ОПЕРАЦИОННАЯ СИСТЕМА АЛЬТ СЕРВЕР ВИРТУАЛИЗАЦИИ 9.2

Описание функциональных характеристик

СОДЕРЖАНИЕ

1 ОБЩИЕ СВЕДЕНИЯ ОБ ОС АЛЬТ СЕРВЕР ВИРТУАЛИЗАЦИИ 9.2

1.1 Краткое описание возможностей

Операционная система «Альт Сервер Виртуализации» (далее – ОС «Альт Сервер Виртуализации»), представляет собой совокупность интегрированных программ, созданных на основе ОС «Linux», и обеспечивает обработку, хранение и передачу информации в круглосуточном режиме эксплуатации.

ОС «Альт Сервер Виртуализации» – серверный дистрибутив, нацеленный на предоставление функций виртуализации в корпоративной инфраструктуре. Дистрибутив включает в себя средства виртуализации:

- вычислений (ЦПУ и память);
- сети;
- хранения данных.

Управление системой виртуализации возможно через командный интерфейс, веб-интерфейс, с использованием API.

ОС «Альт Сервер Виртуализации» представляет собой решение уровня предприятия, позволяющее осуществить миграцию на импортозамещающее программное и аппаратное обеспечение.

ОС «Альт Сервер Виртуализации» обладает следующими функциональными характеристиками:

- обеспечивает возможность обработки, хранения и передачи информации;
- обеспечивает возможность функционирования в многозадачном режиме (одновременное выполнение множества процессов);
- обеспечивает возможность масштабирования системы: возможна эксплуатация ОС как на одной ПЭВМ, так и в информационных системах различной архитектуры;
- обеспечивает многопользовательский режим эксплуатации;
- обеспечивает поддержку мультипроцессорных систем;
- обеспечивает поддержку виртуальной памяти;
- обеспечивает поддержку запуска виртуальных машин;
- обеспечивает сетевую обработку данных, в том числе разграничение доступа к сетевым пакетам.

ОС «Альт Сервер Виртуализации» предоставляет 4 основных типа установки:

- «Базовый гипервизор». Включает в себя поддержку виртуализации KVM на уровне ядра Linux, утилиты запуска виртуальных машин qemu и унифицированный интерфейс создания

и настройки виртуального окружения libvirt. Устанавливается на отдельно стоящий сервер или группу независимых серверов. Для управления используются интерфейс командной строки virsh или графическое приложение virt-manager на рабочей станции администратора.

- «Кластер серверов виртуализации на основе проекта PVE». Устанавливается на группу серверов (до 32 штук). Предназначено для управления виртуальным окружением KVM и контейнерами LXC, виртуальным сетевым окружением и хранилищем данных. Для управления используется интерфейс командной строки, а также веб-интерфейс. Возможна интеграция с корпоративными системами аутентификации (AD, LDAP и другие на основе PAM).
- «Облачная виртуализация уровня предприятия на основе проекта OpenNebula». Для использования необходим 1 или 3 и более серверов управления (могут быть виртуальными), и группа серверов для запуска виртуальных окружений KVM или контейнеров LXC. Возможна интеграция с корпоративными системами аутентификации.
- «Контейнерная виртуализация». К использования предлагаются Docker, Podman или LXC/LXD. Для построения кластера и управления контейнерами возможно использование Kubernetes.

ОС «Альт Сервер Виртуализации» поддерживает клиент-серверную архитектуру и может обслуживать процессы как в пределах одной компьютерной системы, так и процессы на других ПЭВМ через каналы передачи данных или сетевые соединения.

1.2 Структура программных средств

ОС «Альт Сервер Виртуализации» состоит из набора компонентов предназначенных для реализации функциональных задач необходимых пользователям (должностным лицам для выполнения определённых должностными инструкциями, повседневных действий) и поставляется в виде дистрибутива и комплекта эксплуатационной документации.

В структуре ОС «Альт Сервер Виртуализации» можно выделить следующие функциональные элементы:

- ядро ОС;
- системные библиотеки;
- утилиты и драйверы;
- средства обеспечения информационной безопасности;
- системные приложения;
- средства обеспечения облачных и распределенных вычислений, средства виртуализации и системы хранения данных;
- системы мониторинга и управления;
- средства подготовки исполнимого кода;

- средства версионного контроля исходного кода;
- библиотеки подпрограмм (SDK);
- среды разработки, тестирования и отладки;
- интерактивные рабочие среды;
- программные серверы;
- веб-серверы;
- системы управления базами данных;
- командные интерпретаторы.

Ядро ОС «Альт Сервер Виртуализации» управляет доступом к оперативной памяти, сети, дисковым и прочим внешним устройствам. Оно запускает и регистрирует процессы, управляет разделением времени между ними, реализует разграничение прав и определяет политику безопасности, обойти которую, не обращаясь к нему, нельзя.

Ядро работает в режиме «супервизора», позволяющем ему иметь доступ сразу ко всей оперативной памяти и аппаратной таблице задач. Процессы запускаются в «режиме пользователя»: каждый жестко привязан ядром к одной записи таблицы задач, в которой, в числе прочих данных, указано, к какой именно части оперативной памяти этот процесс имеет доступ. Ядро постоянно находится в памяти, выполняя системные вызовы – запросы от процессов на выполнение этих подпрограмм.

Системные библиотеки – наборы программ (пакетов программ), выполняющие различные функциональные задачи и предназначенные для динамического подключения к работающим программам, которым необходимо выполнение этих задач.

ОС «Альт Сервер Виртуализации» предоставляет набор дополнительных служб, востребованных в инфраструктуре виртуализации любой сложности и архитектуры:

- сервер сетевой файловой системы NFS;
- распределённая сетевая файловая система CEPH;
- распределённая сетевая файловая система GlusterFS;
- поддержка iSCSI как в качестве клиента, так и создание сервера;
- сетевые службы DNS и DHCP;
- виртуальный сетевой коммутатор Open vSwitch;
- служба динамической маршрутизации bird с поддержкой протоколов BGP, OSPF и др.;
- сетевой балансировщик нагрузки HAProxy, keepallived;
- веб-серверы Apache и Nginx.

В ОС «Альт Сервер Виртуализации» входят агенты мониторинга (Zabbix, telegraf, Prometheus) и архивирования (Bacula, UrBackup), которые могут использоваться совместно с сервисами на ОС «Альт Сервер».

2 ЗАГРУЗКА ОПЕРАЦИОННОЙ СИСТЕМЫ

2.1 Настройка загрузки

Вызов ОС «Альт Сервер Виртуализации», установленной на жесткий диск, происходит автоматически и выполняется после запуска ПЭВМ и отработки набора программ BIOS. ОС «Альт Сервер Виртуализации» вызывает специальный загрузчик.

Загрузчик настраивается автоматически и включает в свое меню все системы, установку которых на ПЭВМ он определил. Поэтому загрузчик также может использоваться для вызова других ОС, если они установлены на компьютере.

Примечание. При наличии на компьютере нескольких ОС (или при наличии нескольких вариантов загрузки), оператор будет иметь возможность выбрать необходимую ОС (вариант загрузки). В случае если пользователем ни один вариант не был выбран, то по истечении заданного времени будет загружена ОС (вариант загрузки), заданные по умолчанию.

При стандартной установке ОС «Альт Сервер Виртуализации» в начальном меню загрузчика доступны несколько вариантов загрузки (Рис. 1): обычная загрузка, загрузка с дополнительными параметрами (например, «recovery mode» – загрузка с минимальным количеством драйверов), загрузка в программу проверки оперативной памяти (memtest).

По умолчанию, если не были нажаты управляющие клавиши на клавиатуре, загрузка ОС «Альт Сервер Виртуализации» продолжится автоматически после небольшого времени ожидания (обычно несколько секунд). Нажав клавишу <Enter>, можно начать загрузку немедленно.

Для выбора дополнительных параметров загрузки нужно выбрать пункт «Дополнительные параметры для ALT Virtualization Server 9.2».

Для выполнения тестирования оперативной памяти нужно выбрать пункт «Memtest86+-5.01».

Нажатием клавиши <E> можно вызвать редактор параметров загрузчика GRUB и указать параметры, которые будут переданы ядру ОС при загрузке.

Примечание. Если при установке системы был установлен пароль на загрузчик, потребуется ввести имя пользователя «boot» и заданный на шаге «Установка загрузчика» пароль.

В процессе загрузки ОС «Альт Сервер Виртуализации» пользователь может следить за информацией процесса загрузки, которая отображает этапы запуска различных служб и программных серверов в виде отдельных строк (Рис. 2), на экране монитора.

Варианты загрузки

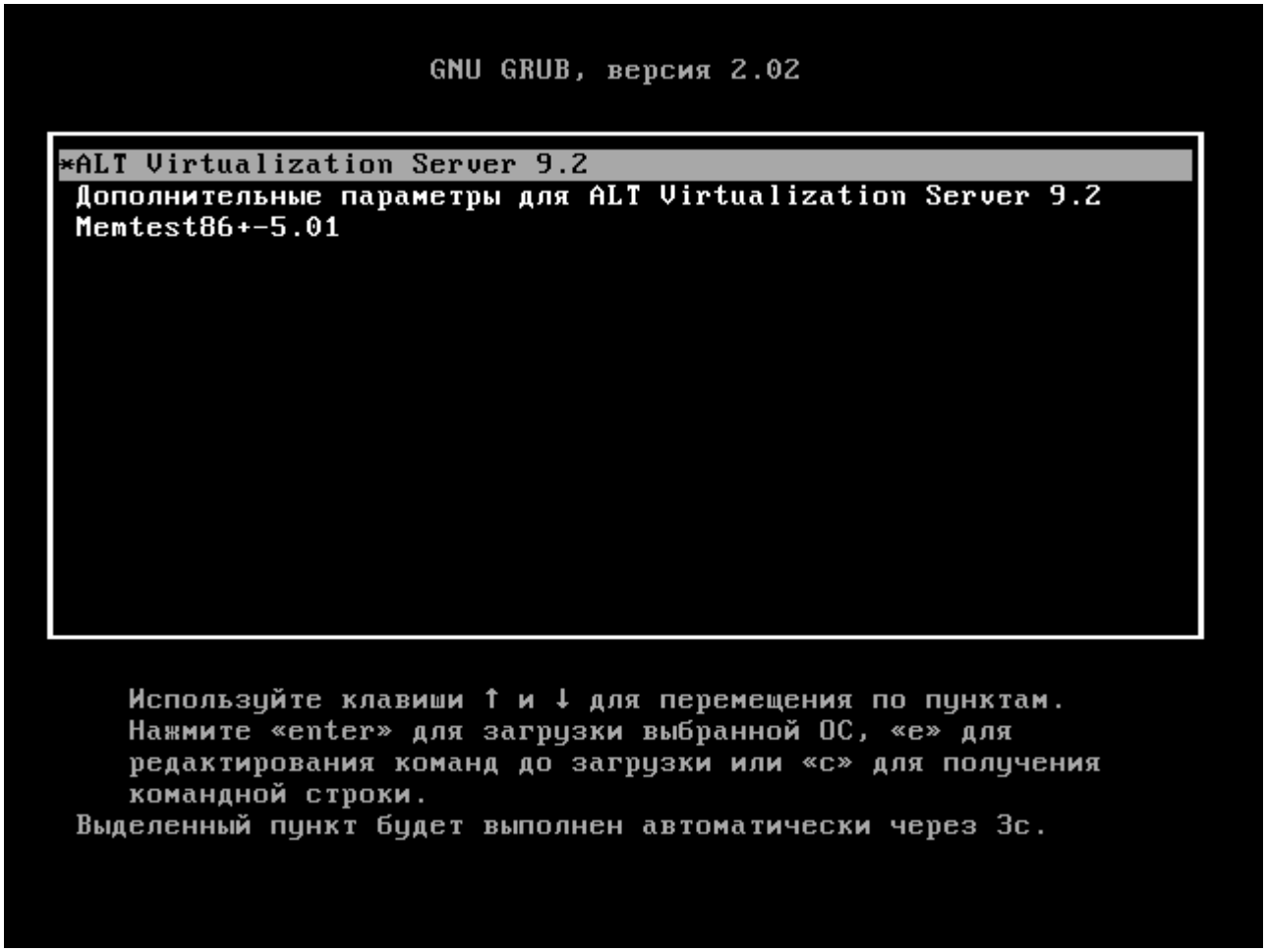

Рис. 1

Загрузка ОС

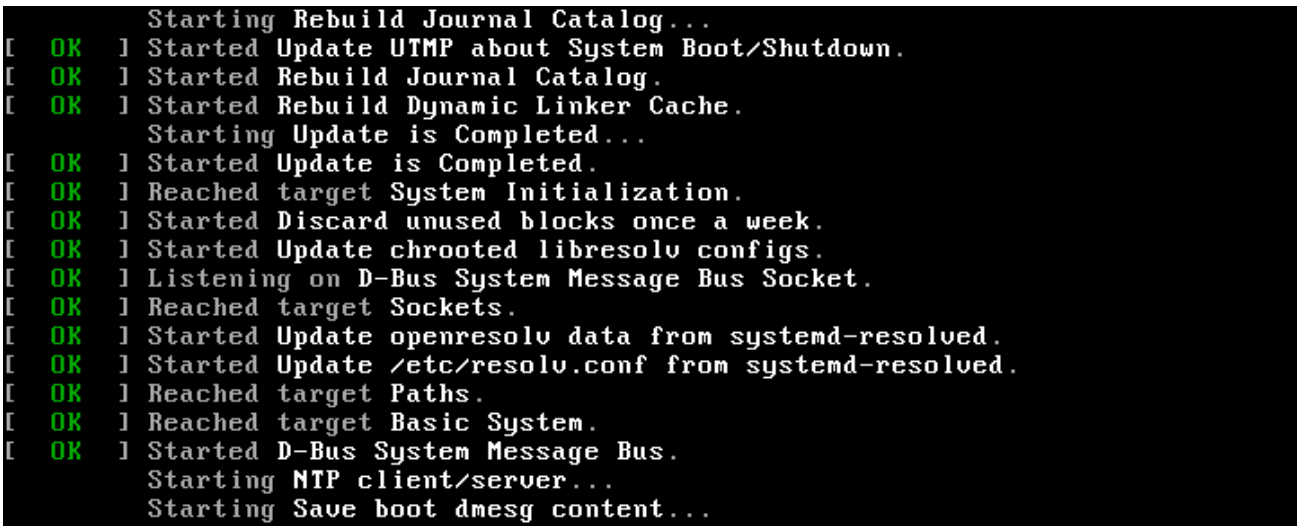

Рис. 2

При этом каждая строка начинается словом вида [XXXXXXX] (FAILED или OK), являющегося признаком нормального или ненормального завершения этапа загрузки. Слово

XXXXXXX=FAILED (авария) свидетельствует о неуспешном завершении этапа загрузки, что требует вмешательства и специальных действий администратора системы.

Загрузка ОС может занять некоторое время, в зависимости от производительности компьютера. Основные этапы загрузки операционной системы – загрузка ядра, подключение (монтирование) файловых систем, запуск системных служб – периодически могут дополняться проверкой файловых систем на наличие ошибок. В этом случае время ожидания может занять больше времени, чем обычно.

2.2 Получение доступа к зашифрованным разделам

В случае если был создан шифрованный раздел, потребуется вводить пароль при обращении к этому разделу.

Например, если был зашифрован домашний раздел `/home`, то для того, чтобы войти в систему, потребуется ввести пароль этого раздела и затем нажать <Enter>.

Примечание. Если не ввести пароль за отведенный промежуток времени, то загрузка системы завершится ошибкой. В этом случае следует перезагрузить систему, нажав для этого два раза <Enter>, а затем клавиши <Ctrl>+<Alt>+<Delete>.

2.3 Вход и работа в системе в консольном режиме

Стандартная установка ОС «Альт Сервер Виртуализации» включает базовую систему, работающую в консольном режиме.

При загрузке в консольном режиме работа загрузчика ОС «Альт Сервер Виртуализации» завершается запросом на ввод логина и пароля учетной записи (Рис. 3).

Запрос на ввод логина

Рис. 3

Примечание. После загрузки будут показаны имя и IP адрес компьютера (Рис. 5), а также адрес доступа к панели управления (если были установлены OpenNebula или PVE)

IP адрес компьютера и адрес панели управления PVE

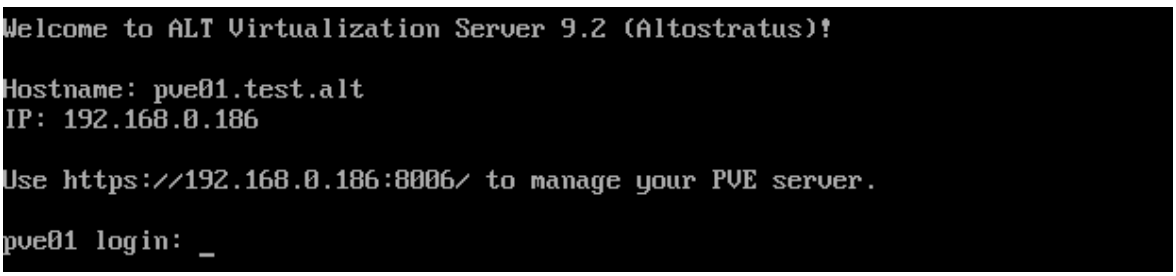

Рис. 4

Для дальнейшего входа в систему необходимо ввести логин и пароль учетной записи пользователя.

В случае успешного прохождения процедуры аутентификации и идентификации будет выполнен вход в систему. ОС «Альт Сервер Виртуализации» перейдет к штатному режиму работы и предоставит дальнейший доступ к консоли (Рис. 5).

Приглашение для ввода команд

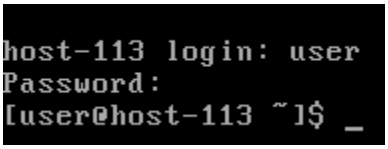

Рис. 5

2.4 Виртуальная консоль

В процессе работы ОС «Альт Сервер Виртуализации» активно несколько виртуальных консолей. Каждая виртуальная консоль доступна по одновременному нажатию клавиш <Ctrl>, <Alt> и функциональной клавиши с номером этой консоли от <F1> до <F6>.

На первых шести виртуальных консолях (от <Ctrl>+<Alt>+<F1> до <Ctrl>+<Alt>+<F6>) пользователь может зарегистрироваться и работать в текстовом режиме. Двенадцатая виртуальная консоль (<Ctrl>+<Alt>+<F12>) выполняет функцию системной консоли – на нее выводятся сообщения о происходящих в системе событиях.

3 OPENNEBULA

OpenNebula – это платформа облачных вычислений для управления разнородными инфраструктурами распределенных центров обработки данных. Платформа OpenNebula управляет виртуальной инфраструктурой центра обработки данных для создания частных, общедоступных и гибридных реализаций инфраструктуры как службы.

Облачная архитектура определяется 3-мя элементами: хранилищем данных, сетью и системой виртуализации.

OpenNebula состоит из следующих компонентов:

- Сервер управления (Front-end) – на нём выполняются сервисы OpenNebula;
- Серверы с виртуальными машинами;
- Хранилище данных – содержит образы ВМ;
- Физическая сеть – обеспечивает связь между хранилищем данных, серверами с ВМ, поддерживает VLAN-ы для ВМ, а также управление сервисами OpenNebula.

Примечание. Компоненты OpenNebula будут установлены в систему, если при установке дистрибутива выбрать профиль «Вычислительный узел Opennebula KVM» или «Сервер Opennebula».

3.1 Планирование ресурсов

3.1.1 Сервер управления

Минимальные требования к серверу управления показаны в таблице 1.

Таблица 1 – Минимальные требования к серверу управления

Ресурс	Минимальное значение
Оперативная память	2 ГБ
CPU	1 CPU (2 ядра)
Диск	100 ГБ
Сеть	2 интерфейса

Максимальное количество серверов (узлов виртуализации), управляемых одним сервером управления, зависит от инфраструктуры, особенно от производительности хранилища. Обычно рекомендуется не управлять более чем 500 серверами из одной точки, хотя существуют примеры с более чем 1000 серверами.

3.1.2 Серверы виртуализации

Серверы виртуализации – это физические машины, на которых выполняются виртуальные машины. Подсистема виртуализации – это компонент, который отвечает за связь с гипервизором,

установленным на узлах, и выполнение действий, необходимых для каждого этапа жизненного цикла виртуальной машины (ВМ).

Серверы (узлы) виртуализации имеют следующие характеристики и их рекомендованные значения:

- CPU – в обычных условиях каждое ядро, предоставляемое ВМ, должно быть реальным ядром физического процессора. Например, для обслуживания 40 ВМ с двумя процессорами в каждой, облако должно иметь 80 физических ядер. При этом они могут быть распределены по разным серверам: 10 серверов с восемью ядрами или 5 серверов с 16 ядрами на каждом. В случае перераспределения недостаточных ресурсов используются атрибуты CPU и VCPU: CPU определяет физические ядра, выделенные для ВМ, а VCPU – виртуальные ядра для гостевой ОС;
- Память – по умолчанию, OpenNebula не предоставляет памяти для гостевых систем больше, чем есть на самом деле. Желательно рассчитывать объём памяти с запасом в 10% на гипервизор. Например, для 45 ВМ с 2 ГБ памяти на каждой, необходимо 90 ГБ физической памяти. Важным параметром является количество физических серверов: каждый сервер должен иметь 10% запас для работы гипервизора, так, 10 серверов с 10 ГБ памяти на каждом могут предоставить по 9 ГБ для виртуальных машин и смогут обслужить 45 машин из этого примера (10% от 10 ГБ = 1 ГБ на гипервизор).

3.1.3 Хранилище данных

OpenNebula работает с двумя видами данных в хранилище: образцами виртуальных машин и образами (дисками) самих ВМ.

В хранилище образов (Images Datastore) OpenNebula хранит все зарегистрированные образы, которые можно использовать для создания ВМ.

Системное хранилище (System Datastore) – используется для хранения дисков виртуальных машин, работающих в текущий момент. Образы дисков перемещаются, или клонируются, в хранилище образов или из него при развертывании и отключении ВМ, при подсоединении или фиксировании мгновенного состояния дисков.

Одним из основных способов управления хранилищем данных является ограничение хранилища, доступного для пользователей, путем определения квот по максимальному количеству ВМ, а также максимального объема энергозависимой памяти, который может запросить пользователь, и обеспечения достаточного пространства хранения системных данных и образов, отвечающего предельным установленным квотам. OpenNebula позволяет администратору добавлять хранилища системных данных и образов.

Планирование хранилища – является критически важным аспектом, поскольку от него зависит производительность облака. Размер хранилищ сильно зависит от базовой технологии.

Например, при использовании Ceph для среднего по размеру облака, необходимо взять как минимум 3 сервера в следующей конфигурации: 5 дисков по 1 ТБ, 16 ГБ памяти, 2 CPU по 4 ядра в каждом и как минимум 2 сетевые карты.

3.1.4 Сетевая инфраструктура

Сетевая инфраструктура должна быть спланирована так, чтобы обеспечить высокую надёжность и пропускную способность. Рекомендуется использовать 2 сетевых интерфейса на сервере управления и по 4 на каждом сервере виртуализации (публичный, внутренний, для управления и для связи с хранилищем).

3.2 Запуск сервера управления OpenNebula

3.2.1 Установка пароля для пользователя oneadmin

При установке OpenNebula система автоматически создает нового пользователя oneadmin, все дальнейшие действия по управлению OpenNebula необходимо выполнять от этого пользователя.

Примечание. Файл `/var/lib/one/.one/one_auth` будет создан со случайно сгенерированным паролем. Необходимо поменять этот пароль перед запуском OpenNebula.

Для установки пароля для пользователя oneadmin необходимо выполнить команду:

```
# passwd oneadmin
```

Теперь зайдя под пользователем oneadmin, следует заменить содержимое `/var/lib/one/.one/one_auth`. Он должен содержать следующее: `oneadmin: <пароль>`. Например:

```
$ echo "oneadmin:mypassword" > ~/.one/one_auth
```

3.2.2 Настройка MySQL (MariaDB) для хранения конфигурации

По умолчанию OpenNebula работает с SQLite. Если планируется использовать OpenNebula с MySQL, следует настроить данную конфигурацию перед первым запуском OpenNebula, чтобы избежать проблем с учетными данными oneadmin и serveradmin.

Примечание. Задать пароль root для mysql и настройки безопасности:

```
# mysql_secure_installation
```

Создать нового пользователя, предоставить ему привилегии в базе данных opennebula (эта база данных будет создана при первом запуске OpenNebula) и настроить уровень изоляции:

```
$ mysql -u root -p
Enter password:

MariaDB > GRANT ALL PRIVILEGES ON opennebula.* TO 'oneadmin' IDENTIFIED BY
'<thepassword>';
Query OK, 0 rows affected (0.003 sec)
```

```
MariaDB > SET GLOBAL TRANSACTION ISOLATION LEVEL READ COMMITTED;
Query OK, 0 rows affected (0.001 sec)

MariaDB > quit
```

Перед запуском сервера OpenNebula в первый раз необходимо настроить параметры доступа к базе данных в конфигурационном файле `/etc/one/oned.conf`:

```
#DB = [ BACKEND = "sqlite" ]

# Sample configuration for MySQL
DB = [ BACKEND = "mysql",
       SERVER  = "localhost",
       PORT    = 0,
       USER    = "oneadmin",
       PASSWD  = "<thepassword>",
       DB_NAME = "opennebula",
       CONNECTIONS = 50 ]
```

3.2.3 Запуск OpenNebula

Для запуска OpenNebula необходимо выполнить следующие команды:

```
# systemctl start opennebula
# systemctl start opennebula-sunstone
```

3.2.4 Проверка установки

После запуска OpenNebula в первый раз, следует проверить, что команды могут подключаться к демону OpenNebula. Это можно сделать в командной строке или в графическом интерфейсе пользователя: Sunstone.

В командной строке:

```
$ oneuser show
USER 0 INFORMATION
ID              : 0
NAME            : oneadmin
GROUP           : oneadmin
PASSWORD        : 3bc15c8aae3e4124dd409035f32ea2fd6835efc9
AUTH_DRIVER     : core
ENABLED         : Yes
USER TEMPLATE
TOKEN_PASSWORD="ec21d27e2fe4f9ed08a396cbd47b08b8e0a4ca3c"
VMS USAGE & QUOTAS
```

```
VMS USAGE & QUOTAS - RUNNING
DATASTORE USAGE & QUOTAS
NETWORK USAGE & QUOTAS
IMAGE USAGE & QUOTAS
```

Также можно попробовать войти в веб-интерфейс Sunstone. Для этого необходимо перейти по адресу http://<внешний адрес>:9869. Если все в порядке, будет предложена страница входа (Рис. 6).

Страница авторизации opennebula-sunstone

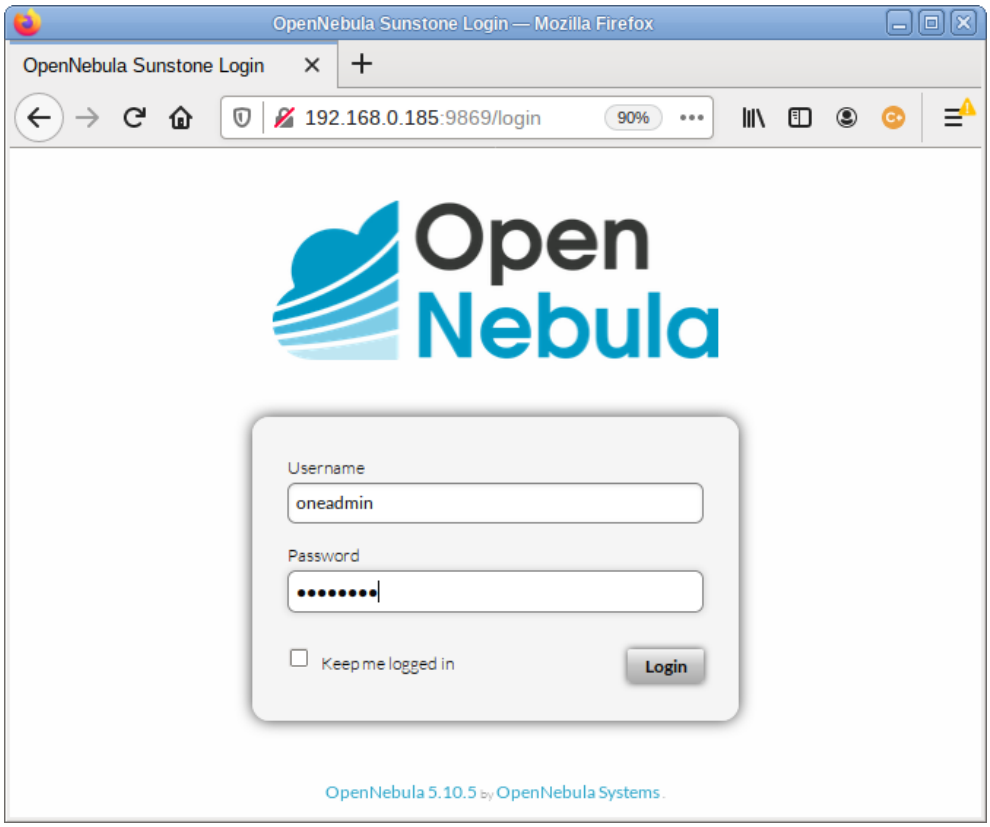

Рис. 6

Необходимо ввести в соответствующие поля имя пользователя (oneadmin) и пароль пользователя (тот, который находится в файле `/var/lib/one/.one/one_auth`).

После входа в систему будет доступна панель инструментов (Рис. 7).

Для смены языка интерфейса необходимо в левом меню выбрать пункт «Settings», и на открывшейся странице в выпадающем списке «Language» выбрать пункт «Russian (ru_RU)» (Рис. 8). Язык интерфейса будет изменён на русский (Рис. 9).

Панель инструментов opennebula-sunstone

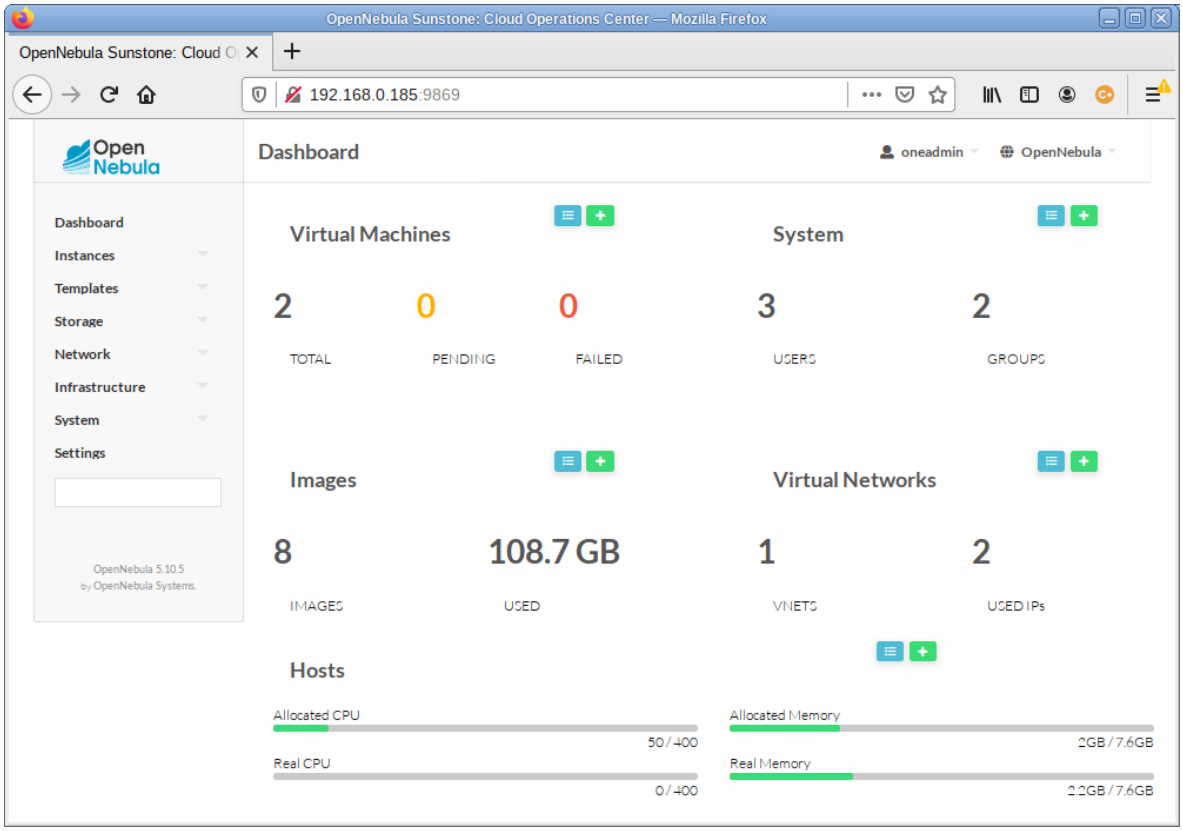

Рис. 7

Выбор языка интерфейса

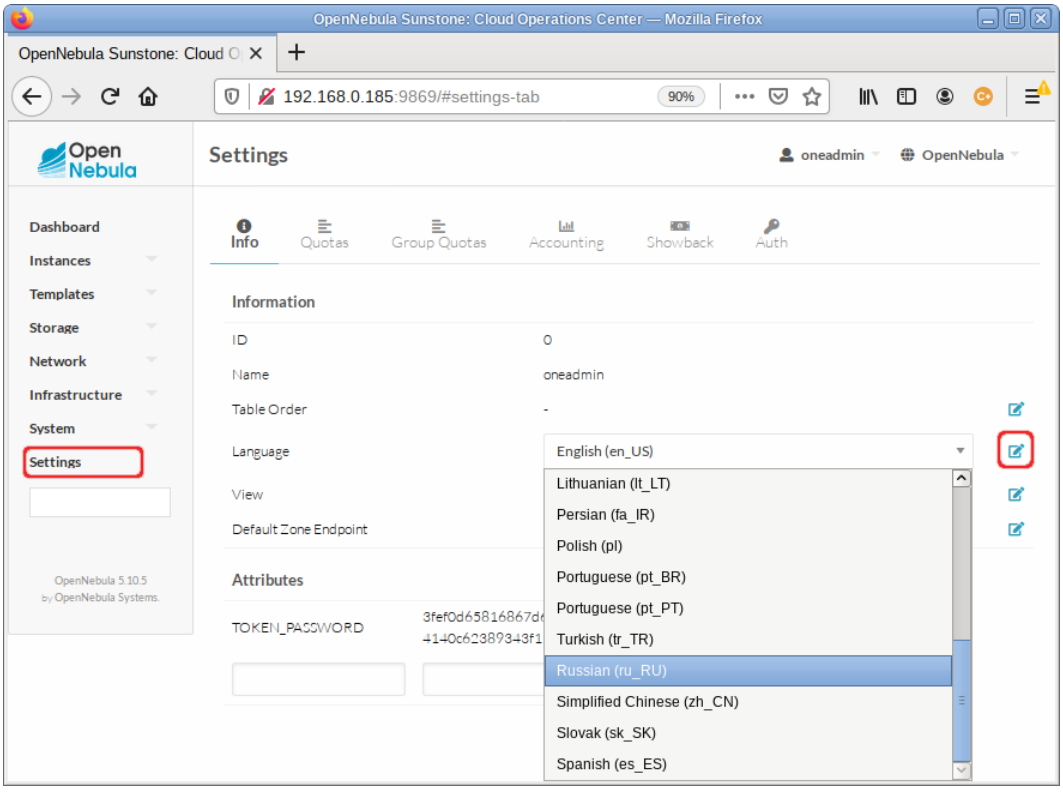

Рис. 8

Панель инструментов opennebula-sunstone с русским языком интерфейса

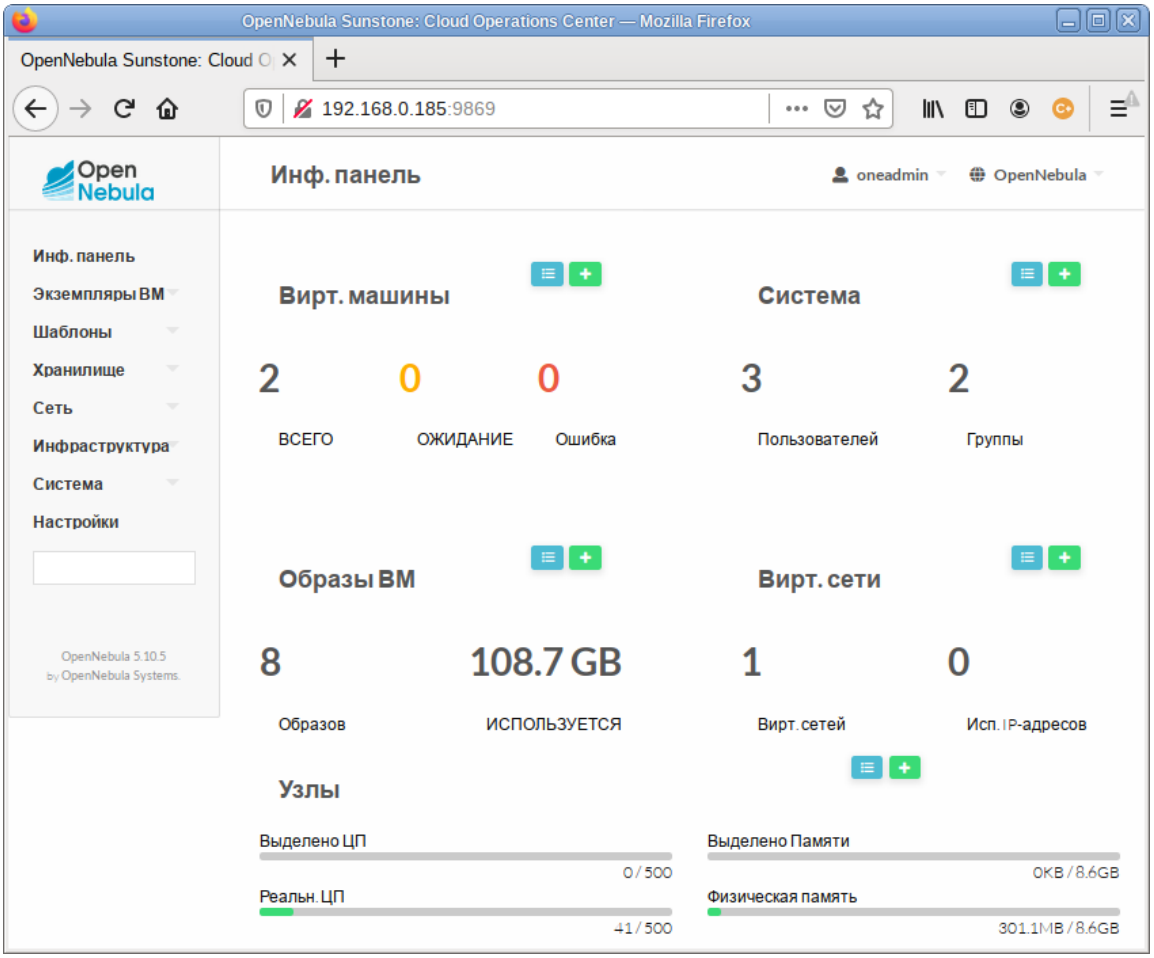

Рис. 9

3.2.5 Ключи для доступа по SSH

Сервер управления OpenNebula подключается к хостам гипервизора по SSH. Необходимо распространить открытый ключ пользователя oneadmin со всех машин в файл `/var/lib/one/.ssh/authorized_keys` на всех машинах.

При установке сервера управления OpenNebula ключ SSH был сгенерирован и добавлен в авторизованные ключи. Необходимо синхронизировать id_rsa, id_rsa.pub и authorized_keys сервера управления и узлов. Кроме того, следует создать файл known_hosts и также синхронизировать его с узлами. Чтобы создать файл known_hosts, необходимо выполнить следующую команду (от пользователя oneadmin на сервере управления) со всеми именами узлов и именем сервера управления в качестве параметров:

```
$ ssh-keyscan <сервер управления> <узел1> <узел2> <узел3> ... >> /var/lib/one/.ssh/known_hosts
```

Далее необходимо нужно скопировать каталог `/var/lib/one/.ssh` на все узлы. Самый простой способ – установить временный пароль для oneadmin на всех хостах и скопировать каталог с сервера управления:

```
$ scp -rp /var/lib/one/.ssh <узел1>:/var/lib/one/
$ scp -rp /var/lib/one/.ssh <узел2>:/var/lib/one/
$ scp -rp /var/lib/one/.ssh <узел3>:/var/lib/one/
...
```

После этого необходимо убедиться, что при подключении от сервера управления под пользователем oneadmin к узлам и самому серверу управления, а также от узлов к серверу управления пароль не запрашивается.

Если требуется дополнительный уровень безопасности, можно хранить закрытый ключ только на сервере управления, а не копировать его на весь гипервизор. Таким образом, пользователь oneadmin в гипервизоре не сможет получить доступ к другим гипервизорам. Это достигается путем изменения `/var/lib/one/.ssh/config` на сервере управления и добавления параметра `ForwardAgent` к хостам гипервизора для пересылки ключа:

```
$ cat /var/lib/one/.ssh/config
 Host host1
    User oneadmin
    ForwardAgent yes
 Host host2
    User oneadmin
    ForwardAgent yes
```

3.2.6 Конфигурация сети

Сервисам, работающим на сервере управления, необходим доступ к узлам с целью управления гипервизорами и их мониторинга, а также для передачи файлов образов. Для этой цели рекомендуется использовать выделенную сеть.

Примечание. Настройка сети необходима только на серверах с виртуальными машинами. Точное имя ресурсов (br0, br1 и т.д.) значения не имеет, но важно, чтобы мосты и сетевые карты имели одно и то же имя на всех узлах.

3.3 Добавление узлов в OpenNebula

Регистрация узла в интерфейсе OpenNebula может быть выполнена в командной строке или в графическом пользовательском интерфейсе Sunstone.

3.3.1 Добавление узла в OpenNebula-Sunstone

Для добавления узла, необходимо в левом меню выбрать «Инфраструктура» → «Узлы» и на загруженной странице нажать кнопку «+» (Рис. 10).

Далее необходимо указать тип виртуализации, заполнить поле «Имя хоста» (можно ввести Ip-адрес узла, или его имя) и нажать кнопку «Создать» (Рис. 11).

Добавление узла в OpenNebula-Sunstone

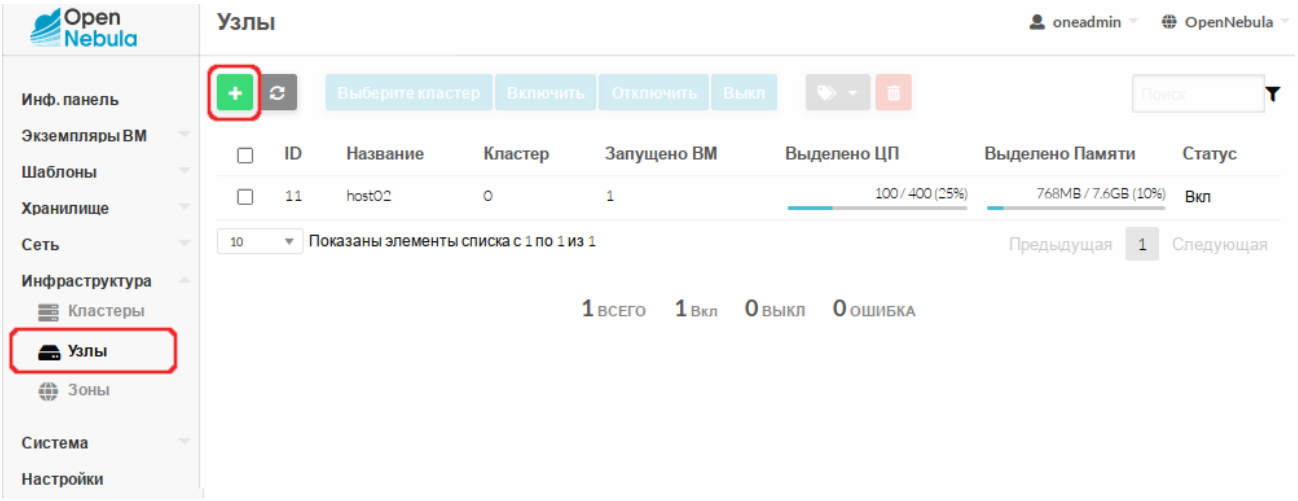

Рис. 10

Добавление узла в OpenNebula-Sunstone

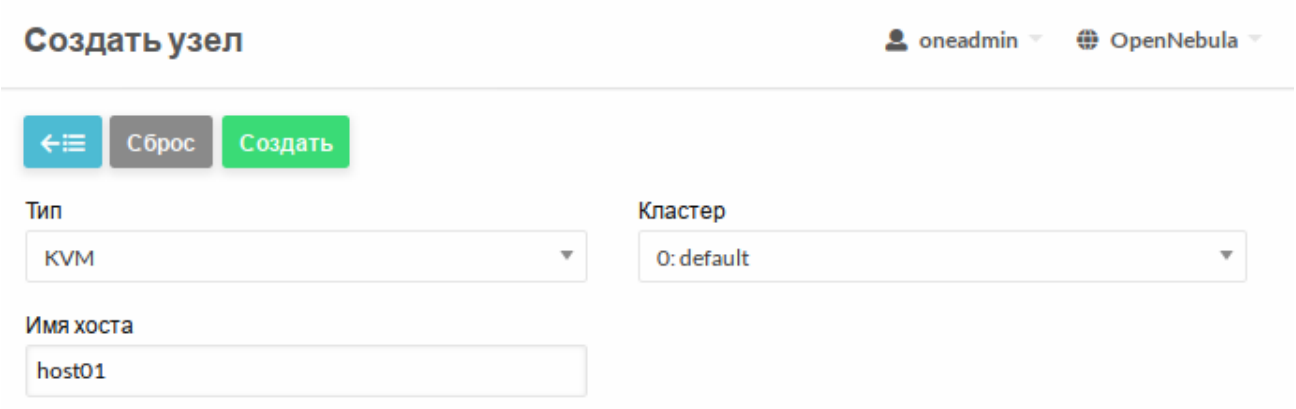

Рис. 11

Затем следует вернуться к списку узлов и убедиться, что узел перешел в состояние «Вкл» (это должно занять от 20 секунд до 1 минуты, можно нажать кнопку «Обновить» для обновления состояния) (Рис. 12).

