.00%, avg=416137.94, stdev=1390.75, samples=14399

   iops        : min=46750, max=56586, avg=52016.15, stdev=173.86, samples=14399

  lat (usec)   : 50=0.01%, 100=0.01%, 250=0.18%, 500=35.61%, 750=31.64%

  lat (usec)   : 1000=19.02%

  lat (msec)   : 2=9.77%, 4=3.65%, 10=0.10%, 20=0.02%, 50=0.01%

  cpu          : usr=4.67%, sys=10.17%, ctx=62764594, majf=0, minf=292

  IO depths    : 1=0.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=100.0%, 32=0.0%, >=64=0.0%

     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%

     complete  : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.1%, 32=0.0%, 64=0.0%, >=64=0.0%

     issued rwts: total=109174512,46786253,0,0 short=0,0,0,0 dropped=0,0,0,0
```

### Сравнение метрик для 4 ВМ (профилем 100% записи)

Тестирование проводилось с одинаковым профилем нагрузки для всех 4 ВМ. Нагрузка запускалась с отдельной ВМ по команде:

```
fio --output=/root/fileio_4VM$(date +"%Y.%d.%m_%H.%M.%S").log \
    --client=172.19.28.196 --remote-config /root/fio_w.fio \
    --client=172.19.28.197 --remote-config /root/fio_w.fio \
    --client=172.19.28.193 --remote-config /root/fio_w.fio \
    --client=172.19.28.194 --remote-config /root/fio_w.fio
```

Содержимое профиля:

```
[global]
direct=1
ioengine=libaio
continue_on_error=none
group_reporting=1
rw=write
blocksize=64k
ramp_time=10
time_based=1
runtime=900

[LOAD]
filename=/dev/vda
cpus_allowed=1-8
numjobs=4
iodepth=8
```

### Сравнение метрик для 4 ВМ с репликацией (100% записи)

```
hostname=shb-app04, be=0, 64-bit, os=Linux, arch=x86-64, fio=fio-3.35, flags=1

hostname=shb-app03, be=0, 64-bit, os=Linux, arch=x86-64, fio=fio-3.35, flags=1

hostname=shb-app02, be=0, 64-bit, os=Linux, arch=x86-64, fio=fio-3.35, flags=1

hostname=shb-app01, be=0, 64-bit, os=Linux, arch=x86-64, fio=fio-3.35, flags=1

<shb-app02> LOAD: (g=0): rw=write, bs=(R) 64.0KiB-64.0KiB, (W) 64.0KiB-64.0KiB, (T) 64.0KiB-
64.0KiB, ioengine=libaio, iodepth=8

<shb-app01> LOAD: (g=0): rw=write, bs=(R) 64.0KiB-64.0KiB, (W) 64.0KiB-64.0KiB, (T) 64.0KiB-
64.0KiB, ioengine=libaio, iodepth=8

<shb-app02> ...

<shb-app01> ...

<shb-app04> LOAD: (g=0): rw=write, bs=(R) 64.0KiB-64.0KiB, (W) 64.0KiB-64.0KiB, (T) 64.0KiB-
64.0KiB, ioengine=libaio, iodepth=8

<shb-app03> LOAD: (g=0): rw=write, bs=(R) 64.0KiB-64.0KiB, (W) 64.0KiB-64.0KiB, (T) 64.0KiB-
64.0KiB, ioengine=libaio, iodepth=8

<shb-app04> ...

<shb-app03> ...

<shb-app03> Starting 4 processes

<shb-app04> Starting 4 processes

<shb-app01> Starting 4 processes

<shb-app02> Starting 4 processes

<shb-app03>

LOAD: (groupid=0, jobs=4): err= 0: pid=1841: Tue Mar 31 20:53:55 2026

  write: IOPS=21.6k, BW=1349MiB/s (1414MB/s)(1186GiB/900002msec); 0 zone resets

    slat (nsec): min=1879, max=1889.1k, avg=8399.52, stdev=6872.65

    clat (usec): min=42, max=121676, avg=1473.48, stdev=601.05

     lat (usec): min=361, max=121683, avg=1481.88, stdev=601.04

    clat percentiles (usec):

