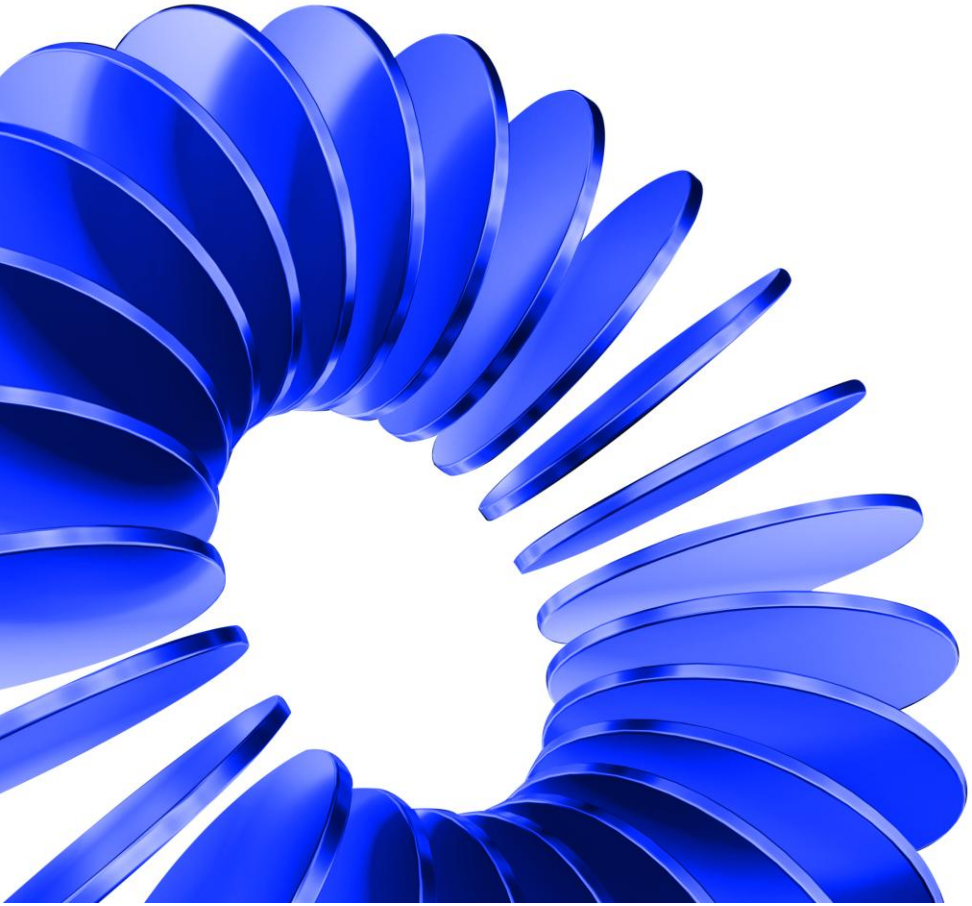
ОБЩЕСТВО С ОГРАНИЧЕННОЙ
ОТВЕТСТВЕННОСТЬЮ «РУСБИТЕХ-АСТРА»

СИСТЕМА ХРАНЕНИЯ ДАННЫХ TROK

РУКОВОДСТВО ПО НАСТРОЙКЕ S3



Москва, 2026г.

СОДЕРЖАНИЕ

1. ОБЩАЯ ИНФОРМАЦИЯ

Руководство описывает настройку доступа по протоколу S3 к хранилищу SDS.

Использование стандартизированного REST-API обеспечивает интеграцию решений для резервного копирования, аналитики и мультимедиа с единым объектным хранилищем без установки дополнительных драйверов и без сложных процедур настройки.

SDS — наиболее удобная основа для S3, поскольку программно отделяет функции хранения от аппаратуры, обеспечивая независимость от поставщика, горизонтальное масштабирование и богатый набор сервисов данных.