Добавление узла в OpenNebula-Sunstone

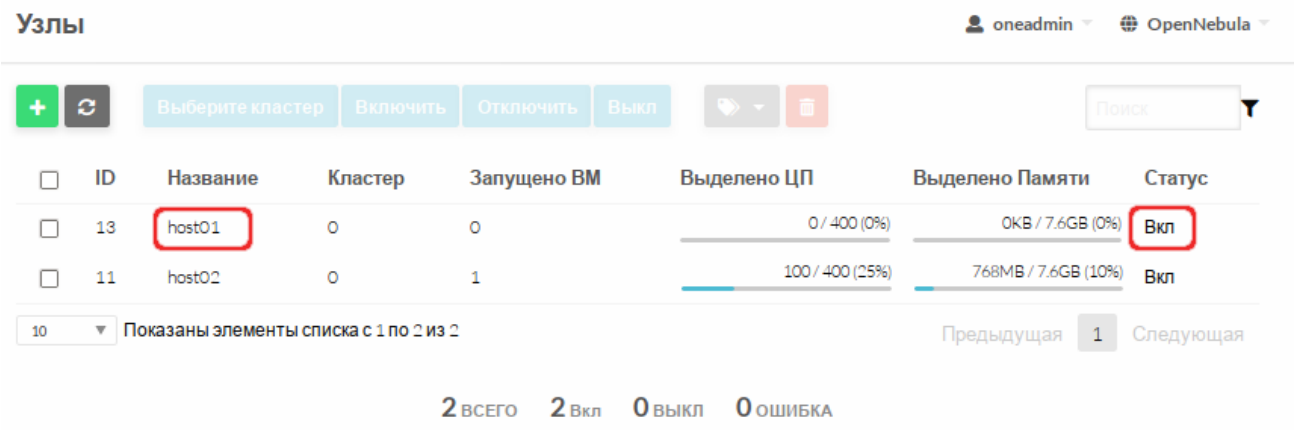

Рис. 12

3.3.2 Работа с узлами в командной строке

onehost – это инструмент управления узлами в OpenNebula. Описание всех доступных опций утилиты onehost можно получить, выполнив команду:

```
$ man onehost
```

Для добавления узла в облако, необходимо выполнить следующую команду от oneadmin на сервере управления:

```
$ onehost create host01 -im kvm -vm kvm
ID: 1
```

Список узлов можно просмотреть, выполнив команду:

```
$ onehost list
  ID NAME          CLUSTER    TVM      ALLOCATED_CPU      ALLOCATED_MEM STAT
   1 host01        default      0       0 / 400 (0%)     0K / 7.6G (0%) on
```

Примечание. Если возникли проблемы с добавлением узла, то скорее всего неправильно настроен ssh. Ошибки можно просмотреть в `/var/log/one/oned.log`.

Для указания узла можно использовать его ID или имя. Например, удаление узла с указанием ID или имени:

```
$ onehost delete 1
$ onehost delete host01
```

Изменение статуса узла:

```
$ onehost disable host01 // деактивировать узел
$ onehost enable host01 // активировать узел
$ onehost offline host01 // полностью выключить узел
```

Просмотр информации об узле:

```
$ onehost show 1
```

Информация об узле содержит:

- общую информацию, включая имя и драйверы, используемые для взаимодействия с ним;
- информацию об объёме (Host Shares) процессора и памяти;
- информацию о локальном хранилище данных (Local System Datastore), если хост настроен на использование локального хранилища данных;
- информацию мониторинга;
- активных ВМ на узле.

3.4 Управление пользователями

OpenNebula включает полную систему управления пользователями и группами.

Ресурсы, к которым пользователь может получить доступ в OpenNebula, контролируются системой разрешений. По умолчанию только владелец ресурса может использовать и управлять им. Пользователи могут делиться ресурсами, предоставляя разрешения на использование или управление другим пользователям в своей группе или любому другому пользователю в системе.

oneuser – инструмент командной строки для управления пользователями в OpenNebula.

При установке OpenNebula создаются две административные учетные записи (oneadmin и serveradmin), и две группы (oneadmin и users):

```
$ oneuser list
  ID NAME                GROUP    AUTH            VMS     MEMORY         CPU
   1 serveradmin         oneadmin server_c    0 /   -      0M /   0.0 /   -
   0 oneadmin            oneadmin core             -          -          -
$ onegroup list
  ID NAME                USERS       VMS            MEMORY          CPU
   1 users                   0   0 /   -      0M /       -  0.0 /    -
   0 oneadmin                2         -                -           -
```

Создание нового пользователя:

```
$ oneuser create <user_name> <password>
```

По умолчанию новый пользователь будет входить в группу users. Изменить группу пользователя:

```
$ oneuser chgrp <user_name> oneadmin
```

Что бы удалить пользователя из группы, необходимо переместить его обратно в группу users.

Временно отключить пользователя:

```
$ oneuser disable <user_name>
```

Включить отключённого пользователя:

```
$ oneuser enable <user_name>
```

Удалить пользователя:

```
$ oneuser delete <user_name>
```

onegroup – инструмент командной строки для управления группами в OpenNebula.

Создание новой группы:

```
$ onegroup create group_name
ID: 100
```

Новая группа получила идентификатор 100, чтобы отличать специальные группы от созданных пользователем.

После создания группы может быть создан связанный пользователь-администратор. По умолчанию этот пользователь сможет создавать пользователей в новой группе.

Пример создания новой группы с указанием, какие ресурсы могут быть созданы пользователями группы (по умолчанию VM+IMAGE+TEMPLATE):

```
$ onegroup create --name testgroup \
--admin_user testgroup-admin --admin_password somestr \
--resources TEMPLATE+VM
```

При выполнении данной команды также будет создан администратор группы.

Сделать существующего пользователя администратором группы:

```
$ onegroup addadmin <groupid_list> <userid>
```

Все операции с пользователями можно производить в веб-интерфейсе (Рис. 13, Рис. 14).

Управление пользователями в OpenNebula-Sunstone

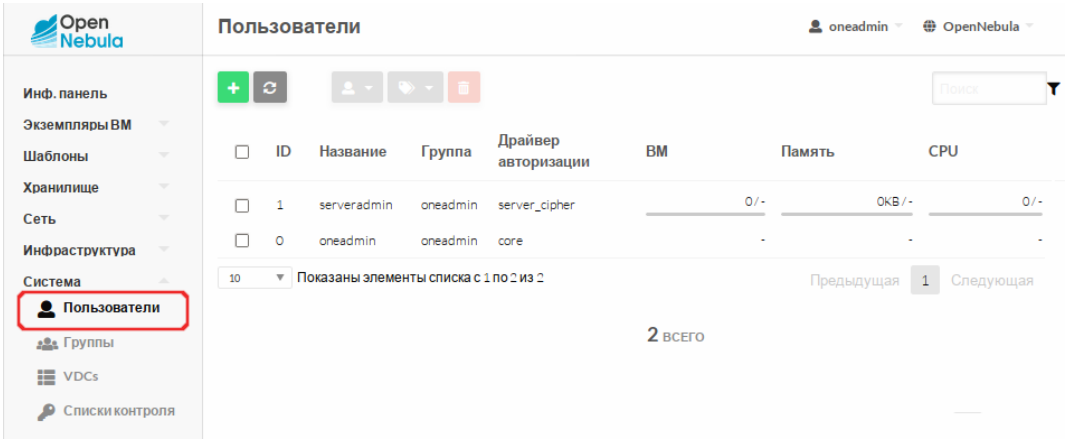

Рис. 13

Создание группы в OpenNebula-Sunstone

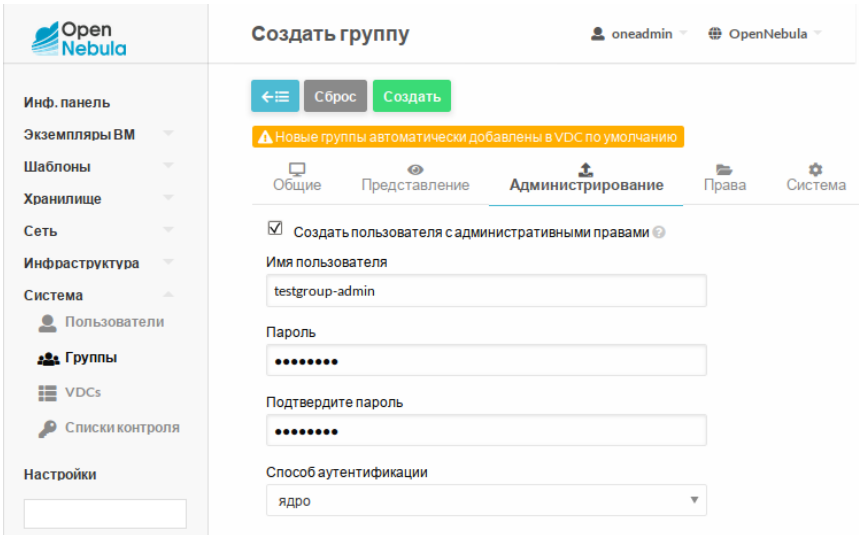

Рис. 14

Созданный пользователь может аутентифицироваться в веб-интерфейсе OpenNebula и изменить настройки (изменить язык интерфейса, пароль, добавить ssh-ключ для доступа на ВМ и т.д.) (Рис. 15).

Панель пользователя в OpenNebula-Sunstone

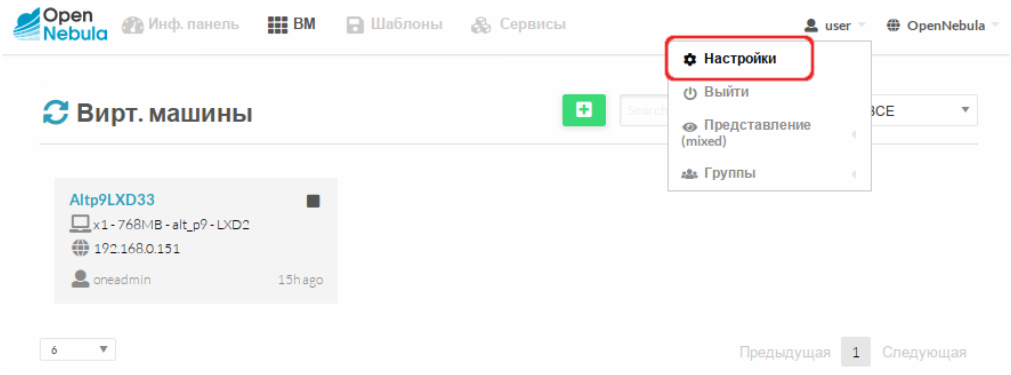

Рис. 15

3.5 Работа с хранилищами в OpenNebula

По умолчанию в OpenNebula созданы хранилище образов (Images), системное (System) и файлов (Files).

onedatastore – инструмент управления хранилищами в OpenNebula. Описание всех доступных опций утилиты onedatastore можно получить, выполнив команду:

```
$ man onedatastore
```

Вывести список хранилищ данных можно, выполнив команду:

```
$ onedatastore list
  ID NAME                SIZE AVA CLUSTERS IMAGES TYPE DS      TM      STAT
   2 files           104.8G 32% 0             1 fil  fs      ssh     on
   1 default         104.8G 32% 0             8 img  fs      ssh     on
   0 system               - -   0             0 sys  -       ssh     on
```

Информация о хранилище:

```
$ onedatastore show default
```

Создавать, включать, отключать, удалять и просматривать информацию о хранилищах можно в веб-интерфейсе (Рис. 16).

Работа с хранилищами в OpenNebula-Sunstone

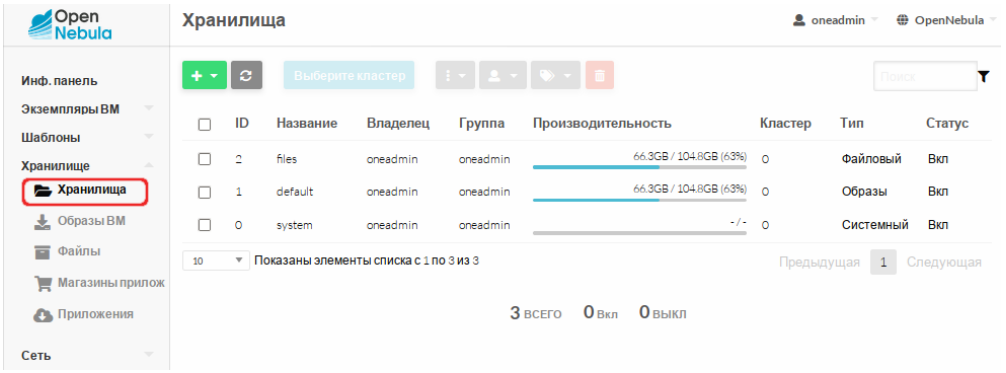

Рис. 16

3.6 Работа с образами в OpenNebula

Система хранилищ позволяет пользователям настраивать/устанавливать образы, которые могут быть образами ОС или данных, для использования в ВМ. Данные образы могут использоваться несколькими ВМ одновременно, а также предоставляться другим пользователями.

Типы образов для дисков ВМ (хранятся в хранилище образов):

- OS – образ загрузочного диска;
- CDROM – файл образа, содержащий CDROM. Эти образы предназначены только для чтения. В каждом шаблоне ВМ, можно использовать только один образ данного типа;
- DATABLOCK – файл образа, содержащий блок данных, создаваемый как пустой блок.

Типы файлов (хранятся в файловом хранилище):

- KERNEL – файл, который будет использоваться в качестве ядра ВМ (kernels);
- RAMDISK – файл, для использования в качестве виртуального диска;
- CONTEXT – файл для включения в контекстный CD-ROM.

Образы могут работать в двух режимах:

- persistent (постоянные) – изменения, внесенные в такие образы, будут сохранены после завершения работы ВМ. В любой момент времени может быть только одна ВМ, использующая постоянный образ.
- non-persistent (непостоянный) – изменения не сохранятся после завершения работы ВМ. Непостоянные образы могут использоваться несколькими ВМ одновременно, поскольку каждая из них будет работать со своей собственной копией.

Управлять образами можно, используя команду `oneimage`. Также управлять образами можно в веб-интерфейсе, на вкладке «Образы ВМ» (Рис. 17).

Управление образами в OpenNebula-Sunstone

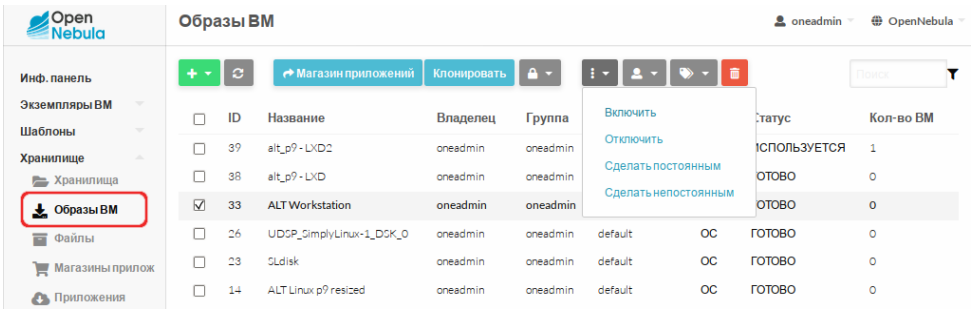

Рис. 17

3.6.1 Работа с образами в OpenNebula

Для создания образа ОС, необходимо подготовить ВМ и извлечь её диск.

3.6.1.1 Создание образов дисков

Создать образ типа CDROM с установочным ISO-образом.

Для этого перейти в раздел «Хранилище» → «Образы ВМ», на загруженной странице нажать «+» и выбрать пункт «Создать» (Рис. 18).

Создание образа в OpenNebula-Sunstone

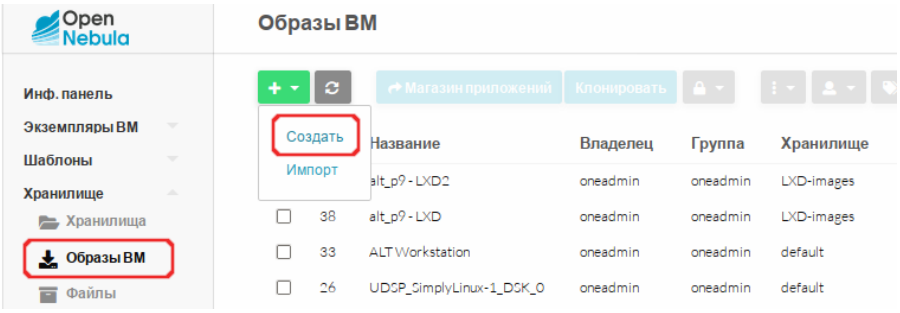

Рис. 18

В открывшемся окне заполнить поле «Название», выбрать в выпадающем списке «Тип» значение «CD-ROM только для чтения», выбрать хранилище, отметить в «Расположение образа» пункт «Путь на сервере OpenNebula» («Path in OpenNebula Server»), указать путь к файлу (.iso) и нажать кнопку «Создать» (Рис. 19).

Создание образа типа CD-ROM

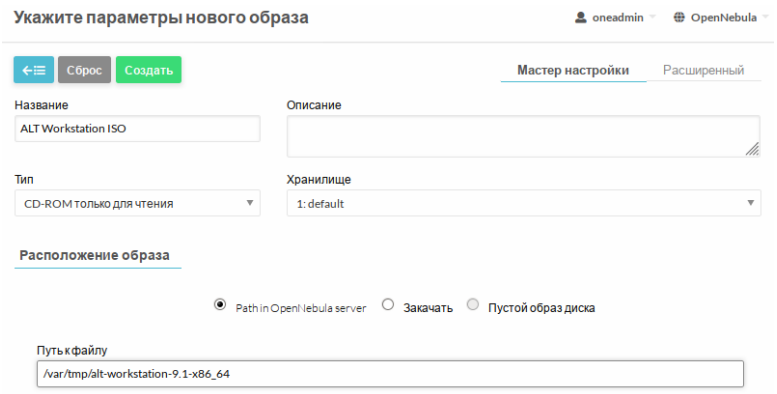

Рис. 19

Примечание. ISO-образ должен быть загружен в папку, к которой имеет доступ пользователь oneadmin.

Создать пустой образ диска, на который будет установлена операционная система.

Для этого создать новый образ. Заполнить поле «Название», в выпадающем списке «Тип» выбрать значение «Generic storage datablock», в выпадающем списке «Этот образ является постоянным» выбрать значение «Да», выбрать хранилище, в разделе «Расположение образа» выбрать пункт «Пустой образ диска», установить размер выбранного блока, например 45GB, в разделе «Расширенные настройки» указать драйвер «qcow2» и нажать кнопку «Создать» (Рис. 20).

Создание диска

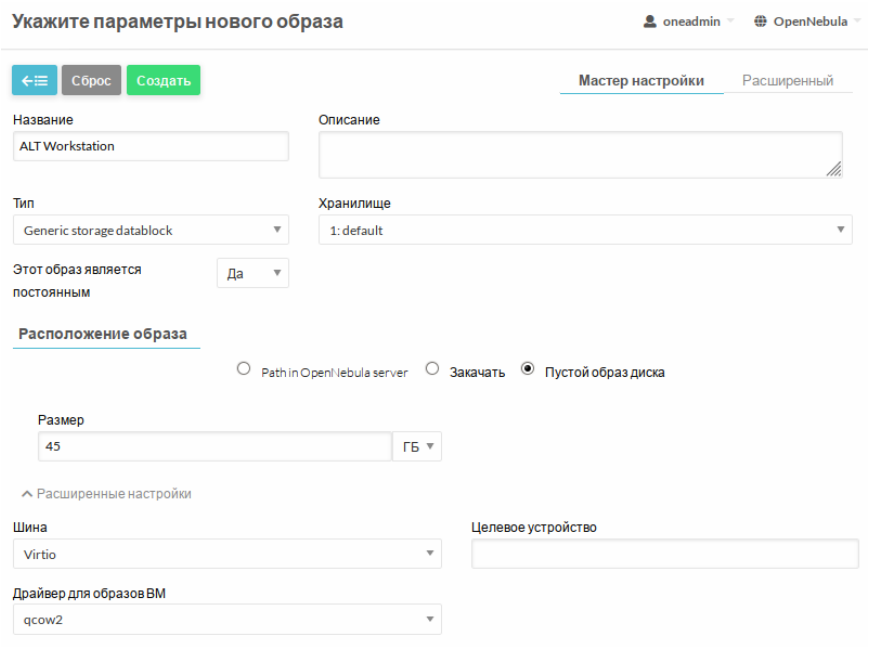

Рис. 20

Эти же действия можно выполнить в командной строке.

Создать образ типа CDROM в хранилище данных по умолчанию (ID = 1):

```
$ oneimage create -d 1 --name "ALT Workstation ISO" --path /var/tmp/alt-workstation-9.1-x86_64.iso --type CDROM
ID: 31
```

Создать пустой образ диска (тип образа – DATABLOCK, размер 45 ГБ, драйвер qcow2):

```
$ oneimage create -d 1 --name "ALT Workstation" --type DATABLOCK --size 45G --persistent --driver qcow2
ID: 33
```

3.6.1.2 Создание шаблона ВМ

Создание шаблона в командной строке:

1) Создать файл template со следующим содержимым:

```
NAME = "ALT Workstation"
CONTEXT = [
  NETWORK = "YES",
  SSH_PUBLIC_KEY = "$USER[SSH_PUBLIC_KEY]" ]
CPU = "0.25"
DISK = [
  IMAGE = "ALT Workstation ISO",
  IMAGE_UNAME = "oneadmin" ]
DISK = [
  DEV_PREFIX = "vd",
  IMAGE = "ALT Workstation",
  IMAGE_UNAME = "oneadmin" ]
```

```
GRAPHICS = [
  LISTEN = "0.0.0.0",
  TYPE = "SPICE" ]
HYPERVISOR = "kvm"
INPUTS_ORDER = ""
LOGO = "images/logos/alt.png"
MEMORY = "1024"
MEMORY_UNIT_COST = "MB"
NIC = [
  NETWORK = "VirtNetwork",
  NETWORK_UNAME = "oneadmin",
  SECURITY_GROUPS = "0" ]
NIC_DEFAULT = [
  MODEL = "virtio" ]
OS = [
  BOOT = "disk1,disk0" ]
SCHED_REQUIREMENTS = "ID=\"0\""
```

2) Создать шаблон:

```
$ onetemplate create template
ID: 22
```

Ниже рассмотрен пример создания шаблона в веб-интерфейсе.

Для создания шаблона ВМ, необходимо в левом меню выбрать «Шаблоны» → «ВМ» и на загруженной странице нажать кнопку «+» и выбрать пункт «Создать».

На вкладке «Общие» необходимо указать параметры процессора, оперативной памяти, а также гипервизор (Рис. 21).

Создание шаблона ВМ. Вкладка «Общие»

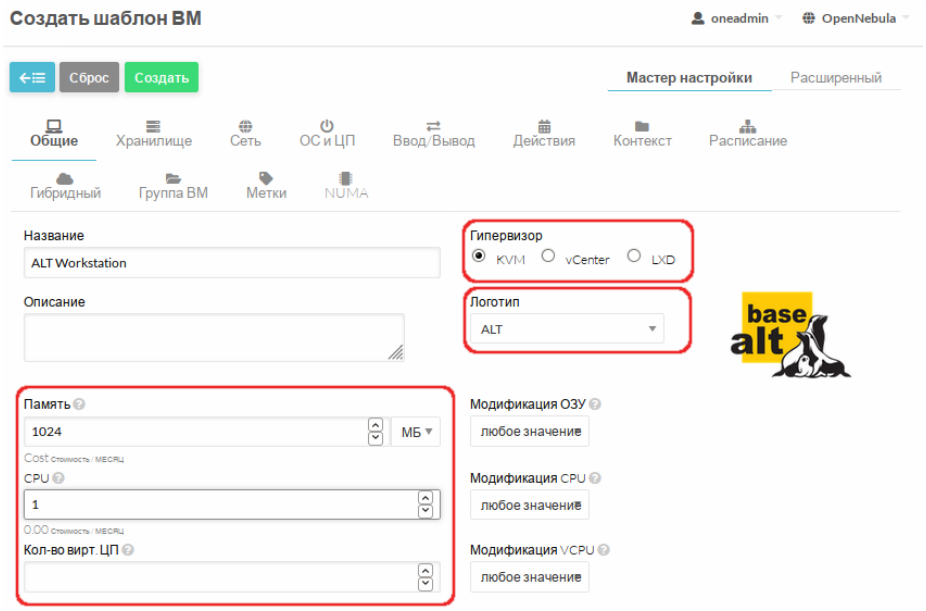

Рис. 21

На вкладке «Хранилище» необходимо указать ранее созданный пустой диск (DATABLOCK), в разделе «Расширенные настройки» в выпадающем списке «Шина» выбрать «Virtio». Далее следует добавить новый диск и указать диск с установщиком ОС (Рис. 22).

Создание шаблона ВМ. Вкладка «Хранилище»

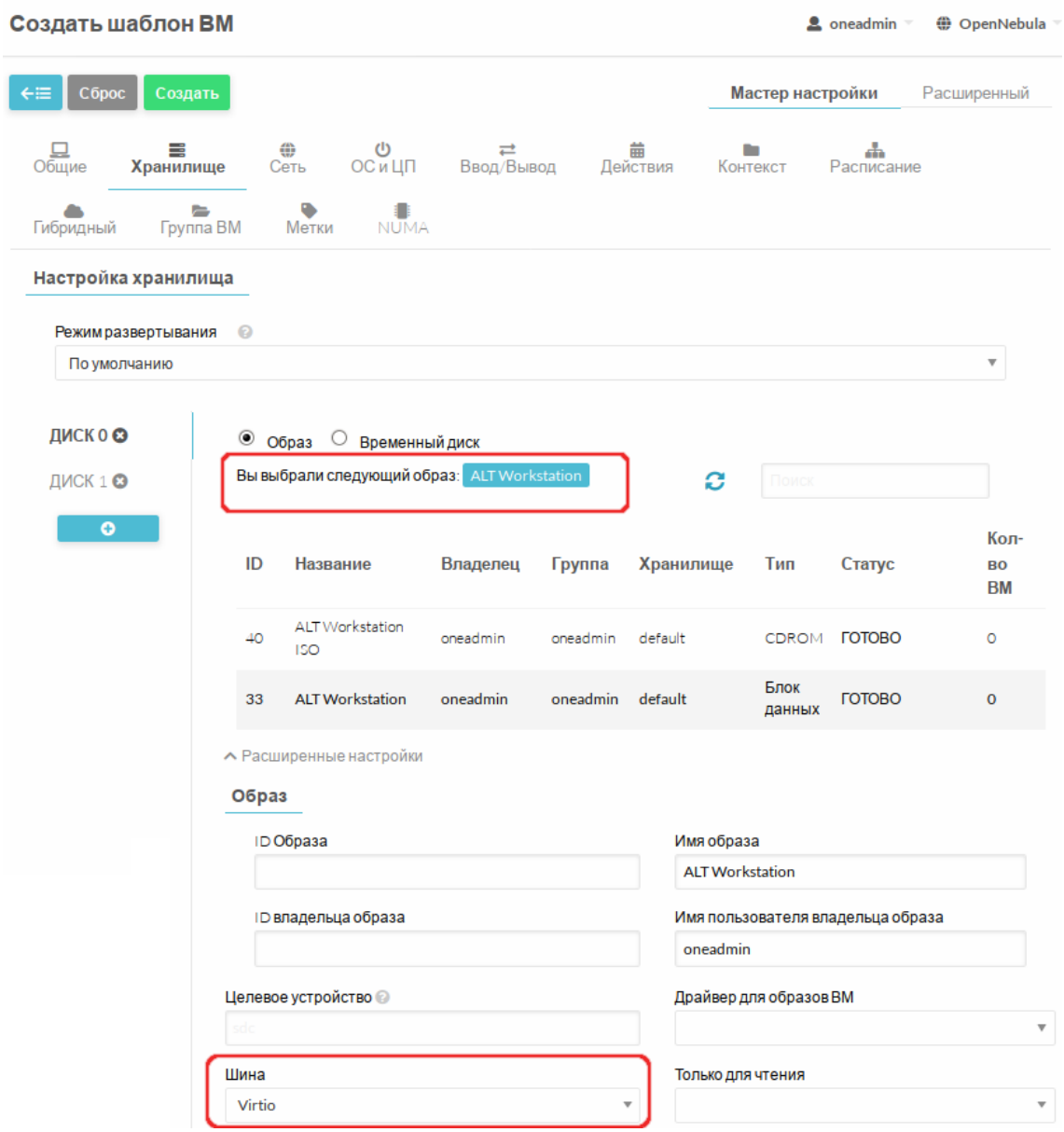

Рис. 22

На вкладке «Сеть» в поле «Default hardware model to emulate for all NICs» следует указать Virtio и если необходимо выбрать сеть (Рис. 23).

На вкладке «ОС и ЦПУ» необходимо указать архитектуру устанавливаемой системы и выбрать порядок загрузки. Можно установить в качестве первого загрузочного устройства – пустой диск (DATABLOCK), а в качестве второго – CDROM (Рис. 24). При такой последовательности загрузочных устройств при пустом диске загрузка произойдёт с CDROM, а в дальнейшем, когда ОС будет уже установлена на диск, загрузка будет осуществляться с него.

Создание шаблона ВМ. Вкладка «Сеть»

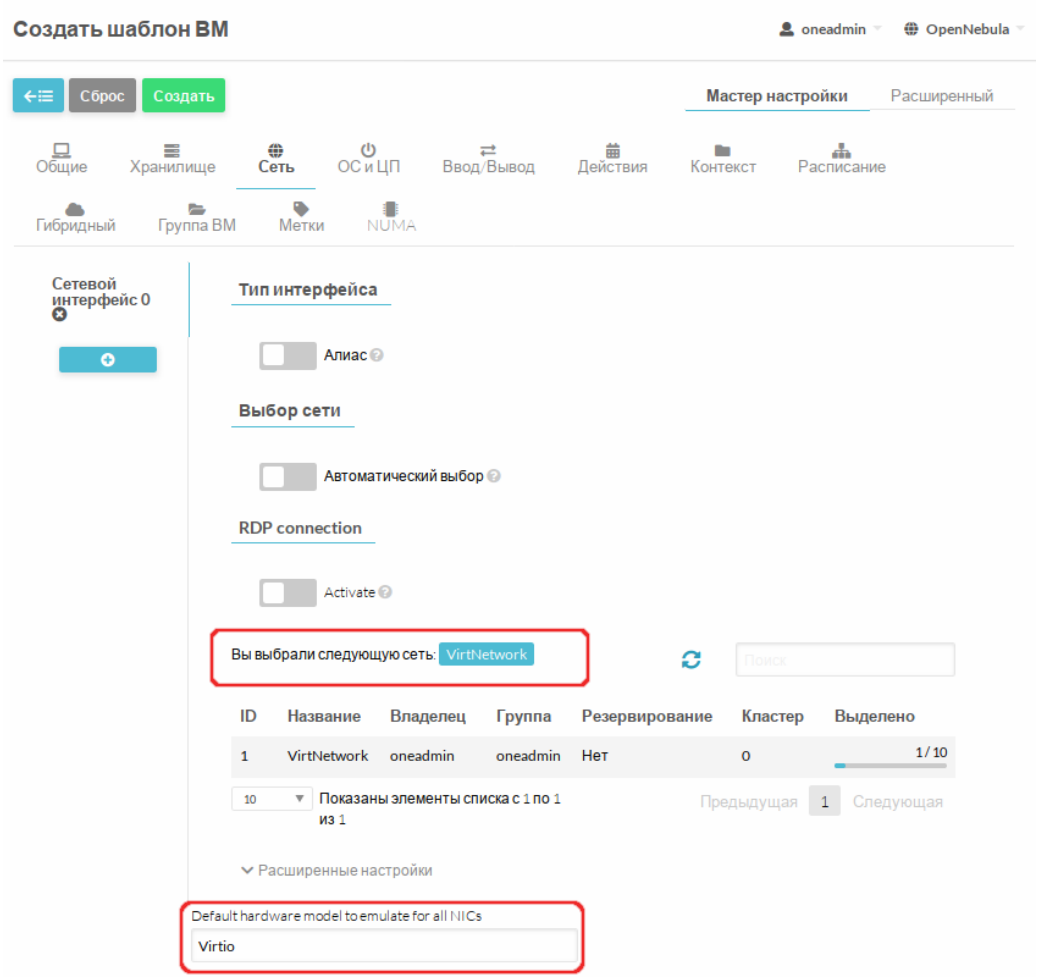

Рис. 23

Создание шаблона ВМ. Вкладка «ОС и ЦПУ»

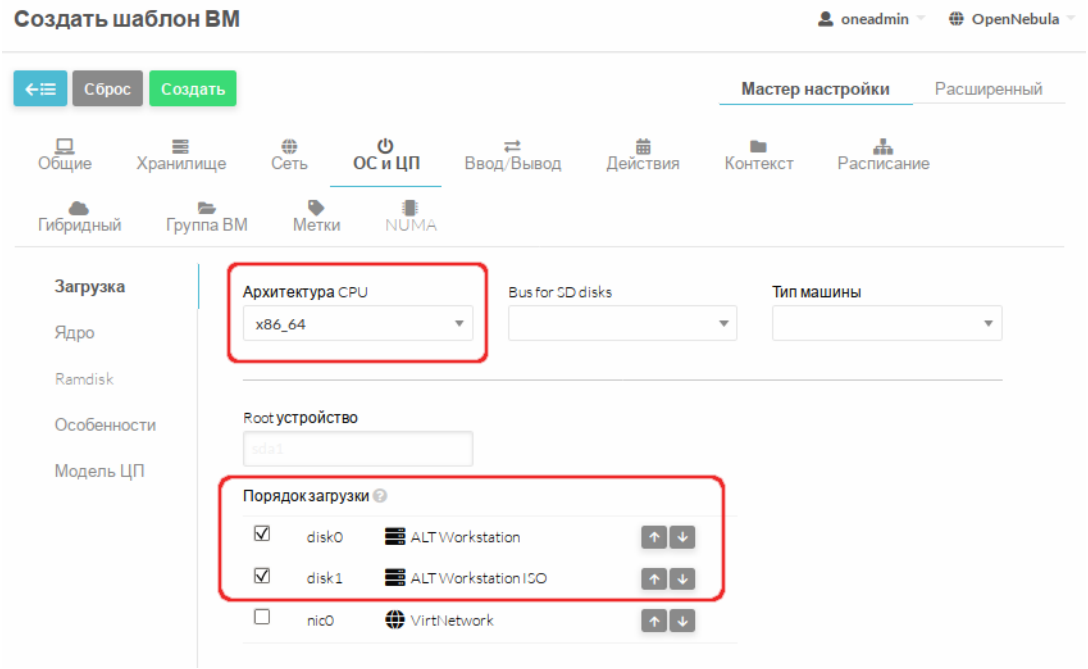

Рис. 24

На вкладке «Ввод/Вывод» следует включить «SPICE» (Рис. 25).

Создание шаблона ВМ. Вкладка «Ввод/Вывод»

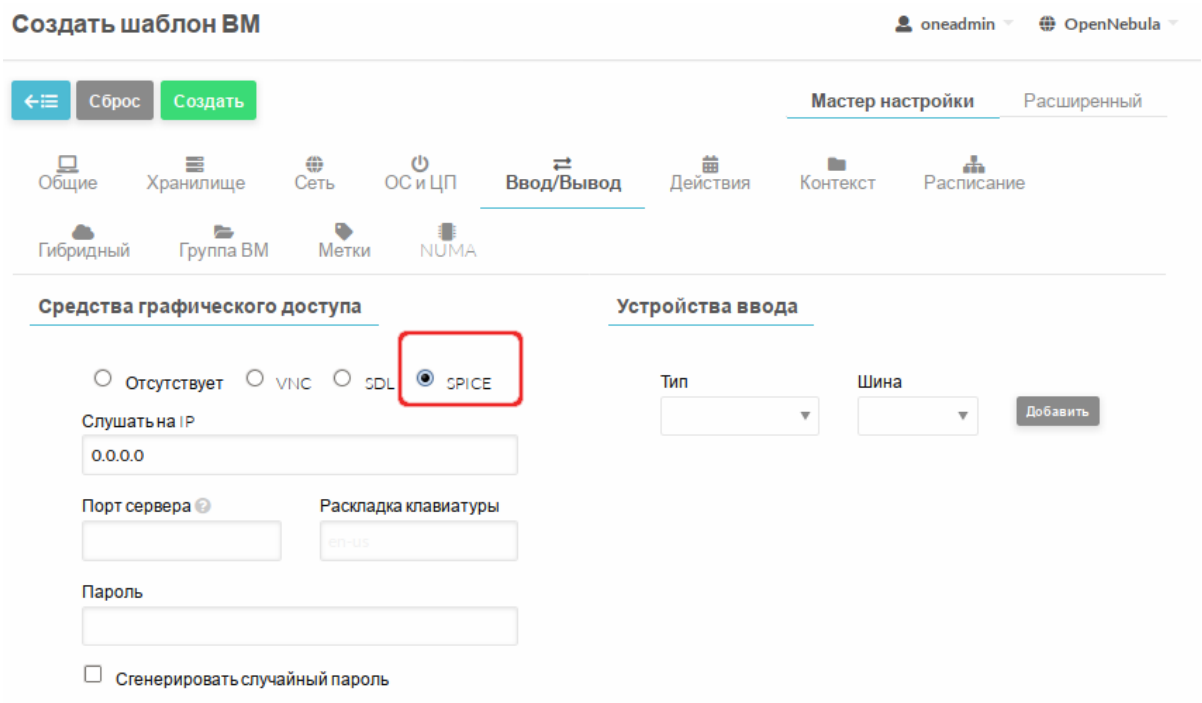

Рис. 25

На вкладке «Контекст» необходимо включить параметр «Использовать сетевое задание контекста», а также авторизацию по RSA-ключам (Рис. 26). Укажите свой открытый SSH (.pub) для доступа к ВМ по ключу. Если оставить это поле пустым, будет использована переменная $USER[SSH_PUBLIC_KEY]

Создание шаблона ВМ. Вкладка «Контекст»

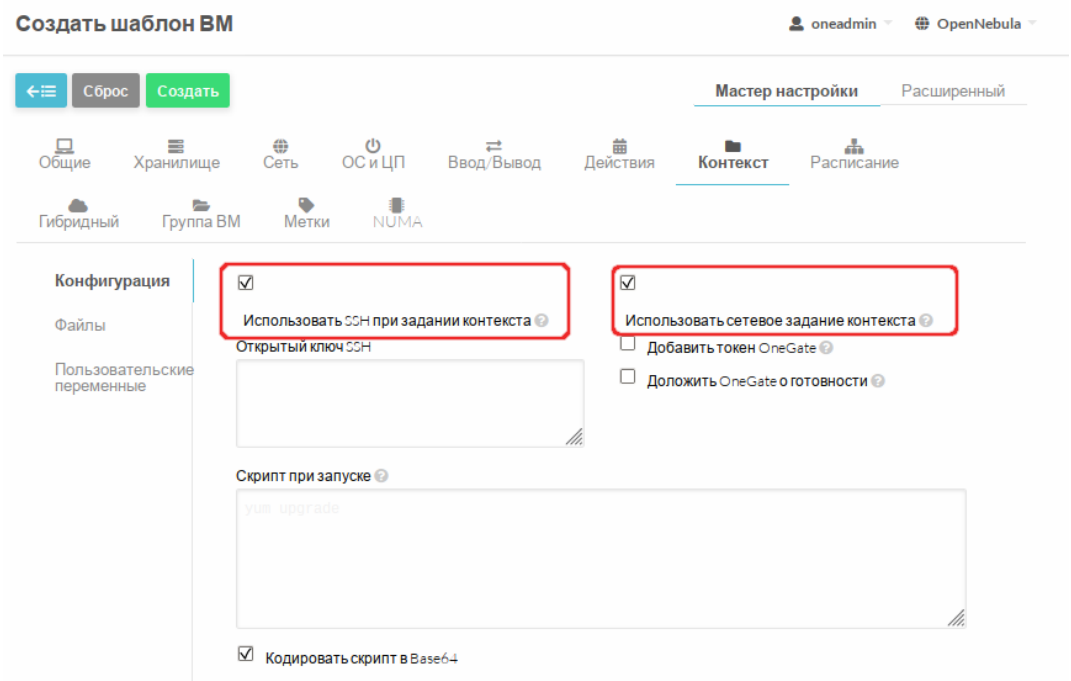

Рис. 26

На вкладке «Расписание» если необходимо можно выбрать кластер/хост, на котором будет размещаться виртуальное окружение (Рис. 27).

Создание шаблона ВМ. Вкладка «Расписание»

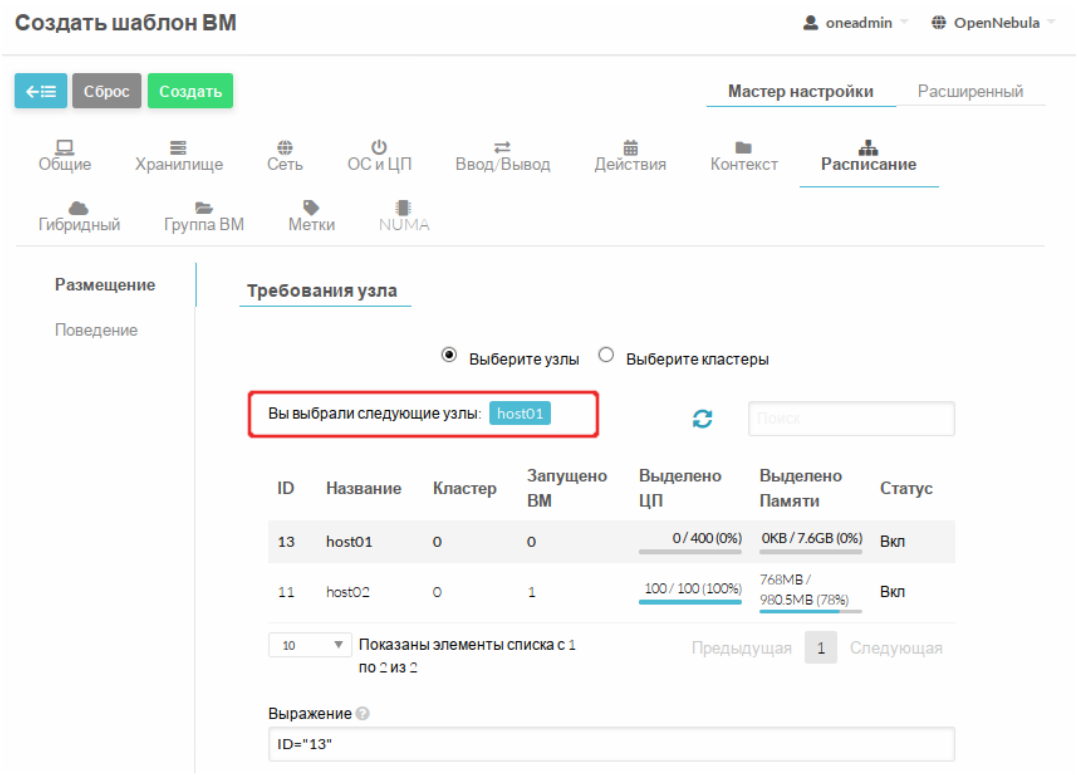

Рис. 27

Для создания шаблона ВМ нажать кнопку «Создать».