     |  1.00th=[  676],  5.00th=[  799], 10.00th=[  889], 20.00th=[ 1012],

     | 30.00th=[ 1139], 40.00th=[ 1254], 50.00th=[ 1385], 60.00th=[ 1500],

     | 70.00th=[ 1663], 80.00th=[ 1860], 90.00th=[ 2180], 95.00th=[ 2442],

     | 99.00th=[ 3097], 99.50th=[ 3392], 99.90th=[ 6849], 99.95th=[ 7504],

     | 99.99th=[ 9372]
```

```
   bw (  MiB/s): min=  993, max= 1571, per=100.00%, avg=1349.84, stdev=17.81, samples=7200
   iops        : min=15889, max=25140, avg=21596.77, stdev=285.03, samples=7200
  lat (usec)   : 50=0.01%, 100=0.01%, 250=0.01%, 500=0.03%, 750=2.79%
  lat (usec)   : 1000=15.78%
  lat (msec)   : 2=66.88%, 4=14.31%, 10=0.21%, 20=0.01%, 50=0.01%
  lat (msec)   : 100=0.01%, 250=0.01%
  cpu          : usr=2.94%, sys=4.86%, ctx=9947360, majf=0, minf=145
  IO depths    : 1=0.0%, 2=0.0%, 4=0.0%, 8=100.0%, 16=0.0%, 32=0.0%, >=64=0.0%
     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     complete  : 0=0.0%, 4=100.0%, 8=0.1%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     issued rwts: total=0,19424516,0,0 short=0,0,0,0 dropped=0,0,0,0
     latency   : target=0, window=0, percentile=100.00%, depth=8
```

Run status group 0 (all jobs):

```
WRITE: bw=1349MiB/s (1414MB/s), 1349MiB/s-1349MiB/s (1414MB/s-1414MB/s), io=1186GiB (1273GB),
run=900002-900002msec
```

Disk stats (read/write):

```
vda: ios=48/19637416, merge=0/0, ticks=6/28627207, in_queue=28627212, util=100.00%
<shb-app04>
LOAD: (groupid=0, jobs=4): err= 0: pid=2494: Tue Mar 31 20:53:55 2026
  write: IOPS=21.6k, BW=1353MiB/s (1419MB/s)(1189GiB/900001msec); 0 zone resets
    slat (usec): min=2, max=6510, avg= 9.09, stdev= 7.26
    clat (usec): min=82, max=138479, avg=1468.26, stdev=604.92
     lat (usec): min=344, max=138485, avg=1477.35, stdev=604.99
    clat percentiles (usec):
     |  1.00th=[  676],  5.00th=[  799], 10.00th=[  889], 20.00th=[ 1012],
     | 30.00th=[ 1123], 40.00th=[ 1237], 50.00th=[ 1369], 60.00th=[ 1500],
     | 70.00th=[ 1663], 80.00th=[ 1860], 90.00th=[ 2147], 95.00th=[ 2442],
     | 99.00th=[ 3097], 99.50th=[ 3392], 99.90th=[ 6849], 99.95th=[ 7504],
     | 99.99th=[ 9241]
   bw (  MiB/s): min= 1031, max= 1559, per=100.00%, avg=1353.85, stdev=17.52, samples=7200
```

```
 iops        : min=16499, max=24952, avg=21661.09, stdev=280.28, samples=7200
lat (usec)   : 100=0.01%, 250=0.01%, 500=0.03%, 750=2.78%, 1000=15.99%
lat (msec)   : 2=66.96%, 4=14.02%, 10=0.21%, 20=0.01%, 50=0.01%
lat (msec)   : 100=0.01%, 250=0.01%
cpu          : usr=3.28%, sys=4.83%, ctx=10025100, majf=0, minf=145
IO depths    : 1=0.0%, 2=0.0%, 4=0.0%, 8=100.0%, 16=0.0%, 32=0.0%, >=64=0.0%
   submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
   complete  : 0=0.0%, 4=100.0%, 8=0.1%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
   issued rwts: total=0,19483964,0,0 short=0,0,0,0 dropped=0,0,0,0
   latency   : target=0, window=0, percentile=100.00%, depth=8
```

Run status group 0 (all jobs):