Служба trok-s3-gw (S3-совместимый шлюз) разворачивается как системный сервис на каждом узле хранения, имеющем локальную реплику соответствующего DRBD-ресурса.

Для корректного функционирования trok-s3-gw в составе кластера рекомендуется предварительно создать один или несколько DRBD-томов. Том данных – содержит пользовательские объекты, метаданные IAM (учётные записи, ключи доступа) и каталог для версионирования. Допускается использование одного тома с логическим разделением по каталогам либо выделение отдельных томов для каждой категории; в типовой конфигурации предполагается единый общий том.

Минимально рекомендуемый объём тома определяется планируемым объёмом пользовательских данных; метаданные IAM занимают незначительный объём.

Все DRBD-ресурсы должны быть заранее созданы и полностью синхронизированы между всеми узлами кластера.

1.1. Архитектура

Все три узла кластера являются равноправными шлюзовыми (gateway) узлами и в равной степени могут стать владельцами ресурса. На одном из них, а именно на узле <узел_шлюза_1>, размещён контроллер TROK (оркестратор DRBD). Однако это размещение не наделяет данный узел статусом выделенного

владельца S3-сервиса. Активный владелец (active owner) определяется автоматически на основе алгоритмов выбора и может оказаться на любом из трёх узлов. Резервный (spare) узел в данной конфигурации не предусмотрен: DRBD-ресурс развёрнут на всех трёх узлах одновременно (параметр place-count = 3), что обеспечивает тройную избыточность без выделения отдельного узла для горячего резерва.

Рисунок 1 – Модель взаимодействия подсистем S3 и SDS

2. НАСТРОЙКА S3-ШЛЮЗА НА БАЗЕ DRBD И DRBD-REACTOR

2.1. Подготовка к интеграции S3 и SDS

Базовые условия:

– Ресурсы предварительно созданы с использованием TROK, выступающего в роли DRBD-оркестратора.

– На каждом узле развернуты и функционируют службы Satellite.

– На одном или нескольких узлах развернута и активна служба Controller.

– Доступны и запущены утилиты уровня узла (инструменты контроля состояния и диагностики).

– drbdadm – управление и статус DRBD-ресурсов.

– drbd-reactorctl – управление наблюдателями и политиками drbd-reactor.

– lvm – управление физическими, групповыми и логическими томами.

– Обеспечены доступы и сетевая связность (адресация, маршрутизация, открытые порты); подготовлена требуемая дисковая разметка/конфигурация.

Перед началом работ обратите внимание: следующие системные имена и пути являются фиксированными и не подлежат изменению:

– trok-s3-gw@.service – шаблон системного юнита шлюза; экземпляры именуются по имени ресурса.

– /etc/trok-s3-gw.d/ – каталог конфигурации шлюза.

– /etc/drbd-reactor.d/ – каталог конфигурации drbd-reactor.

– ocf:heartbeat:Filesystem – агент OCF для монтирования файловой системы.

– ocf:heartbeat:IPaddr2 – агент OCF для управления виртуальным IP-адресом.

Убедитесь, что у вас есть SSH-доступ с правами администратора (root/sudo) ко всем трём узлам. На каждом узле откройте и проверьте доступность следующих TCP-портов:

Таблица 1 – Порты, используемые для настройки

Порт	Описание

22/tcp	SSH
7000–7999/tcp	DRBD-репликация (в процессе настройки выделяется по одному порту на каждый DRBD-ресурс)
7070/tcp	S3 API
7080/tcp	Admin API
7071/tcp	WebUI S3
50000/tcp	сервис TROK Controller
50002/tcp	сервис TROK Worker
80/tcp	Веб-сервер (Nginx) с HTTPS
443/tcp/	HTTPS (веб-сервер Nginx)
2049/tcp	NFS (при использовании NFS-транспорта; при необходимости могут потребоваться дополнительные порты, например rpcbind/mountd)
3260/tcp	iSCSI (при использовании iSCSI-транспорта)
445/tcp 139/tcp	SMB/CIFS (Samba; при использовании устаревшего NetBIOS могут дополнительно потребоваться 137/udp, 138/udp, 139/tcp)
4420/tcp	NVMe-oF (NVMe/TCP)

2.2. Предварительная настройка операционной системы

В первую очередь проверьте синхронизацию системного времени. Даже небольшое рассогласование часов вызывает ошибки подписи S3 (clock skew) и нарушает согласованность кластера. Настройте NTP на всех узлах с помощью chrony и проверьте состояние синхронизации.