3.6.1.3 Создание ВМ

Для инициализации установки ОС из созданного шаблона в левом меню следует выбрать пункт «Шаблоны» → «ВМ», выбрать шаблон и нажать кнопку «Создать ВМ» (Рис. 28).

Создание экземпляра ВМ из шаблона

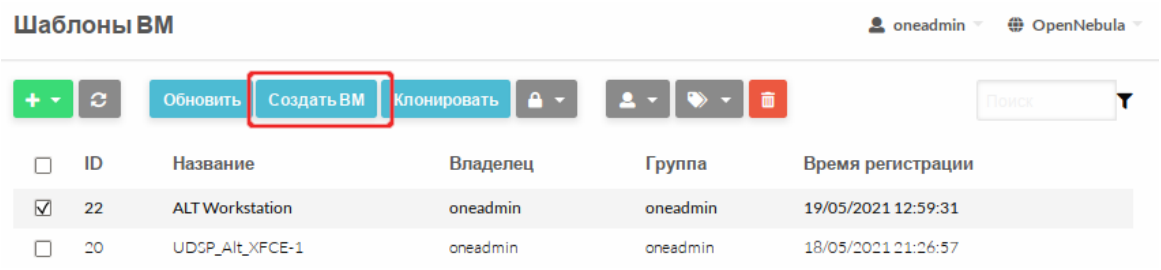

Рис. 28

В открывшемся окне необходимо указать имя ВМ и нажать кнопку «Создать ВМ» (Рис. 29).

Инициализация установки ОС из шаблона

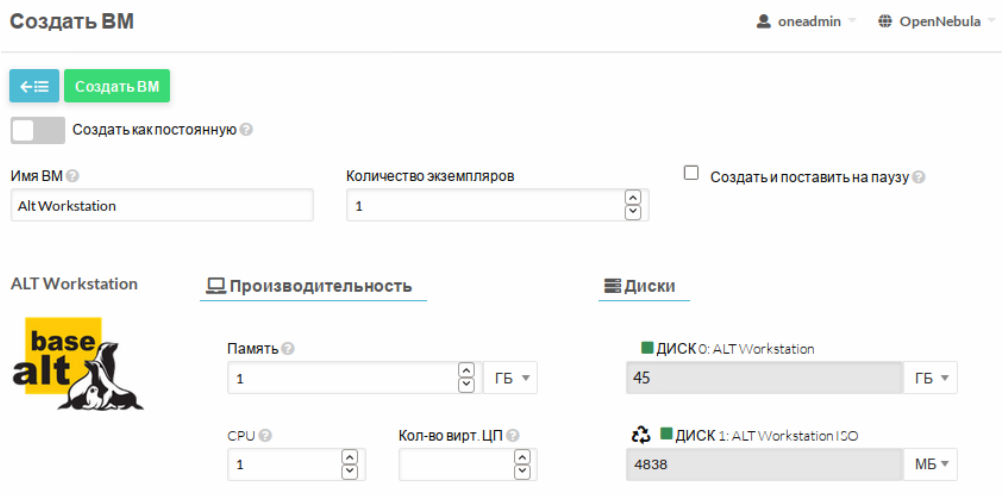

Рис. 29

Создание экземпляра ВМ из шаблона в командной строке:

```
$ onetemplate instantiate 22
VM ID: 94
```

3.6.1.4 Подключение к ВМ и установка ОС

Примечание. Процесс создания ВМ может занять несколько минут. Следует дождаться статуса – «ЗАПУЩЕНО» («RUNNING»).

Подключиться к ВМ можно как из веб-интерфейса Sunstone, раздел «Экземпляры ВМ» → «ВМ» выбрать ВМ и подключиться по SPICE (Рис. 30).

Подключение к ВМ

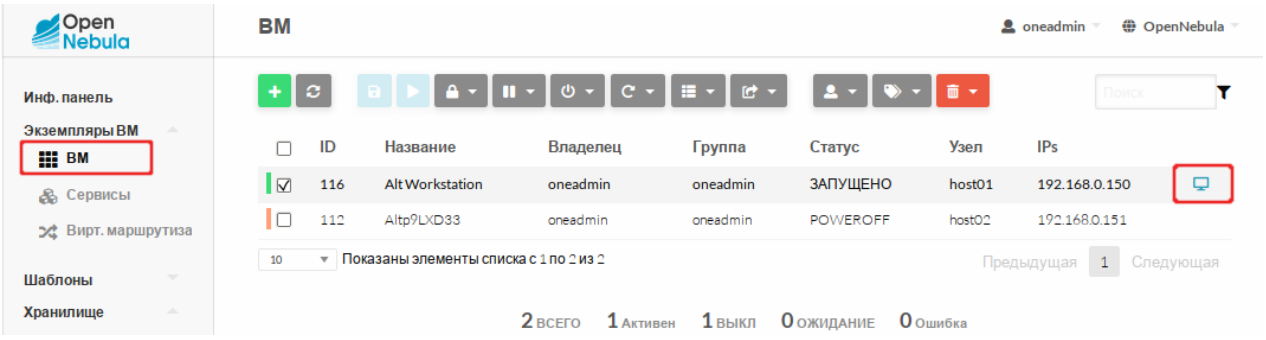

Рис. 30

Так и используя, любой клиент SPICE:

```
spice://192.168.0.190:5994
```

где 192.168.0.190 адрес хоста с ВМ, а 94 идентификатор ВМ (номер порта 5900 + 94)

Далее необходимо провести установку системы (Рис. 31).

Установка ВМ

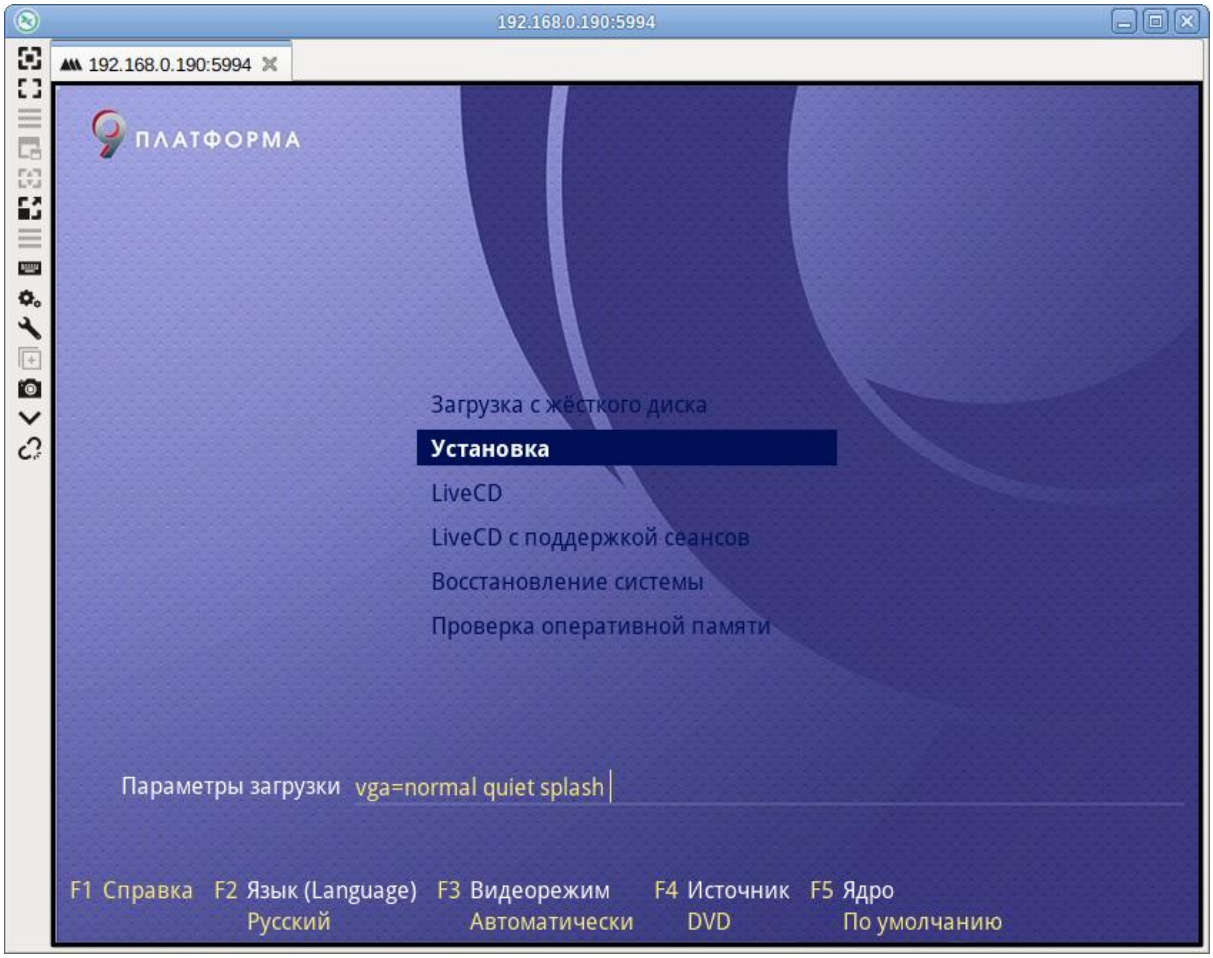

Рис. 31

3.6.1.5 Настройка контекстуализации

OpenNebula использует метод, называемый контекстуализацией, для отправки информации на ВМ во время загрузки. Контекстуализация позволяет установить или переопределить данные ВМ, имеющие неизвестные значения или значения по умолчанию (имя хоста, IP-адрес, .ssh/authorized_keys).

Пример настройки контекстуализации на установленной ОС Альт:

1) Подключиться к ВМ через SPICE или по ssh.

2) Установить пакет opennebula-context:

```
# apt-get update && apt-get install opennebula-context
```

3) Переключиться на systemd-networkd:

– установить пакет systemd-timesyncd:

```
# apt-get install systemd-timesyncd
```

– создать файл автонастройки всех сетевых интерфейсов по DHCP `/etc/systemd/network/lan.network` со следующим содержимым:

```
[Match]
Name = *
```

```
[Network]
DHCP = ipv4
```

– переключиться с etcnet/NetworkManager на systemd-networkd:

```
# systemctl disable network NetworkManager && systemctl enable systemd-networkd systemd-timesyncd
```

4) Перезагрузить систему.

После перезагрузки доступ в систему будет возможен по ssh-ключу, ВМ будет назначен IP-адрес, который OpenNebula через механизм IPAM (подсистема IP Address Management) выделит из пула адресов.

3.6.1.6 Создание образа типа ОС

После завершения установки системы следует выключить и удалить ВМ. Диск находится в состоянии «Persistent», поэтому все внесенные изменения будут постоянными.

Для удаления ВМ в левом меню следует выбрать пункт «Экземпляры ВМ» → «ВМ», выбрать ВМ и нажать кнопку «Уничтожить» (Рис. 32).

Удаление ВМ

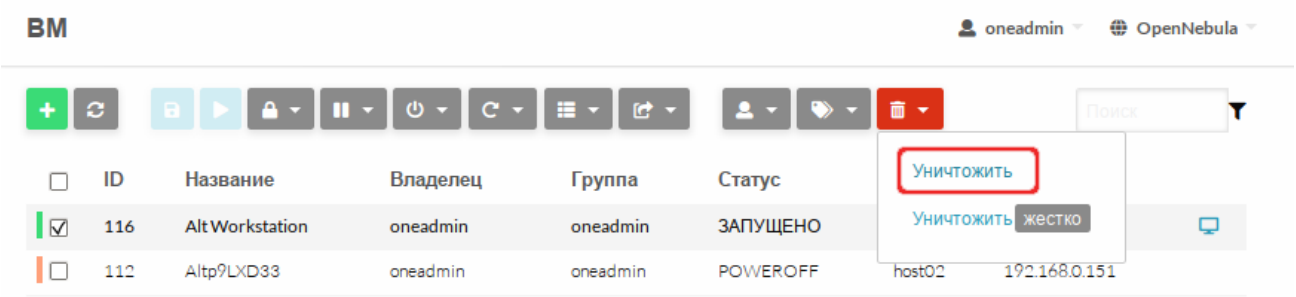

Рис. 32

Примечание. Удаление ВМ в командной строке:

```
$ onevm terminate 94
```

Затем перейти в «Хранилище» → «Образы ВМ», выбрать образ с установленной ОС (ALT Workstation) и изменить тип блочного устройства с «Блок данных» на «ОС» и состояние на «Не постоянный» (Рис. 33).

Примечание. Изменить тип блочного устройства на OS и состояние на Non Persistent в командной строке:

```
$ oneimage chtype 33 OS
$ oneimage nonpersistent 33
```

Образ готов. Далее можно использовать как имеющийся шаблон, так и создать новый на основе образа диска «ALT Workstation».

Изменение типа блочного устройства

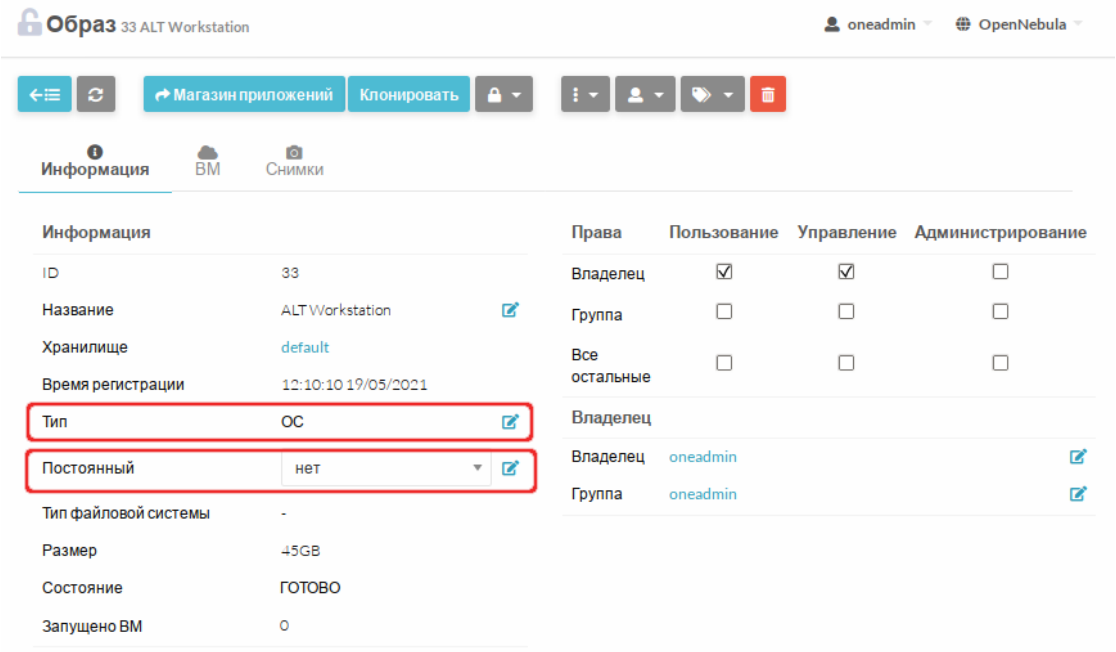

Рис. 33

3.6.2 Использование магазина приложений OpenNebula

Для загрузки приложения из магазина необходимо перейти в «Хранилище» → «Магазины приложений», выбрать «OpenNebula Public» → «Приложения». Появится список доступных приложений (Рис. 34).

Магазин приложений OpenNebula

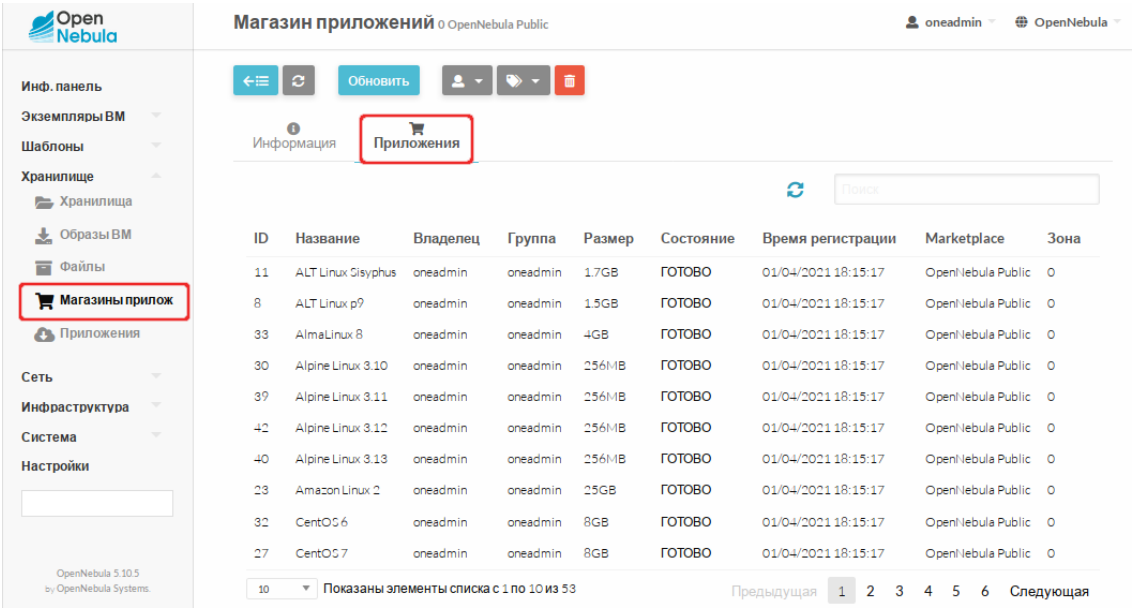

Рис. 34

Каждое приложение содержит образ и шаблон.

Чтобы импортировать приложение, необходимо его выбрать и нажать кнопку «Import into Datastore» (Рис. 35).

Импорт приложения из магазина приложений OpenNebula

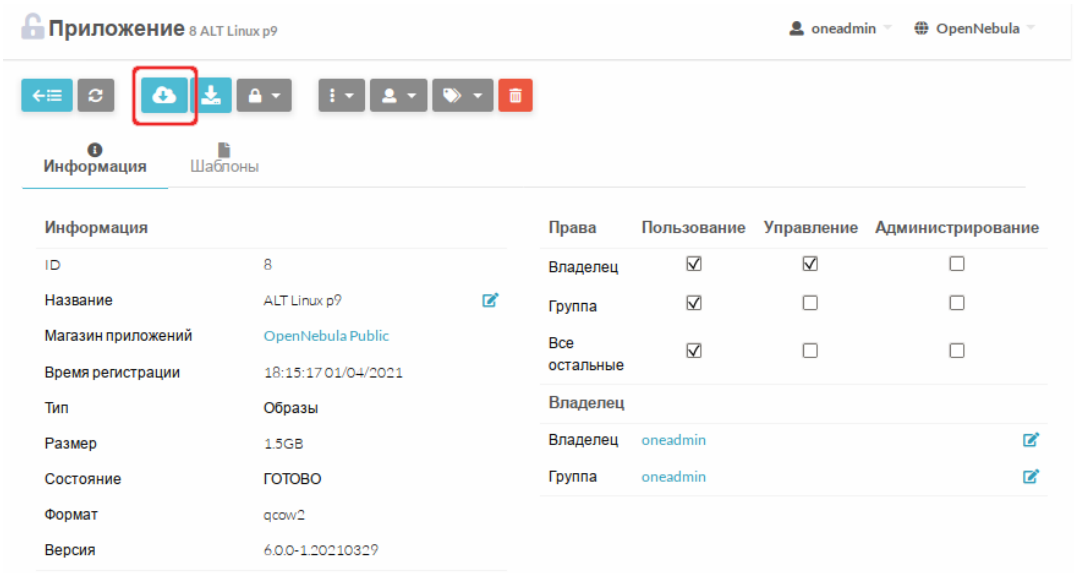

Рис. 35

В открывшемся окне указать имя для образа и шаблона, выбрать хранилище и нажать кнопку «Загрузить» (Рис. 36).

Изменение типа блочного устройства

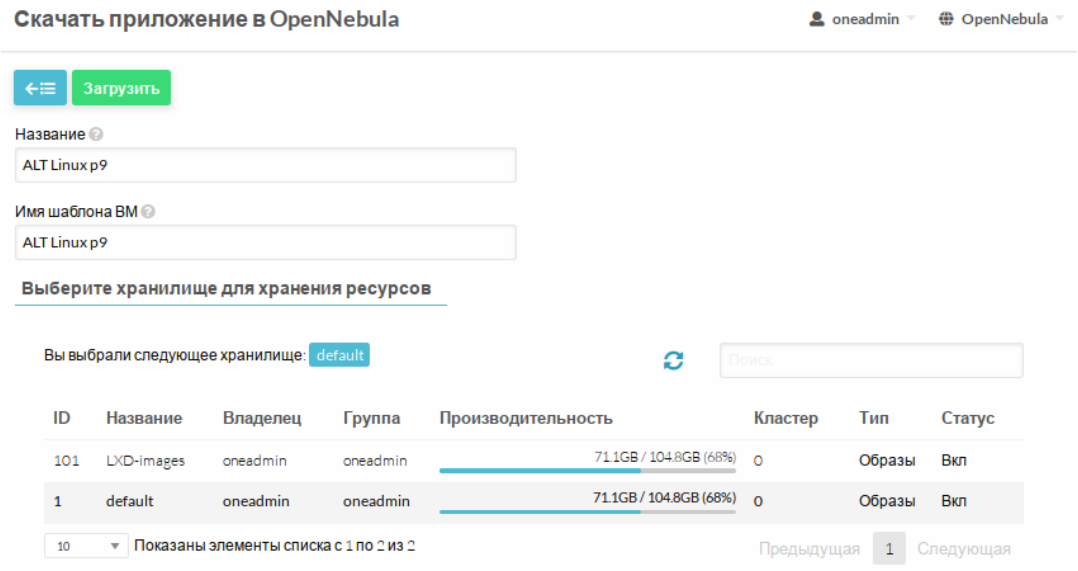

Рис. 36

Настройка образов, загруженных из магазина приложений:

1) изменить состояние образа на «Постоянный» (необходимо дождаться состояния «Готово»);
2) настроить шаблон;
3) создать на основе шаблона ВМ;
4) подключиться к ВМ. Установить/настроить необходимые компоненты;
5) удалить ВМ;
6) изменить состояние образа на «Не постоянный»;

7) далее можно создать новые шаблоны на основе этого образа или использовать существующий.

3.7 Установка и настройка LXD

LXD – это гипервизор LXC контейнеров.

Примечание. Для работы с LXD в Opennebula должна быть настроена пара хранилищ (хранилище образов и системное) НЕ типа qcow2 (например, shared или ssh).

Перед добавлением хоста типа LXD на сервер OpenNebula следует настроить узел LXD.

3.7.1 Настройка узла OpenNebula LXD

Для создания узла типа LXD, при установке дистрибутива нужно выбрать профиль «Вычислительный узел Opennebula LXD» (Рис. 37). В установленной системе следует настроить доступ по SSH (см. раздел Ключи для доступа по SSH).

Установка сервера контейнеризации LXD

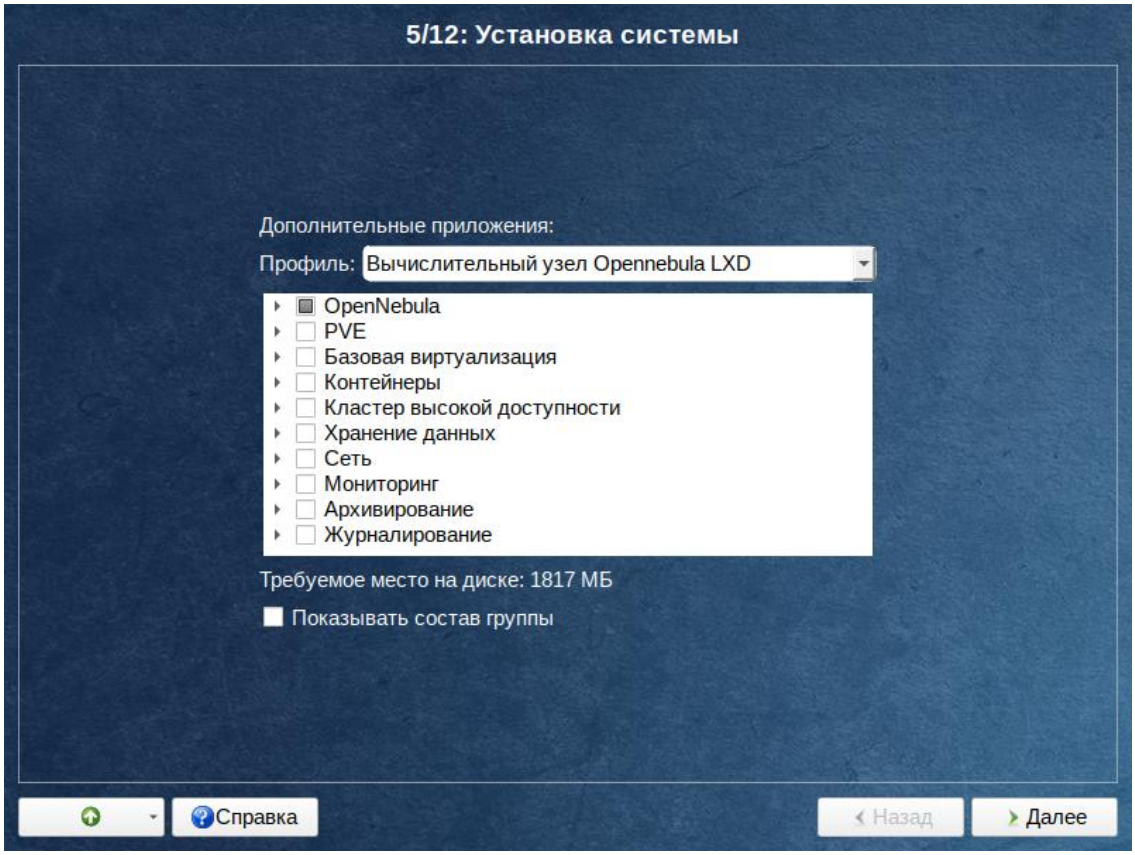

Рис. 37

Примечание. В уже установленной системе можно установить пакет opennebula-node-lxd:

```
# apt-get install opennebula-node-lxd
```

И инициализировать lxd, выполнив команду из файла `/usr/share/doc/opennebula-node-lxd-5.10.5/README.opennebula-lxd`

3.7.2 Добавление узла типа LXD в OpenNebula

Для добавления узла типа LXD на сервере OpenNebula, необходимо в левом меню выбрать «Инфраструктура» → «Узлы» и на загруженной странице нажать кнопку «+».

Далее необходимо указать тип виртуализации – LXD, заполнить поле «Имя хоста» (можно ввести Ip-адрес узла, или его имя) и нажать кнопку «Создать» (Рис. 38).

Добавление узла типа LXD в OpenNebula-Sunstone

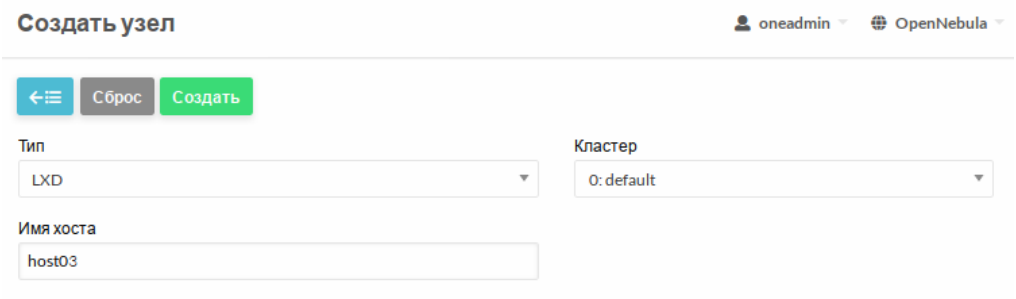

Рис. 38

Затем следует вернуться к списку узлов и убедиться, что узел перешел в состояние ВКЛ (это должно занять от 20 секунд до 1 минуты).

Примечание. Для добавления узла типа LXD в командной строке:

```
$ onehost create host03 -im lxd -vm lxd
ID: 3
```

3.7.3 Скачивание шаблона контейнера из магазина приложений

Для загрузки контейнера из магазина необходимо перейти в «Хранилище» → «Магазины приложений, выбрать «Linux Containers» → «Приложения» (Рис. 39).

Магазин приложений OpenNebula

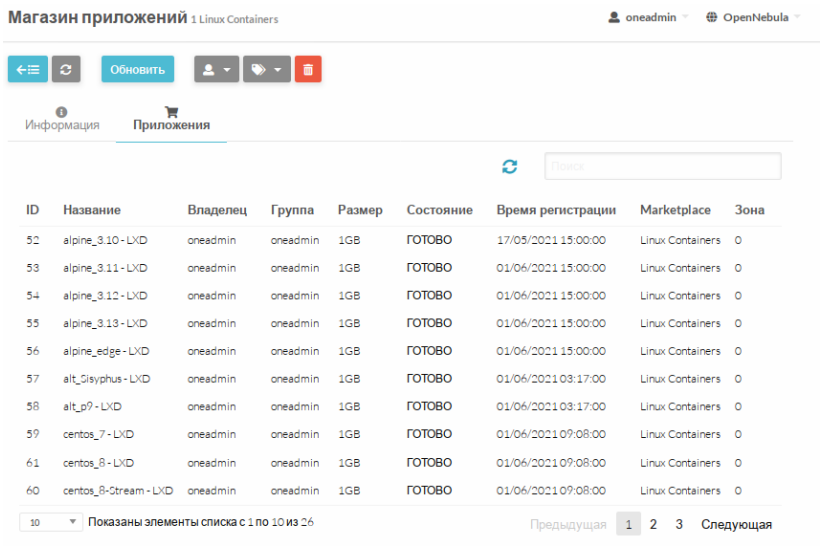

Рис. 39

Выбрать LXD образ. Чтобы импортировать контейнер, необходимо его выбрать и нажать кнопку «Import into Datastore» (Рис. 40).

Импорт контейнера из магазина приложений OpenNebula

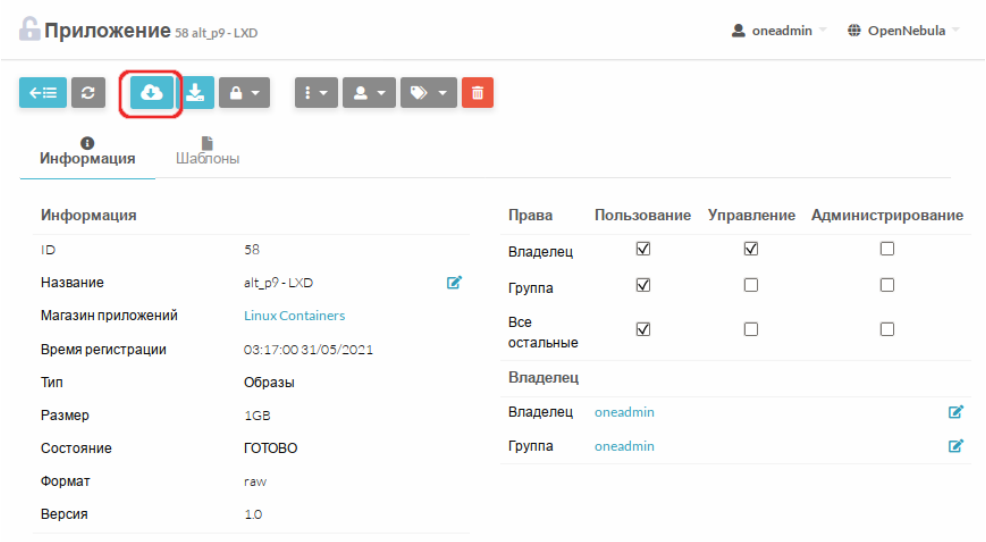

Рис. 40

Каждый контейнер содержит образ и шаблон.

В открывшемся окне указать название для образа и шаблона, выбрать хранилище и нажать кнопку «Загрузить» (Рис. 41).

Импорт контейнера из магазина приложений OpenNebula

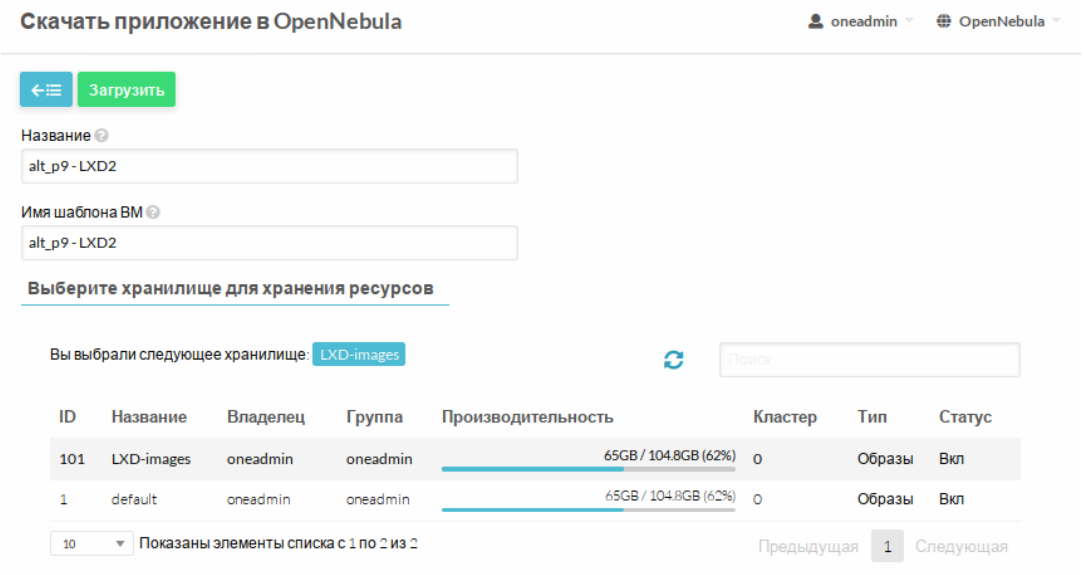

Рис. 41

Из полученного шаблона можно разворачивать контейнеры (ВМ в терминологии Opennebula). Процесс разворачивания контейнера из шаблона такой же, как и процесс разворачивания ВМ из шаблона (Рис. 42).

Разворачивание контейнера из шаблона

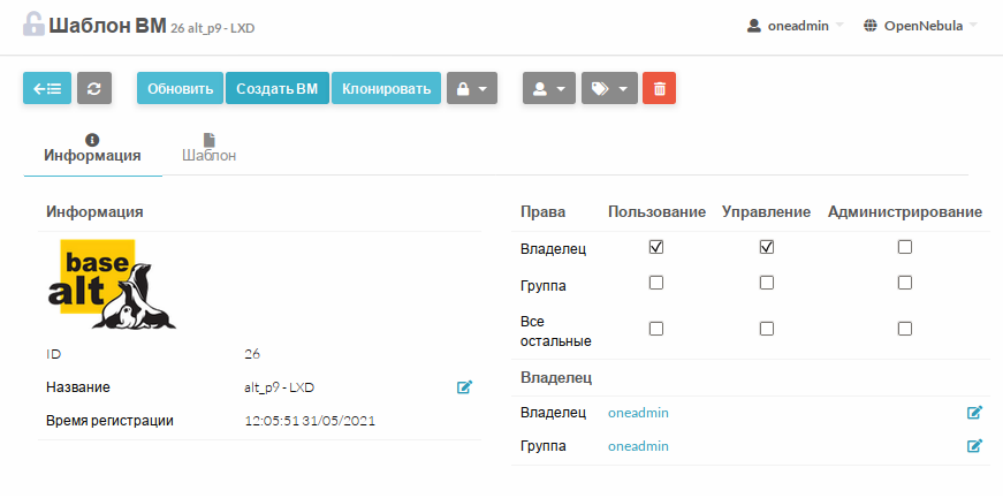

Рис. 42

3.8 Настройка отказоустойчивого кластера

В данном разделе рассмотрена настройка отказоустойчивого кластера (High Available, HA) для основных служб OpenNebula: core (oned), scheduler (mm_sched).

OpenNebula использует распределенный консенсусный протокол Raft, для обеспечения отказоустойчивости и согласованности состояний между службами. Алгоритм консенсуса построен на основе двух концепций:

- Состояние системы – данные, хранящиеся в таблицах базы данных (пользователи, списки управления доступом или виртуальные машины в системе).
- Журнал (Log) – последовательность операторов SQL, которые последовательно применяются к базе данных OpenNebula на всех серверах для изменения состояния системы.

Чтобы сохранить согласованное представление о системе на всех серверах, изменения состояния системы выполняются через специальный узел, лидер или ведущий (Leader). Leader периодически посылает запросы (heartbeats) другим серверам, ведомым (Follower), чтобы сохранить свое лидерство. Если Leader не может послать запрос, Follower-серверы продвигаются к кандидатам и начинают новые выборы.

Каждый раз, когда система изменяется (например, в систему добавляется новая ВМ), Leader обновляет журнал и реплицирует запись у большинства Follower, прежде чем записывать её в базу данных. Таким образом, увеличивается задержка операций с БД, но состояние системы безопасно реплицируется, и кластер может продолжить свою работу в случае отказа узла.

Для настройки High Available требуется:

- нечетное количество серверов (рекомендуемый размер развертывания – 3 или 5 серверов, что обеспечивает отказоустойчивость при отказе 1 или 2 серверов соответственно);
- рекомендуется идентичная конфигурация серверов;

- идентичная программная конфигурация серверов (единственное отличие – это поле SERVER_ID в `/etc/one/oned.conf`);
- рекомендуется использовать подключение к базе данных одного типа (MySQL);
- серверы должны иметь беспарольный доступ для связи друг с другом;
- плавающий IP, который будет газначен лидеру;
- общая файловая система.

Добавлять дополнительные серверы или удалять старые можно после запуска кластера.

В данном примере показана настройка HA кластера из трех серверов:

- 192.168.0.186 opennebula
- 192.168.0.187 server02
- 192.168.0.188 server03
- 192.168.0.200 Floating IP

3.8.1 Первоначальная конфигурация Leader

Запустить сервис OpenNebula и добавить локальный сервер в существующую или новую зону (в примере зона с ID 0):

```
$ onezone list
C   ID NAME             ENDPOINT                                        FED_INDEX
*    0 OpenNebula       http://localhost:2633/RPC2                      -1

$ onezone server-add 0 --name opennebula --rpc http://192.168.0.186:2633/RPC2

$ onezone show 0
ZONE 0 INFORMATION
ID                : 0
NAME              : OpenNebula


ZONE SERVERS
ID NAME            ENDPOINT
 0 opennebula      http://192.168.0.186:2633/RPC2

HA & FEDERATION SYNC STATUS
ID NAME            STATE      TERM       INDEX      COMMIT     VOTE  FED_INDEX
 0 opennebula      solo       0          -1         0          -1    -1

ZONE TEMPLATE
ENDPOINT="http://localhost:2633/RPC2"
```

Остановить сервис opennebula и обновить конфигурацию SERVER_ID в файле /etc/one/oned.conf:

```
FEDERATION = [
    MODE          = "STANDALONE",
    ZONE_ID       = 0,
    SERVER_ID     = 0, # изменить с -1 на 0 (0 — это ID сервера)
    MASTER_ONED   = ""
]
```

Включить Raft-обработчики, чтобы добавить плавающий IP-адрес в кластер (файл /etc/one/oned.conf):

```
RAFT_LEADER_HOOK = [
    COMMAND = "raft/vip.sh",
    ARGUMENTS = "leader enp0s3 192.168.0.200/24"
]

# Executed when a server transits from leader->follower
RAFT_FOLLOWER_HOOK = [
    COMMAND = "raft/vip.sh",
    ARGUMENTS = "follower enp0s3 192.168.0.200/24"
]
```

Запустить сервис OpenNebula и проверить зону:

```
$ onezone show 0
ZONE 0 INFORMATION
ID                : 0
NAME              : OpenNebula


ZONE SERVERS
ID NAME            ENDPOINT
 0 opennebula      http://192.168.0.186:2633/RPC2

HA & FEDERATION SYNC STATUS
ID NAME            STATE      TERM       INDEX      COMMIT     VOTE  FED_INDEX
 0 opennebula      leader     1          5          5          0     -1

ZONE TEMPLATE
ENDPOINT="http://localhost:2633/RPC
```

Сервер opennebula стал Leader-сервером, так же ему присвоен плавающий адрес (Floating IP):

```
$ ip -o a sh enp0s3
```

```
2: enp0s3 inet 192.168.0.186/24 brd 192.168.0.255 scope global enp0s3\
valid_lft forever preferred_lft forever
2: enp0s3 inet 192.168.0.200/24 scope global secondary enp0s3\ valid_lft
forever preferred_lft forever
2: enp0s3 inet6 fe80::a00:27ff:fec7:38e6/64 scope link \ valid_lft forever
preferred_lft forever
```

3.8.2 Добавление дополнительных серверов

Примечание. Данная процедура удалит полностью базу на сервере и заменит её актуальной с Leader-сервера.

Примечание. Рекомендуется добавлять по одному хосту за раз, чтобы избежать конфликта в базе данных.

На Leader создать полную резервную копию актуальной базы и перенести её на другие серверы вместе с файлами из каталога `/var/lib/one/.one/`:

```
$ onedb backup -u oneadmin -d opennebula -p oneadmin
MySQL dump stored in /var/lib/one/mysql_localhost_opennebula_2021-6-23_13:43:21.sql
Use 'onedb restore' or restore the DB using the mysql command:
mysql -u user -h server -P port db_name < backup_file

$ scp /var/lib/one/mysql_localhost_opennebula_2021-6-23_13\:43\:21.sql <ip>:/tmp

$ ssh <ip> rm -rf /var/lib/one/.one
$ scp -r /var/lib/one/.one/ <ip>:/var/lib/one/
```

Остановить сервис OpenNebula на Follower-хостах и восстановить скопированную базу:

```
$ onedb restore -f -u oneadmin -p oneadmin -d opennebula
/tmp/mysql_localhost_opennebula_2021-6-23_13\:43\:21.sql
MySQL DB opennebula at localhost restored.
```

Перейти на Leader-сервер и добавить в зону новые хосты (рекомендуется добавлять серверы по-одному):

```
$ onezone server-add 0 --name server02 --rpc http://192.168.0.187:2633/RPC2
```

Проверить зону на Leader-сервере:

```
$ onezone show 0
ZONE 0 INFORMATION
ID                : 0
NAME              : OpenNebula


ZONE SERVERS
ID NAME            ENDPOINT
 0 opennebula      http://192.168.0.186:2633/RPC2
 1 server02        http://192.168.0.187:2633/RPC2
```

```
HA & FEDERATION SYNC STATUS
ID NAME            STATE      TERM       INDEX      COMMIT     VOTE  FED_INDEX
 0 opennebula      leader     4          22         22         0     -1
 1 server02        error      -          -          -          -     -

ZONE TEMPLATE
ENDPOINT="http://localhost:2633/RPC2"
```

Новый сервер находится в состоянии ошибки, так как OpenNebula на новом сервере не запущена. Следует запомнить идентификатор сервера, в данном случае он равен 1.

Переключиться на добавленный Follower-сервер и обновить конфигурацию SERVER_ID в файле `/etc/one/oned.conf`, (указать в качестве SERVER_ID значение из предыдущего шага). Включить Raft-обработчики, как это было выполнено на Leader.

Запустить сервис OpenNebula на Follower-сервере и проверить на Leader-сервере состояние зоны:

```
$ onezone show 0
ZONE 0 INFORMATION
ID                : 0
NAME              : OpenNebula


ZONE SERVERS
ID NAME            ENDPOINT
 0 opennebula      http://192.168.0.186:2633/RPC2
 1 server02        http://192.168.0.187:2633/RPC2

HA & FEDERATION SYNC STATUS
ID NAME            STATE      TERM       INDEX      COMMIT     VOTE  FED_INDEX
 0 opennebula      leader     4          28         28         0     -1
 1 server02        follower   4          28         28         0     -1

ZONE TEMPLATE
ENDPOINT="http://localhost:2633/RPC2""
```

Повторить данные действия, чтобы добавить третий сервер в кластер.

Примечание. Добавлять серверы в кластер, следует только в случае нормальной работы кластера (работает Leader, а остальные находятся в состоянии Follower). Если в состоянии Error присутствует хотя бы один сервер, необходимо это исправить.