```
WRITE: bw=1353MiB/s (1419MB/s), 1353MiB/s-1353MiB/s (1419MB/s-1419MB/s), io=1189GiB (1277GB),
run=900001-900001msec
```

Disk stats (read/write):

```
 vda: ios=48/19699912, merge=0/0, ticks=6/28612467, in_queue=28612473, util=100.00%

<shb-app01>

LOAD: (groupid=0, jobs=4): err= 0: pid=5519: Tue Mar 31 20:53:55 2026

  write: IOPS=23.9k, BW=1494MiB/s (1567MB/s)(1313GiB/900001msec); 0 zone resets

    slat (nsec): min=1876, max=1137.8k, avg=8741.76, stdev=5333.28

    clat (usec): min=53, max=137626, avg=1328.68, stdev=541.24

     lat (usec): min=343, max=137641, avg=1337.42, stdev=541.01

    clat percentiles (usec):

     |  1.00th=[  594],  5.00th=[  717], 10.00th=[  799], 20.00th=[  922],

     | 30.00th=[ 1037], 40.00th=[ 1139], 50.00th=[ 1254], 60.00th=[ 1369],

     | 70.00th=[ 1500], 80.00th=[ 1663], 90.00th=[ 1926], 95.00th=[ 2180],

     | 99.00th=[ 2737], 99.50th=[ 3032], 99.90th=[ 6652], 99.95th=[ 7308],

     | 99.99th=[ 9110]

   bw (  MiB/s): min= 1165, max= 1742, per=100.00%, avg=1495.27, stdev=22.19, samples=7197

   iops        : min=18639, max=27871, avg=23923.86, stdev=354.99, samples=7197

  lat (usec)   : 100=0.01%, 250=0.01%, 500=0.13%, 750=6.65%, 1000=20.00%

  lat (msec)   : 2=65.07%, 4=7.97%, 10=0.17%, 20=0.01%, 50=0.01%

  lat (msec)   : 100=0.01%, 250=0.01%

  cpu          : usr=2.66%, sys=4.66%, ctx=10463983, majf=0, minf=145

  IO depths    : 1=0.0%, 2=0.0%, 4=0.0%, 8=100.0%, 16=0.0%, 32=0.0%, >=64=0.0%

     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%

     complete  : 0=0.0%, 4=100.0%, 8=0.1%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%

     issued rwts: total=0,21519902,0,0 short=0,0,0,0 dropped=0,0,0,0

     latency   : target=0, window=0, percentile=100.00%, depth=8
```

Run status group 0 (all jobs):

```
WRITE: bw=1494MiB/s (1567MB/s), 1494MiB/s-1494MiB/s (1567MB/s-1567MB/s), io=1313GiB (1410GB),
run=900001-900001msec
```

Disk stats (read/write):

```
vda: ios=49/21766245, merge=0/0, ticks=6/28542522, in_queue=28542528, util=100.00%

<shb-app02>

LOAD: (groupid=0, jobs=4): err= 0: pid=1956: Tue Mar 31 20:53:55 2026

  write: IOPS=23.7k, BW=1483MiB/s (1555MB/s)(1303GiB/900001msec); 0 zone resets

    slat (nsec): min=1942, max=5631.6k, avg=8433.89, stdev=5784.47

    clat (usec): min=59, max=138674, avg=1339.31, stdev=551.66

     lat (usec): min=332, max=138682, avg=1347.74, stdev=551.38

    clat percentiles (usec):

     |  1.00th=[  594],  5.00th=[  717], 10.00th=[  799], 20.00th=[  930],

     | 30.00th=[ 1045], 40.00th=[ 1156], 50.00th=[ 1270], 60.00th=[ 1385],

     | 70.00th=[ 1516], 80.00th=[ 1680], 90.00th=[ 1942], 95.00th=[ 2212],

     | 99.00th=[ 2737], 99.50th=[ 3032], 99.90th=[ 6652], 99.95th=[ 7308],

     | 99.99th=[ 8848]