Откройте терминал с правами администратора. Обновите список доступных пакетов и установите chrony – службу синхронизации системного времени (если она ранее не была установлена):

```
sudo apt-get update && sudo apt-get install -y chrony
```

Запустите службу и включите ее автозапуск:

```
sudo systemctl enable --now chrony
```

Проверьте сводку состояния синхронизации:

```
chronyc tracking
```

Просмотрите список источников времени и их статус:

```
chronyc sources
```

Синхронизация в Chrony происходит полностью автоматически в фоновом режиме. Для принудительной настройки можете использовать команду:

```
sudo chronyc -a makestep
```

Важно. В результате выполнения команды время может «прыгнуть» назад/вперёд – это влияет на таймеры, логи и некоторые сервисы. Лучше выполнять в окно обслуживания.

Проверьте, запущена ли служба drbd. Если служба не активна, выполните её запуск:

```
sudo systemctl is-active --quiet drbd || sudo systemctl start drbd
```

При необходимости просмотрите подробный статус службы для подтверждения корректного запуска:

```
sudo systemctl status drbd --no-pager -l
```

Проверьте, включена ли служба в автозагрузку. Если автозапуск не активирован, включите его:

```
sudo systemctl is-enabled --quiet drbd || sudo systemctl enable drbd
```

Подтвердите, что служба активна и включена в автозагрузку:

```
sudo systemctl is-active drbd

sudo systemctl is-enabled drbd
```

2.3. Подготовка файловой системы

Ниже представлена последовательность шагов по переводу DRBD-ресурса, содержащего данные для trok-s3-gw, в состояние Primary, созданию файловой системы и подготовке структуры каталогов. При использовании нескольких ресурсов повторите операции для каждого.

Откройте терминал с правами администратора. Проверьте список DRBD-ресурсов и их текущее состояние:

```
sudo drbdadm status
```

Пример вывода:

```
<имя_ресурса> role:Secondary
  disk:UpToDate
  <имя_узла_2> role:Secondary
    peer-disk:UpToDate
  <имя_узла_3> role:Secondary
    peer-disk:UpToDate
```

Переведите ресурс в состояние Primary на узле, где будет выполняться форматирование:

```
sudo drbdadm primary <имя_ресурса>
```

Убедитесь, что роль ресурса изменилась на Primary:

```
sudo drbdadm status
```

Создайте файловую систему на томе DRBD-ресурса (операция выполняется один раз на узле с ролью Primary):

```
sudo mkfs -t ext4 /dev/drbd/by-res/<имя_ресурса>/<номер_тома>
```

Здесь <имя_ресурса> – имя DRBD-ресурса.

Важно: создание файловой системы удалит все данные на выбранном томе.

На каждом узле кластера создайте каталог для монтирования DRBD-тома:

```
sudo mkdir -p /mnt/<имя_ресурса>/<номер_тома>
```

Смонтируйте том на активном узле (где ресурс в состоянии Primary):

```
sudo mount /dev/drbd/by-res/<имя_ресурса>/<номер_тома> /mnt/<имя_ресурса>/<номер_тома>
```

Создайте внутри тома структуру каталогов, соответствующую конфигурации: data, versioning и iam. Если <директория_монтирования> – это /mnt/<имя_ресурса>/<номер_тома>, выполните команду:

```
sudo mkdir -p <директория_монтирования>/data <директория_монтирования>/versioning <директория_монтирования>/iam
```