Проверка работоспособности кластера:

```
$ onezone show 0
ZONE 0 INFORMATION
```

```
ID                : 0
NAME              : OpenNebula


ZONE SERVERS
ID NAME            ENDPOINT
 0 opennebula      http://192.168.0.186:2633/RPC2
 1 server02        http://192.168.0.187:2633/RPC2
 2 server03        http://192.168.0.188:2633/RPC2

HA & FEDERATION SYNC STATUS
ID NAME            STATE      TERM       INDEX      COMMIT     VOTE  FED_INDEX
 0 opennebula      leader     4          35         35         0     -1
 1 server02        follower   4          35         35         0     -1
 2 server03        follower   4          35         35         0     -1

ZONE TEMPLATE
ENDPOINT="http://localhost:2633/RPC2"
```

Если какой-либо хост находится в состоянии ошибки, следует проверить журнал (`/var/log/one/oned.log`), как в текущем Leader (если он есть), так и в хосте, который находится в состоянии Error. Все сообщения Raft будут регистрироваться в этом файле.

3.8.3 Добавление и удаление серверов

Команда добавления сервера:

```
onezone server-add <zoneid>
```

Параметры:

- -n, --name – имя сервера зоны;
- -r, --rpc – конечная точка RPC сервера зоны;
- -v, --verbose – подробный режим;
- --user name – имя пользователя, используемое для подключения к OpenNebula;
- --password password – пароль для аутентификации в OpenNebula;
- --endpoint endpoint – URL конечной точки интерфейса OpenNebula xmlrpc.

Команда удаления сервера:

```
onezone server-del <zoneid> <serverid>
```

Параметры:

- -v, --verbose – подробный режим;
- --user name – имя пользователя, используемое для подключения к OpenNebula;
- --password password – пароль для аутентификации в OpenNebula;
- --endpoint endpoint – URL конечной точки интерфейса OpenNebula xmlrpc.

3.8.4 Восстановление сервера

Если Follower -сервер в течение некоторого времени не работает, он может выпасть из окна восстановления. Чтобы восстановить этот сервер необходимо:

- Leader: создать резервную копию БД и скопировать её на отказавший сервер (повторно использовать предыдущую резервную копию нельзя).
- Follower: остановить OpenNebula.
- Follower: восстановить резервную копию БД.
- Follower: запустить OpenNebula.
- Leader: сбросить отказавший Follower, выполнив команду:

```
onezone server-reset <zone_id> <server_id_of_failed_follower>
```

3.8.5 Sunstone

Есть несколько способов развертывания Sunstone в среде HA. Базовым является Sunstone, работающий на каждом интерфейсном узле OpenNebula. Клиенты используют только один сервер – Leader с плавающим IP.

4 СРЕДСТВО УПРАВЛЕНИЯ ВИРТУАЛЬНЫМИ ОКРУЖЕНИЯМИ PVE

4.1 Краткое описание возможностей

PVE – средство управления виртуальными окружениями на базе гипервизора KVM и системы контейнерной изоляции LXC. Основными компонентами среды являются:

- физические серверы, на которых выполняются процессы гипервизора KVM, и процессы, работающие в контейнерах LXC;
- хранилища данных, в которых хранятся образы установочных дисков, образы дисков виртуальных машин KVM и файлы, доступные из контейнеров LXC;
- виртуальные сетевые коммутаторы, к которым подключаются сетевые интерфейсы виртуальных окружений.

PVE состоит из веб-интерфейса, распределенного хранилища данных конфигурации виртуальных окружений, а также утилит конфигурирования, работающих в командной строке. Все эти инструменты предназначены только для управления средой выполнения виртуальных окружений. За формирование среды отвечают компоненты системы, не входящие непосредственно в состав PVE. В первую очередь это сетевая и дисковая подсистемы, а также механизмы аутентификации.

4.1.1 Веб-интерфейс

Веб-интерфейс PVE предназначен для решения следующих задач:

- создание, удаление, настройка виртуальных окружений;
- управление физическими серверами;
- мониторинг активности виртуальных окружений и использования ресурсов среды;
- фиксация состояний (snapshot-ы), создание резервных копий и шаблонов виртуальных окружений, восстановление виртуальных окружений из резервных копий.

Кроме решения пользовательских задач, веб-интерфейс PVE можно использовать еще и для встраивания в интегрированные системы управления – например, в панели управления хостингом. Для этого он имеет развитый RESTful API с JSON в качестве основного формата данных.

Для аутентификации пользователей веб-интерфейса можно использовать как собственные механизмы PVE, так и PAM. Использование PAM дает возможность, например, интегрировать PVE в домен аутентификации.

Так как используется кластерная файловая система (pmxcfs), можно подключиться к любому узлу для управления всем кластером. Каждый узел может управлять всем кластером.

Веб-интерфейс PVE доступен по адресу https://<внешний-ip-адрес>:8006. Потребуется пройти аутентификацию (логин по умолчанию: root, пароль указывается в процессе установки).

Пользовательский интерфейс PVE (Рис. 43) состоит из четырех областей:

– заголовок – верхняя часть. Показывает информацию о состоянии и содержит кнопки для наиболее важных действий;
– дерево ресурсов – с левой стороны. Дерево навигации, где можно выбирать конкретные объекты;
– панель контента – центральная часть. Здесь отображаются конфигурация и статус выбранных объектов;
– панель журнала – нижняя часть. Отображает записи журнала для последних задач. Чтобы получить более подробную информацию или прервать выполнение задачи, следует дважды щелкнуть левой клавишей мыши по записи журнала.

Веб-интерфейс PVE

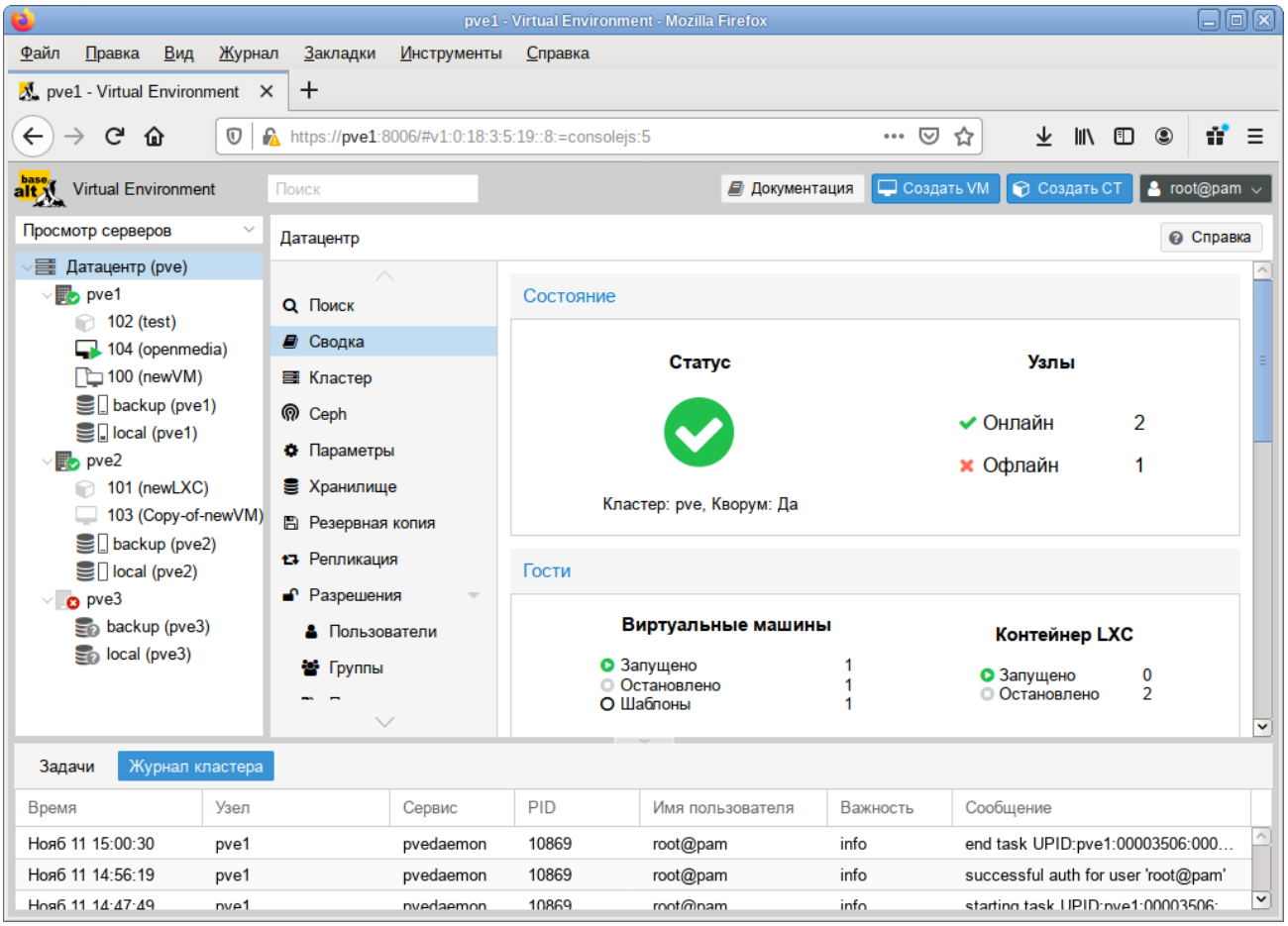

Рис. 43

4.1.2 Хранилище данных

В случае локальной установки PVE для размещения данных виртуальных окружений можно дополнительно ничего не настраивать и использовать локальную файловую систему

сервера. Но в случае кластера из нескольких серверов имеет смысл настроить сетевую или распределенную сетевую файловую систему, обеспечивающую параллельный доступ к файлам со всех серверов, на которых выполняются процессы виртуальных окружений. В качестве таких файловых систем могут выступать, например, NFS или CEPH.

4.1.3 Сетевая подсистема

В отличие от хранилища данных, сетевая подсистема требует специальной настройки даже в случае локальной установки PVE. Это обусловлено тем, что сетевые интерфейсы виртуальных окружений должны подключаться к какому-то виртуальному же устройству, обеспечивающему соединение с общей сетью передачи данных. Перед началом настройки сети следует принять решение о том, какой тип виртуализации (LXC/KVM) и какой тип подключения будет использоваться (маршрутизация/мост).

4.2 Установка и настройка PVE

Примечание. Компоненты PVE будут установлены в систему, если при установке дистрибутива выбрать профиль «Виртуальное окружение Proxmox». Также при установке дистрибутива, необходимо настроить Ethernet-мост vmbr0 и при заполнении поля с именем компьютера указать полное имя с доменом.

Все остальные настройки можно делать в веб-интерфейсе см. «Создание кластера PVE».

4.2.1 Установка PVE

Если развертывание PVE происходит в уже установленной системе на базе Девятой платформы, достаточно любым штатным способом (apt-get, aptitude, synaptic) установить пакет `pve-manager` (все необходимые компоненты при этом будут установлены автоматически):

```
# apt-get update
# apt-get install pve-manager
```

Также следует убедиться в том, что пакет systemd обновлен до версии, находящейся в репозитории P9.

4.2.2 Настройка сетевой подсистемы

На всех узлах кластера необходимо настроить Ethernet-мост.

Примечание. Мосту должно быть назначено имя vmbr0 и оно должно быть одинаково на всех узлах.

Примечание. При использовании дистрибутива Альт Сервер Виртуализации интерфейс vmbr0 создаётся и настраивается в процессе установки системы.

4.2.2.1 Настройка Ethernet-моста в командной строке

Перед настройкой Ethernet-моста (далее – моста) с помощью etcnet сначала необходимо убедиться, что установлен пакет `bridge-utils`. Etcnet использует утилиту `brctl` для

настройки моста, и, если утилита не установлена, то при перезапуске системы сеть станет недоступна. Если интерфейсы, входящие в состав моста, являются единственными физически подключенными и настройка моста происходит с удаленного узла через эти интерфейсы, то требуется соблюдать осторожность, т.к. эти интерфейсы перестанут быть доступны. В случае ошибки в конфигурации потребуется физический доступ к серверу. Для страховки, перед перезапуском сервиса network можно открыть еще одну консоль и запустить там, например, команду: `sleep 500 && reboot`.

Для настройки Ethernet-моста с именем vmbr0, следует выполнить следующие команды:

```
# mkdir /etc/net/ifaces/vmbr0
# cp /etc/net/ifaces/enp0s3/* /etc/net/ifaces/vmbr0/
# rm -f /etc/net/ifaces/enp0s3/{i,r}*
# cat <<EOF > /etc/net/ifaces/vmbr0/options
BOOTPROTO=static
CONFIG_WIRELESS=no
CONFIG_IPV4=yes
HOST='enp0s3'
ONBOOT=yes
TYPE=bri
EOF
```

Имя интерфейса, обозначенного здесь как `enp0s3`, следует указать в соответствии с реальной конфигурацией сервера.

IP-адрес для интерфейса будет взят из `ipv4address`.

В опции HOST файла `options` нужно указать те интерфейсы, которые будут входить в мост. Если в него будут входить интерфейсы, которые до этого имели IP-адрес (например, `enp0s3`), то этот адрес должен быть удален (например, можно закомментировать содержимое файла `/etc/net/ifaces/enp0s3/ipv4address`).

Для того, чтобы изменения вступили в силу необходим перезапуск сервиса `network`:

```
# systemctl restart network
```

При старте сети сначала поднимаются интерфейсы, входящие в мост, затем сам мост (автоматически).

4.2.2.2 Настройка Ethernet-моста в веб-интерфейсе

При установленных пакетах alterator-net-eth и alterator-net-bridge, для настройки Ethernet-моста можно воспользоваться веб-интерфейсом центра управления системой.

Примечание. Должен также быть установлен пакет alterator-fbi и запущены сервисы ahttpd и alteratord:

```
# apt-get install alterator-fbi
# systemctl start ahttpd
# systemctl start alteratord
```

Веб-интерфейс доступен по адресу https://ip-address:8080.

Для настройки Ethernet-моста необходимо выполнить следующие действия:

1) в группе «Сеть» выбрать пункт «Ethernet-интерфейсы»;
2) удалить IP-адрес и шлюз по умолчанию (Рис. 44) и нажать кнопку «Создать сетевой мост»;
3) в открывшемся окне (Рис. 45), задать имя моста vmbr0, выбрать сетевой интерфейс в списке доступных интерфейсов («Available interfaces»), переместить его в список «Участники» («Members») и нажать кнопку «Ok»;
4) настроить сетевой интерфейс vmbr0: ввести имя компьютера, задать IP-адрес и нажать кнопку «Добавить», ввести адрес шлюза по умолчанию и DNS-сервера (Рис. 46).

Настройка сети в веб-интерфейсе

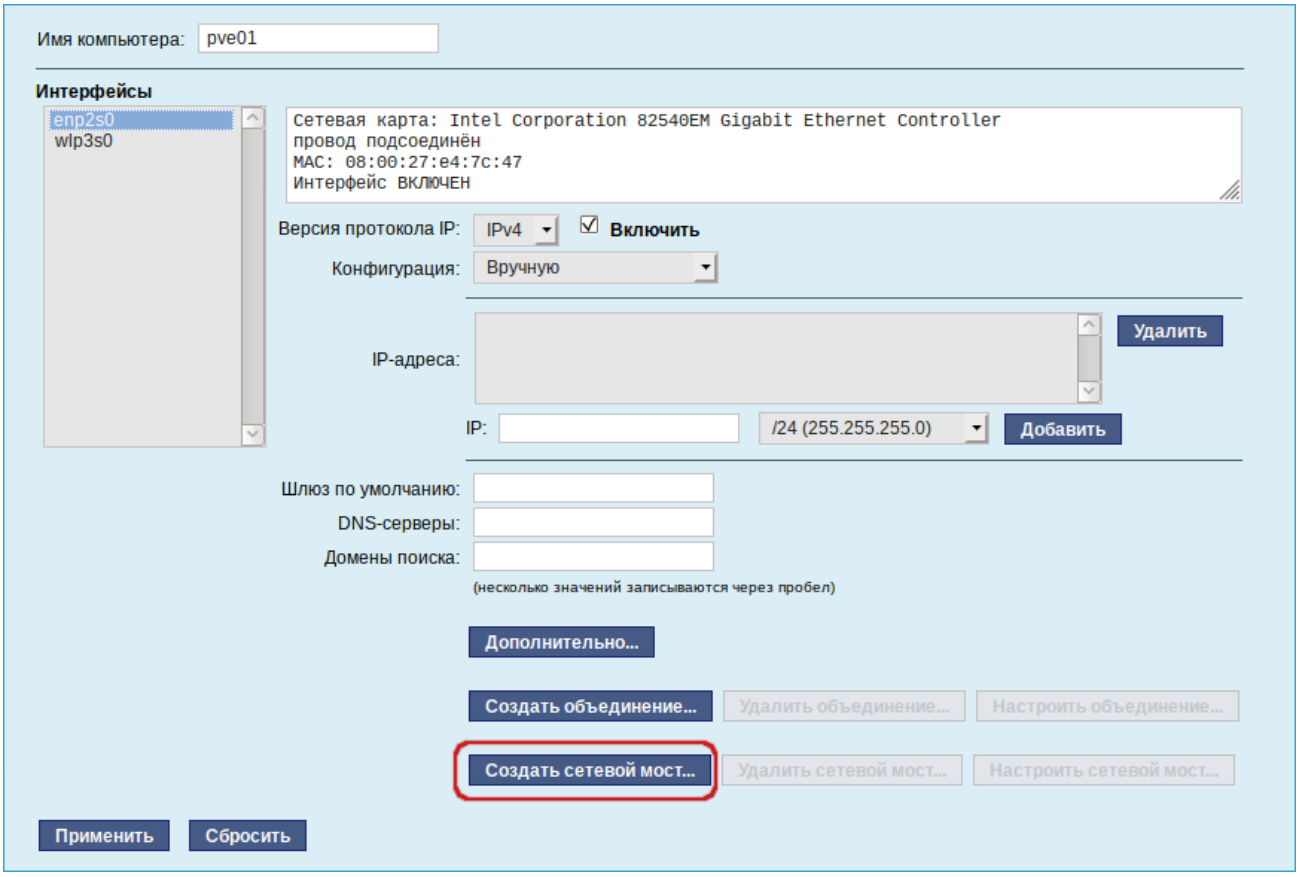

Рис. 44

Выбор сетевого интерфейса

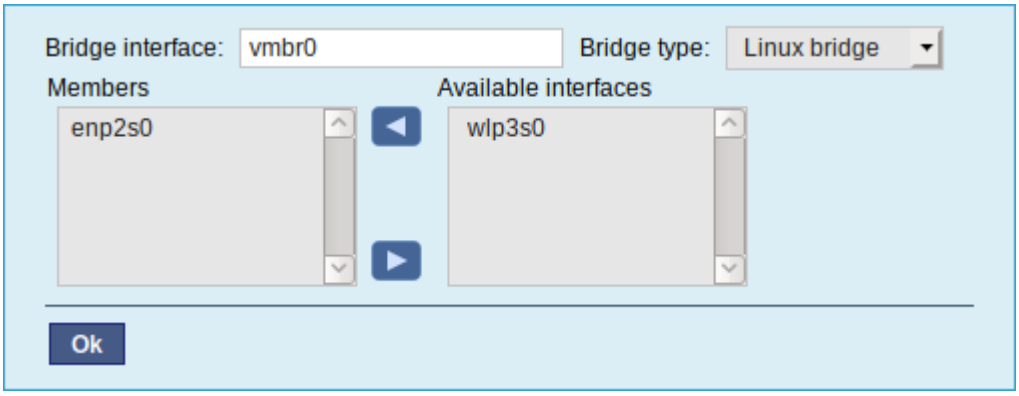

Рис. 45

Настройка параметров сетевого интерфейса vmbr0

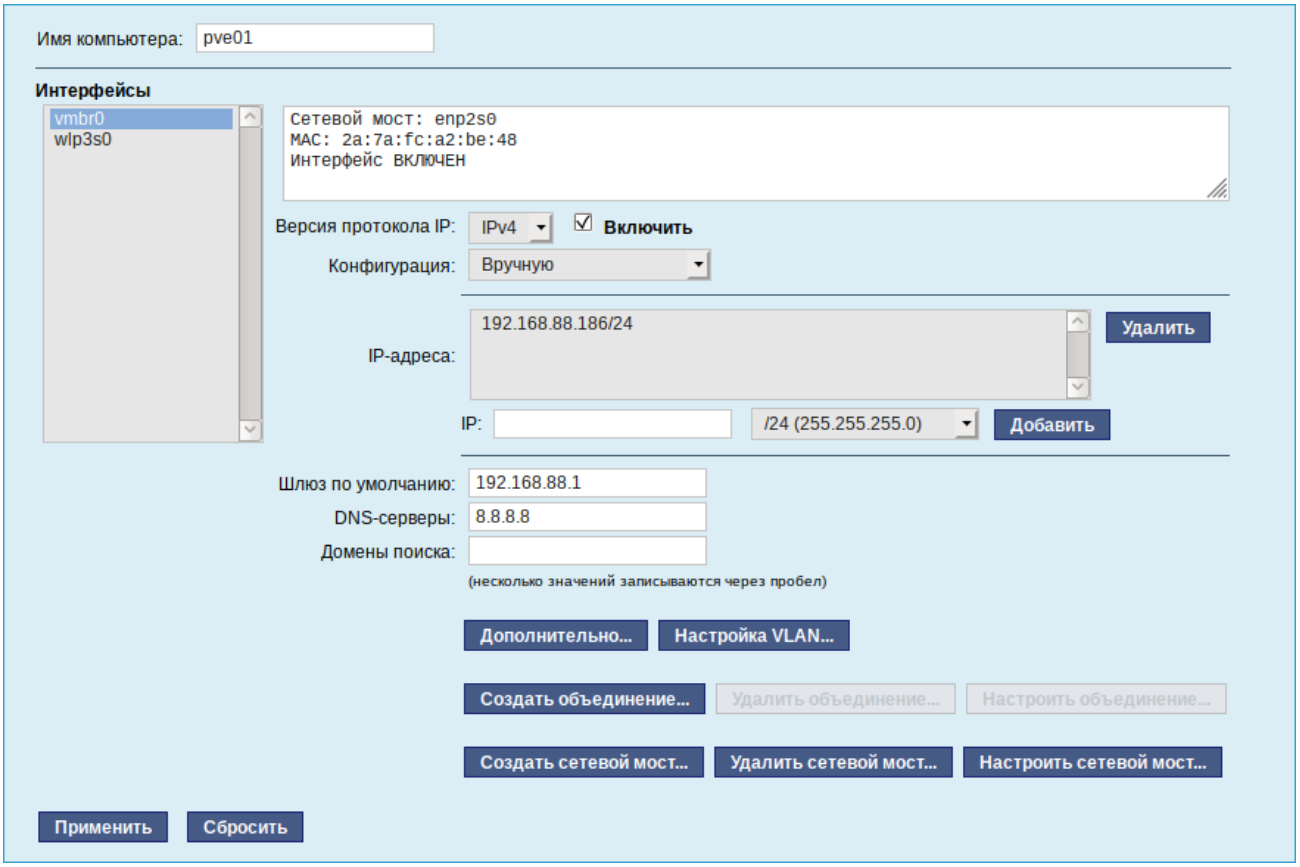

Рис. 46

4.3 Создание кластера PVE

Рекомендации:

- все узлы должны иметь возможность подключаться друг к другу через UDP порты 5404 и 5405;
- дата и время должны быть синхронизированы;
- между узлами используется SSH туннель на 22 TCP порту;

- если необходимо обеспечение высокой доступности (High Availability), необходимо иметь как минимум три узла для организации кворума. На всех узлах должна быть установлена одна версия PVE.
- рекомендуется использовать выделенный сетевой адаптер для трафика кластера, особенно если используется общее хранилище.

В настоящее время создание кластера может быть выполнено либо в консоли (вход через ssh), либо в веб-интерфейсе («Датацентр» → «Кластер»).

Примечание. PVE при создании кластера включает парольную аутентификацию для root в sshd. В целях повышения безопасности, после включения всех серверов в кластер, рекомендуется отключить парольную аутентификацию для root в sshd:

```
# control sshd-permit-root-login without_password
# systemctl restart sshd
```

При добавлении в кластер нового сервера, можно временно включить парольную аутентификацию:

```
# control sshd-permit-root-login enabled
# systemctl restart sshd
```

А после того, как сервер будет добавлен, снова отключить.

Кластеры используют ряд определенных портов (Табл. 2) для различных функций. Важно обеспечить доступность этих портов и отсутствие их блокировки межсетевыми экранами.

Таблица 2 – Используемые порты

Порт	Функция
TCP 8006	Веб-интерфейс PVE
TCP 5900-5999	Доступ к консоли VNC
TCP 3128	Доступ к консоли SPICE
TCP 22	SSH доступ
UDP 5404, 5405	Широковещательный CMAN для применения настроек кластера

Кластер не создается автоматически на только что установленном узле PVE. Он должен быть создан через интерфейс командной строки с любого из узлов PVE, который будет частью кластера. После того как кластер создан и узлы добавлены в этот кластер, большая часть управления может выполняться через графический интерфейс.

Для правильного функционирования кластера необходимо как минимум три узла. С тремя узлами возможен кворум, который позволяет кластерам обеспечивать работоспособность и функционирование должным образом.

4.3.1 Настройка узлов кластера

PVE должен быть установлен на всех узлах. Следует убедиться, что каждый узел установлен с окончательным именем хоста и IP-конфигурацией. Изменение имени хоста и IP-адреса невозможно после создания кластера.

Необходимо обеспечить взаимно однозначное прямое и обратное преобразование имен для всех узлов кластера. Крайне желательно использовать DNS, в крайнем случае, можно обойтись соответствующими записями в локальных файлах `/etc/hosts`:

```
# echo "192.168.0.186 pve01.test.alt pve01" >> /etc/hosts
# echo "192.168.0.90 pve02.test.alt pve02" >> /etc/hosts
# echo "192.168.0.70 pve03.test.alt pve03" >> /etc/hosts
```

Примечание. В PVE это можно сделать в панели управления: выбрать узел, перейти в «Система» → «Hosts», добавить все узлы, которые будут включены в состав кластера (Рис. 47).

Редактирование записей в файле /etc/hosts

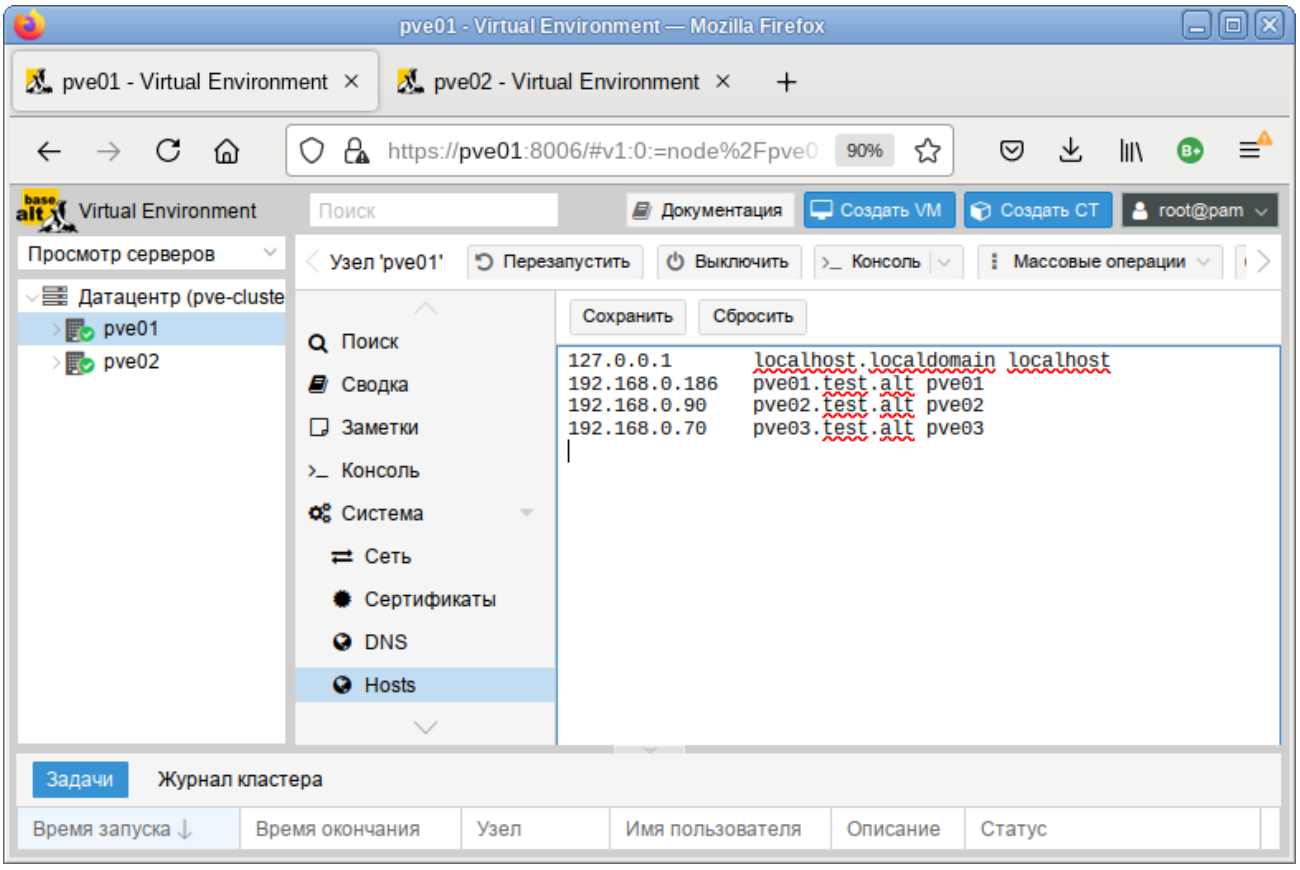

Рис. 47

Примечание. Имя машины не должно присутствовать в файле `/etc/hosts` разрешающимся в 127.0.0.1.

4.3.2 Создание кластера в веб-интерфейсе

Веб-интерфейс PVE доступен по адресу https://<имя-компьютера>:8006. Потребуется пройти аутентификацию (логин по умолчанию: root, пароль указывается в процессе установки) (Рис. 48).

Аутентификация в веб-интерфейсе PVE

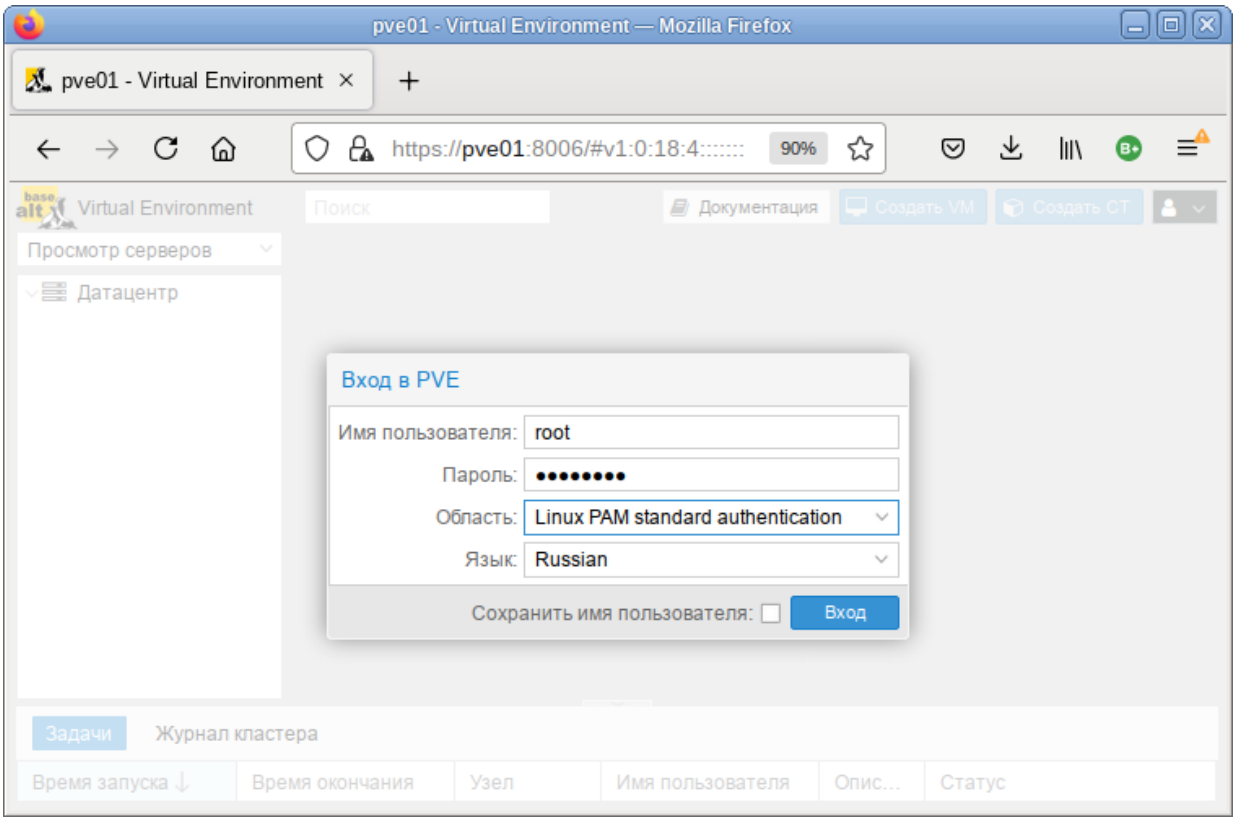

Рис. 48

Для создания кластера необходимо выполнить следующие действия:

1) в панели управления любого узла кластера выбрать «Датацентр» → «Кластер» и нажать кнопку «Создать кластер» (Рис. 49);
2) в открывшемся окне, задать название кластера, выбрать IP-адрес интерфейса, на котором узел кластера будет работать, и нажать кнопку «Создать» (Рис. 50);
3) при успешном создании кластера будет выведена соответствующая информация (Рис. 51).

Создание кластера в веб-интерфейсе

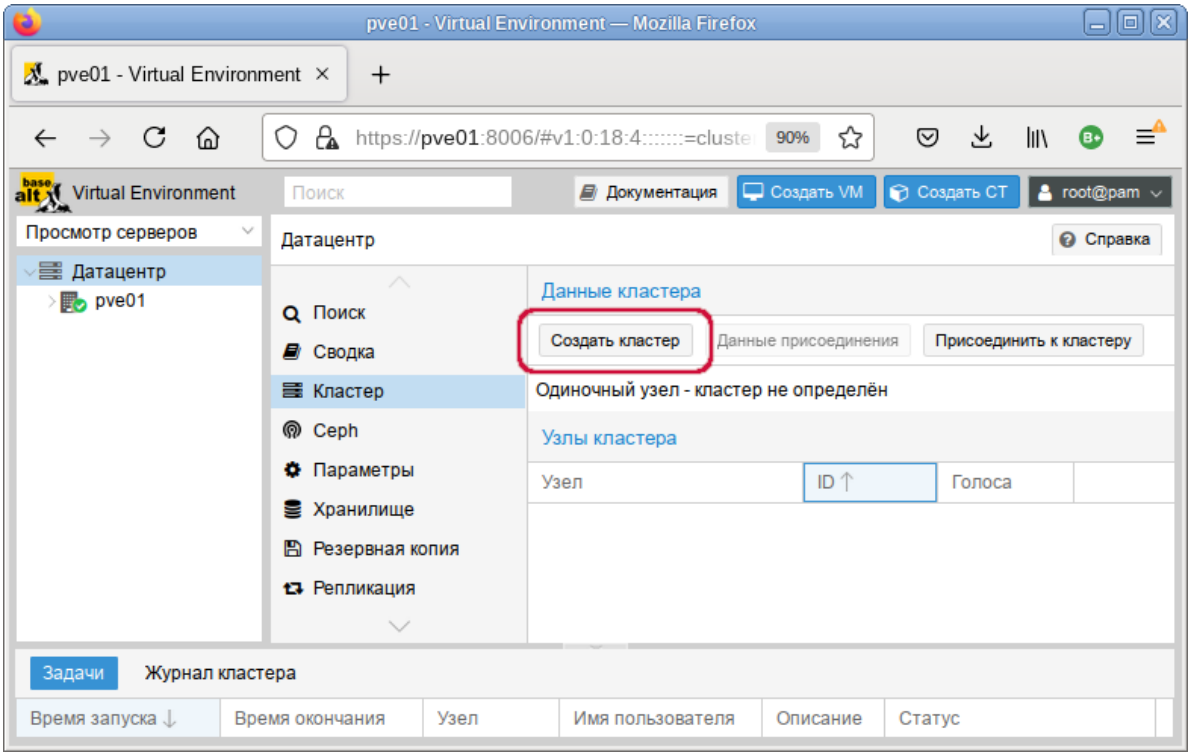

Рис. 49

Создание кластера в веб-интерфейсе. Название кластера

Рис. 50

Информация о создании кластера

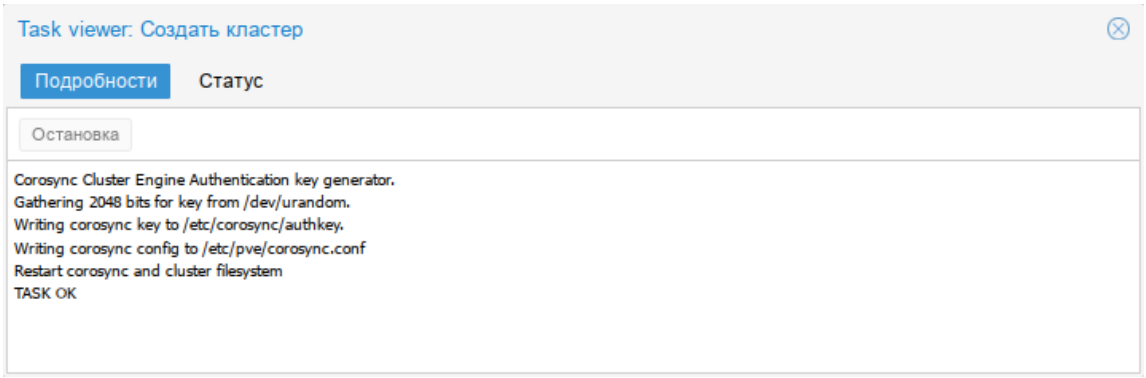

Рис. 51

Для добавления узла в кластер необходимо выполнить следующие действия:

1) в панели управления узла, на котором был создан кластер, выбрать «Датацентр» → «Кластер» и нажать кнопку «Данные присоединения» (Рис. 52);
2) в открывшемся окне, скопировать данные присоединения (поле «Данные присоединения») (Рис. 53);
3) перейти в панель управления узла, который следует присоединить к кластеру. Выбрать «Датацентр» → «Кластер» и нажать кнопку «Присоединить к кластеру» (Рис. 54);
4) в поле «Данные» вставить скопированные данные присоединения, в поле «Пароль» ввести пароль пользователя root первого узла (Рис. 55) и нажать кнопку «Присоединение» («Join <имя кластера>»). Поля «Адрес сервера» и «Отпечаток» будут заполнены автоматически;
5) через несколько минут, после завершения репликации всех настроек, узел будет подключен к кластеру (Рис. 56).