   bw (  MiB/s): min= 1147, max= 1725, per=100.00%, avg=1483.79, stdev=21.21, samples=7200

   iops        : min=18364, max=27608, avg=23740.04, stdev=339.39, samples=7200

  lat (usec)   : 100=0.01%, 250=0.01%, 500=0.13%, 750=6.47%, 1000=19.58%

  lat (msec)   : 2=65.17%, 4=8.47%, 10=0.18%, 20=0.01%, 50=0.01%

  lat (msec)   : 100=0.01%, 250=0.01%

  cpu          : usr=2.54%, sys=4.51%, ctx=10202160, majf=0, minf=145

  IO depths    : 1=0.0%, 2=0.0%, 4=0.0%, 8=100.0%, 16=0.0%, 32=0.0%, >=64=0.0%

     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%

     complete  : 0=0.0%, 4=100.0%, 8=0.1%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%

     issued rwts: total=0,21354159,0,0 short=0,0,0,0 dropped=0,0,0,0

     latency   : target=0, window=0, percentile=100.00%, depth=8
```

Run status group 0 (all jobs):

```
WRITE: bw=1483MiB/s (1555MB/s), 1483MiB/s-1483MiB/s (1555MB/s-1555MB/s), io=1303GiB (1399GB),
run=900001-900001msec
```

Disk stats (read/write):

```
vda: ios=48/21592660, merge=0/0, ticks=6/28551417, in_queue=28551423, util=100.00%

All clients: (groupid=0, jobs=4): err= 0: pid=0: Tue Mar 31 20:53:55 2026

  write: IOPS=90.9k, BW=5679Mi (5955M)(4992GiB/900002msec); 0 zone resets

    slat (nsec): min=1876, max=6510.8k, avg=8662.77, stdev=6328.99

    clat (usec): min=42, max=138674, avg=1399.10, stdev=578.12

     lat (usec): min=332, max=138682, avg=1407.76, stdev=578.02

   bw (  MiB/s): min= 4337, max= 6598, per=100.00%, avg=5682.76, stdev=19.79, samples=28797

   iops        : min=69391, max=105571, avg=90921.76, stdev=316.60, samples=28797

  lat (usec)   : 50=0.01%, 100=0.01%, 250=0.01%, 500=0.08%, 750=4.77%

  lat (usec)   : 1000=17.93%

  lat (msec)   : 2=65.97%, 4=11.05%, 10=0.19%, 20=0.01%, 50=0.01%

  lat (msec)   : 100=0.01%, 250=0.01%

  cpu          : usr=2.86%, sys=4.71%, ctx=40638603, majf=0, minf=580

  IO depths    : 1=0.0%, 2=0.0%, 4=0.0%, 8=100.0%, 16=0.0%, 32=0.0%, >=64=0.0%

     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%

     complete  : 0=0.0%, 4=100.0%, 8=0.1%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%

     issued rwts: total=0,81782541,0,0 short=0,0,0,0 dropped=0,0,0,0
```

**Сравнение метрик для 4 ВМ без репликации (100% записи)**

Run status group 0 (all jobs):

```
WRITE: bw=2055MiB/s (2155MB/s), 2055MiB/s-2055MiB/s (2155MB/s-2155MB/s), io=1806GiB (1939GB),
run=900001-900001msec
```

Disk stats (read/write):

```
vda: ios=48/29912850, merge=0/0, ticks=7/28215545, in_queue=28215552, util=100.00%

<shb-app03>

LOAD: (groupid=0, jobs=4): err= 0: pid=1872: Tue Mar 31 21:17:02 2026

  write: IOPS=24.3k, BW=1518MiB/s (1592MB/s)(1334GiB/900002msec); 0 zone resets

    slat (nsec): min=1858, max=3022.6k, avg=7011.31, stdev=5409.78

    clat (usec): min=45, max=18725, avg=1309.85, stdev=444.93

     lat (usec): min=222, max=18732, avg=1316.86, stdev=444.77

    clat percentiles (usec):