Установите права доступа на созданные каталоги:

```
sudo chmod 0777 <директория_монтирования>/data <директория_монтирования>/versioning <директория_монтирования>/iam
```

После создания структуры каталогов размонтируйте том:

```
sudo umount /mnt/<имя_ресурса>/<номер_тома>
```

Переведите ресурс обратно в состояние Secondary (если не планируете сразу настраивать drbd-reactor):

```
sudo drbdadm secondary <имя_ресурса>
```

2.4. Настройка drbd-reactor для trok-s3-gw

Посмотреть содержимое шаблона unit-файла сервиса trok-s3-gw:

```
systemctl cat trok-s3-gw@.service
```

Шаблон юнита systemd (unit template) устанавливается вместе с пакетом; ручное создание unit-файла не требуется. Вручную создаётся только конфигурационный файл экземпляра (instance).

Заполните файл /etc/trok-s3-gw.d/<имя_экземпляра>.conf, который должен присутствовать на всех gateway-узлах, актуальной информацией. Используйте команду:

```
sudo tee /etc/trok-s3-gw.d/<имя_экземпляра>.conf >/dev/null <<EOF
VGW_BACKEND=posix
VGW_BACKEND_ARG=<директория_монтирования>/data
VGW_VERSIONING_DIR=<директория_монтирования>/versioning
VGW_PORT=<плавающий_виртуальный_IP_кластера>:7070
VGW_ADMIN_PORT=<плавающий_виртуальный_IP_кластера>:7080
VGW_WEBUI_PORT=<плавающий_виртуальный_IP_кластера>:7071
VGW_WEBUI_NO_TLS=true
VGW_HEALTH=/health
VGW_REGION=<регион_S3>
VGW_CORS_ALLOW_ORIGIN=http://<плавающий_виртуальный_IP_кластера>:7071
VGW_IAM_DIR=<STATE_ROOT>/iam
ROOT_ACCESS_KEY_ID=<имя_пользователя>
ROOT_SECRET_ACCESS_KEY=<секрет>
EOF
```

Выберите подходящее значение переменной VGW_REGION; в качестве примера можно использовать ru-central-1.

VGW_WEBUI_NO_TLS=true отключает TLS для WebUI. Используйте это только в изолированном окружении.

Для эксплуатации выберите один из вариантов:

Первый вариант – терминировать TLS на самом шлюзе (и для S3 на <плавающий_виртуальный_IP_кластера>:7070, и для WebUI):

Укажите пути к сертификату и ключам:

– VGW_CERT=/etc/ssl/certs/your-cert.pem;

– VGW_KEY=/etc/ssl/private/your-key.pem;

– VGW_ADMIN_CERT=/etc/ssl/certs/your-cert.pem;

– VGW_ADMIN_CERT_KEY=/etc/ssl/private/your-key.pem.

Уберите из конфига строку VGW_WEBUI_NO_TLS=true.

Задайте HTTPS в VGW_CORS_ALLOW_ORIGIN (например, https://IP:7071).

Пример с локально выпущенным самоподписанным сертификатом:

```
sudo mkdir -p /etc/ssl/certs /etc/ssl/private
sudo openssl req -x509 -newkey rsa:2048 -nodes \
  -keyout /etc/ssl/private/your-key.pem \
```

```
  -out /etc/ssl/certs/your-cert.pem \
  -days 365 -subj "/CN=RU"
```

Пример конфигурации:

```
sudo tee /etc/trok-s3-gw.d/default.conf >/dev/null <<EOF
VGW_BACKEND=posix
VGW_BACKEND_ARG=/mnt/res1/0/data
VGW_VERSIONING_DIR=/mnt/res1/0/versioning
VGW_PORT=192.168.0.167:7070
VGW_ADMIN_PORT=192.168.0.167:7080
VGW_WEBUI_PORT=192.168.0.167:7071
VGW_HEALTH=/health
VGW_REGION=ru-central-1
VGW_CORS_ALLOW_ORIGIN=https://192.168.0.167:7071
VGW_IAM_DIR=/mnt/res1/0/iam
ROOT_ACCESS_KEY_ID=test
ROOT_SECRET_ACCESS_KEY=test
VGW_CERT=/etc/ssl/certs/your-cert.pem
VGW_KEY=/etc/ssl/private/your-key.pem
VGW_ADMIN_CERT=/etc/ssl/certs/your-cert.pem
VGW_ADMIN_CERT_KEY=/etc/ssl/private/your-key.pem
EOF
```