Создание кластера в веб-интерфейсе. Получить данные присоединения

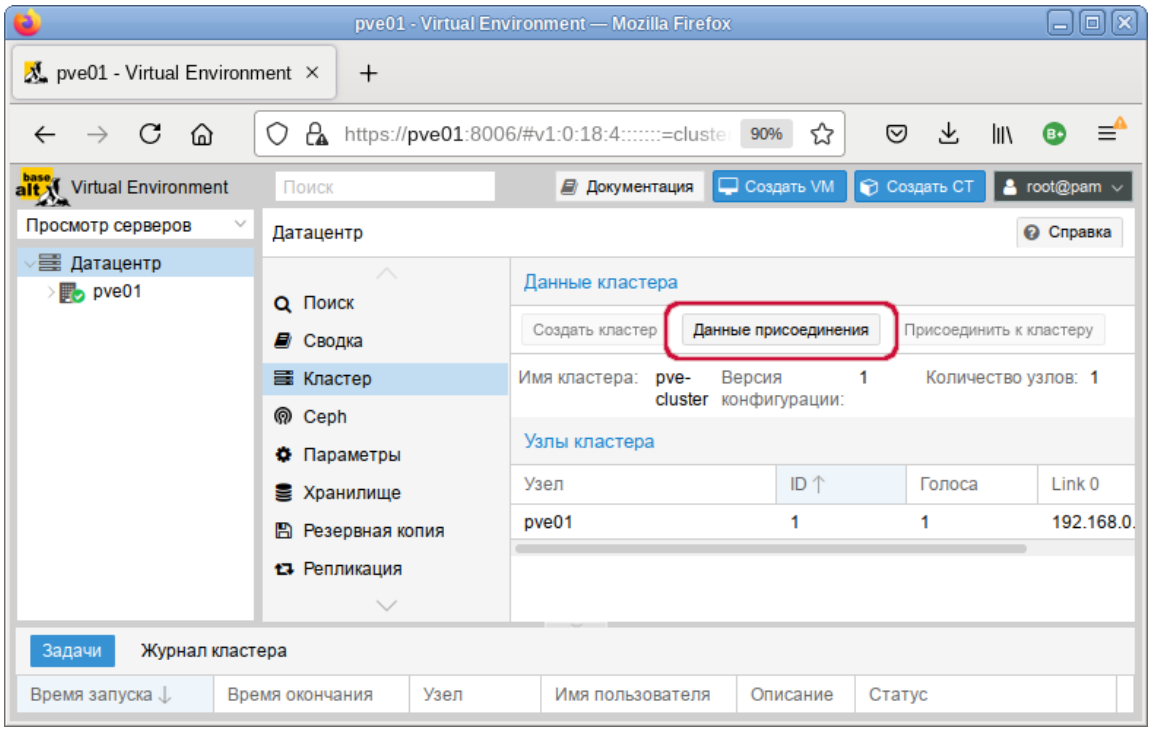

Рис. 52

Создание кластера в веб-интерфейсе. Данные присоединения

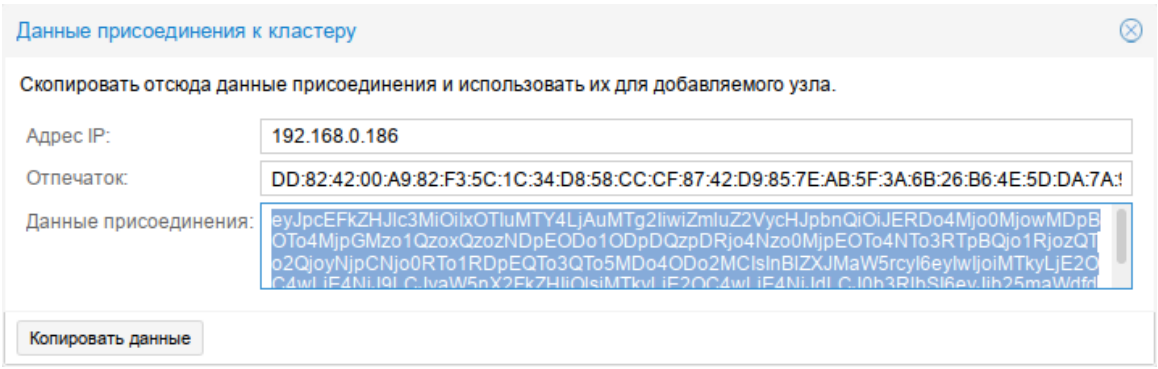

Рис. 53

Узел, который следует присоединить к кластеру

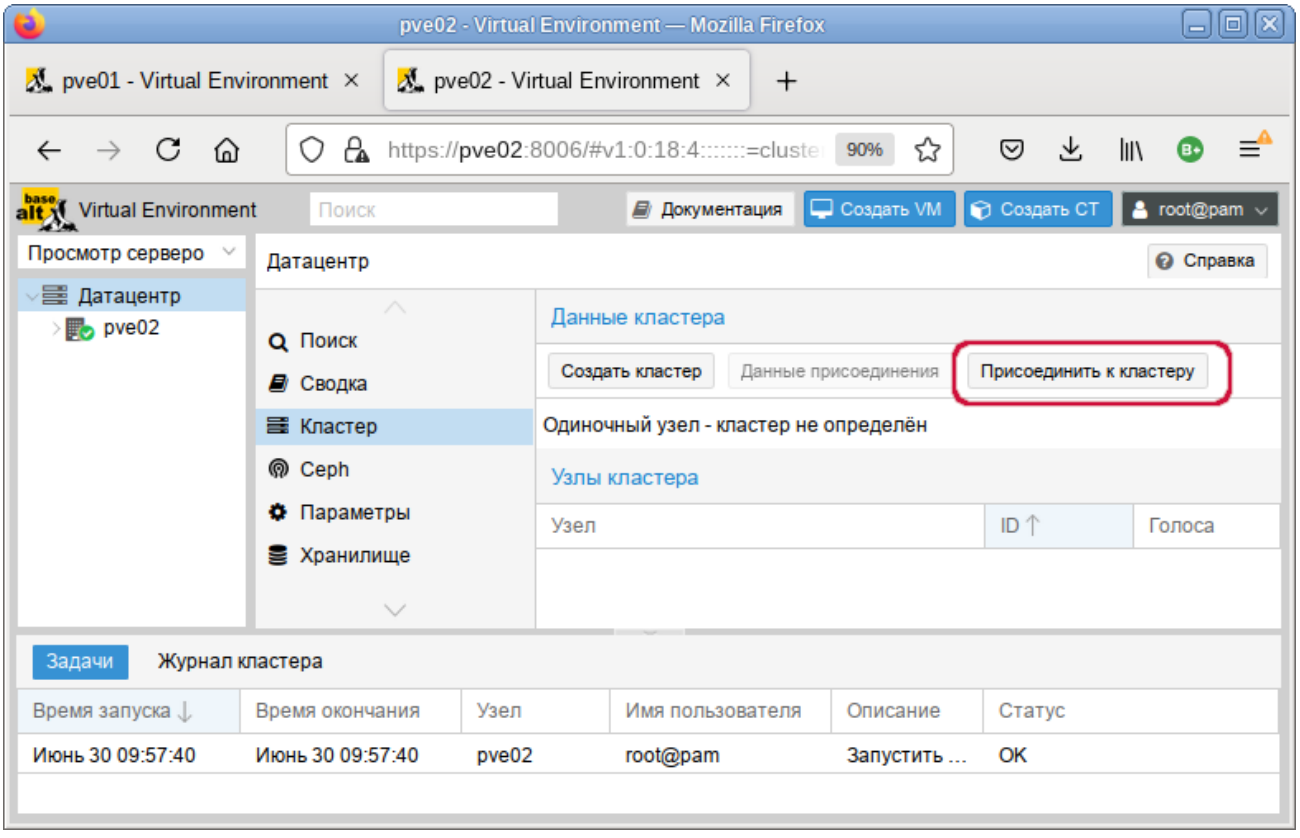

Рис. 54

Присоединение узла к кластеру

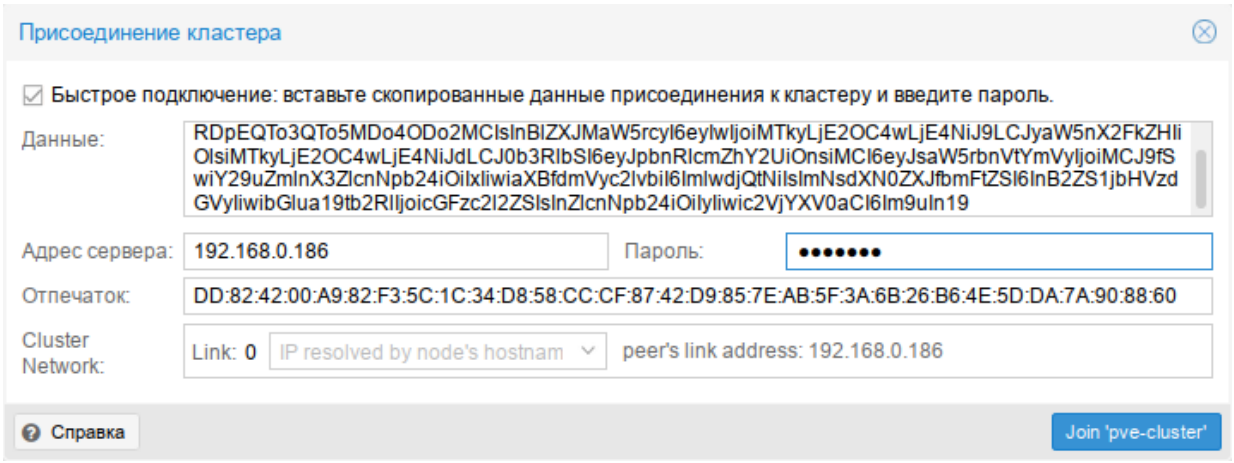

Рис. 55

Узлы кластера в веб-интерфейсе

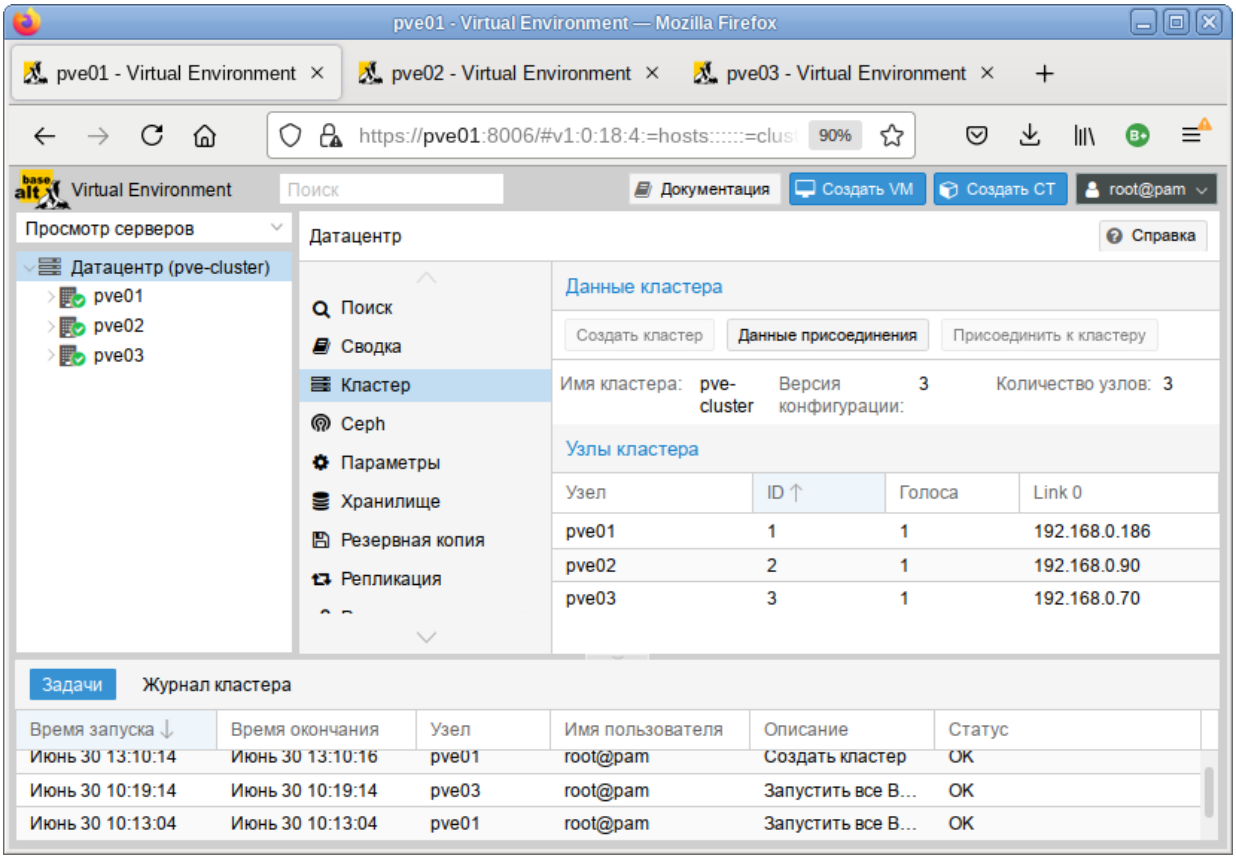

Рис. 56

Сразу после инициализации кластера в пределах каждого из узлов доступно одно локальное хранилище данных (Рис. 57).

PVE кластер может состоять из двух и более серверов. Максимальное количество узлов в кластере – 32.

Веб-интерфейс PVE

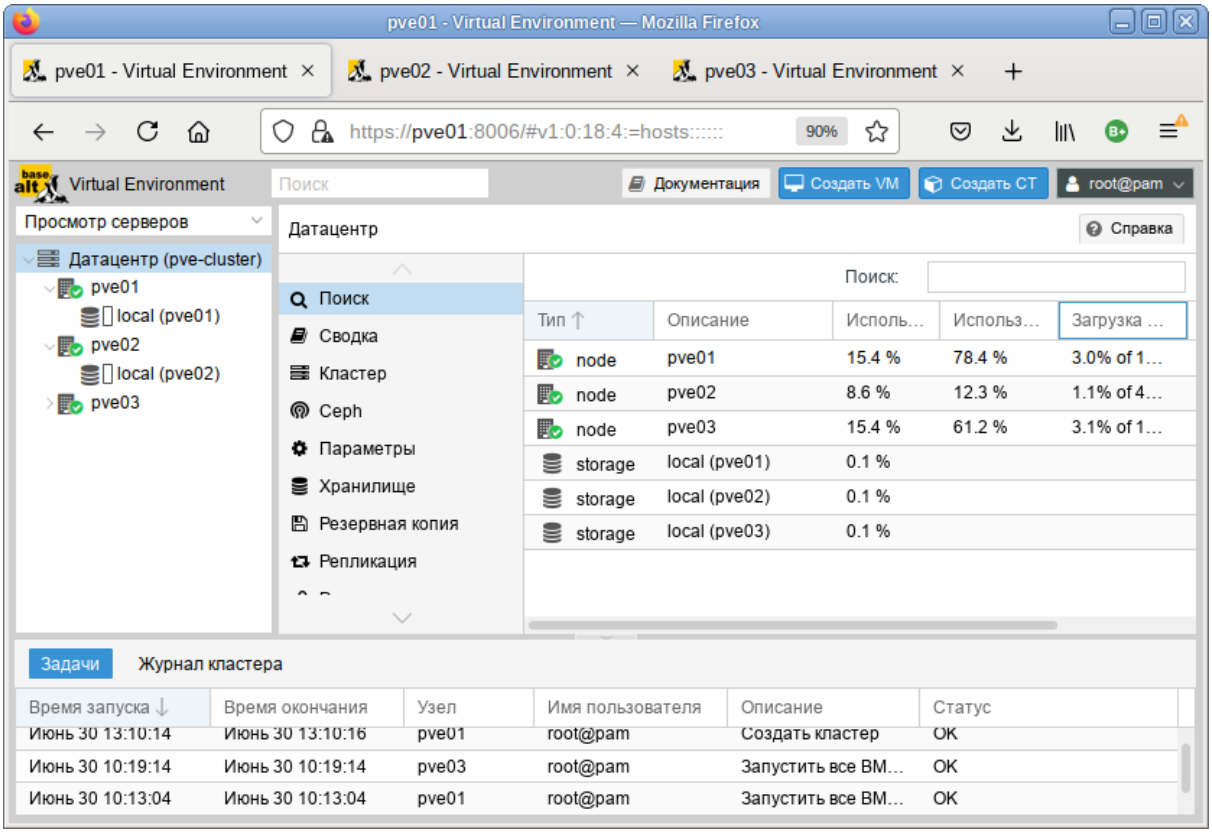

Рис. 57

4.3.3 Создание кластера в консоли

Команда, создания кластера:

```
# pvecm create <cluster_name>
```

На «головном» узле кластера необходимо выполнить команде инициализации конкретного кластера PVE, в данном примере – «pve-cluster»:

```
# pvecm create pve-cluster
```

Проверка:

```
# pvecm status
Cluster information
-------------------
Name:             pve-cluster
Config Version:   1
Transport:        knet
Secure auth:      on

Quorum information
------------------
```

```
Date:             Thu Jul  8 14:21:44 2021
Quorum provider:  corosync_votequorum
Nodes:            1
Node ID:          0x00000001
Ring ID:          1.d6
Quorate:          Yes

Votequorum information
----------------------
Expected votes:   1
Highest expected: 1
Total votes:      1
Quorum:           1
Flags:            Quorate

Membership information
----------------------
    Nodeid      Votes Name
0x00000001          1 192.168.0.186 (local)
```

Команда создания кластера создает файл настройки `/etc/pve/corosync.conf`. По мере добавления узлов в кластер файл настройки будет автоматически пополняться информацией об узлах.

Команда, для добавления узла в кластер:

```
# pvecm add <existing_node_in_cluster>
```

где `existing_node_in_cluster` – адрес уже добавленного узла (рекомендуется указывать самый первый).

Для добавления узлов в кластер, необходимо на «подчиненных» узлах выполнить команде:

```
# pvecm add pve01
```

где `pve01` – имя или IP-адрес «головного» узла.

При добавлении узла в кластер, потребуется ввести пароль администратора главного узла кластера:

```
# pvecm add pve01
Please enter superuser (root) password for 'pve01': ***
Establishing API connection with host 'pve01'
```

```
Login succeeded.
Request addition of this node
Join request OK, finishing setup locally
stopping pve-cluster service
backup old database to '/var/lib/pve-cluster/backup/config-1625747072.sql.gz'
waiting for quorum...OK
(re)generate node files
generate new node certificate
merge authorized SSH keys and known hosts
generated new node certificate, restart pveproxy and pvedaemon services
successfully added node 'pve03' to cluster.
```

После добавления узлов в кластер, файл настройки кластера в `/etc/pve/corosync.conf` должен содержать информацию об узлах кластера.

4.3.4 Удаление узла из кластера

Перед удалением узла из кластера необходимо переместить все ВМ с этого узла, а также убедиться, что нет никаких локальных данных или резервных копий, которые необходимо сохранить.

Для удаления узла из кластера необходимо выполнить следующие шаги:

1) войти в узел кластера не подлежащий удалению;

2) ввести команду `pvecm nodes`, чтобы определить идентификатор узла, который следует удалить:

```
# pvecm nodes
Membership information
----------------------
    Nodeid      Votes Name
         1          1 pve01 (local)
         2          1 pve02
         3          1 pve03
```

3) выключить подлежащий удалению узел (в данном случае pve02);

4) удалить узел из кластера, выполнив команду:

```
# pvecm delnode pve02
```

5) проверить, что узел удален (команда отобразит список узлов кластера без удаленного узла):

```
# pvecm nodes
Membership information
----------------------
    Nodeid      Votes Name
         1          1 pve01 (local)
         3          1 pve03
```

Если необходимо вернуть удаленный узел обратно в кластер, следует выполнить следующие действия:

- переустановить PVE на этом узле (это гарантирует, что все секретные ключи кластера/ssh и любые данные конфигурации будут уничтожены);
- присоединиться к кластеру.

4.3.5 Кластерная файловая система PVE (pmxcfs)

Кластерная файловая система PVE (pmxcfs) – это файловая система на основе базы данных для хранения файлов конфигурации виртуальных окружений, реплицируемая в режиме реального времени на все узлы кластера с помощью `corosync`. Эта файловая система используется для хранения всех конфигурационных файлов связанных с PVE. Хотя файловая система хранит все данные в базе данных на диске, копия данных находится в оперативной памяти. Это накладывает ограничение на максимальный размер, который в настоящее время составляет 30 МБ. Но этого вполне достаточно, чтобы хранить конфигурации нескольких тысяч виртуальных машин.

Преимущества файловой системы:

- мгновенная репликация и обновление конфигурации на все узлы в кластере;
- исключается вероятность дублирования идентификаторов виртуальных машин;
- в случае развала кворума в кластере, файловая система становится доступной только для чтения.

Все файлы и каталоги принадлежат пользователю root и имеют группу www-data. Только root имеет права на запись, но пользователи из группы www-data могут читать большинство файлов. Файлы в каталогах `/etc/pve/priv/` и `/etc/pve/nodes/${NAME}/priv/` доступны только root.

Файловая система смонтирована в `/etc/pve/`.

4.4 Системы хранения

Образы ВМ могут храниться в одном или нескольких локальных хранилищах или в общем хранилище, например NFS или iSCSI (NAS, SAN). Ограничений нет, можно настроить столько пулов хранения, сколько необходимо.

В кластерной среде PVE наличие общего хранилища не является обязательным, однако, оно делает управление хранением более простой задачей. Преимущества общего хранилища:

- миграция ВМ в реальном масштабе времени;
- плавное расширение пространства хранения с множеством узлов;
- централизованное резервное копирование;
- многоуровневое кэширование данных;
- централизованное управление хранением.

4.4.1 Типы хранилищ в PVE

Существует два основных класса типов хранения:

- файловые хранилища – хранят данные в виде файлов. Технологии хранения на уровне файлов обеспечивают доступ к полнофункциональной файловой системе (POSIX). В целом они более гибкие, чем любое хранилище на уровне блоков, и позволяют хранить контент любого типа;
- блочное хранилище – позволяет хранить большие необработанные образы. Обычно в таких хранилищах невозможно хранить другие файлы (ISO-образы, резервные копии, и т.д.). Большинство современных реализаций хранилищ на уровне блоков поддерживают моментальные снимки и клонирование. RADOS и GlusterFS являются распределенными системами, реплицирующими данные хранилища на разные узлы.

Хранилищами данных удобно управлять через веб-интерфейс. В качестве бэкенда хранилищ PVE может использовать:

- сетевые файловые системы, в том числе распределенные: NFS, CEPH, GlusterFS;
- локальные системы управления дисковыми томами: LVM, ZFS;
- удаленные (iSCSI) и локальные дисковые тома;
- локальные дисковые каталоги.

Доступные типы хранилищ приведены в табл. 3.

Ряд хранилищ и формат образа Qemu qcow2 поддерживают тонкую настройку. Если активирована тонкая настройка в хранилище будут записаны только те блоки, которые фактически использует гостевая система.

Например, была создана ВМ с жестким диском объемом 32 ГБ, а после установки ОС гостевой системы корневая файловая система ВМ стала содержать 3 ГБ данных. В этом случае только 3 ГБ записывается в хранилище, даже если ВМ видит жесткий диск 32 ГБ. Таким образом, тонкая настройка позволяет создавать образы дисков, размер которых превышает доступные в настоящее время блоки хранения. Можно создавать большие образы дисков для ВМ, а при необходимости добавлять больше дисков в свое хранилище без изменения размера файловых систем ВМ.

Все типы хранилищ, которые поддерживают функцию «Снимки», также поддерживают тонкую настройку.

Т а б л и ц а 3 – Доступные типы хранилищ

Хранилище	PVE тип	Уровень	Общее (shared)	Снимки (snapshots)
ZFS (локальный)	zfspool	файл	нет	да
Каталог	dir	файл	нет	нет (возможны в формате qcow2)
NFS	nfs	файл	да	нет (возможны в формате qcow2)
CIFS	cifs	файл	да	нет (возможны в формате qcow2)
GlusterFS	glusterfs	файл	да	нет (возможны в формате qcow2)
CephFS	cephfs	файл	да	да
LVM	lvm	блок	нет	нет
LVM-thin	lvmthin	блок	нет	да
iSCSI/kernel	iscsi	блок	да	нет
iSCSI/libiscsi	iscsidirect	блок	да	нет
Ceph/RBD	rbd	блок	да	да
ZFS over iSCSI	zfs	блок	да	да
Proxmox Backup	pbs	файл/блок	да	-

4.4.2 Конфигурация хранилища

Вся связанная с PVE конфигурация хранилища хранится в одном текстовом файле `/etc/pve/storage.cfg`. Поскольку этот файл находится в `/etc/pve/`, он автоматически распространяется на все узлы кластера. Таким образом, все узлы имеют одинаковую конфигурацию хранилища.

Совместное использование конфигурации хранилища имеет смысл для общего хранилища, поскольку одно и то же «общее» хранилище доступно для всех узлов. Но также полезно для локальных типов хранения. В этом случае такое локальное хранилище доступно на всех узлах, но оно физически отличается и может иметь совершенно разное содержимое.

Каждое хранилище имеет <тип> и уникально идентифицируется своим <STORAGE_ID>. Конфигурация хранилища выглядит следующим образом:

```
<type>: <STORAGE_ID>
        <property> <value>
        <property> <value>
        ...
```

Строка `<type>: <STORAGE_ID>` определяет хранилище, затем следует список свойств.

Пример файла `/etc/pve/storage.cfg`:

```
# cat /etc/pve/storage.cfg
dir: local
     path /var/lib/vz
     content images,rootdir,iso,snippets,vztmpl
```

```
        maxfiles 0
nfs: nfs-storage
        export /export/storage
        path /mnt/nfs-vol
        server 192.168.0.105
        content images,iso,backup,vztmpl
        options vers=3,nolock,tcp
```

В данном случае файл содержит одно специальное хранилище local, которое ссылается на каталог `/var/lib/vz` и NFS хранилище nfs-storage.

Несколько параметров являются общими для разных типов хранилищ (табл. 4).

Т а б л и ц а 4 – Параметры хранилищ

Свойство	Описание
nodes	Список узлов кластера, где это хранилище можно использовать/доступно. Можно использовать это свойство, чтобы ограничить доступ к хранилищу
content	Хранилище может поддерживать несколько типов контента. Не все типы хранилищ поддерживают все типы контента. Это свойство указывает, для чего используется это хранилище. Доступные опции: – images – образы виртуальных дисков; – rootdir – данные контейнеров; – vztmpl – шаблоны контейнеров; – backup – резервные копии; – iso – ISO образы; – snippets – файлы сниппетов
shared	Пометить хранилище как общее
disable	Отключить хранилище
maxfiles	Устарело, используйте вместо этого prune-backups. Максимальное количество файлов резервных копий на ВМ
prune-backups	Варианты хранения резервных копий
format	Формат по умолчанию (raw|qcow2|vmdk)

Сетевые файловые системы подходят для организации распределенных хранилищ данных, доступных со всех узлов кластера. Остальные могут использоваться только в качестве локальных хранилищ, доступных в пределах одного узла.

4.4.3 Работа с хранилищами в PVE

4.4.3.1 Использование командной строки

Утилита `pvesm` (PVE Storage Manager), позволяет выполнять общие задачи управления хранилищами.

Команды добавления (подключения) хранилища:

```
pvesm add <TYPE> <STORAGE_ID> <OPTIONS>
```

```
pvesm add dir <STORAGE_ID> --path <PATH>
pvesm add nfs <STORAGE_ID> --path <PATH> --server <SERVER> --export
<EXPORT>
pvesm add lvm <STORAGE_ID> --vgname <VGNAME>
pvesm add iscsi <STORAGE_ID> --portal <HOST[:PORT]> --target <TARGET>
```

Отключить хранилище:

```
pvesm set <STORAGE_ID> --disable 1
```

Включить хранилище:

```
pvesm set <STORAGE_ID> --disable 0
```

Для того чтобы изменить/установить опции хранилища можно, выполнить команды:

```
pvesm set <STORAGE_ID> <OPTIONS>
pvesm set <STORAGE_ID> --shared 1
pvesm set local --format qcow2
pvesm set <STORAGE_ID> --content iso
```

Удалить хранилище (при этом никакие данные не удаляются, удаляется только конфигурация хранилища):

```
pvesm remove <STORAGE_ID>
```

Команда выделения тома:

```
pvesm alloc <STORAGE_ID> <VMID> <name> <size> [--format <raw|qcow2>]
```

Выделить том 4 ГБ в локальном хранилище (имя будет сгенерировано):

```
pvesm alloc local <VMID> '' 4G
```

Освободить место (уничтожает все данные тома):

```
pvesm free <VOLUME_ID>
```

Вывести список хранилищ:

```
pvesm status
```

Вывести список содержимого хранилища:

```
pvesm list <STORAGE_ID> [--vmid <VMID>]
```

4.4.3.2 Добавление хранилища в веб-интерфейсе PVE

Хранилища, которые могут быть добавлены через веб-интерфейс PVE (Рис. 58):

- Локальные хранилища:
  - Каталог – хранение на существующей файловой системе;
  - LVM – локальные устройства, такие как, FC, DRBD и т.д.;
  - ZFS;
- Сетевые хранилища:
  - LVM – сетевая поддержка с iSCSI target;

- NFS;
- CIFS;
- GlusterFS;
- iSCSI;
- CephFS;
- RBD;
- ZFS over iSCSI.

Выбор типа добавляемого хранилища

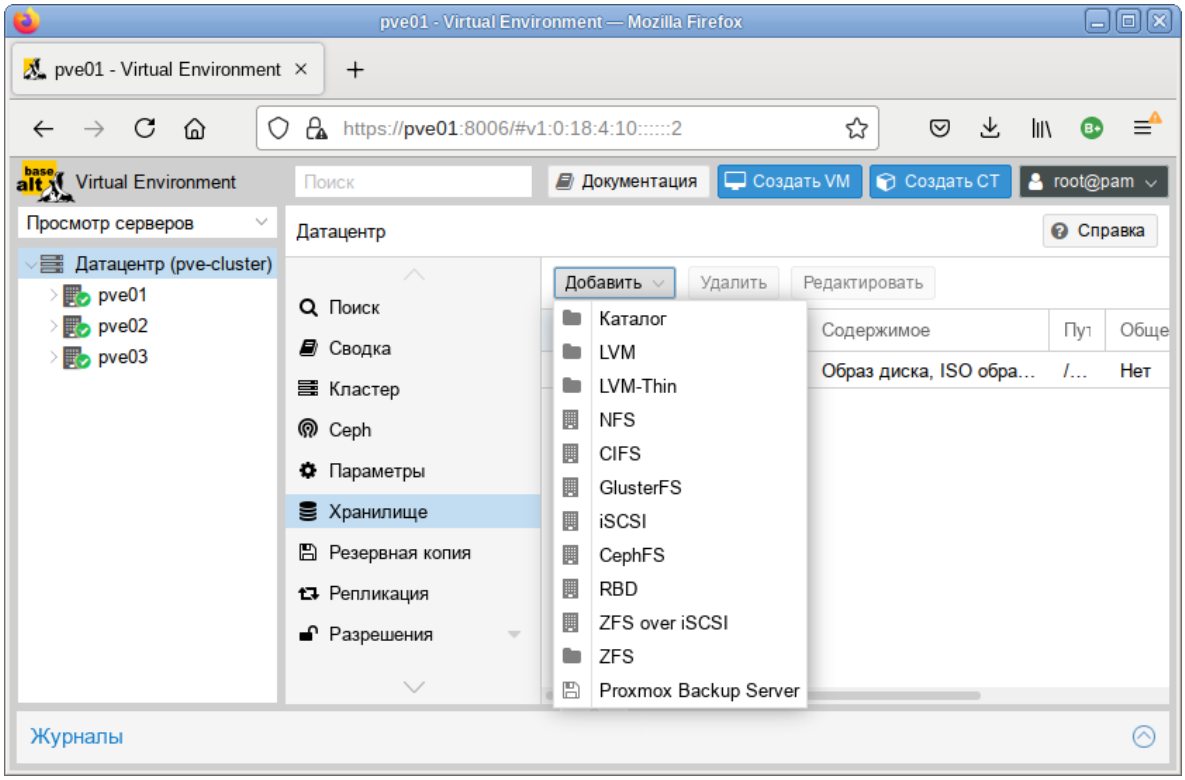

Рис. 58

При создании каждому хранилищу данных присваивается роль или набор ролей. Например, хранение контейнеров, образов виртуальных дисков, файлов .iso и так далее. Список возможных ролей зависит от бэкенда хранилища. Например, для NFS или каталога локальной файловой системы доступны любые роли или наборы ролей (*Рис. 59*), а хранилища на базе CEPH можно использовать только для хранения ISO-образов или шаблонов контейнеров.

Выбор ролей для хранилища

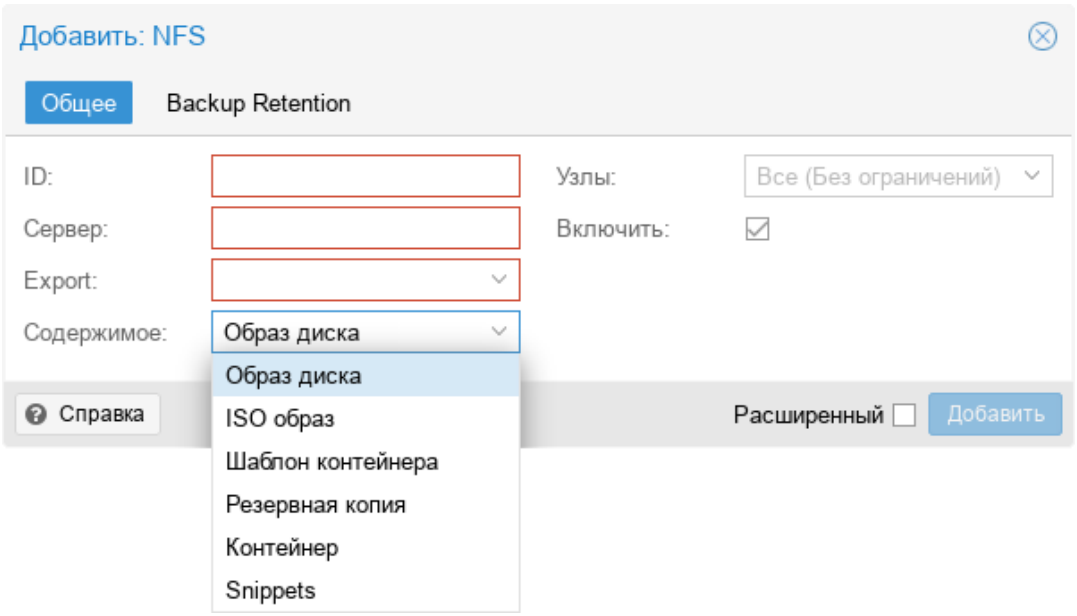

Рис. 59

4.4.3.3 Каталог (Directory)

PVE может использовать локальные каталоги или локально смонтированные общие ресурсы для организации хранилища. Каталог – это хранилище на уровне файлов, поэтому в нем можно хранить данные любого типа, например образы виртуальных дисков, контейнеры, шаблоны, ISO-образы или файлы резервных копий. Для хранения данных разного типа, используется предопределенная структура каталогов (табл. 5). Эта структура используется на всех файловых хранилищах.

Т а б л и ц а 5 – Структура каталогов.

Тип данных	Подкаталог
Образы дисков ВМ	images/<VMID>/
ISO образы	template/iso/
Шаблоны контейнеров	template/cache/
Резервные копии	dump/
Snippets	snippets/

Для создания нового хранилища типа «Каталог» необходимо выбрать «Датацентр» → «Хранилище» («Datacenter» → «Storage»), нажать кнопку «Добавить» («Add») и в выпадающем меню выбрать пункт «Каталог» («Directory») (Рис. 58). На Рис. 60 показан диалог создания хранилища local-iso типа «Каталог» для хранения ISO-образов и шаблонов контейнеров, которое будет смонтировано в каталог `/mnt/iso`.

Добавление хранилища «Каталог»

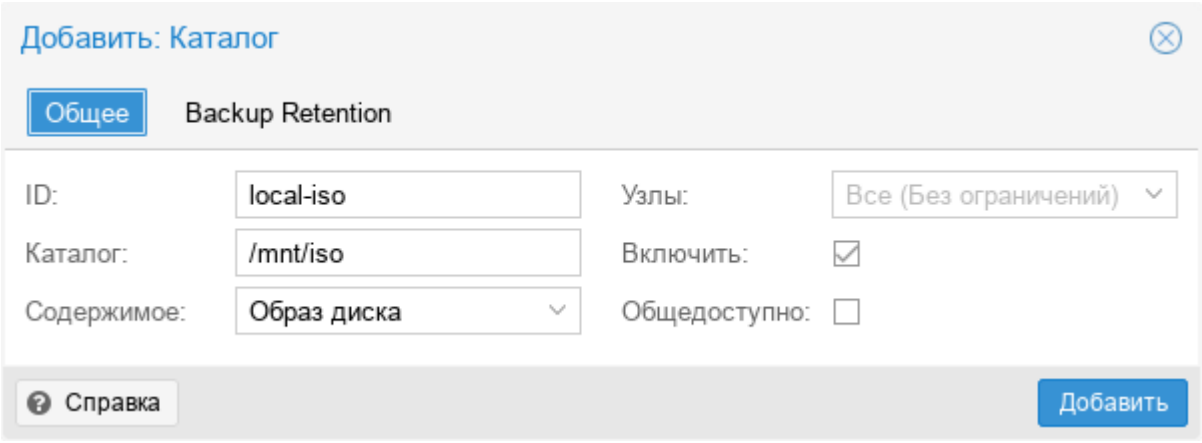

Рис. 60

Данное хранилище поддерживает все общие свойства хранилищ и дополнительно свойство `path` для указания каталога. Это должен быть абсолютный путь к файловой системе.

Пример файла конфигурации (`/etc/pve/storage.cfg`):

```
dir: backup
     path /mnt/backup
     content backup
     prune-backups keep-last=7
     shared 0
```

Данная конфигурация определяет пул хранения резервных копий. Этот пул может использоваться для хранения последних 7 резервных копий на виртуальную машину. Реальный путь к файлам резервных копий – `/mnt/backup/dump/...`

Хранилище «Каталог» использует четко определенную схему именования образов ВМ:

`VM-<VMID>-<NAME>.<FORMAT>`

где:

<VMID> – ID виртуальной машины;

<NAME> – произвольное имя (ascii) без пробелов. По умолчанию используется disk- [N], где [N] заменяется целым числом.

<FORMAT> – определяет формат образа (raw|qcow2|vmdk).

Пример:

```
# ls /var/lib/vz/images/101
vm-101-disk-0.qcow2  vm-101-disk-1.qcow2
```

При создании шаблона ВМ все образы дисков ВМ переименовываются, чтобы указать, что они теперь доступны только для чтения и могут использоваться в качестве базового образа для клонов:

base-<VMID>-<NAME>.<FORMAT>

4.4.3.4 NFS

Хранилище NFS аналогично хранению каталогов и файлов на диске, с дополнительным преимуществом совместного хранения и миграции в реальном времени. Свойства хранилища NFS во многом совпадают с хранилищем типа «Каталог». Структура каталогов и соглашение об именах файлов также одинаковы. Основным преимуществом является то, что можно напрямую настроить свойства сервера NFS, и общий ресурс будет монтироваться автоматически.

Данное хранилище поддерживает все общие свойства хранилищ кроме флага `shared`, который всегда установлен. Кроме того, для настройки NFS используются следующие свойства:

- server – IP-адрес сервера или DNS-имя. Предпочтительнее использовать IP-адрес вместо DNS-имени (чтобы избежать задержек при поиске DNS);
- export – совместный ресурс с сервера NFS (список можно просмотреть, выполнив команду `pvesm scan nfs <server>`);
- path – локальная точка монтирования (по умолчанию `/mnt/pve/<STORAGE_ID>/`);
- options – параметры монтирования NFS.

Пример файла конфигурации (`/etc/pve/storage.cfg`):

```
nfs: nfs-storage
     export /export/storage
     path /mnt/nfs-vol
     server 192.168.0.105
     content images,iso,backup,vztmpl
     options vers=3,nolock,tcp
```

Примечание. Для возможности монтирования NFS хранилища должны быть запущены службы rpcbind и nfslock:

```
# systemctl enable --now rpcbind
# systemctl enable --now nfslock
```

На *Рис. 61* показано присоединение хранилища NFS с именем nfs-storage с удаленного сервера 192.168.0.105.

После ввода IP-адреса удаленного сервера, доступные ресурсы будут автоматически просканированы и будут отображены в выпадающем меню «Export». В данном примере обнаруженная в блоке диалога точка монтирования – `/export/storage`.

Пример добавления хранилища NFS в командной строке с помощью утилиты `pvesm`:

```
# pvesm add nfs nfs-storage --path /mnt/nfs-vol --server 192.168.0.105
--options vers=3,nolock,tcp --export /export/storage --content
images,iso,vztmpl,backup
```

Создание хранилища NFS

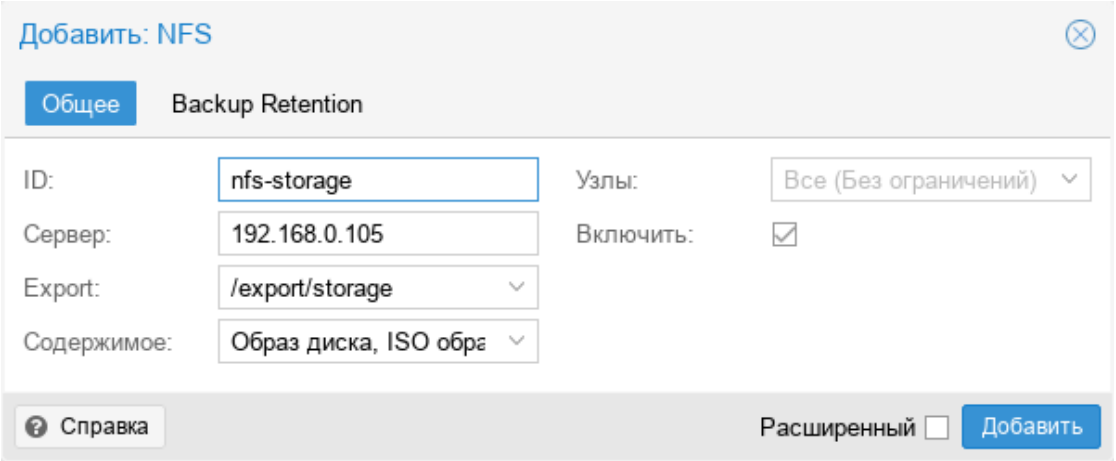

Рис. 61

Получить список совместных ресурсов с сервера NFC:

```
# pvesm nfsscan <server>
```

4.4.3.5 CIFS

Хранилище CIFS расширяет хранилище типа «Каталог», поэтому ручная настройка монтирования CIFS не требуется.

Примечание. Для возможности просмотра общих ресурсов на каждом узле кластера необходимо установить пакет `samba-client`.

Данное хранилище поддерживает все общие свойства хранилищ кроме флага `shared`, который всегда установлен. Кроме того, для настройки CIFS используются следующие свойства:

- server – IP-адрес сервера или DNS-имя. Предпочтительнее использовать IP-адрес вместо DNS-имени (чтобы избежать задержек при поиске DNS);
- share – совместный ресурс с сервера CIFS (список можно просмотреть, выполнив команду `pvesm scan cifs <server>`);
- username – имя пользователя для хранилища CIFS (необязательно, по умолчанию «guest»);
- password – пароль пользователя (необязательно). Пароль будет сохранен в файле, доступном только для чтения root-пользователю (`/etc/pve/priv/<STORAGE_ID>.cred`);
- domain – устанавливает домен пользователя (рабочую группу) для этого хранилища (необязательно);
- smbversion – версия протокола SMB (необязательно, по умолчанию 3);
- path – локальная точка монтирования (по умолчанию `/mnt/pve/<STORAGE_ID>/`).

Пример файла конфигурации (`/etc/pve/storage.cfg`):

```
cifs: newCIFS
      path /mnt/pve/newCIFS
      server 192.168.0.105
      share smb_data
```

Получить список совместных ресурсов с сервера CIFS можно, выполнив команду:

```
# pvesm cifsscan <server> [--username <username>] [--password]
```

Команда добавления общего ресурса в качестве хранилища:

```
# pvesm add cifs <storagename> --server <server> --share <share> [--username <username>] [--password]
```

На *Рис. 62* показано присоединение хранилища CIFS с именем newCIFS с удаленного сервера 192.168.0.105.

Добавление CIFS хранилища

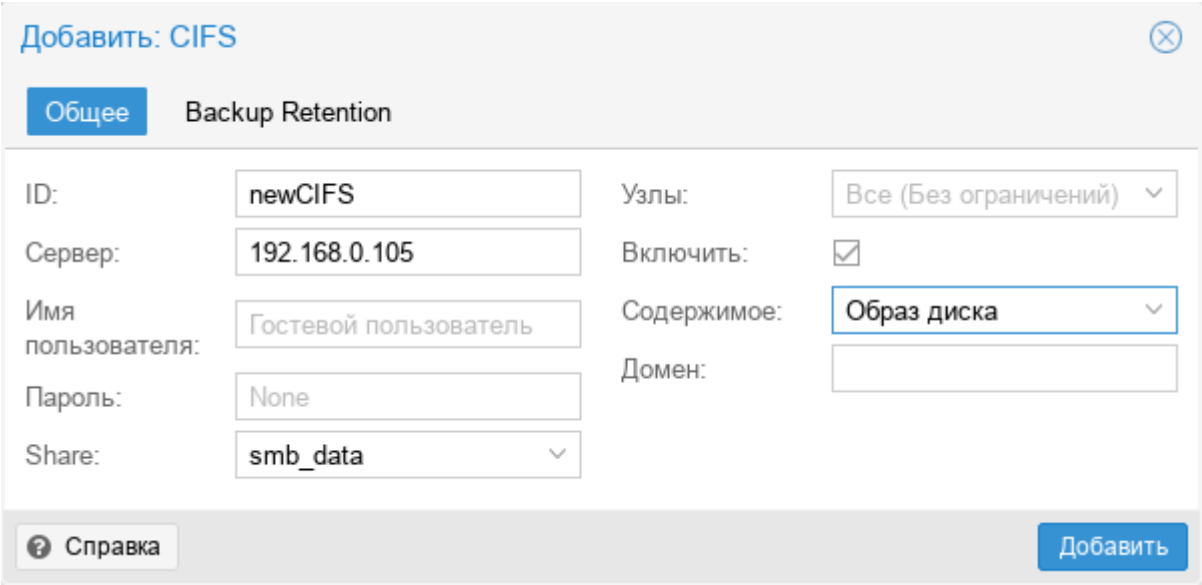

Рис. 62

После ввода IP-адреса удаленного сервера, доступные ресурсы будут автоматически просканированы и будут отображены в выпадающем меню «Share».

Примечание. При создании CIFS хранилища в веб-интерфейсе, PVE предполагает, что удаленный сервер поддерживает CIFS v3. В веб-интерфейсе нет возможности указать версию CIFS, поэтому в случае, если удалённый сервер поддерживает версии CIFS отличные от v3, создать хранилище можно в командной строке, например:

```
# pvesm add cifs newCIFS --server 192.168.0.105 --share smb_data --smbversion 2.1
```

4.4.3.6 GlusterFS

GlusterFS – это масштабируемая сетевая файловая система. Система использует модульную конструкцию, работает на аппаратном оборудовании и может обеспечить высокодоступное корпоративное хранилище при низких затратах. Такая система способна масштабироваться до нескольких петабайт и может обрабатывать тысячи клиентов.

Данное хранилище поддерживает все общие свойства хранилищ, и дополнительно используются следующие свойства:

- server – IP-адрес или DNS-имя сервера GlusterFS;
- server2 – IP-адрес или DNS-имя резервного сервера GlusterFS;
- volume – том GlusterFS;
- transport – транспорт GlusterFS: tcp, unix или rdma.

Пример файла конфигурации (`/etc/pve/storage.cfg`):

```
glusterfs: gluster-01
            server 192.168.0.105
            server2 192.168.0.110
            volume glustervol
            content images,iso
```

На *Рис. 63* показано присоединение хранилища GlusterFS с именем gluster-01 с удаленного сервера 192.168.0.105.

Создание хранилища GlusterFS

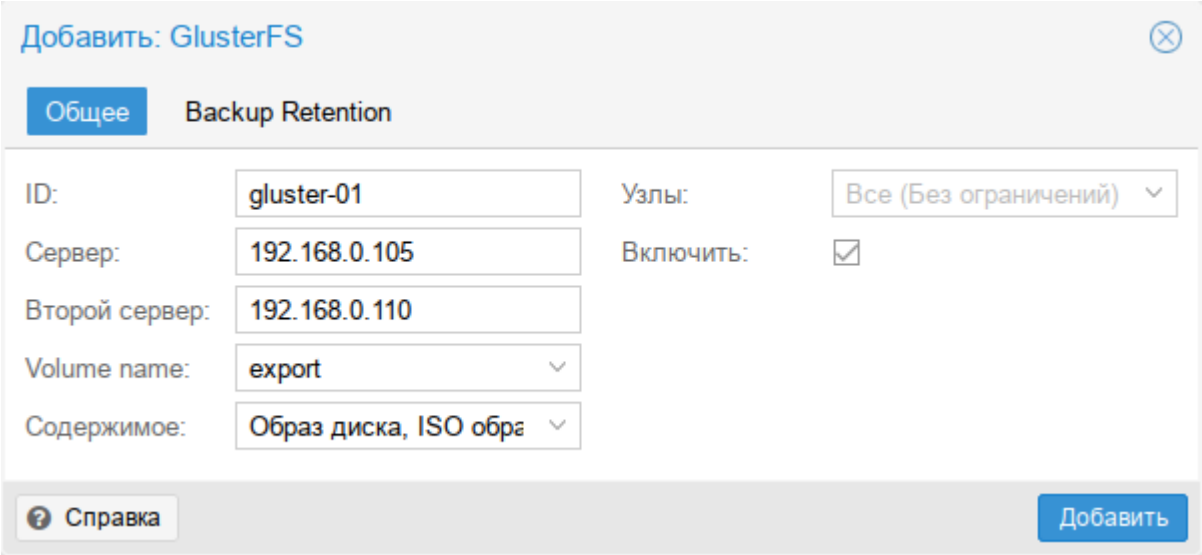

Рис. 63

4.4.3.7 Локальный ZFS

Примечание. Для работы с локальным ZFS хранилищем должен быть установлен модуль ядра `kernel-modules-zfs-std-def`. Включить модуль:

```
# modprobe zfs
```

Чтобы не вводить эту команду после перезагрузки, следует раскомментировать строку: `#zfs` в файле `/etc/modules-load.d/zfs.conf`.