     |  1.00th=[  603],  5.00th=[  750], 10.00th=[  832], 20.00th=[  971],

     | 30.00th=[ 1074], 40.00th=[ 1172], 50.00th=[ 1254], 60.00th=[ 1352],

     | 70.00th=[ 1483], 80.00th=[ 1614], 90.00th=[ 1844], 95.00th=[ 2040],

     | 99.00th=[ 2474], 99.50th=[ 2671], 99.90th=[ 4359], 99.95th=[ 7046],

     | 99.99th=[ 9765]

   bw (  MiB/s): min= 1356, max= 1680, per=100.00%, avg=1519.00, stdev=12.32, samples=7200

   iops        : min=21704, max=26888, avg=24303.47, stdev=197.11, samples=7200

  lat (usec)   : 50=0.01%, 100=0.01%, 250=0.01%, 500=0.22%, 750=4.78%

  lat (usec)   : 1000=18.62%

  lat (msec)   : 2=70.73%, 4=5.54%, 10=0.10%, 20=0.01%

  cpu          : usr=2.89%, sys=4.35%, ctx=11134123, majf=0, minf=145

  IO depths    : 1=0.0%, 2=0.0%, 4=0.0%, 8=100.0%, 16=0.0%, 32=0.0%, >=64=0.0%

     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%

     complete  : 0=0.0%, 4=100.0%, 8=0.1%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%

     issued rwts: total=0,21859541,0,0 short=0,0,0,0 dropped=0,0,0,0

     latency   : target=0, window=0, percentile=100.00%, depth=8
```

Run status group 0 (all jobs):

```
WRITE: bw=1518MiB/s (1592MB/s), 1518MiB/s-1518MiB/s (1592MB/s-1592MB/s), io=1334GiB (1433GB),
run=900002-900002msec
```

Disk stats (read/write):

```
vda: ios=48/22103432, merge=0/0, ticks=8/28655302, in_queue=28655310, util=100.00%

<shb-app04>

LOAD: (groupid=0, jobs=4): err= 0: pid=2527: Tue Mar 31 21:17:02 2026

  write: IOPS=24.2k, BW=1514MiB/s (1588MB/s)(1331GiB/900003msec); 0 zone resets

    slat (usec): min=2, max=4721, avg= 7.58, stdev= 5.15

    clat (usec): min=115, max=18122, avg=1312.37, stdev=444.96

     lat (usec): min=220, max=18127, avg=1319.96, stdev=444.85

    clat percentiles (usec):

     |  1.00th=[  611],  5.00th=[  750], 10.00th=[  832], 20.00th=[  971],

     | 30.00th=[ 1074], 40.00th=[ 1172], 50.00th=[ 1254], 60.00th=[ 1369],

     | 70.00th=[ 1483], 80.00th=[ 1614], 90.00th=[ 1844], 95.00th=[ 2040],

     | 99.00th=[ 2474], 99.50th=[ 2671], 99.90th=[ 4228], 99.95th=[ 7046],

     | 99.99th=[ 9765]

   bw (  MiB/s): min= 1332, max= 1685, per=100.00%, avg=1515.39, stdev=12.43, samples=7200

   iops        : min=21316, max=26968, avg=24245.70, stdev=198.91, samples=7200

  lat (usec)   : 250=0.01%, 500=0.22%, 750=4.64%, 1000=18.55%

  lat (msec)   : 2=70.90%, 4=5.59%, 10=0.10%, 20=0.01%

  cpu          : usr=3.20%, sys=4.23%, ctx=11118543, majf=0, minf=145

  IO depths    : 1=0.0%, 2=0.0%, 4=0.0%, 8=100.0%, 16=0.0%, 32=0.0%, >=64=0.0%

     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%

     complete  : 0=0.0%, 4=100.0%, 8=0.1%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%

     issued rwts: total=0,21808298,0,0 short=0,0,0,0 dropped=0,0,0,0

     latency   : target=0, window=0, percentile=100.00%, depth=8
```

Run status group 0 (all jobs):

```
WRITE: bw=1514MiB/s (1588MB/s), 1514MiB/s-1514MiB/s (1588MB/s-1588MB/s), io=1331GiB (1429GB),
run=900003-900003msec
```

Disk stats (read/write):