Второй вариант – терминировать TLS на reverse-proxy перед <плавающий_виртуальный_IP_кластера>:

Настройте прокси на HTTPS и замените http:// на https:// в VGW_CORS_ALLOW_ORIGIN и во всех клиентских проверках/URL.

Смените владельца файла на root и группу на root:

```
sudo chown root:root /etc/trok-s3-gw.d/<имя_экземпляра>.conf
```

Установите режим доступа 0640 (восьмерично: rw-r-----): владельцу – чтение и запись; группе – только чтение; прочим пользователям – никаких прав.

```
sudo chmod 0640 /etc/trok-s3-gw.d/<имя_экземпляра>.conf
```

Один и тот же файл /etc/trok-s3-gw.d/<имя_экземпляра>.conf должен присутствовать на всех gateway-узлах.

После создания перезагрузите конфигурацию systemd:

```
sudo systemctl daemon-reload
```

Проверьте ручной запуск сервиса с указанием имени экземпляра (в примере default) на активном узле после монтирования тома:

```
   sudo              mount              /dev/drbd/by-res/<имя_ресурса>/<номер_тома>
/mnt/<имя_ресурса>/<номер_тома>

   sudo systemctl start trok-s3-gw@default.service

   sudo systemctl status trok-s3-gw@default.service
```

Остановите сервис и размонтируйте том перед настройкой drbd-reactor:

```
   sudo systemctl stop trok-s3-gw@<имя_конфигурации>

   sudo umount /mnt/drbd
```

Создайте конфигурационный файл промоутера на каждом узле:

```
   sudo touch /etc/drbd-reactor.d/trok-s3-gw.toml
```

Откройте файл и добавьте следующую конфигурацию (подставьте свои имена ресурсов, номера томов, плавающий IP и т.д.):

```
   [[promoter]]
   id = "trok-s3-gw"

   [promoter.resources]

     [promoter.resources.<имя_ресурса>]
       on-drbd-demote-failure = "reboot-immediate"
       runner = "systemd"
       stop-services-on-exit = true

       start = [
         """ocf:heartbeat:Filesystem          fs_state          device=/dev/drbd/by-
res/<имя_ресурса>/<номер_тома>            directory=/mnt/<имя_ресурса>/<номер_тома>
fstype=ext4""",
         """ocf:heartbeat:IPaddr2     vip     ip=<плавающий_виртуальный_IP_кластера>
cidr_netmask=24""",
         """trok-s3-gw@<имя_экземпляра>.service"""
       ]

       on-drbd-demote-failure = "reboot"
```

Разбор конфигурации:

– **[[promoter]] id = "trok-s3-gw"**

Открывает новый блок описания промоутера с идентификатором trok-s3-gw (используется в логах/метриках и позволяет иметь несколько независимых промоутеров). Позволяет определить набор правил и действий, которые drbd-reactor будет выполнять при изменении роли ресурса (promote/demote).

– **[promoter.resources]**

Блок описания DRBD-ресурсов, которыми управляет данный промоутер.

– **[promoter.resources.<имя_ресурса>]**

Блок настроек конкретного DRBD-ресурса. Вместо <имя_ресурса> укажите реальное имя ресурса, для которого настраиваются действия при переходах ролей.

– **on-drbd-demote-failure = "reboot-immediate"**

Если при переключении ресурса DRBD из Primary в Secondary (demote) произойдёт ошибка, узел будет немедленно перезагружён для предотвращения split-brain и восстановления кластерного состояния.