Позволяет получить доступ к локальным пулам ZFS (или файловым системам ZFS внутри таких пулов). Данное хранилище поддерживает все общие свойства хранилищ, кроме того, для настройки ZFS используются следующие свойства:

- pool – пул/файловая система ZFS;
- blocksize – размер блока;
- sparse – использовать тонкую инициализацию ZFS.

Пул ZFS поддерживает следующие типы RAID:

- RAID-0 (Single Disk) – размер такого пула – сумма емкостей всех дисков. RAID0 не добавляет избыточности, поэтому отказ одного диска делает том не пригодным для использования (минимально требуется один диск);
- пул RAID-1 (Mirror) – данные зеркалируются на все диски (минимально требуется два диска);
- пул RAID-10 – сочетание RAID0 и RAID1 (минимально требуется четыре диска);
- пул RAIDZ-1 – вариация RAID-5, одинарная четность (минимально требуется три диска);
- пул RAIDZ-2 – вариация на RAID-5, двойной паритет (минимально требуется четыре диска);
- пул RAIDZ-3 – разновидность RAID-5, тройная четность (минимально требуется пять дисков).

Пример файла конфигурации (`/etc/pve/storage.cfg`):

```
zfspool: vmdata
        pool vmdata
        content images,rootdir
        mountpoint /vmdata
        nodes pve03
```

Возможные типы содержимого: `rootdir` (данные контейнера), `images` (образ виртуального диска в формате raw или subvol).

Используется следующая схема именования образов дисков ВМ:

- vm-<VMID>-<NAME> – образ ВМ;
- base-<VMID>-<NAME> – шаблон образа ВМ (только для чтения);
- subvol-<VMID>-<NAME> – файловая система ZFS для контейнеров.

4.4.3.7.1 Создание локального хранилища ZFS в веб-интерфейсе

Для создания локального хранилища ZFS в веб-интерфейсе, следует выбрать узел, на котором будет создано хранилище, в разделе «Диски» («Disks») выбрать пункт «ZFS» и нажать кнопку «Создать: ZFS» (*Рис. 64*).

В открывшемся окне (*Рис. 65*) следует задать параметры ZFS хранилища: имя хранилища, выбрать диски, уровень RAID и нажать кнопку «Создать».

Статус пула можно просмотреть выбрав его в списке и нажав кнопку «Подробно» (Рис. 66).

Добавление ZFS хранилища

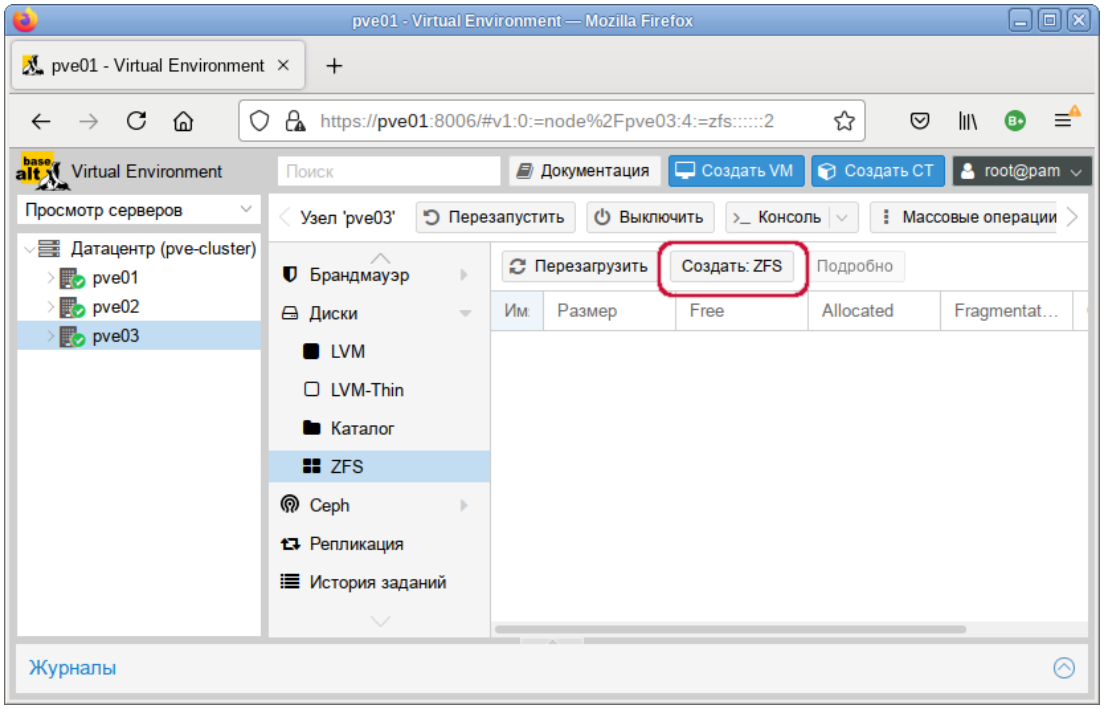

Рис. 64

Параметры ZFS хранилища

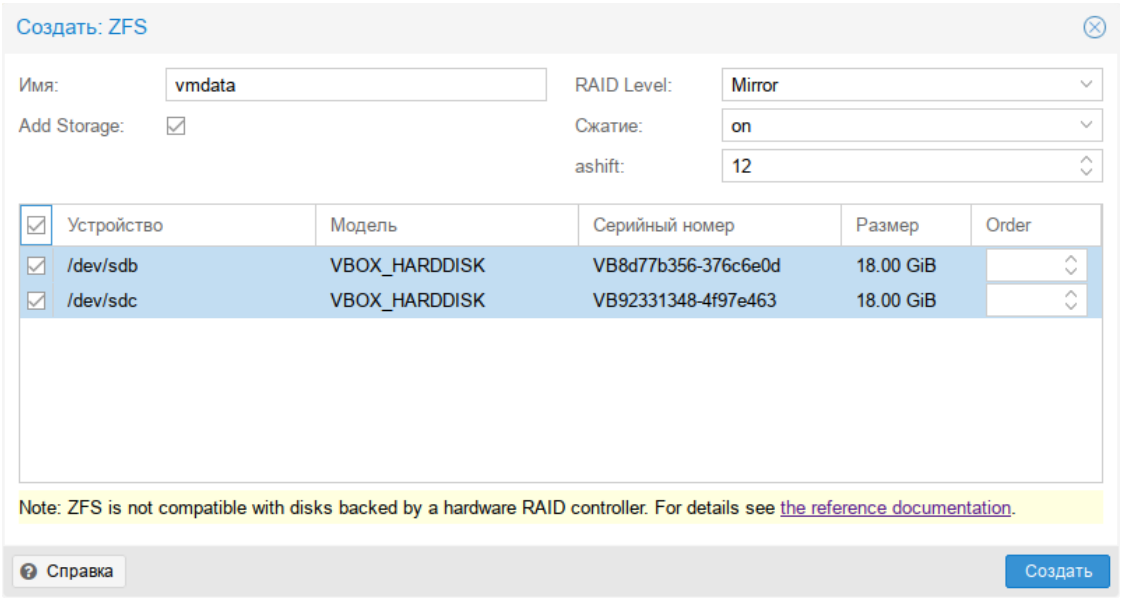

Рис. 65

Локальное ZFS хранилище

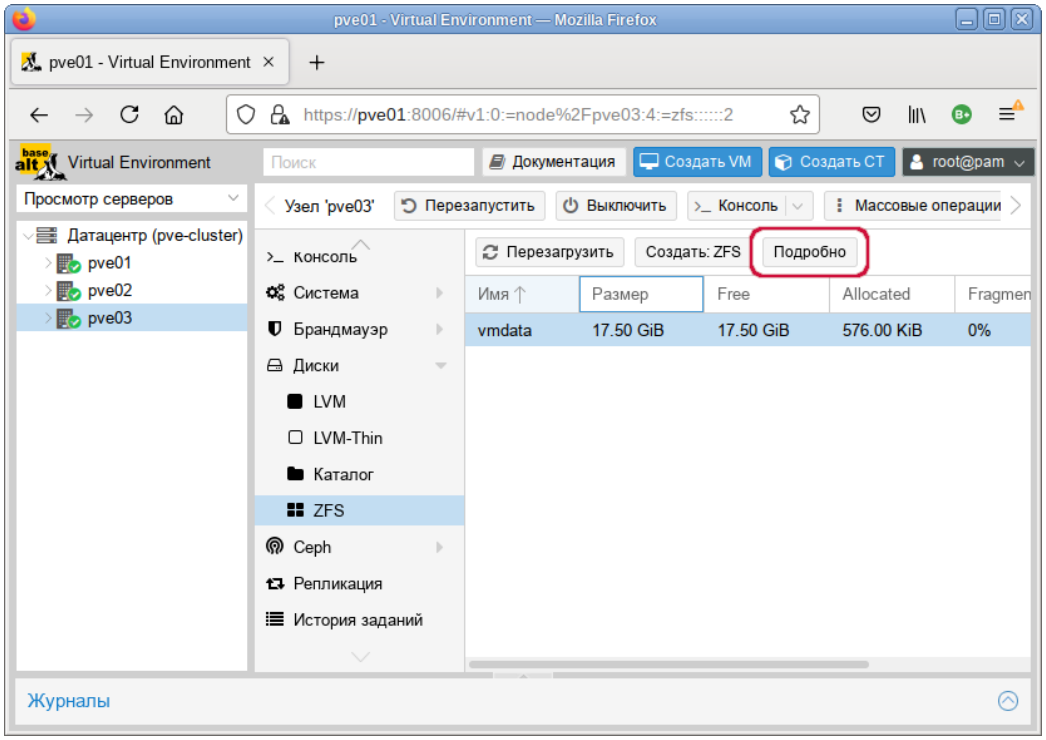

Рис. 66

Для того чтобы внести изменения в настройки ZFS хранилища следует выбрать «Датацентр» → «Хранилище» («Datacenter» → «Storage»), затем нужное хранилище и нажать кнопку «Редактировать» (*Рис. 67*). В открывшемся окне (*Рис. 68*) можно изменить тип содержимого контейнера, включить/отключить хранилище, включить дисковое резервирование.

Выбор хранилища для редактирования

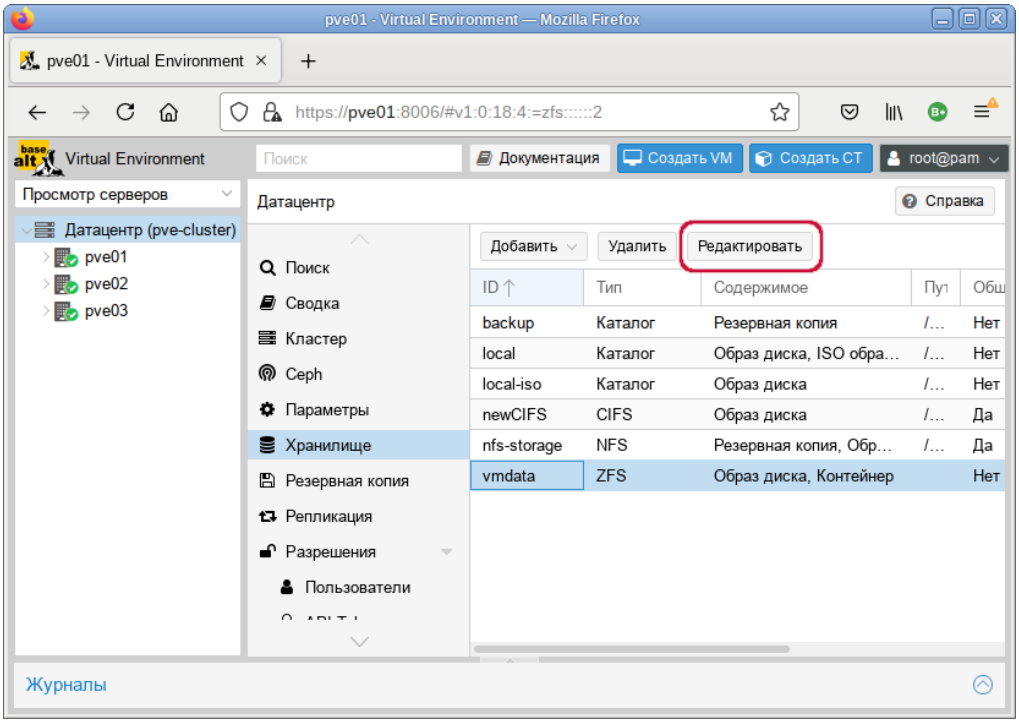

Рис. 67

Редактирование ZFS хранилища

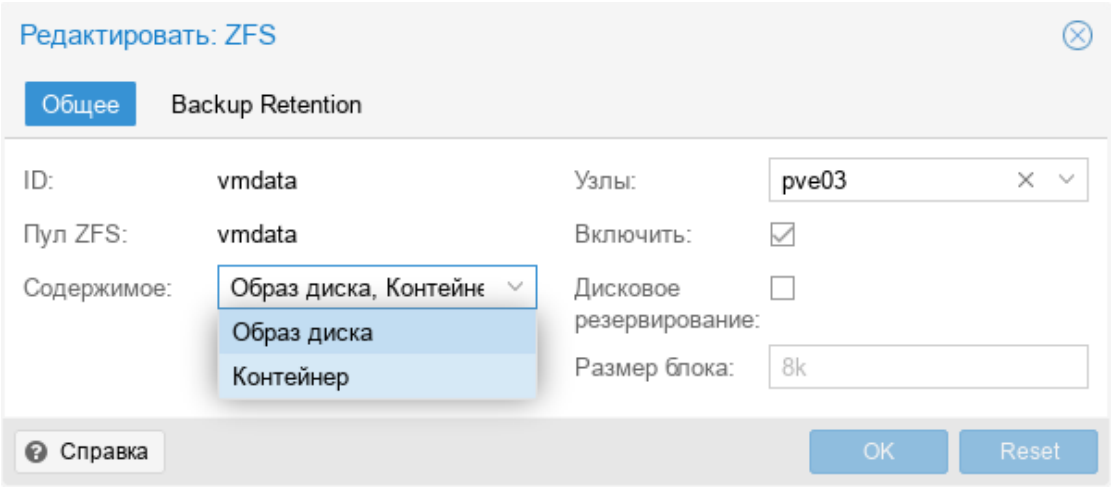

Рис. 68

4.4.3.7.2 Администрирование ZFS

Основными командами для управления ZFS являются zfs и zpool.

Для создания нового пула необходим как минимум один пустой диск.

Создание нового пула RAID-0 (минимум 1 диск):

```
zpool create -f -o ashift=12 <pool> <device1> <device2>
```

Создание нового пула RAID-1 (минимум 2 диска):

```
zpool create -f -o ashift=12 <pool> mirror <device1> <device2>
```

Создание нового пула RAID-10 (минимум 4 диска):

```
zpool create -f -o ashift=12 <pool> mirror <device1> <device2> mirror
<device3> <device4>
```

Создание нового пула RAIDZ-1 (минимум 3 диска):

```
zpool create -f -o ashift=12 <pool> raidz1 <device1> <device2>
<device3>
```

Создание нового пула RAIDZ-2 (минимум 4 диска):

```
zpool create -f -o ashift=12 <pool> raidz2 <device1> <device2>
<device3> <device4>
```

Смена неисправного устройства:

```
zpool replace -f <pool> <old device> <new device>
```

Включить сжатие:

```
zfs set compression=on <pool>
```

Получить список доступных ZFS файловых систем:

```
# pvesm zfsscan
```

Пример создания RAID1(mirror) с помощью zfs:

```
# zpool create -f vmdata mirror sdb sdc
```

Просмотреть созданные в системе пулы:

```
# zpool list
NAME     SIZE  ALLOC   FREE  CKPOINT  EXPANDSZ   FRAG   CAP  DEDUP  HEALTH  ALTROOT
vmdata  17,5G   492K  17,5G        -         -     0%    0%  1.00x  ONLINE  -
```

Просмотреть статус пула:

```
# zpool status
  pool: vmdata
 state: ONLINE
  scan: none requested
config:

      NAME        STATE     READ WRITE CKSUM
      vmdata      ONLINE       0     0     0
        mirror-0  ONLINE       0     0     0
          sdb     ONLINE       0     0     0
          sdc     ONLINE       0     0     0

errors: No known data errors
```

4.4.3.8 LVM

LVM (Logical Volume Management) это система управления дисковым пространством. Позволяет логически объединить несколько дисковых пространств (физические тома) в одно, и уже из этого пространства (дисковой группы или группы томов – VG), можно выделять разделы (логические тома –LV), доступные для работы.

Использование LVM групп обеспечивает лучшую управляемость. Логические тома можно легко создавать/удалять/перемещать между физическими устройствами хранения. Если база хранения для группы LVM доступна на всех PVE узлах (например, ISCSI LUN) или репликах (например, DRBD), то все узлы имеют доступ к образам ВМ, и возможна live-миграция.

Данное хранилище поддерживает все общие свойства хранилищ, кроме того, для настройки LVM используются следующие свойства:

- vgname – имя группы томов LVM (должно указывать на существующую группу томов);
- base – базовый объем;
- saferemove – обнуление данных при удалении LV. При удалении тома это гарантирует, что все данные будут удалены;
- saferemove_throughput – очистка пропускной способности (значение параметра `cstream -t`).

Пример файла конфигурации (`/etc/pve/storage.cfg`):

```
lvm: vg
```

```
vgname vg
content rootdir,images
nodes pve03
shared 0
```

Возможные типы содержимого: `rootdir` (данные контейнера), `images` (образ виртуального диска в формате raw).

4.4.3.8.1 Создание локального хранилища LVM в веб-интерфейсе

Примечание. Для создания локального LVM хранилища в веб-интерфейсе необходимо чтобы в системе имелся хотя бы один пустой жесткий диск.

Для создания локального LVM хранилища в веб-интерфейсе, следует выбрать узел, на котором будет создано хранилище, в разделе «Диски» («Disks») выбрать пункт «LVM» и нажать кнопку «Создать: Volume Group» (*Рис. 69*). В открывшемся окне (*Рис. 70*) следует выбрать диск, задать имя группы томов, отметить пункт «Add Storage» (если этот пункт не отмечен будет создана только группа томов) и нажать кнопку «Создать».

Пункт «LVM» в разделе «Диски»

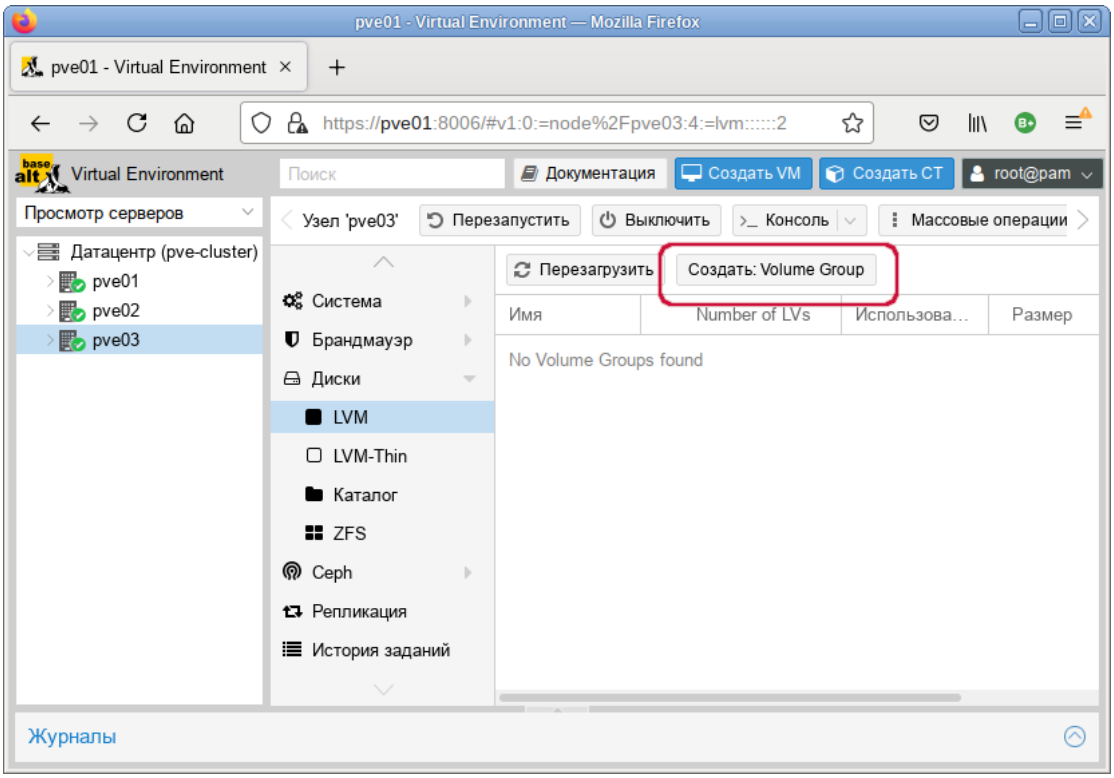

Рис. 69

Создание группы томов

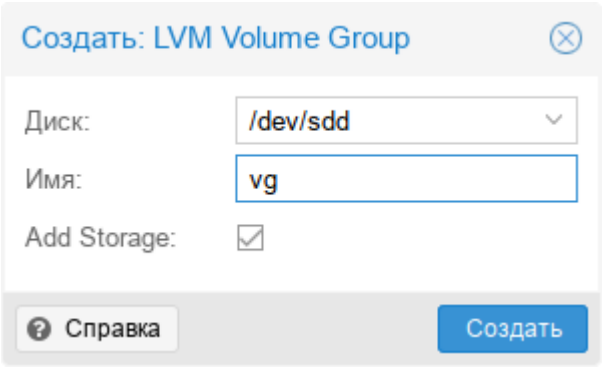

Рис. 70

Для того чтобы внести изменения в настройки LVM хранилища следует выбрать «Дата-центр» → «Хранилище» («Datacenter» → «Storage»), затем нужное хранилище и нажать кнопку «Редактировать». В открывшемся окне (*Рис. 71*) можно изменить тип содержимого контейнера, включить/отключить хранилище.

Редактирование LVM хранилища

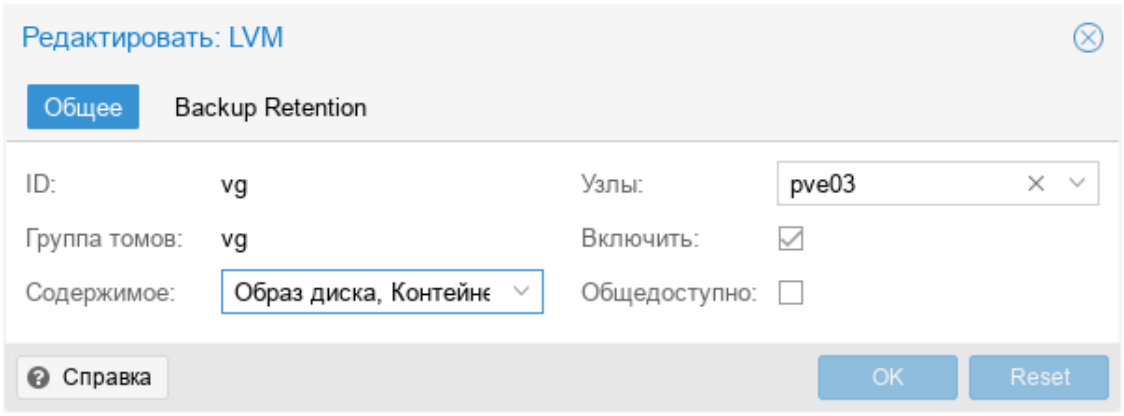

Рис. 71

4.4.3.8.2 Создание хранилища LVM в командной строке

Пример создания LVM хранилища на пустом диске /dev/sdd:

1) Создать физический том (PV):

```
# pvcreate /dev/sdd
  Physical volume "/dev/sdd" successfully created.
```

2) Создать группу томов (VG) с именем vg:

```
# vgcreate vg /dev/sdd
  Volume group "vg" successfully created
```

3) Создать логические тома (LV):

```
# lvcreate -n lv01 -L 10G vg
  Logical volume "lv01" created.
# lvcreate -n lv02 -L 5G vg
```

```
  Logical volume "lv02" created.
```

4) Показать информацию о физических томах:

```
# pvs
PV         VG         Fmt  Attr PSize   PFree
  /dev/sdd   vg         lvm2 a--  <18,00g <3,00g
```

5) Показать информацию о группах томов:

```
# vgs
  VG         #PV #LV #SN Attr   VSize   VFree
  vg           1   2   0 wz--n- <18,00g <3,00g
```

6) Показать информацию о логических томах:

```
# lvs
  LV        VG        Attr       LSize  Pool Origin Data%  Meta%  Move
Log Cpy%Sync Convert
  lv01      vg        -wi-a----- 10,00g
  lv02      vg        -wi-a-----  5,00g
```

7) Получить список доступных PVE групп томов:

```
# pvesm lvmscan
vg
```

8) Создать LVM хранилище с именем myspace:

```
# pvesm add lvm myspace --vgname vg --nodes pve03
```

4.4.3.9 LVM-thin

LVM-thin (thin provision) – это возможность использовать какое-либо внешнее блочное устройство в режиме только для чтения как основу для создания новых логических томов LVM. Такие разделы при создании уже будут выглядеть так, будто они заполнены данными исходного блочного устройства. Операции с томами изменяются налету таким образом, что чтение данных выполняется с исходного блочного устройства (или с тома, если данные уже отличаются), а запись – на том.

Такая возможность может быть полезна, например, при создании множества однотипных ВМ или для решения других аналогичных задач, т.е. задач, где нужно получить несколько изменяемых копий одних и тех же исходных данных.

Данное хранилище поддерживает все общие свойства хранилищ, кроме того, для настройки LVM-thin используются следующие свойства:

- vgname – имя группы томов LVM (должно указывать на существующую группу томов);
- thinpool – название тонкого пула LVM.

Пример файла конфигурации (`/etc/pve/storage.cfg`):

```
lvmthin: vmstore
        thinpool vmstore
        vgname vmstore
        content rootdir,images
        nodes pve03
```

Возможные типы содержимого: `rootdir` (данные контейнера), `images` (образ виртуального диска в формате raw).

LVM thin является блочным хранилищем, но полностью поддерживает моментальные снимки и клоны. Новые тома автоматически инициализируются с нуля.

Тонкие пулы LVM не могут совместно использоваться несколькими узлами, поэтому их можно использовать только в качестве локального хранилища.

4.4.3.9.1 Создание локального хранилища LVM-Thin в веб-интерфейсе

Примечание. Для создания локального LVM-Thin хранилища в веб-интерфейсе необходимо чтобы в системе имелся хотя бы один пустой жесткий диск.

Для создания локального LVM-Thin хранилища в веб-интерфейсе, следует выбрать узел, на котором будет создано хранилище, в разделе «Диски» («Disks») выбрать пункт «LVM-Thin» и нажать кнопку «Создать: Thinpool» (*Рис. 72*). В открывшемся окне (*Рис. 73*) следует выбрать диск, задать имя группы томов, отметить пункт «Add Storage» (если этот пункт не отмечен будет создана только группа томов) и нажать кнопку «Создать».

Пункт «LVM-Thin» в разделе «Диски»

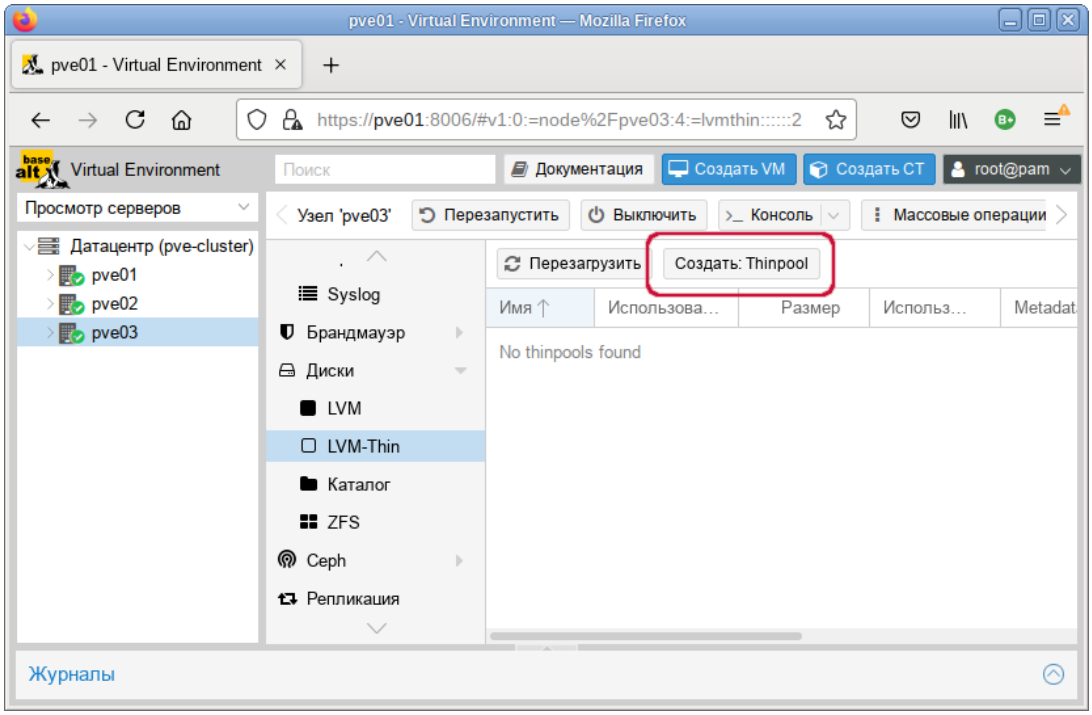

Рис. 72

Создание LVM-Thin хранилища

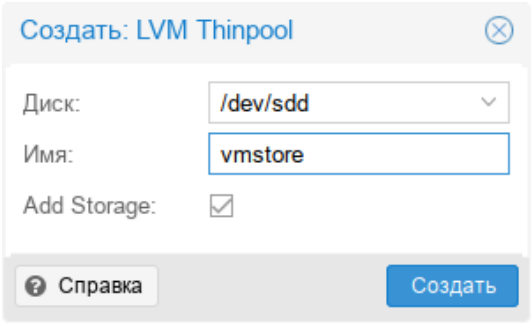

Рис. 73

Для того чтобы внести изменения в настройки LVM-Thin хранилища следует выбрать «Датацентр» → «Хранилище» («Datacenter» → «Storage»), затем нужное хранилище и нажать кнопку «Редактировать». В открывшемся окне (*Рис. 74*) можно изменить тип содержимого контейнера, включить/отключить хранилище.

Редактирование LVM-Thin хранилища

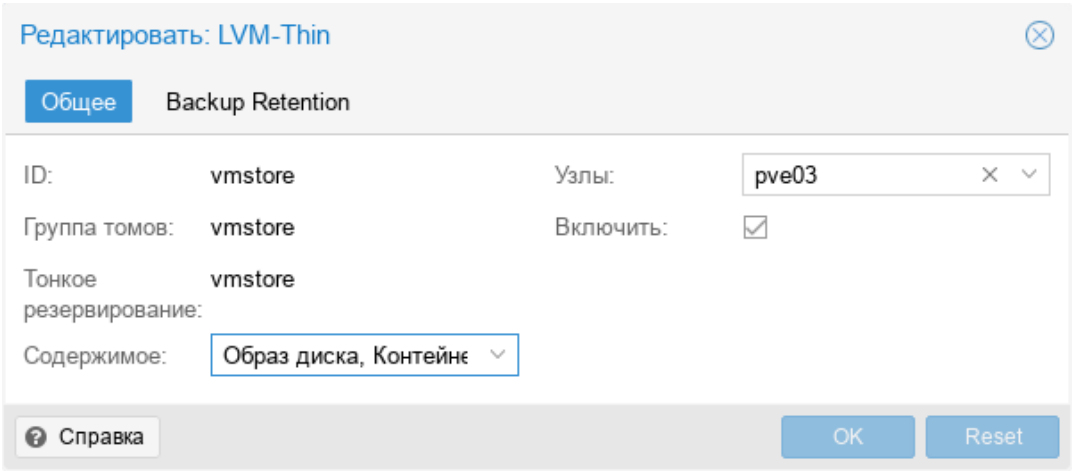

Рис. 74

4.4.3.9.2 Создание хранилища LVM-Thin в командной строке

Для создания и управления пулами LVM-Thin можно использовать инструменты командной строки.

Пул LVM-Thin должен быть создан поверх группы томов.

Команда создания нового тонкого пула LVM (размер 80 ГБ) с именем vmstore (предполагается, что группа томов LVM с именем vg уже существует):

```
# lvcreate -L 80G -T -n vmstore vg
```

Получить список доступных LVM-thin пулов в группе томов vg:

```
# pvesm lvmthinscan vg
vmstore
```

Команда создания LVM-Thin хранилища с именем vmstore на узле pve03:

```
# pvesm add lvmthin vmstore --thinpool vmstore --vgname vg --nodes pve03
```

4.4.3.10 iSCSI

iSCSI (Internet Small Computer System Interface) – широко применяемая технология, используемая для подключения к серверам хранения.

Данное хранилище поддерживает все общие свойства хранилищ, и дополнительно используются следующие свойства:

- portal – IP-адрес или DNS-имя сервера iSCSI;
- target – iSCSI target.

Пример файла конфигурации (`/etc/pve/storage.cfg`):

```
iscsi: test1-iSCSI
     portal 192.168.0.105
     target iqn.2021-7.local.omv:test
     content images
```

Возможные типы содержимого: `images` (образ виртуального диска в формате raw).

iSCSI является типом хранилища блочного уровня и не предоставляет интерфейса управления. Поэтому обычно лучше экспортировать одно большое LUN и установить LVM поверх этого LUN.

Примечание. Для работы с устройством, подключенным по интерфейсу iSCSI, на всех узлах необходимо выполнить команду (должен быть установлен пакет open-iscsi):

```
# systemctl enable --now iscsid
```

На *Рис. 75* показано добавление адресата iSCSI с именем test1-iSCSI, который настроен на удаленном узле 192.168.0.105. Параметр «Использовать LUN напрямую» («Use LUNs directly») – разрешение/запрет прямого применения LUN (параметр должен быть установлен на запрет, разрешение может привести к потере данных).

*Добавление хранилища «*iSCSI*»*

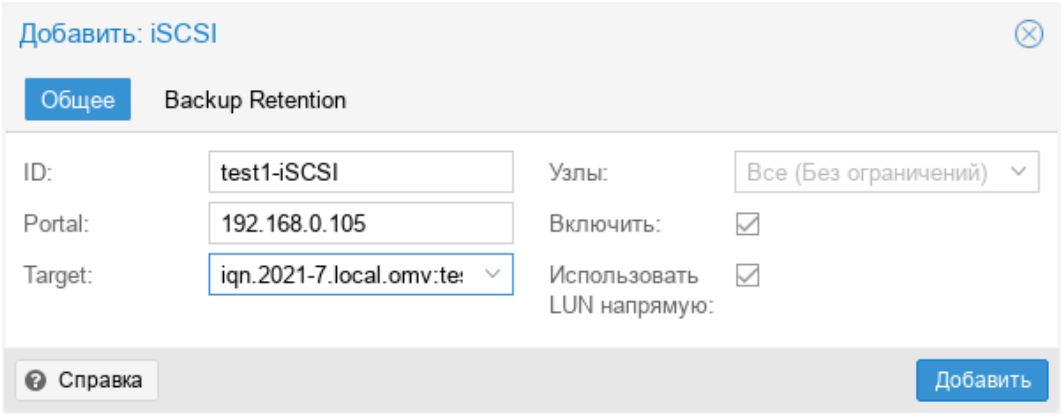

Рис. 75

Посмотреть доступные для подключения iSCSI target-ы:

```
pvesm scan iscsi <HOST[:PORT]>
```

4.4.3.11 Ceph RBD

Хранилище RBD (Rados Block Device) предоставляется распределенной системой хранения Ceph. По своей архитектуре Ceph является распределенной системой хранения. Хранилище RBD может содержать только форматы образов `.raw`.

Данное хранилище поддерживает все общие свойства хранилищ, и дополнительно используются следующие свойства:

- monhost – список IP-адресов демона монитора (только если Ceph не работает на кластере PVE);
- pool – название пула Ceph (rbd);
- username – идентификатор пользователя Ceph (только если Ceph не работает на кластере PVE);
- subdir – подкаталог CephFS для монтирования (по умолчанию /);
- fuse – доступ к CephFS через FUSE (по умолчанию 0).

Пример файла конфигурации (`/etc/pve/storage.cfg`):

```
rbd: new
     content images
     krbd 0
     monhost 192.168.0.105
     pool rbd
     username admin
```

Возможные типы содержимого: `rootdir` (данные контейнера), `images` (образ виртуального диска в формате raw).

На *Рис. 76* показано добавление хранилища RBD.

*Добавление хранилища «*RBD*»*

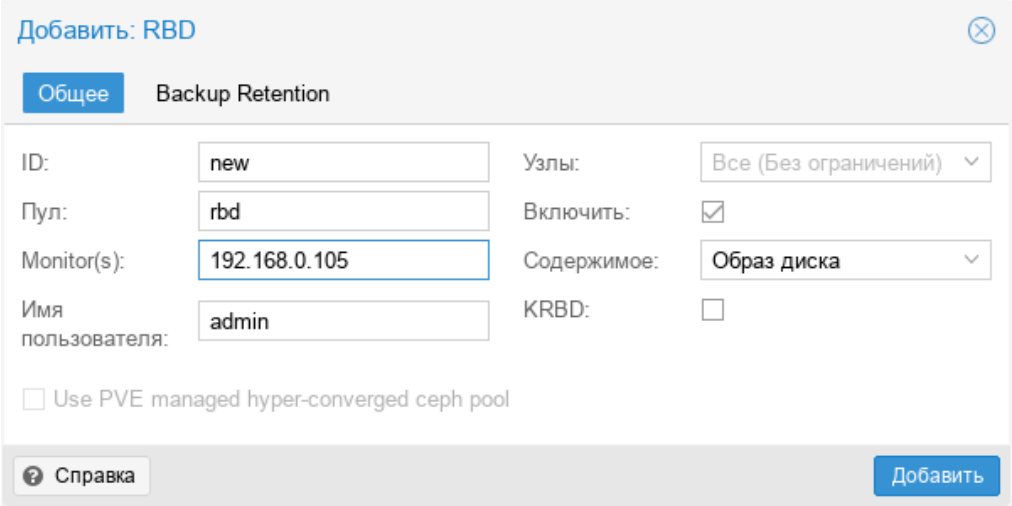

Рис. 76

4.4.3.12 CephFS

CephFS реализует POSIX-совместимую файловую систему, использующую кластер хранения Ceph для хранения своих данных. Поскольку CephFS основывается на Ceph, он разделяет большинство свойств, включая избыточность, масштабируемость, самовосстановление и высокую доступность.

Данное хранилище поддерживает все общие свойства хранилищ, и дополнительно используются следующие свойства:

– monhost – список IP-адресов демона монитора (только если Ceph не работает на кластере PVE);
– path – локальная точка монтирования (по умолчанию используется `/mnt/pve/<STORAGE_ID>/`);
– username – идентификатор пользователя (только если Ceph не работает на кластере PVE);
– krbd – доступ к блочным устройствам через модуль ядра krbd.

Пример файла конфигурации (`/etc/pve/storage.cfg`):

```
cephfs: cephfs-external
     content backup,images
     monhost 192.168.0.105
     username admin
```

Возможные типы содержимого: vztmpl (шаблон контейнера), iso (ISO-образ), backup (резервная копия), snippets (сниппеты).

На *Рис. 77* показано добавление хранилища CephFS.

*Добавление хранилища «*CephFS*»*

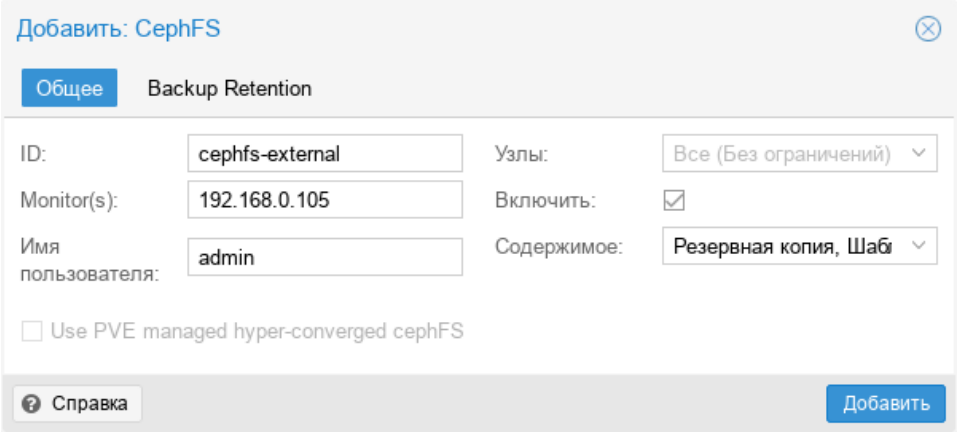

Рис. 77

4.5 Управление ISO образами и шаблонами LXC

Для выгрузки ISO образов и шаблонов в хранилище PVE следует выполнить следующие шаги:

1) выбрать хранилище;

2) перейти на вкладку «ISO Images» или «CT Templates» для отображения сохраненных в хранилище ISO образов (Рис. 78) или шаблонов LXC. Если загруженных файлов не существует, будет доступна только кнопка «Загрузить» («Upload»);

3) нажать кнопку «Загрузить» («Upload») для открытия диалогового блока (Рис. 79);

4) нажать кнопку «Выбрать файл…» («Select File...») для выбора файла для выгрузки с локального компьютера;

5) нажать кнопку «Загрузить» («Upload») для старта выгрузки файла в хранилище.

Локальное хранилище

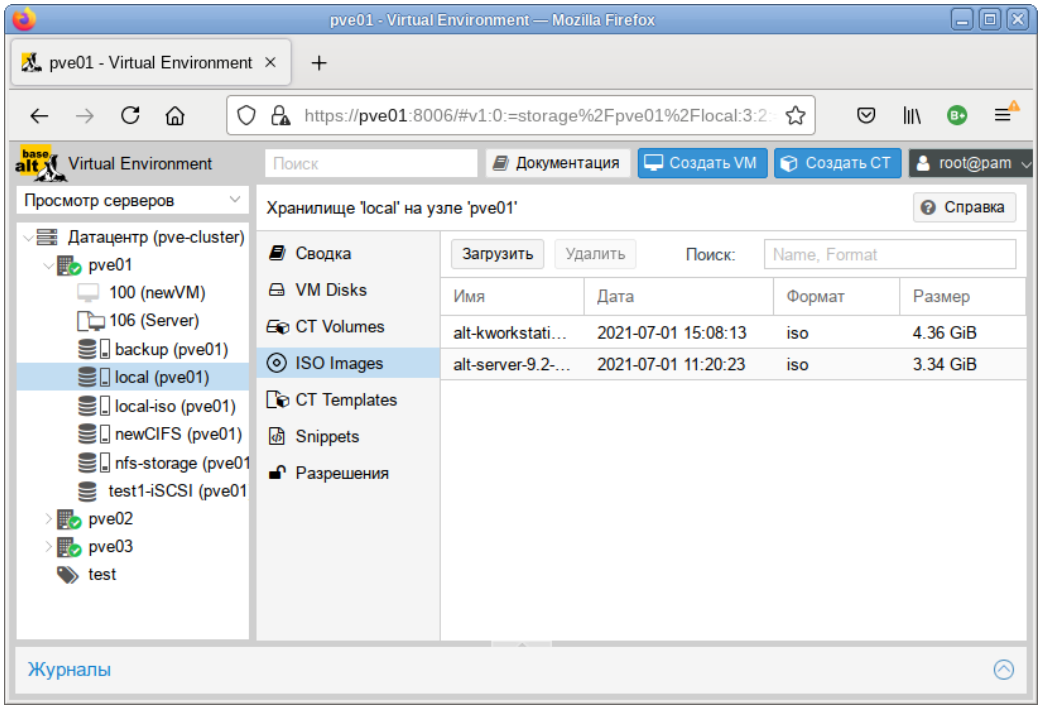

Рис. 78

Выбор образа

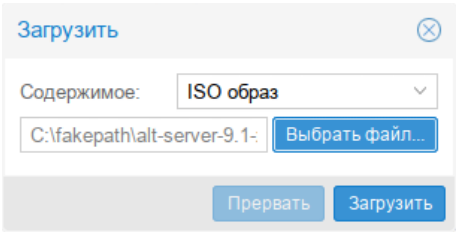

Рис. 79

Для удаления ISO образа или шаблона LXC следует выбрать файл из списка в хранилище (Рис. 78) и нажать кнопку «Удалить» («Remove»).