```
 vda: ios=48/22048814, merge=0/0, ticks=7/28639629, in_queue=28639635, util=100.00%
<shb-app01>
LOAD: (groupid=0, jobs=4): err= 0: pid=5624: Tue Mar 31 21:17:02 2026
  write: IOPS=33.0k, BW=2063MiB/s (2163MB/s)(1813GiB/900001msec); 0 zone resets
    slat (nsec): min=1904, max=1462.5k, avg=7953.35, stdev=5007.35
    clat (usec): min=28, max=17812, avg=960.61, stdev=371.86
     lat (usec): min=180, max=17814, avg=968.56, stdev=371.74
    clat percentiles (usec):
     |  1.00th=[  392],  5.00th=[  494], 10.00th=[  578], 20.00th=[  685],
     | 30.00th=[  766], 40.00th=[  832], 50.00th=[  914], 60.00th=[  996],
     | 70.00th=[ 1090], 80.00th=[ 1221], 90.00th=[ 1385], 95.00th=[ 1549],
     | 99.00th=[ 1893], 99.50th=[ 2089], 99.90th=[ 3097], 99.95th=[ 6521],
     | 99.99th=[ 8717]
   bw (  MiB/s): min= 1743, max= 2336, per=100.00%, avg=2064.31, stdev=19.07, samples=7200
   iops        : min=27894, max=37376, avg=33028.38, stdev=305.09, samples=7200
  lat (usec)   : 50=0.01%, 100=0.01%, 250=0.01%, 500=5.27%, 750=23.16%
  lat (usec)   : 1000=31.86%
  lat (msec)   : 2=39.03%, 4=0.59%, 10=0.07%, 20=0.01%
  cpu          : usr=3.32%, sys=5.59%, ctx=10965122, majf=0, minf=145
  IO depths    : 1=0.0%, 2=0.0%, 4=0.0%, 8=100.0%, 16=0.0%, 32=0.0%, >=64=0.0%
     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     complete  : 0=0.0%, 4=100.0%, 8=0.1%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     issued rwts: total=0,29711059,0,0 short=0,0,0,0 dropped=0,0,0,0
     latency   : target=0, window=0, percentile=100.00%, depth=8
```

Run status group 0 (all jobs):

```
WRITE: bw=2063MiB/s (2163MB/s), 2063MiB/s-2063MiB/s (2163MB/s-2163MB/s), io=1813GiB (1947GB),
run=900001-900001msec
```

Disk stats (read/write):

```
vda: ios=49/30041316, merge=0/0, ticks=6/28196967, in_queue=28196973, util=100.00%
All clients: (groupid=0, jobs=4): err= 0: pid=0: Tue Mar 31 21:17:02 2026
  write: IOPS=114k, BW=7150Mi (7498M)(6285GiB/900003msec); 0 zone resets
    slat (nsec): min=1858, max=4721.9k, avg=7561.12, stdev=5176.50
    clat (usec): min=28, max=18725, avg=1110.52, stdev=439.25
     lat (usec): min=180, max=18732, avg=1118.08, stdev=439.05
   bw (  MiB/s): min= 6197, max= 7962, per=100.00%, avg=7154.74, stdev=15.69, samples=28796
   iops        : min=99158, max=127406, avg=114473.72, stdev=250.98, samples=28796
  lat (usec)   : 50=0.01%, 100=0.01%, 250=0.01%, 500=3.13%, 750=15.17%
  lat (usec)   : 1000=26.13%
  lat (msec)   : 2=52.80%, 4=2.67%, 10=0.08%, 20=0.01%
  cpu          : usr=3.14%, sys=4.89%, ctx=43975897, majf=0, minf=580
  IO depths    : 1=0.0%, 2=0.0%, 4=0.0%, 8=100.0%, 16=0.0%, 32=0.0%, >=64=0.0%
     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     complete  : 0=0.0%, 4=100.0%, 8=0.1%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     issued rwts: total=0,102966594,0,0 short=0,0,0,0 dropped=0,0,0,0
```