– **runner = "systemd"**

Определяет механизм управления прикладными службами и OCF-ресурсами. Значение systemd означает, что запуск/остановка и мониторинг сервисов выполняются стандартным менеджером systemd.

– **stop-services-on-exit = true**

При остановке/выходе промоутера (или при демоуте ресурса) он останавливает все ранее запущенные сервисы и ресурсы, обеспечивая корректную уборку.

– **start = [...]**

Массив действий, выполняемых автоматически при промоуте ресурса в Primary. Порядок критичен: сначала готовится файловая система и точки монтирования, затем запускается прикладной сервис, после чего (опционально) поднимается плавающий IP.

– **"""ocf:heartbeat:Filesystem <имя_fs_ресурса> device=/dev/drbd/by-res/<имя_ресурса>/<номер_тома> directory=/mnt/<имя_ресурса>/<номер_тома> fstype=ext4 run_fsck=no""",**

Монтирует том DRBD по стабильному пути /dev/drbd/by-res/<имя_ресурса>/<номер_тома> в /mnt/drbd с типом файловой системы ext4.

– **""" ocf:heartbeat:IPaddr2 vip ip=<плавающий_виртуальный_IP_кластера> cidr_netmask=24""",**

Поднимает плавающий IP-адрес для клиентского доступа к сервису. Этот адрес управляется кластером и всегда активен только на текущем узле Primary. Убедитесь, что указаны корректные сеть/маска, адрес свободен и маршрутизация настроена на всех узлах одинаково.

– **""" trok-s3-gw@<имя_экземпляра>.service""",**

Запускает экземпляр сервиса S3-шлюза через systemd.

– **on-drbd-demote-failure = "reboot"**

При ошибке демоута DRBD (перехода из Primary в Secondary) узел выполняет штатную перезагрузку (reboot) для предотвращения split-brain и восстановления кластера.

Запустите сервис drbd-reactor и включите автозапуск при старте ОС:

```
systemctl enable --now drbd-reactor
```

Проверьте текущее состояние сервиса (без перехода в постраничный просмотр):

```
systemctl status drbd-reactor --no-pager
```

Проверьте статус инстанса промоутера с id “trok-s3-gw”; если инстанс ещё не создан/не запущен и команда вернёт ошибку, оператор || true подавит её (скрипт продолжит выполнение):

```
sudo drbd-reactorctl status trok-s3-gw || true
```

2.5. Проверка работоспособности

После первоначальной настройки убедитесь в корректной работе DRBD-ресурса и S3-шлюза. Выполните последовательность действий, описанную ниже.

Для подтверждения статуса Primary и наличия смонтированного тома выполните команду:

```
mount | grep /mnt/drbd
```

После выполнения команда должна вернуть строку с точкой монтирования /mnt/drbd.

Для проверки работы сервиса trok-s3-gw и прослушивания портов 7070 (S3) и 7071 (WebUI) выполните команду:

```
ss -tulpn | grep -E '7070|7071'
```

После выполнения в выводе должны отображаться процессы, слушающие указанные порты.

Для проверки доступности S3-API через шлюз выполните команду:

```
aws s3 --endpoint-url http://192.168.10.199:7070 ls
```

После выполнения в выводе должен появиться список бакетов или пустой результат без ошибки.

Откройте в браузере адрес http://192.168.10.199:7071 (при наличии). Убедитесь, что страница WebUI загружается без ошибок.

Для проверки отказоустойчивости смоделируйте отказ активного узла. Используйте один из вариантов:

- Выключите активный узел.
- Переведите ресурс в состояние Secondary вручную.

После выполнения убедитесь в корректном переключении:

- Второй узел автоматически становится Primary.
- Том монтируется на втором узле.
- Сервис trok-s3-gw запускается и слушает порты 7070 и 7071.
- Плавающий IP 192.168.10.199 переходит на второй узел.