ISO образы и шаблоны LXC могут также копироваться через интерфейс командной строки. Если используются только локальные хранилища, эти образы ISO и шаблоны необходимо выгрузить на все узлы в кластере. При общем хранилище можно хранить все образы в одном месте, таким образом, сохраняя пространство локальных хранилищ.

В таблице 6 показаны каталоги для локального хранилища. В таблице 7 показаны каталоги для всех других хранилищ.

Т а б л и ц а 6 – Каталоги локального хранилища

Каталог	Тип шаблона
`/var/lib/vz/template/iso`	ISO образы
`/var/lib/vz/template/cache`	Шаблоны контейнеров LXC

Т а б л и ц а 7 – Каталоги общих хранилищ

Каталог	Тип шаблона
/mnt/pve/<storage_name>/template/iso	ISO образы
/mnt/pve/<storage_name>/template/cache	Шаблоны контейнеров LXC

4.6 Виртуальные машины на базе KVM

4.6.1 Создание виртуальной машины на базе KVM

Прежде чем создать в интерфейсе PVE виртуальную машину (ВМ), необходимо определиться со следующими моментами:

– откуда будет загружен инсталлятор ОС, которая будет установлена внутрь ВМ;
– на каком физическом узле будет выполняться процесс гипервизора kvm;
– на каком хранилище данных будут располагаться образы дисков ВМ.

Все остальные параметры ВМ относятся к конфигурации виртуального компьютера и могут быть определены по ходу процесса создания ВМ (PVE пытается выбрать разумные значения по умолчанию для ВМ).

Чтобы установить ОС на ВМ, расположенную на этом узле, нужно обеспечить возможность загрузки инсталлятора на этой ВМ. Для этого необходимо загрузить ISO-образ инсталлятора в хранилище данных выбранного физического узла или общее хранилище. Это удобно делать через веб-интерфейс (Рис. 78).

Для создания ВМ необходимо нажать кнопку «Создать VM» («Create VM»), расположенную в правом верхнем углу веб-интерфейса PVE (Рис. 80).

Кнопка «Создать VM»

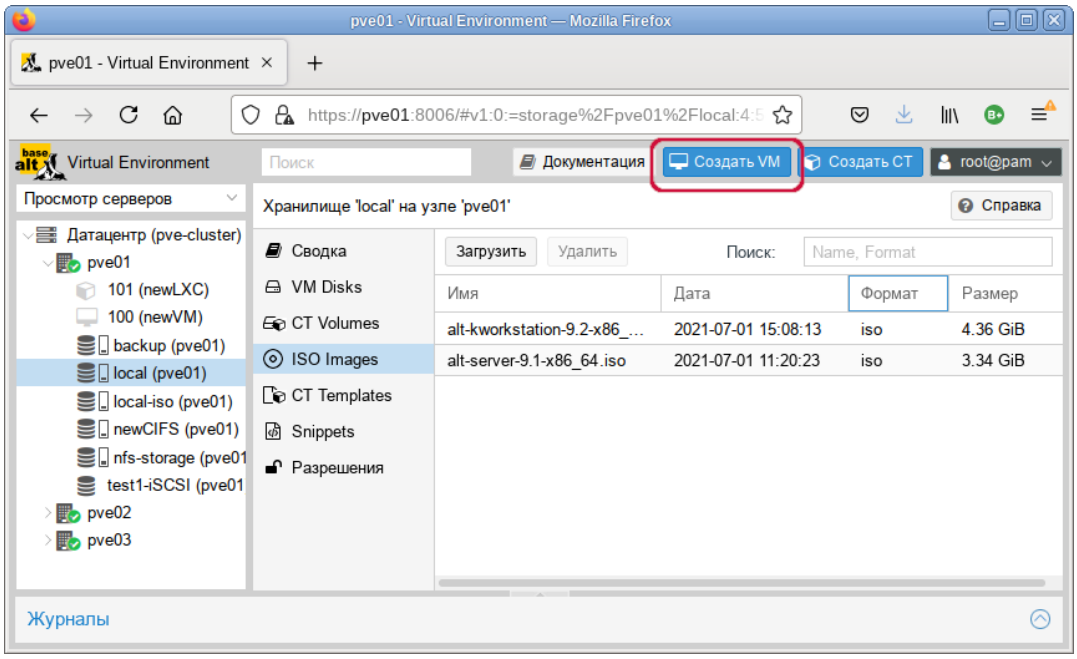

Рис. 80

Процесс создания ВМ оформлен в виде «мастера», привычного для пользователей систем управления ВМ.

На вкладке «Общее» необходимо указать (Рис. 81):

– «Узел» («Node») – физический сервер, на котором будет работать ВМ;
– «VM ID» – идентификатор ВМ в численном выражении. Одно и то же значение идентификатора не может использоваться более чем для одной машины. Поле идентификатора ВМ заполняется автоматически инкрементально: первая созданная ВМ, по умолчанию будет иметь VM ID со значением 100, следующая 101 и так далее;
– «Имя» («Name») – текстовая строка названия ВМ;
– «Пул ресурсов» («Resource pool») – имя пула данной ВМ, к которому она будет относиться. Данное значение не обязательное. Чтобы иметь возможность выбора, этот пул должен быть предварительно создан.

На вкладке «ОС» необходимо указать источник установки ОС, выбрать тип операционной системы для данной ВМ (Рис. 82).

Вкладка «Общее»

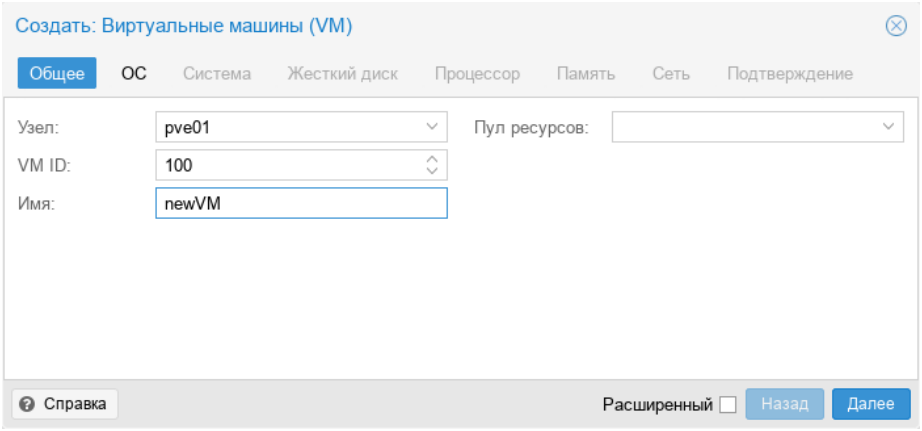

Рис. 81

Вкладка «ОС»

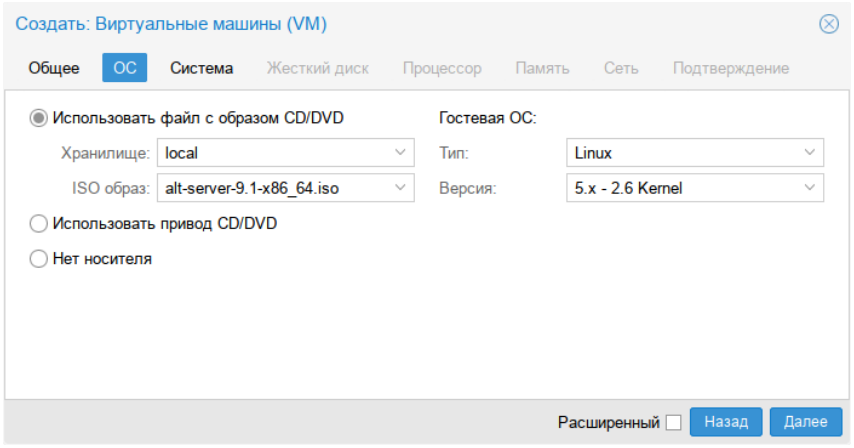

Рис. 82

Возможны следующие варианты источник установки ОС:

– «Использовать файл с образом CD/DVD» («Use CD/DVD disc image file») – выбирает уже выгруженный в хранилище образ ISO (Рис. 83);
– «Использовать привод CD/DVD» («Use physical CD/DVD Drive») – использовать физический диск хоста PVE;
– «Нет носителя» («Do not use any media») – не использовать ISO образ или физический носитель.

Выбор типа гостевой ОС при создании ВМ позволяет PVE оптимизировать некоторые параметры низкого уровня.

Выбор ISO образа

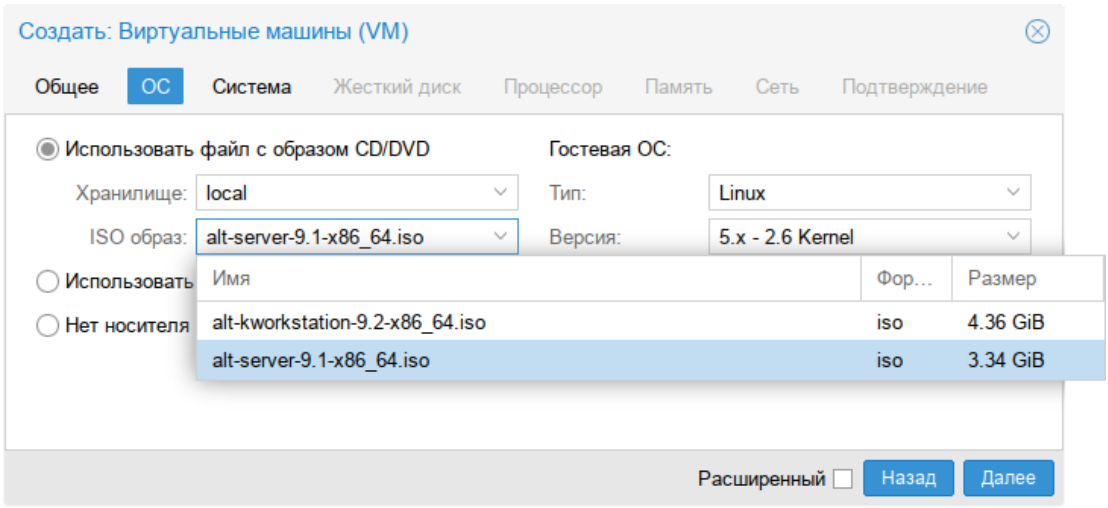

Рис. 83

На следующем этапе (вкладка «Система») можно выбрать видеокарту, контроллер SCSI, указать использовать ли Агент QEMU (Рис. 84).

Вкладка «Система»

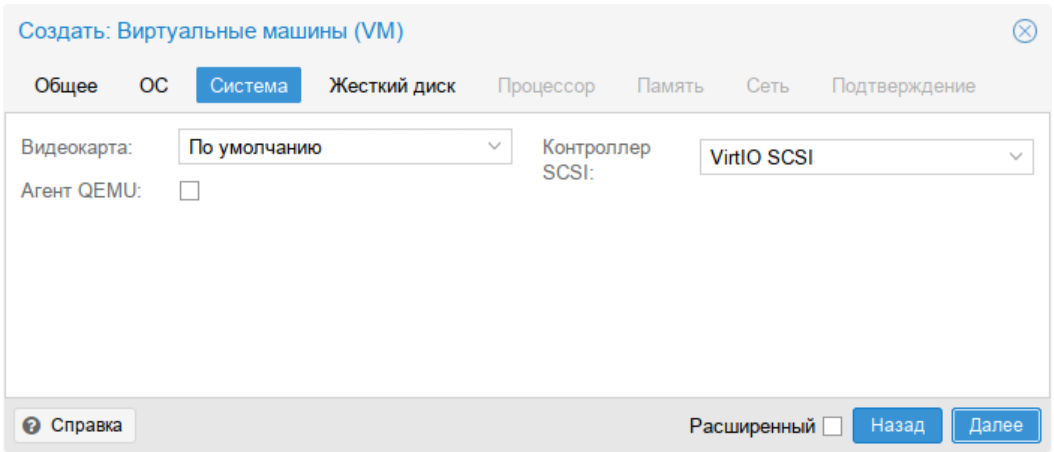

Рис. 84

Вкладка «Жесткий диск» («Hard Disk») содержит следующие настройки (Рис. *85*):

– «Шина/Устройство» («Bus/Device») – тип устройства виртуального диска. Допустимые значения: «IDE», «SATA», «VirtIO Block» и «SCSI» (по умолчанию). Можно также выбрать номер порта;
– «Хранилище» («Storage») – выбор хранилища для размещения данного виртуального диска;
– «Размер диска» («Disk size») (GiB) – размер виртуального диска в гигабайтах;
– «Формат» («Format») – выбирается формат образа виртуального диска. Доступные значения: «Несжатый образ диска (raw)», «Формат образа QEMU (qcow2)» и «Формат образа Vmware (vmdk)». Формат образа RAW является полностью выделяемым (thick-provisioned), т.е. выделяется сразу весь объем образа. QEMU и VMDK поддерживают динамичное выделение пространства (thin-provisioned), т.е. объем растет по мере сохранения данных на виртуальный диск;
– «Кэш» («Cache») – выбор метода кэширования виртуальной машины. По умолчанию выбирается работа без кэширования. Доступные значения: «Direct sync», «Write through», «Write back» и «Writeback (не безопасно)» и «Нет кэша»;
– «Отклонить» («Discard») – делает доступным TRIM, что очищает неиспользуемое пространство образа виртуального диска.

Вкладка «Жесткий диск»

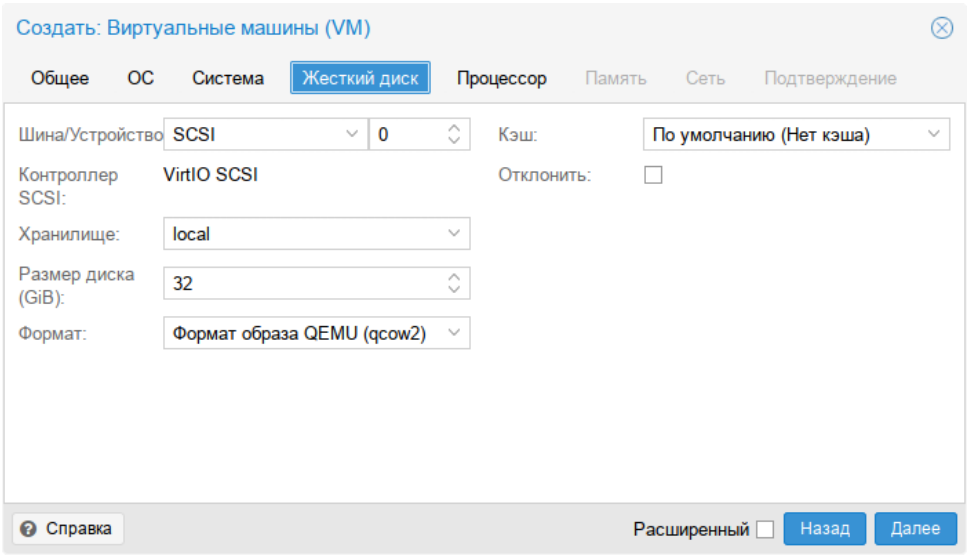

Рис. 85

Максимальный объем виртуального диска – 128ТБ.

На следующем этапе настраивается процессор (CPU) (Рис. 86):

– «Сокеты» («Sockets») – число сокетов ЦПУ для данной ВМ;
– «Ядра» («Cores») – число ядер для данной ВМ;
– «Тип» («Type») – тип процессора.

Вкладка «Процессор»

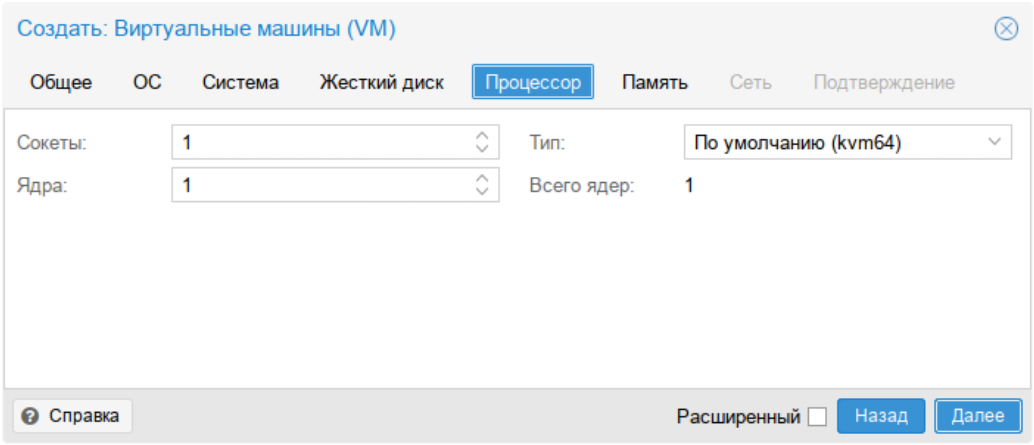

Рис. 86

Максимальное количество виртуальных процессоров в ВМ – 512.

На вкладке «Память» («Memory») (Рис. *87*) необходимо указать объем оперативной памяти выделяемой ВМ.

Вкладка «Память»

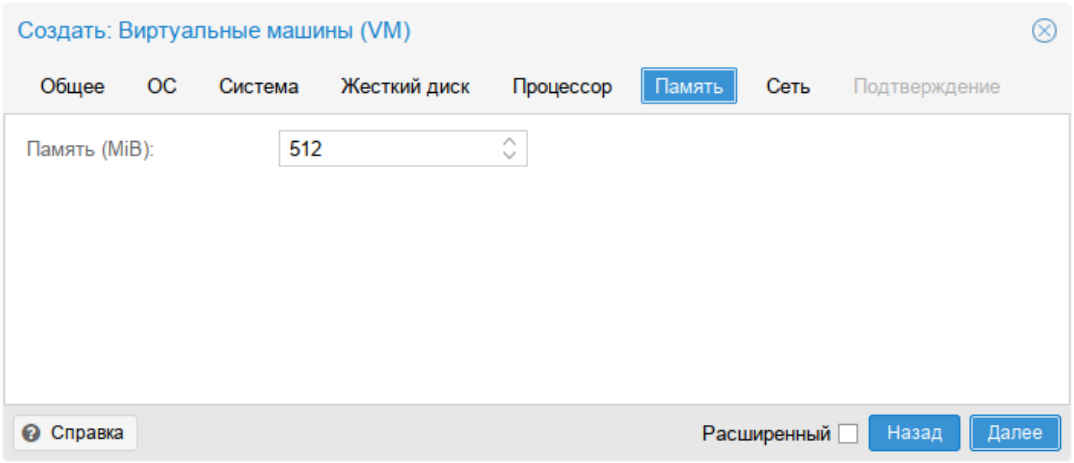

Рис. 87

Максимальное количество памяти, выделяемое ВМ – 2ТБ.

Вкладка «Сеть» («Network») содержит следующие настройки (Рис. 88):

– «Нет сетевого устройства» («No network device») – выбор данного параметра пропускает шаг настройки сетевой среды;
– «Сетевой мост» («Bridged mode») – установка сетевого интерфейса в режиме моста. Это предпочтительный параметр для сетевой среды ВМ. В этом режиме возможно создание множества мостов с виртуальными сетями для создания изолированных сетей в одной и той же платформе, поскольку ВМ не имеют прямого доступа к реальной локальной сетевой среде;
– «Брандмауэр» («Firewall») – разрешает использование для ВМ встроенных межсетевых экранов;

– «Модель» («Model») – тип драйвера сетевого устройства. Для максимальной сетевой производительности ВМ следует выбрать пункт «VirtIO (паравиртуализированно)»;
– «Адрес MAC» («MAC address») – по умолчанию PVE автоматически создает уникальный MAC адрес для сетевого интерфейса. Если есть такая необходимость, можно ввести пользовательский MAC адрес вручную.

Последняя вкладка «Подтверждение» («Confirm») отобразит все введенные или выбранные значения для данной ВМ (Рис. 89). Для создания ВМ следует нажать кнопку «Далее». Если необходимо внести изменения в параметры ВМ, можно перейти по вкладкам назад. Если отметить пункт «Start after created» ВМ будет запущена после создания.

Вкладка «Сеть»

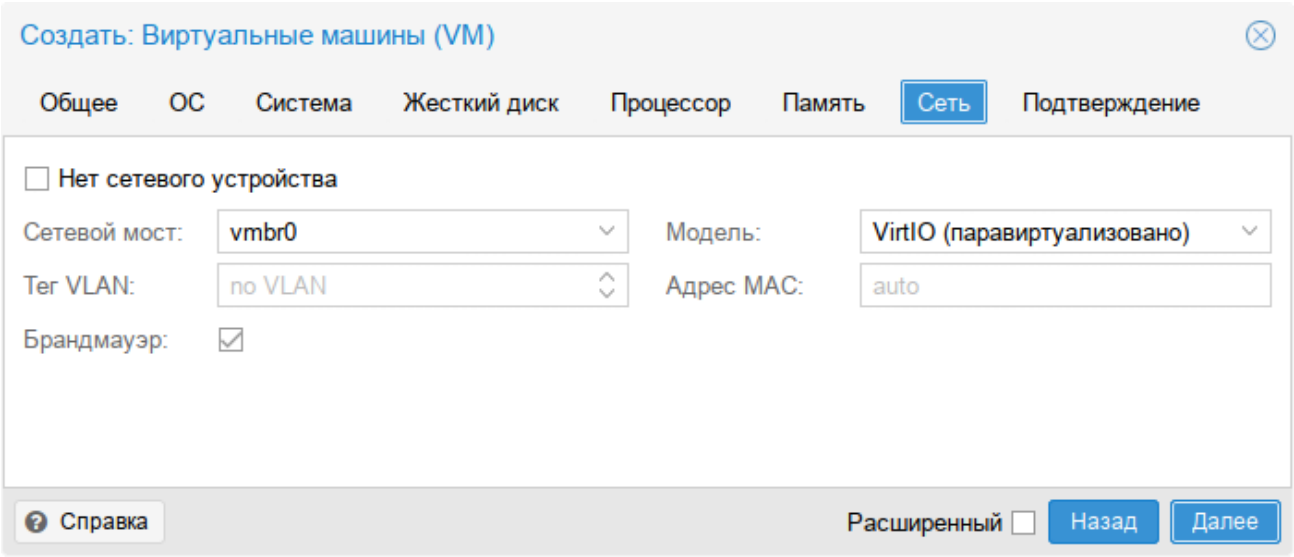

Рис. 88

Вкладка «Подтверждение»

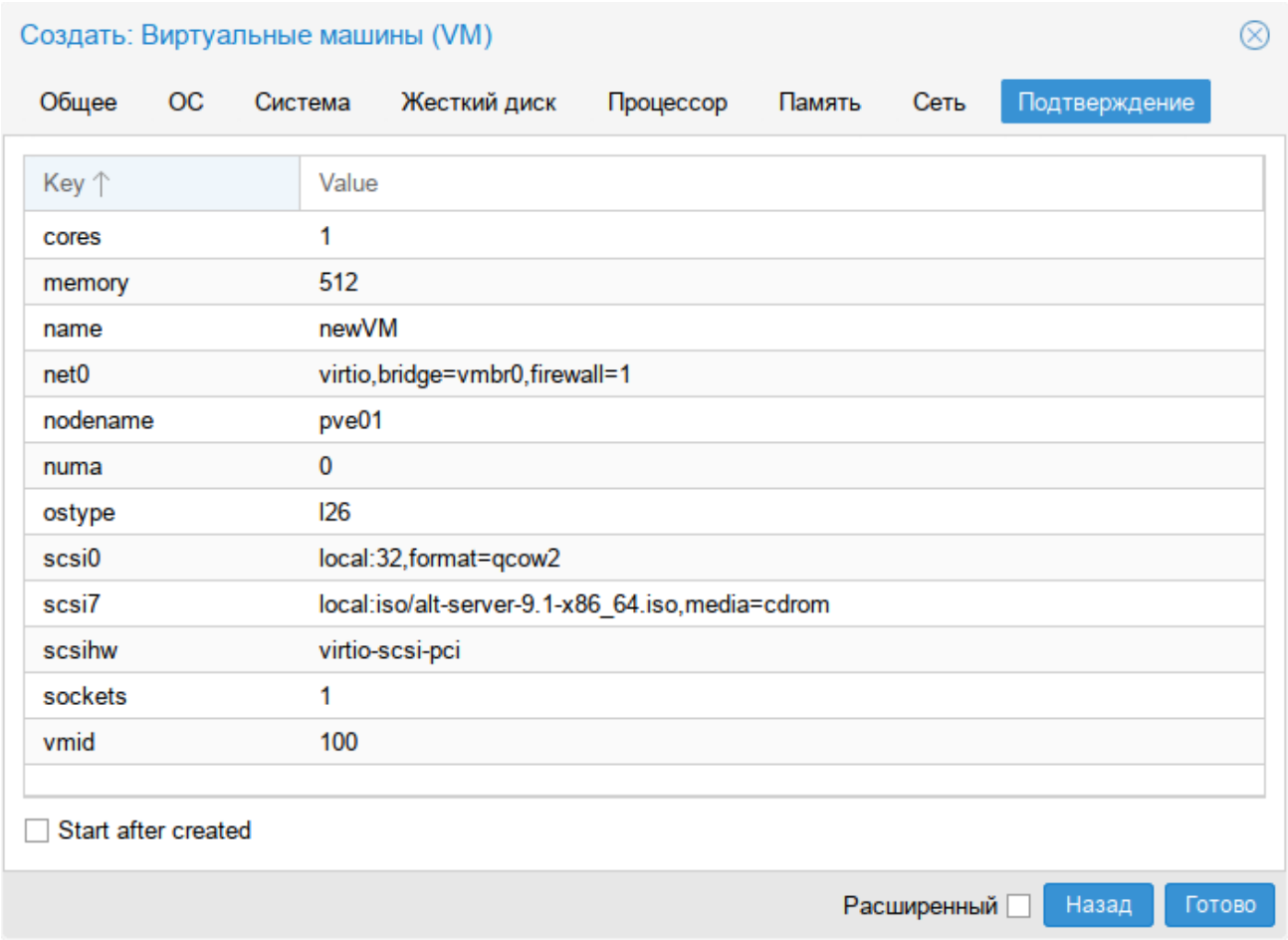

Рис. 89

4.6.2 Внесение изменений в ВМ

Вносить изменения в конфигурацию ВМ можно и после ее создания. Для того чтобы внести изменения в конфигурацию ВМ необходимо выбрать ВМ и перейти на вкладку «Оборудование» («Hardware»). На этой вкладке следует выбрать технические средства (Рис. 90) и нажать кнопку «Редактировать» («Edit») для выполнения изменений. Однако эти изменения могут не вступить в действие сразу, и может потребоваться выключение с последующим включением ВМ для инициализации данных аппаратных изменений.

Оборудование ВМ

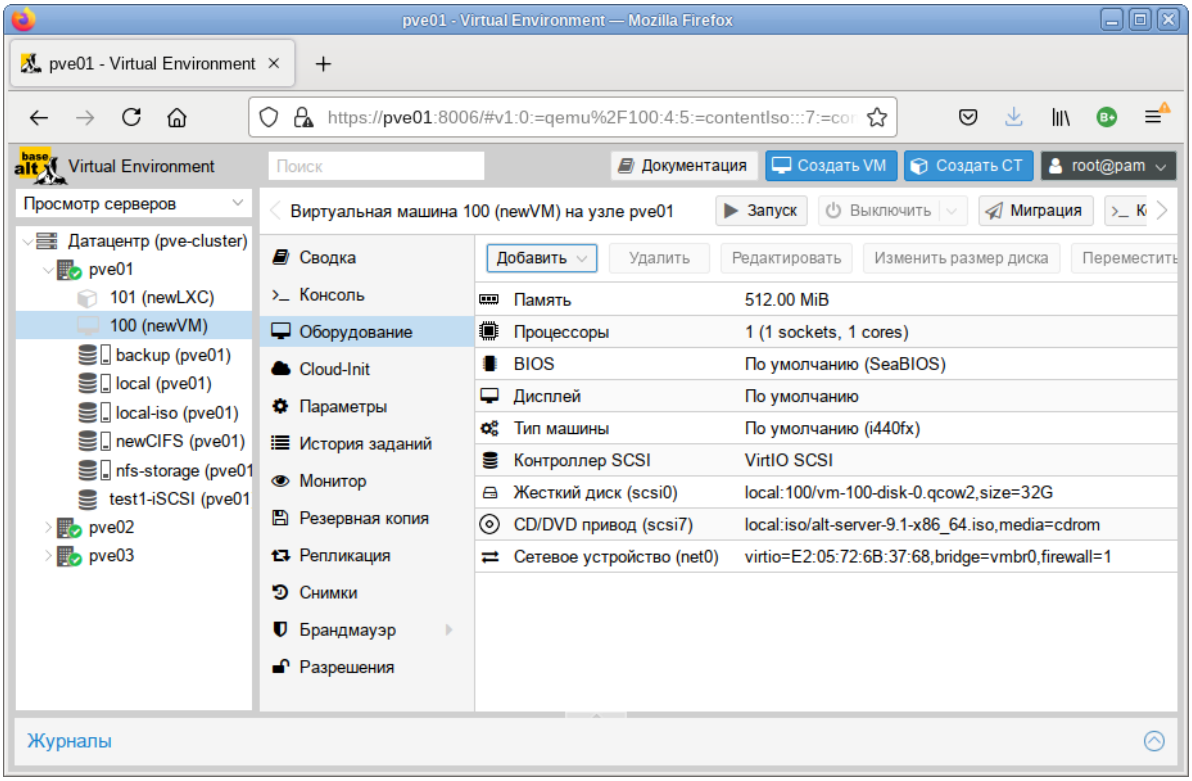

Рис. 90

4.6.3 Управление ВМ с помощью qm

Если веб-интерфейс PVE недоступен, можно управлять ВМ при помощи командной строки (либо через сеанс SSH, либо из консоли noVNC, или зарегистрировавшись на физическом хосте).

`qm` – это инструмент для управления ВМ Qemu/KVM в PVE. Утилиту `qm` можно использовать для создания/удаления ВМ, для контроля выполнения (запуск/остановка/приостановка/возобновление), для установки параметров в соответствующем конфигурационном файле, а также для создания виртуальных дисков.

Чтобы просмотреть доступные для ВМ команды PVE можно выполнить следующую команду:

```
# qm help
```

Примеры использования утилиты `qm`:

– создать ВМ, используя ISO-файл, загруженный в локальное хранилище, с диском IDE 21 ГБ, в хранилище local-lvm:

```
# qm create 300 -ide0 local-lvm:21 -net0 e1000 -cdrom local:iso/alt-server-9.1-x86_64.iso
```

– запуск ВМ:

```
# qm start 300
```

– отправить запрос на отключение, и дождаться остановки ВМ:

```
# qm shutdown 300 && qm wait 300
```

4.6.4 Файлы конфигурации ВМ

Файлы конфигурации ВМ хранятся в файловой системе кластера PVE и доступны по адресу `/etc/pve/qemu-server/<VMID>.conf`. Как и другие файлы, хранящиеся в `/etc/pve/`, они автоматически реплицируются на все другие узлы кластера.

Примечание. VMID < 100 зарезервированы для внутренних целей. VMID должны быть уникальными для всего кластера.

Пример файла конфигурации:

```
boot: order=scsi0;scsi7;net0
cores: 1
memory: 512
name: newVM
net0: virtio=C6:E4:55:03:79:5A,bridge=vmbr0,firewall=1
numa: 0
ostype: l26
scsi0: local:100/vm-100-disk-0.qcow2,size=32G
scsi7: local:iso/alt-server-9.1-x86_64.iso,media=cdrom
scsihw: virtio-scsi-pci
smbios1: uuid=a4ce5cab-f4df-4bde-a19b-6b9f3ebbaddb
sockets: 1
vmgenid: a1109827-0dc7-4d91-8625-1cff91c23d33
```

Файлы конфигурации ВМ используют простой формат: разделенные двоеточиями пары ключ/значение (пустые строки игнорируются, строки, начинающиеся с символа #, рассматриваются как комментарии и также игнорируются):

```
OPTION: value
```

Для применения изменений, которые напрямую вносились в файл конфигурации, необходимо перезапустить ВМ. По этой причине рекомендуется использовать команду `qm` для генерации и изменения этих файлов, либо выполнять такие действия в веб-интерфейсе.

При создании снимка ВМ, конфигурация ВМ во время снимка, сохраняется в этом же файле конфигурации в отдельном разделе. Например, после создания снимка «snapshot» файл конфигурации будет выглядеть следующим образом:

```
bootdisk: scsi0
…
parent: snapshot
…
```

```
vmgenid: a1109827-0dc7-4d91-8625-1cff91c23d33

[snapshot]
bootdisk: scsi0
cores: 1
memory: 512
name: newVM
net0: virtio=C6:E4:55:03:79:5A,bridge=vmbr0,firewall=1
numa: 0
ostype: l26
runningmachine: pc-i440fx-4.1
scsi0: local:100/vm-100-disk-0.qcow2,size=32G
scsi7: local:iso/alt-server-9.1-x86_64.iso,media=cdrom
scsihw: virtio-scsi-pci
smbios1: uuid=a4ce5cab-f4df-4bde-a19b-6b9f3ebbaddb
snaptime: 1595498248
sockets: 1
vmgenid: a1109827-0dc7-4d91-8625-1cff91c23d33
vmstate: local:100/vm-100-state-snapshot.raw
```

Свойство `parent` при этом используется для хранения родительских/дочерних отношений между снимками, а `snaptime` – это отметка времени создания снимка (эпоха Unix).

4.6.5 Запуск и останов ВМ

4.6.5.1 Изменение состояния ВМ в веб-интерфейсе

Для запуска ВМ следует выбрать ее в левой панели; иконка ВМ должна быть серого цвета, обозначая, что ВМ не запущена.

Запустить ВМ можно выбрав в контекстном меню ВМ пункт «Запуск» (*Рис. 91*), либо нажав на кнопку «Запуск» («Start») (*Рис. 92*).

Запущенная ВМ будет обозначена зеленой стрелкой на значке ВМ.

Контекстное меню ВМ

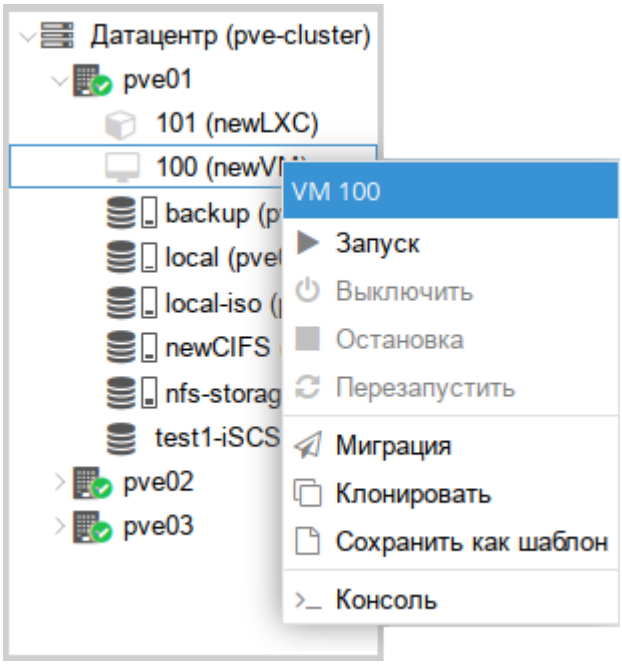

Рис. 91

Кнопки управления состоянием ВМ

Рис. 92

Для запущенной ВМ доступны следующие действия (*Рис. 93*):

– «Пауза» («Pause») – перевод ВМ в спящий режим;
– «Hibernate» – перевод ВМ в ждущий режим;
– «Выключить» («Shutdown») – выключение ВМ;
– «Остановка» («Stop») – остановка ВМ, путем прерывания ее работы;
– «Перезапустить» («Reboot») – перезапустить ВМ.

Контекстное меню запущенной ВМ

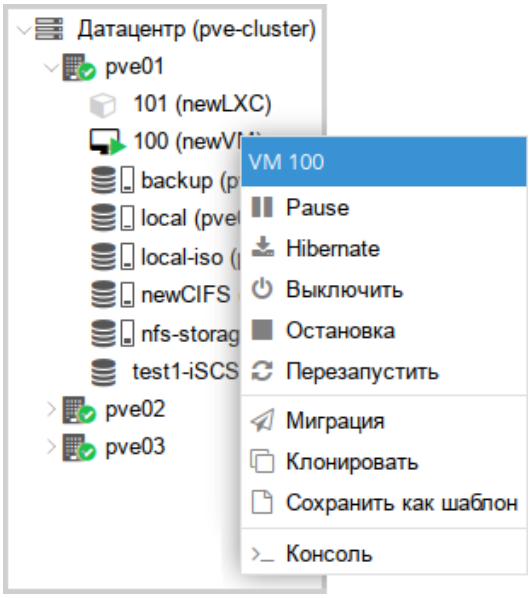

Рис. 93

4.6.5.2 Изменение состояний ВМ в командной строке

Состоянием ВМ можно управлять из командной строки PVE (либо через сеанс SSH, либо из консоли noVNC, или зарегистрировавшись на физическом хосте).

Для запуска ВМ с VM ID 105 необходимо ввести команду:

```
# qm start 105
```

Эта же ВМ может быть остановлена при помощи команды:

```
# qm stop 105
```

4.6.5.3 Автоматический запуск ВМ

Для того чтобы ВМ запускалась автоматически при загрузке хост-системы, необходимо выбрать опцию «Запуск при загрузке» на вкладке «Параметры» требуемой ВМ в веб-интерфейсе или установить ее с помощью следующей команды:

```
# qm set <vmid> -onboot 1
```

Иногда необходимо точно настроить порядок загрузки ВМ, например, если одна из ВМ обеспечивает межсетевой экран или DHCP для других гостевых систем. Для настройки порядка запуска ВМ можно использовать следующие параметры (Рис. 94) (опция «Порядок запуска и выключения» на вкладке «Параметры» требуемой ВМ):

– «Порядок запуска и выключения» («Start/Shutdown order») – определяет приоритет порядка запуска. Для того чтобы ВМ запускалась первой необходимо установить этот параметр в значение 1 (для выключения используется обратный порядок: ВМ машина с порядком запуска 1 будет выключаться последней). Если несколько хостов имеют одинаковый порядок, определенный на хосте, они будут дополнительно упорядочены в порядке возрастания VMID;

– «Задержка запуска» («Startup delay») – определяет интервал (в секундах) между запуском этой ВМ и последующими запусками ВМ;
– «Задержка выключения» («Shutdown timeout») – определяет продолжительность в секундах, в течение которой PVE должен ожидать, пока ВМ не будет в автономном режиме после выдачи команды выключения. По умолчанию это значение установлено равным 180, что означает, что PVE выдаст запрос на отключение и подождет 180 секунд, пока машина не будет в автономном режиме. Если машина все еще находится в сети после истечения времени ожидания, она будет принудительно остановлена.

Настройка порядка запуска и выключения ВМ

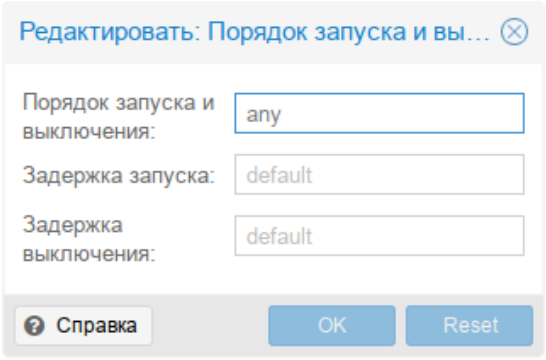

Рис. 94

Примечание. Виртуальные машины, управляемые стеком HA, в настоящее время не поддерживают параметры запуска при загрузке и порядка загрузки.

ВМ без настроенного параметра «Порядок запуска и выключения» всегда будут запускаться после тех, для которых установлен этот параметр. Кроме того, этот параметр может быть применен только для ВМ, работающих на одном хосте, но не для всего кластера.

4.6.6 Управление образами виртуальных дисков

Образ виртуального диска является файлом или группой файлов, в которых ВМ хранит свои данные. В PVE файл настройки ВМ может создаваться повторно и использоваться для подключения образа диска. Однако если образ сам по себе утрачен, он может быть восстановлен только из резервной копии. Существуют различные типы форматов образов виртуальных дисков доступные для применения виртуальными машинами.

4.6.6.1 Поддерживаемые форматы образов

PVE поддерживает форматы виртуальных дисков .raw, .qcow2 и .vmdk. Каждый формат имеет свои преимущества и недостатки. Формат образа обычно выбирается на основании функции ВМ, используемой системы хранения, требующейся производительности и доступного бюджета. В процессе создания виртуального диска в веб-интерфейсе можно выбрать формат диска (*Рис. 95*).

Формат образа qcow2 (QEMU image format) – универсальный формат для выполнения любых задач. Его преимущество в том, что файл с данными будет содержать только реально

занятое место внутри ВМ. Например, если было выделено 40 Гб места, а реально было занято только 2 Гб, то все остальное место будет доступно для других ВМ. По мере сохранения пользователем данных в этой ВМ, такой образ будет постепенно увеличиваться в размере. Формат образа .qcow2 позволяет администратору превышать обеспечение ВМ файлом образа диска .qcow2. В таких средах необходимо постоянно наблюдать за доступным пространством хранения. Все операции ввода-вывода при использовании этого формата программно обрабатываются, что влечет за собой замедление при активной работе с дисковой подсистемой. Если стоит задача развернуть на сервере базу данных, то лучше выбрать формат RAW. Создаваемые из образа qcow2 резервные копии могут восстанавливаться только в NFS или локальный каталог.

Формат образа vmdk (VMware image format) очень распространен в инфраструктуре VMware. Он позволяет выполнить миграцию виртуальной машины VMware в инфраструктуру PVE.

Выбор формата виртуального диска

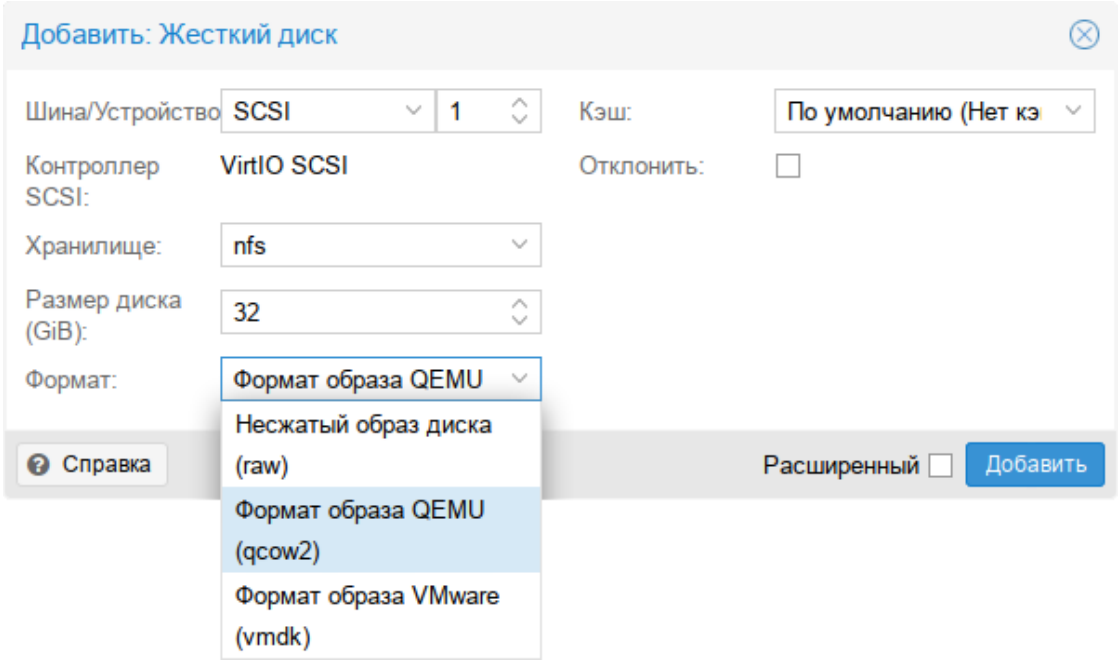

Рис. 95

Формат образа raw – это файл с данными жесткого диска «байт в байт» без сжатия или оптимизации. Его главное преимущество состоит в производительности – это самый быстрый «тип» накопителя, так как гипервизору не нужно его никак обрабатывать. Формат .raw может создавать только файлы образов ВМ фиксированного размера с предварительным выделением (thick-provisioned). Например, создаваемая ВМ с пространством хранения 50 ГБ будет иметь файл образа 50 ГБ. При этом администратор точно знает, сколько пространства используется, поэтому отсутствует возможность неуправляемого выхода за пределы хранения. Формат образа .raw может восстанавливаться практически на любой тип системы хранения.

4.6.6.2 *Управление образами виртуальных дисков*

Для управления образами виртуальных дисков ВМ необходимо выбрать нужную ВМ в меню навигации и перейти на вкладку «Оборудование» («Hardware») (Рис. 96). После выбора образа диска становятся доступными все связанные меню, такие как «Добавить» («Add»), «Отключить» («Remove»), «Редактировать» («Edit»), «Изменить размер диска» («Resize disk»), «Переместить диск» («Move disk»).

Вкладка «Оборудование»

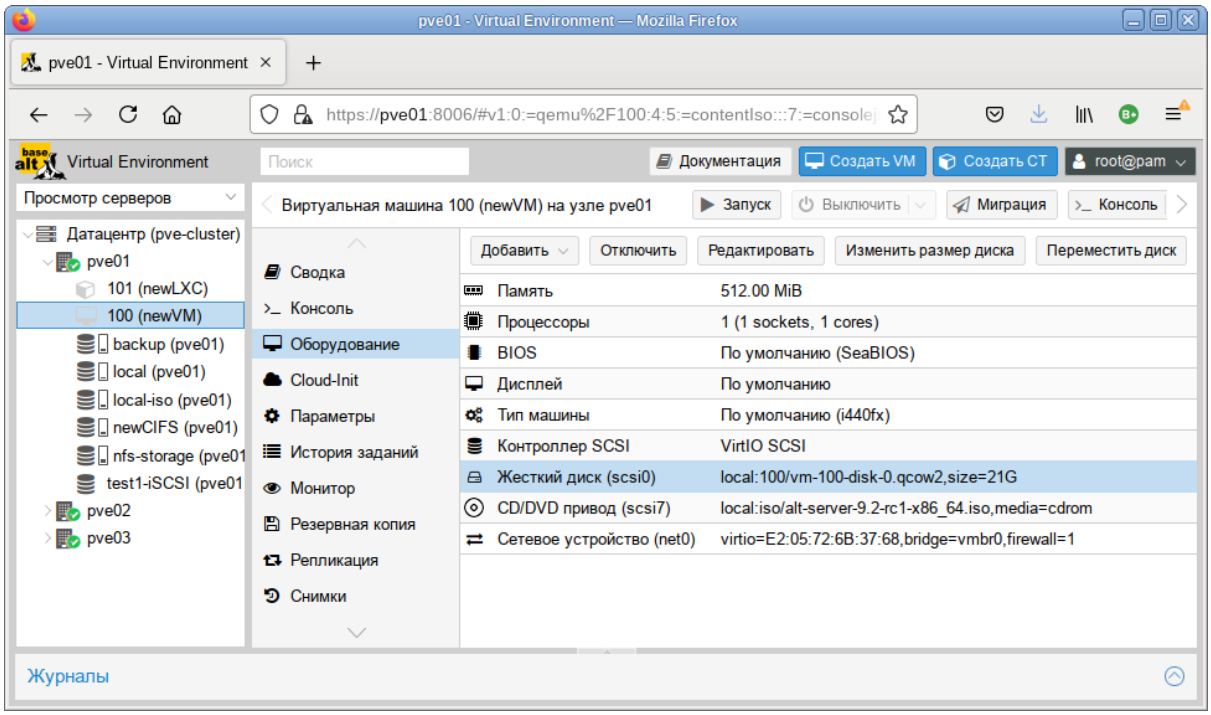

Рис. 96

4.6.6.2.1 Добавление виртуального диска в ВМ

Для добавления образа виртуального диска к ВМ необходимо:

1) перейти на вкладку «Оборудование» («Hardware») (Рис. 96);

2) нажать кнопку «Добавить» («Add») и выбрать в выпадающем списке пункт «Жесткий диск» («Hard Disk») (Рис. 97);

3) после ввода необходимых значений (Рис. 98) нажать кнопку «Добавить» («Add») для завершения добавления виртуального диска. В примере добавляется образ виртуального диска объемом 32 ГБ в виртуальную машину 100.

Кнопка «Добавить»→«Жесткий диск»

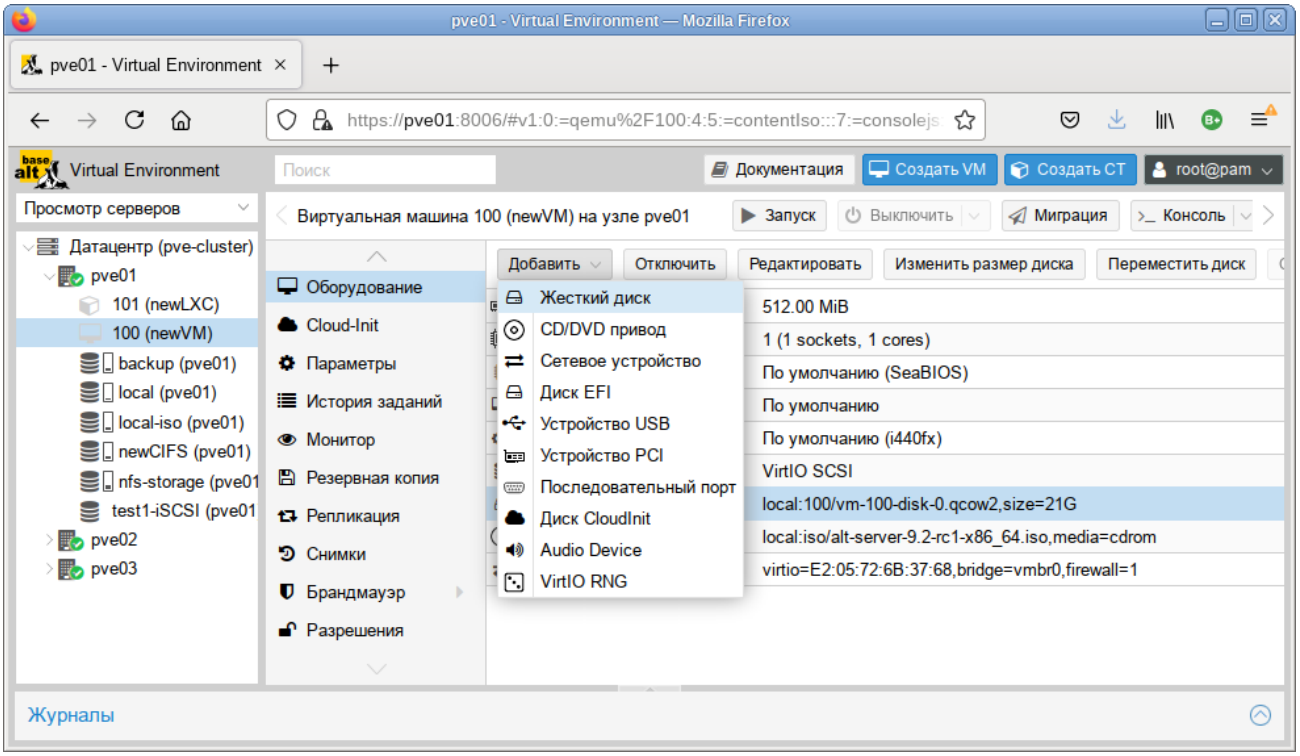

Рис. 97

Опции добавления жесткого диска

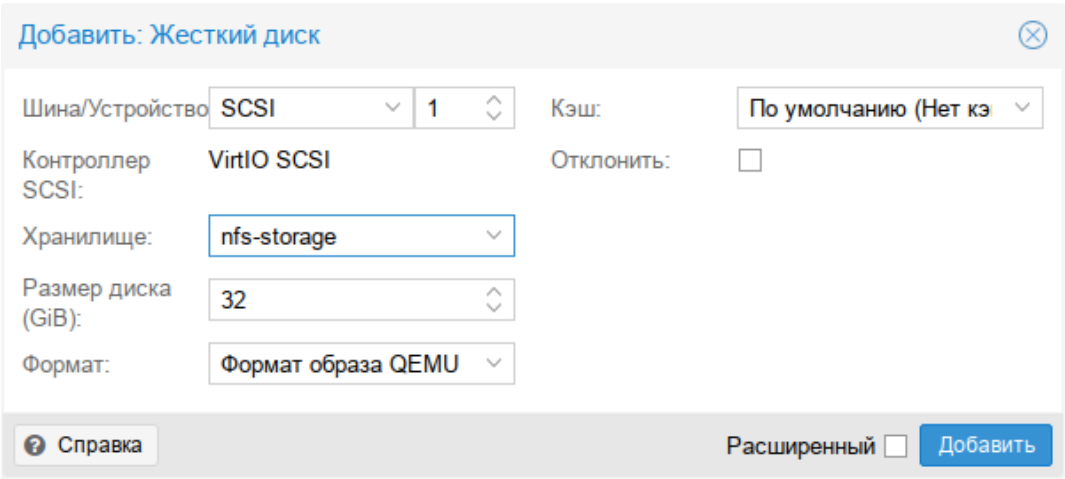

Рис. 98

4.6.6.2.2 Удаление образа виртуального диска

Для удаления образа виртуального диска необходимо:

1) перейти на вкладку «Оборудование» («Hardware») (Рис. 96);
2) выбрать образ диска ВМ;
3) нажать кнопку «Отключить» («Remove»);
4) в окне подтверждения нажать кнопку «Да» («Yes») для подтверждения действия. При этом виртуальный диск будет отсоединен от ВМ, но не удален полностью. Он будет присутствовать в списках как «Неиспользуемый диск» («Unused Disk») (Рис. 99).

«Неиспользуемый диск»

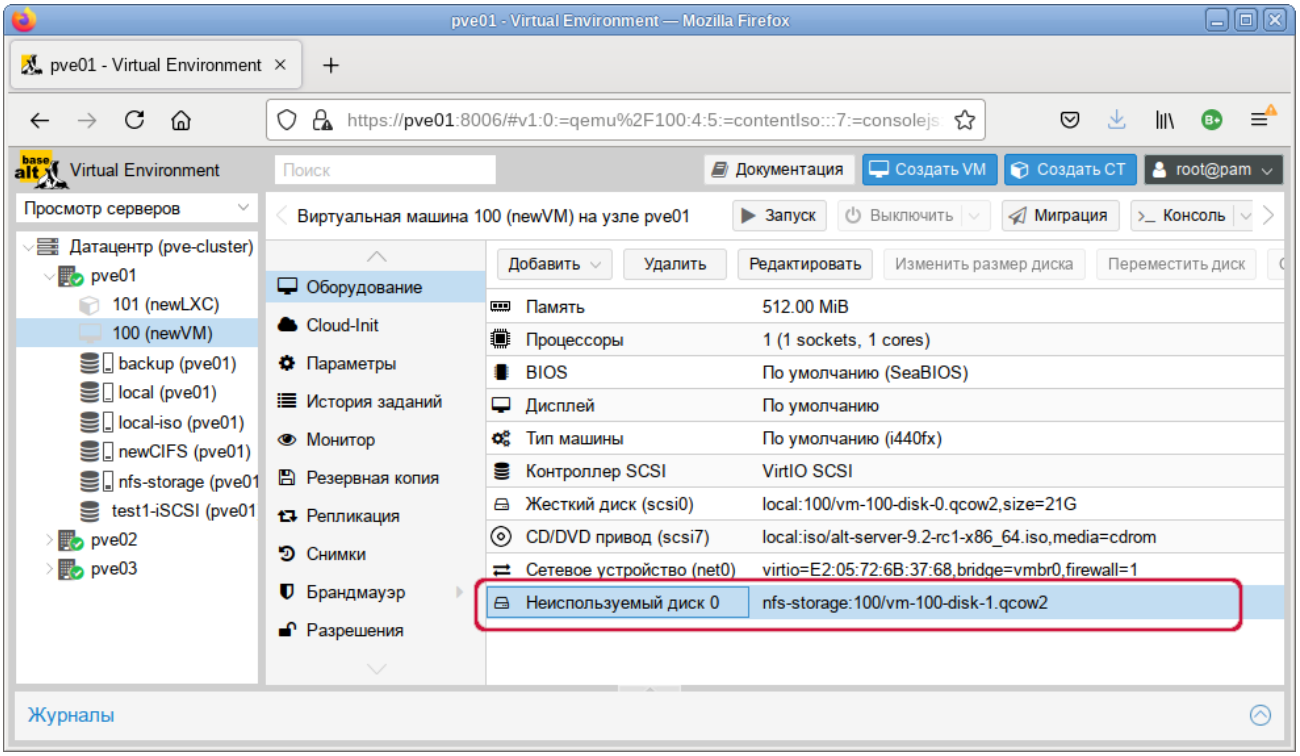

Рис. 99

Чтобы удалить образ диска окончательно, следует выбрать неиспользуемый диск и нажать кнопку «Удалить» («Remove»).

Если образ диска был отключен от ВМ по ошибке, можно повторно подключить его к ВМ, выполнив следующие действия:

1) выбрать неиспользуемый диск;
2) нажать кнопку «Редактировать» («Edit»);
3) в открывшемся диалоговом окне (Рис. 100) изменить, если это необходимо, параметры «Шина/Устройство» («Bus/Device»).
4) нажать кнопку «Добавить» («Add)» для повторного подключения образа диска.

Подключение неиспользуемого диска

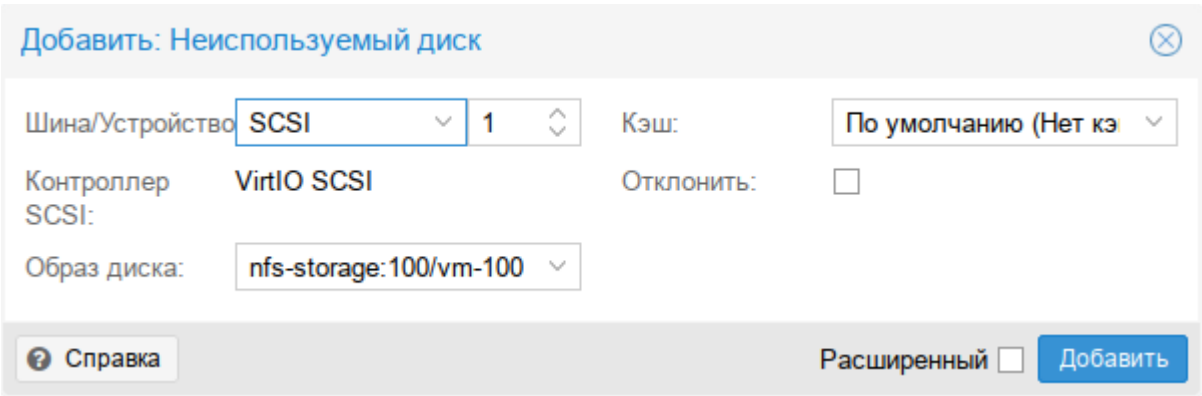

Рис. 100

4.6.6.2.3 Изменение размера диска

Функция изменения размера поддерживает только увеличение размера файла образа виртуального диска. При изменении размера образа виртуального диска изменяется только размер файла образа виртуального диска. После изменения размера файла, разделы жесткого диска должны быть изменены внутри самой ВМ.

Для изменения размера виртуального диска необходимо:

1) перейти на вкладку «Оборудование» («Hardware») (Рис. 96);
2) выбрать образ виртуального диска.
3) нажать кнопку «Изменить размер диска» («Resize disk»);
4) в открывшемся диалоговом окне в поле «Увеличение размера (GiB)» ввести значение увеличения размера диска. Например, если размер существующего диска составляет 20 ГБ, следует ввести 10 для изменения размера диска до 30 ГБ (*Рис. 101*);
5) нажать кнопку «Изменить размер диска» («Resize Disk») для завершения изменения размера.

Изменение размера диска

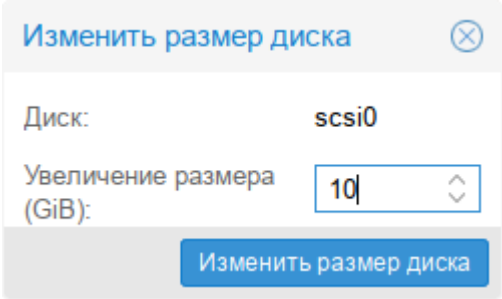

Рис. 101

Команда изменения размера виртуального диска:

```
# qm resize <vm_id> <virtual_disk> +<size>G
```

4.6.6.2.4 Перемещение диска на другое хранилище

Образы виртуального диска могут перемещаться с одного хранилища на другое в пределах одного кластера. Данные шаги выполнят задачу для перемещения образа диска:

1) перейти на вкладку «Оборудование» («Hardware») (Рис. 96);
2) выбрать образ перемещаемого виртуального диска;
3) нажать кнопку «Переместить диск» («Move disk»);
4) в открывшемся диалоговом окне (Рис. 102) выбрать в выпадающем меню «Целевое хранилище» («Target Storage») – хранилище-получатель, место, куда будет перемещен образ виртуального диска;
5) выбрать в выпадающем меню «Формат» («Format») образ диска, если это необходимо. Этот параметр полезен для преобразования образов диска из одного формата в другой;

6) отметить, если это необходимо, пункт «Удалить источник» («Delete source») для удаления образа диска источника после его перемещения в новое хранилище;

7) нажать кнопку «Переместить диск» («Move disk») для запуска перемещения образа диска.

Диалоговое окно перемещения диска

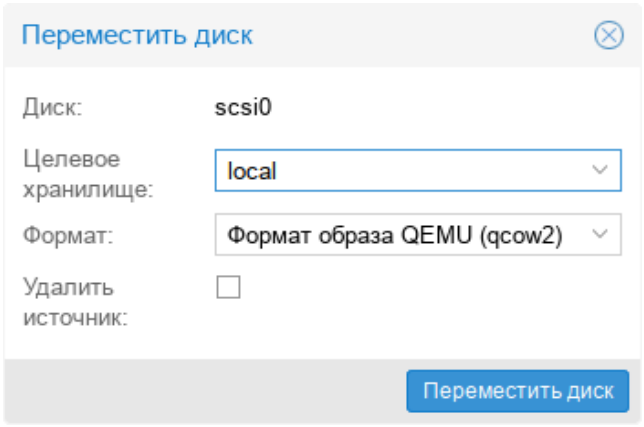

Рис. 102

4.6.6.3 Управление дисками в командной строке

В комплект PVE входит утилита `qemu-img`. `qemu-img` – утилита для манипулирования с образами дисков машин QEMU. Она поддерживает несколько подкоманд:

- create
- commit
- convert
- info

Команда преобразования (конвертирования) vmdk-образа виртуального накопителя VMware под названием test в формат qcow2:

```
# qemu-img convert -f vmdk test.vmdk -O qcow2 test.qcow2
```

Создать образ test в формате RAW, размером 40 ГБ:

```
# qemu-img create -f raw test.raw 40G
```

Изменение размера виртуального диска:

```
# qemu-img resize -f raw test.raw 80G
```

Просмотр информации об образе:

```
# qemu-img info test.raw
```

4.6.6.4 Кэширование виртуального диска

Параметр настройки кэширования доступен в диалоговом окне создания или изменений текущего образа диска ВМ (Рис. 103).

Выбор типа кэширования

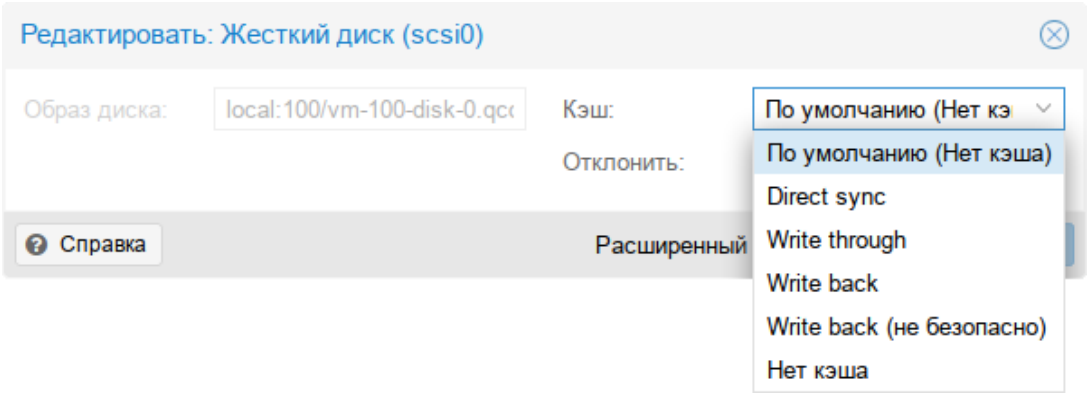

Рис. 103

Доступны следующие виды кэширования:

– «Нет кэша» («No cache») – опция кэширования по умолчанию. При этой опции кэширования на уровне хоста не происходит, однако гостевые ВМ выполняют кэширование отложенной записи. Диск такой ВМ напрямую получает подтверждение от устройства хранения. В случае внезапного отключения питания существует большой риск потери данных. Данный тип кэширования представляет хорошее соотношение производительности и безопасности;

– «Write through» (кэш сквозной записи) – использует «host cache» для чтения. При этом типе кэширования подтверждение на запись выдается только когда данные зафиксированы на устройстве хранения. «Write through» выполняет fsync (синхронизация находящихся данных в памяти с диском) для каждой записи. Это более безопасный режим кеширования, так как возможность потери данных стремится к нулю, однако при этом более медленный;

– «Direct sync» (прямая синхронизация) – хост не производит никакого кэширования, т.е. читает всегда напрямую с блочного устройства и пишет в блочное устройства обязательно дожидаясь подтверждения записи. Прямая синхронизация рекомендуется для ВМ, которые не отправляют запросы на сброс при их необходимости. Это наиболее безопасный кэш, так как данные не будут утрачены при отказе питания, однако он также и самый медленный;

– «Write back» (отложенная запись) – хост выполняет и кэширование чтения, и кэширование записи. Подтверждение на запись диском ВМ выполняется как только данные зафиксированы в кэше хоста вне зависимости от того были они зафиксированы в хранилище или нет. Самый быстрый тип хеширования, но при потере питания пропадают все данные;

– «Write back (не безопасно)» (ненадежная отложенная запись) – аналогично отложенной записи за исключением того, что все сбросы данных полностью игнорируются со стороны гостевой ВМ. Это самый быстрый и небезопасный тип кэширования. Он не должен применяться в промышленных кластерах. Обычно этот кэш используется для ускорения установ-

ки ОС в ВМ. После установки ОС этот кэш должен быть отключен и возвращен в другую более безопасную опцию.

4.7 Создание и настройка контейнера LXC

4.7.1 Создание контейнера в графическом интерфейсе

Перед созданием контейнера можно загрузить шаблоны LXC в хранилище.

Нажать на кнопку «Создать CT» (Рис. 104), это запустит диалог «Создать: Контейнер LXC» (Рис. 105), который предоставляет графический интерфейс для настройки контейнера.

Создание контейнера

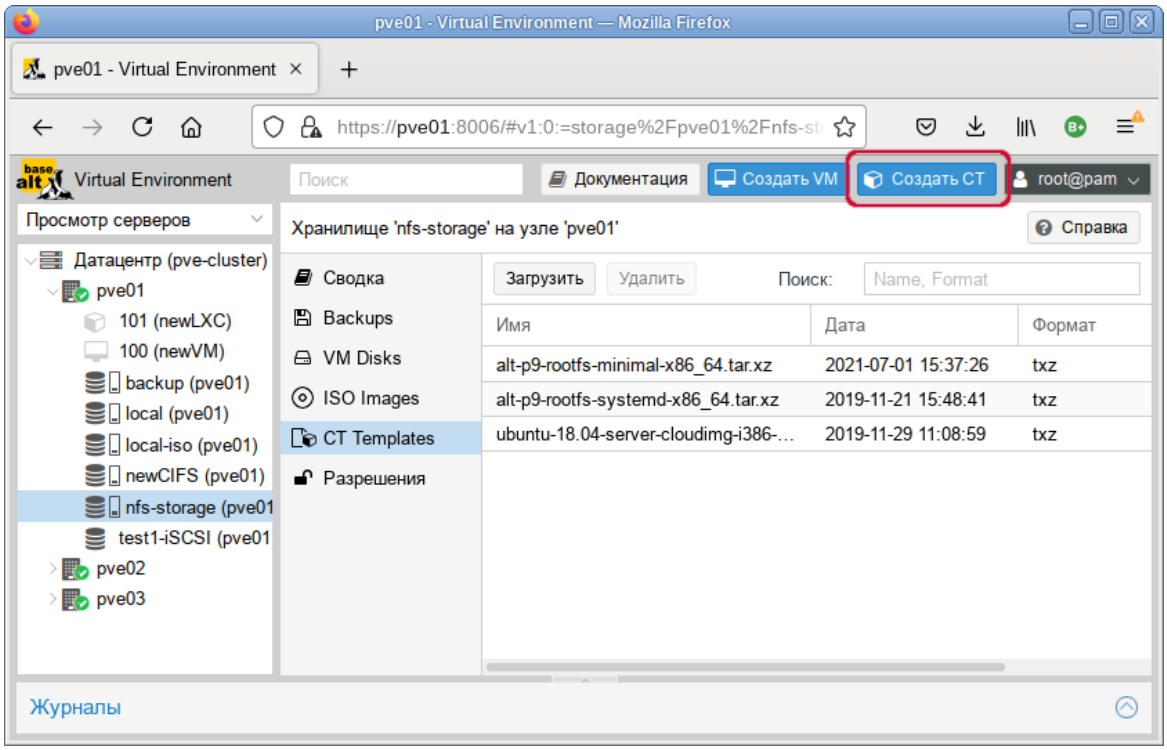

Рис. 104

На первой вкладке «Общее» («General») необходимо указать (Рис. 105):

- «Узел» («Node») – узел назначения для данного контейнера;
- «CT ID» – идентификатор контейнера в численном выражении;
- «Имя хоста» («Hostname») – алфавитно-цифровая строка названия контейнера;
- «Непривилегированный контейнер» – определяет, как будут запускаться процессы контейнера (если процессам внутри контейнера не нужны полномочия администратора, то необходимо снять отметку с этого пункта);
- «Пул ресурсов» («Resource pool») – имя пула данного контейнера, к которому он будет относиться. Данное значение не обязательное. Чтобы иметь возможность выбора этот пул должен быть предварительно создан;
- «Пароль» («Password») – пароль для данного контейнера;

– «Открытый SSH ключ» («SSH public key») – ssh ключ.

Вкладка «Общее» диалога создания контейнера

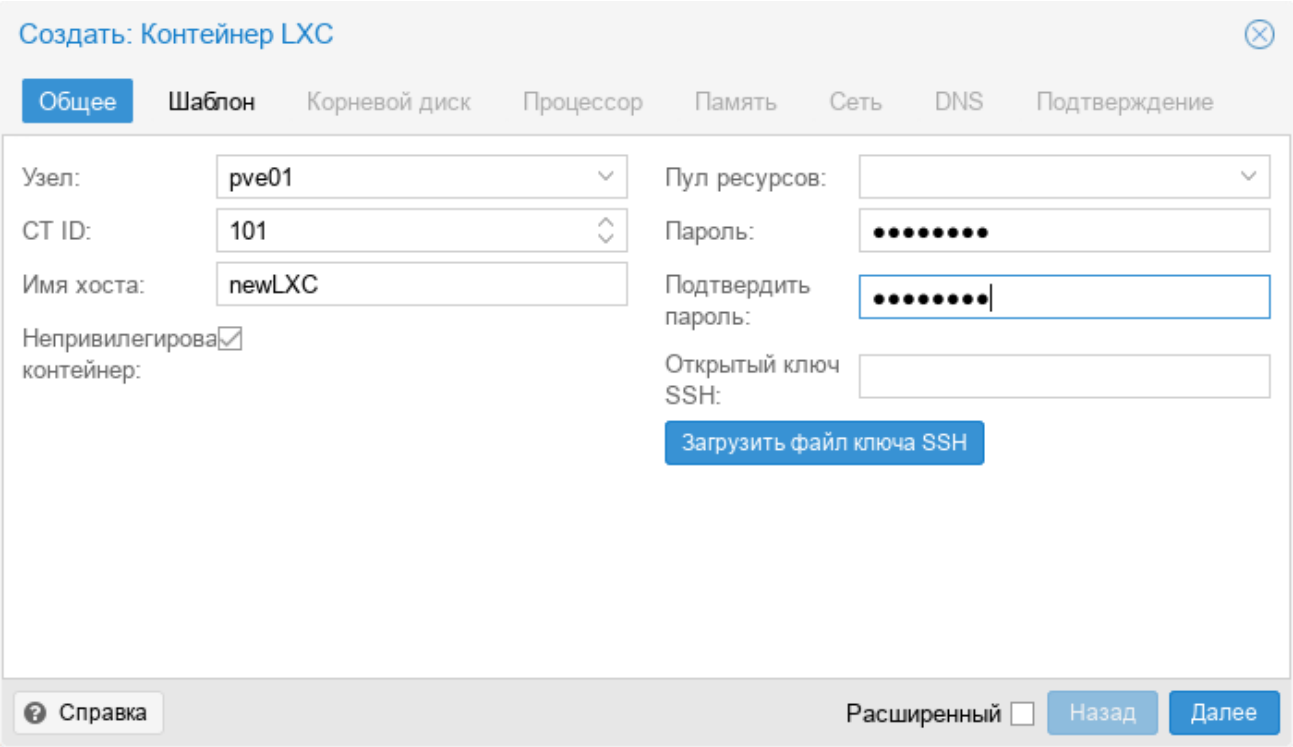

Рис. 105

На вкладке «Шаблон» («Template») настраиваются (Рис. 106):

– «Хранилище» («Storage») – хранилище в котором хранятся шаблоны LXC;

– «Шаблон» («Template») – шаблон контейнера.

На вкладке «Корневой диск» («Root Disk») определяется хранилище, где будут храниться диски контейнера (Рис. 107). Здесь также можно определить размер виртуального диска (не следует выбирать размер диска менее 4 ГБ).

Вкладка «Шаблон» диалога создания контейнера

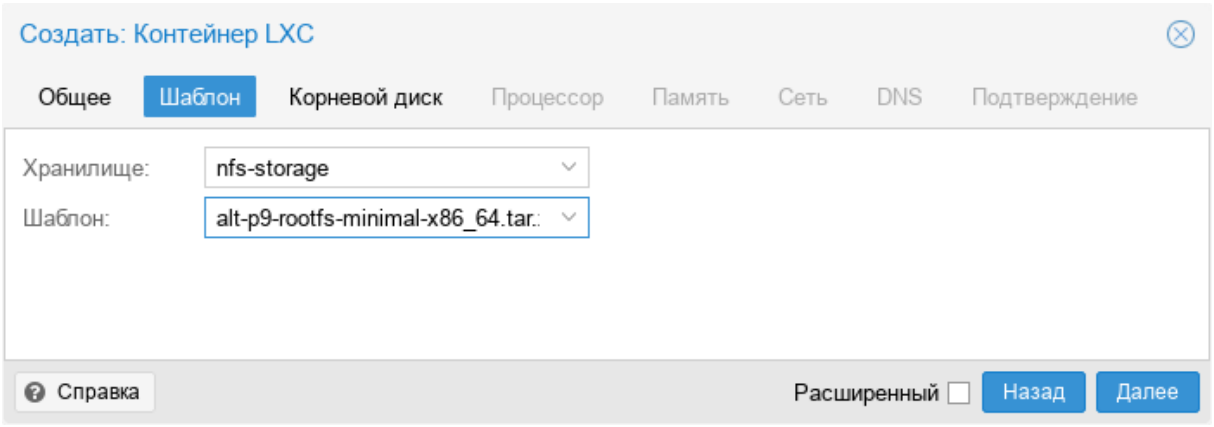

Рис. 106

Вкладка «Корневой диск» диалога создания контейнера

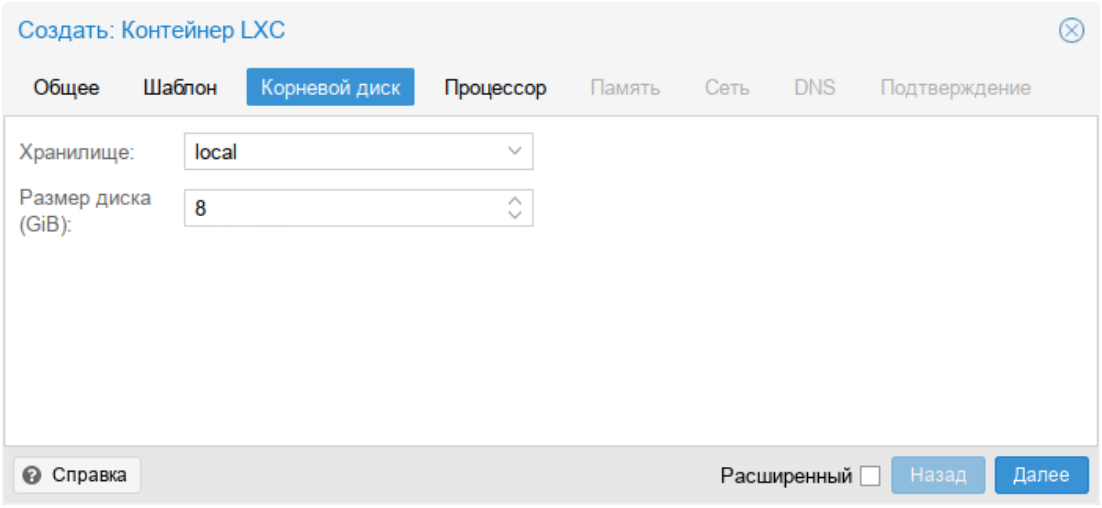

Рис. 107

На вкладке «Процессор» («CPU») определяется количество ядер процессора, которые будут выделены контейнеру (Рис. 108).

Вкладка «Процессор» диалога создания контейнера

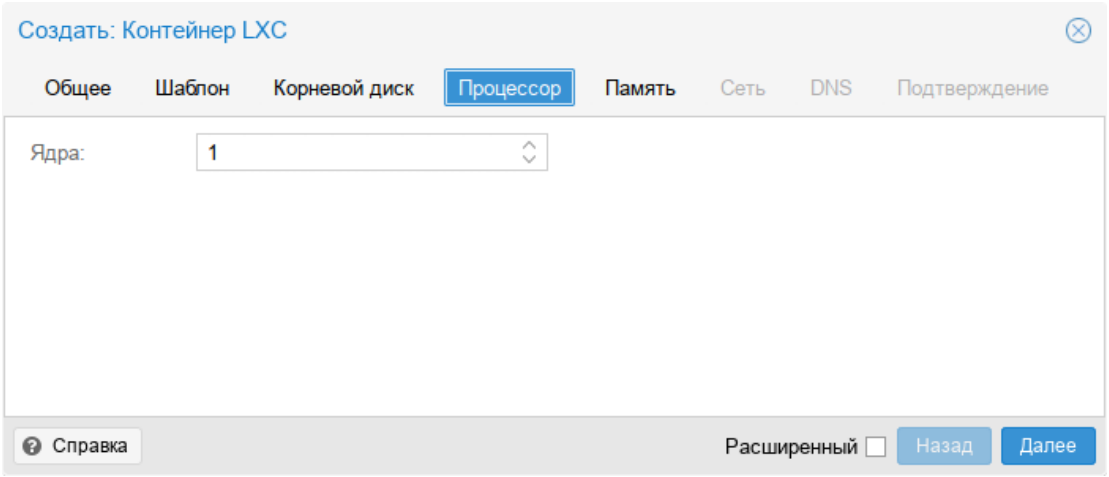

Рис. 108

На вкладке «Память» («Memory») настраиваются (Рис. 109):

– «Память» («Memory») (MiB) – выделяемая память в мегабайтах;
– «Подкачка» («Swap») (MiB) – выделяемое пространство подкачки в мегабайтах.

Вкладка «Память» диалога создания контейнера

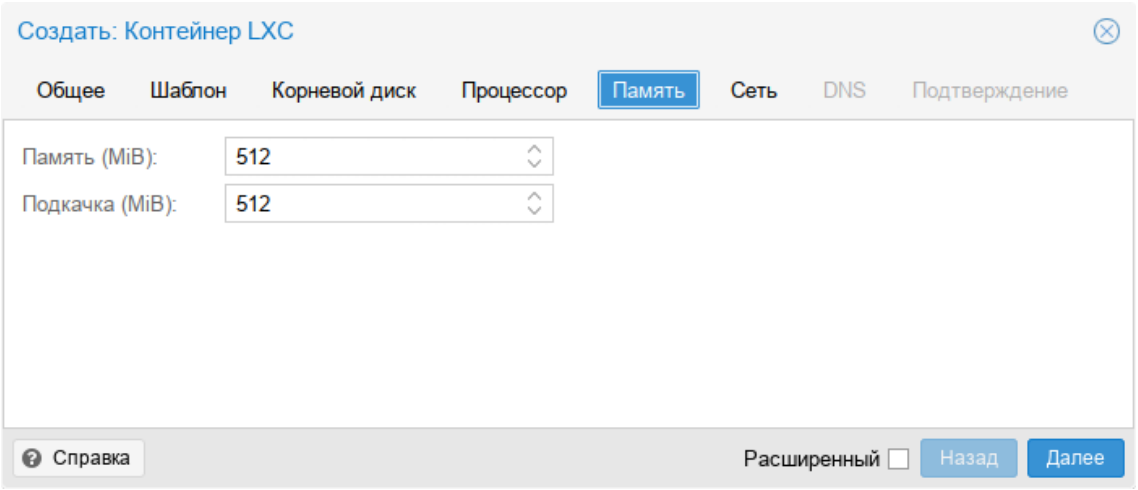

Рис. 109

Вкладка «Сеть» («Network») включает следующие настройки (Рис. *110*):

– «Имя» («Name») – определяет, как будет именоваться виртуальный сетевой интерфейс внутри контейнера; значение по умолчанию eth0;
– «Адрес MAC» («MAC address») – по умолчанию, все MAC адреса для виртуальных сетевых интерфейсов назначаются автоматически. Можно задать определенный MAC адрес необходимый для приложения в данном контейнере;
– «Сетевой мост» («Bridge») – выбор виртуального моста, к которому будет подключаться данный интерфейс; значение по умолчанию установлено в vmbr0;
– «Тег VLAN» («VLAN Tag») – применяется для установки идентификатора VLAN для данного виртуального интерфейса;
– «Ограничение трафика» («Rate limit») (MBit/s) – ограничение пропускной способности сетевой среды (в Мб/с). Для работы без ограничений следует оставить поле пустым;
– «Брандмауэр» («Firewall») – поддержка межсетевого экрана (если пункт отмечен, применяются правила хоста);
– «IPv4/IPv6» – можно настроить и IPv4, и IPv6 для виртуального сетевого интерфейса. IP адреса можно устанавливать вручную или разрешить получать от DHCP-сервера для автоматического назначения IP. IP должен вводиться в соответствии с CIDR (например, 192.168.0.0/24).

Вкладка «Сеть» диалога создания контейнера

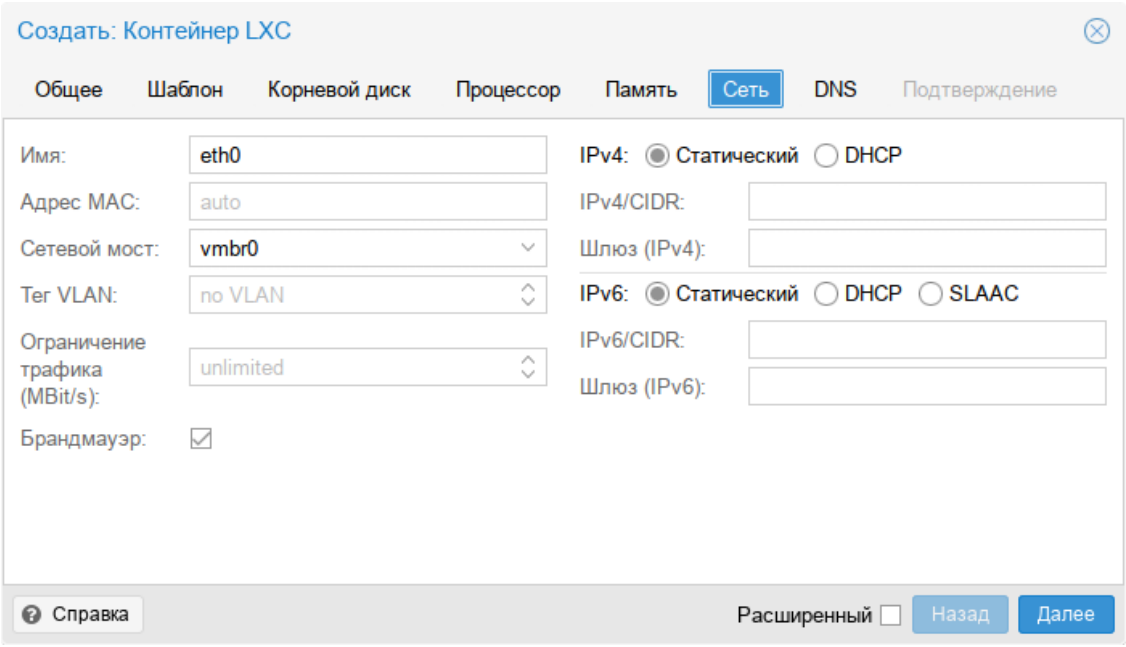

Рис. 110

Вкладка «DNS» содержит настройки (Рис. 111):

– «Домен DNS» («DNS domain») – имя домена (по умолчанию используются параметры хост системы);

– «Серверы DNS» («DNS server») – IP адреса серверов DNS (доступно только при введенном имени домена).

Вкладка «DNS» диалога создания контейнера

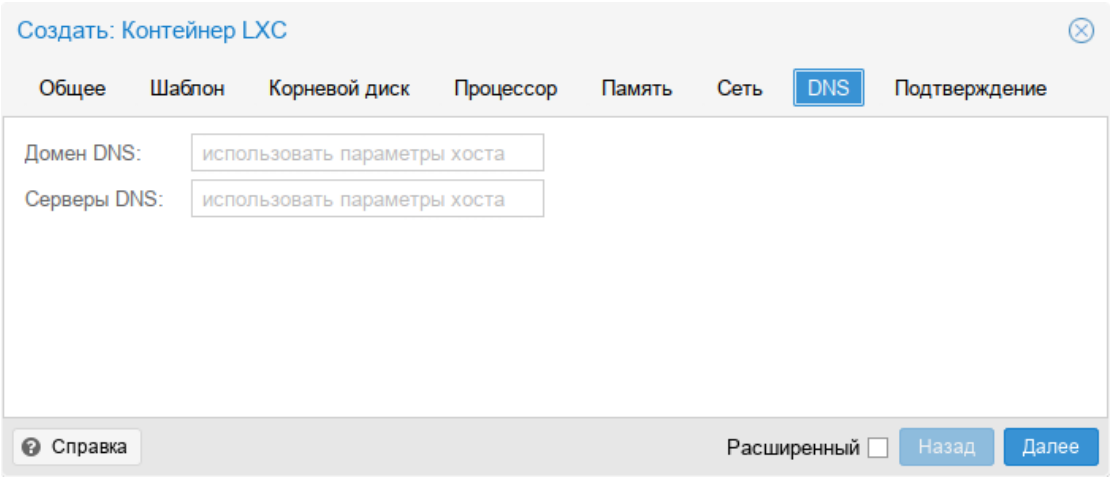

Рис. 111

Во вкладке «Подтверждение» («Confirm») отображаются все введенные или выбранные значения для данного контейнера (Рис. 112). Для создания контейнера необходимо нажать кнопку «Далее» («Finish»). Если необходимо внести изменения в параметры контейнера, можно перейти по вкладкам назад.

Если отметить пункт «Start after created» контейнер будет запущен сразу после создания.

После нажатия кнопки «Далее» во вкладке «Подтверждение», диалог настройки закрывается и в браузере открывается новое окно, которое предлагает возможность наблюдать за построением PVE контейнера LXC из шаблона (Рис. 113).

Вкладка «Подтверждение» диалога создания контейнера

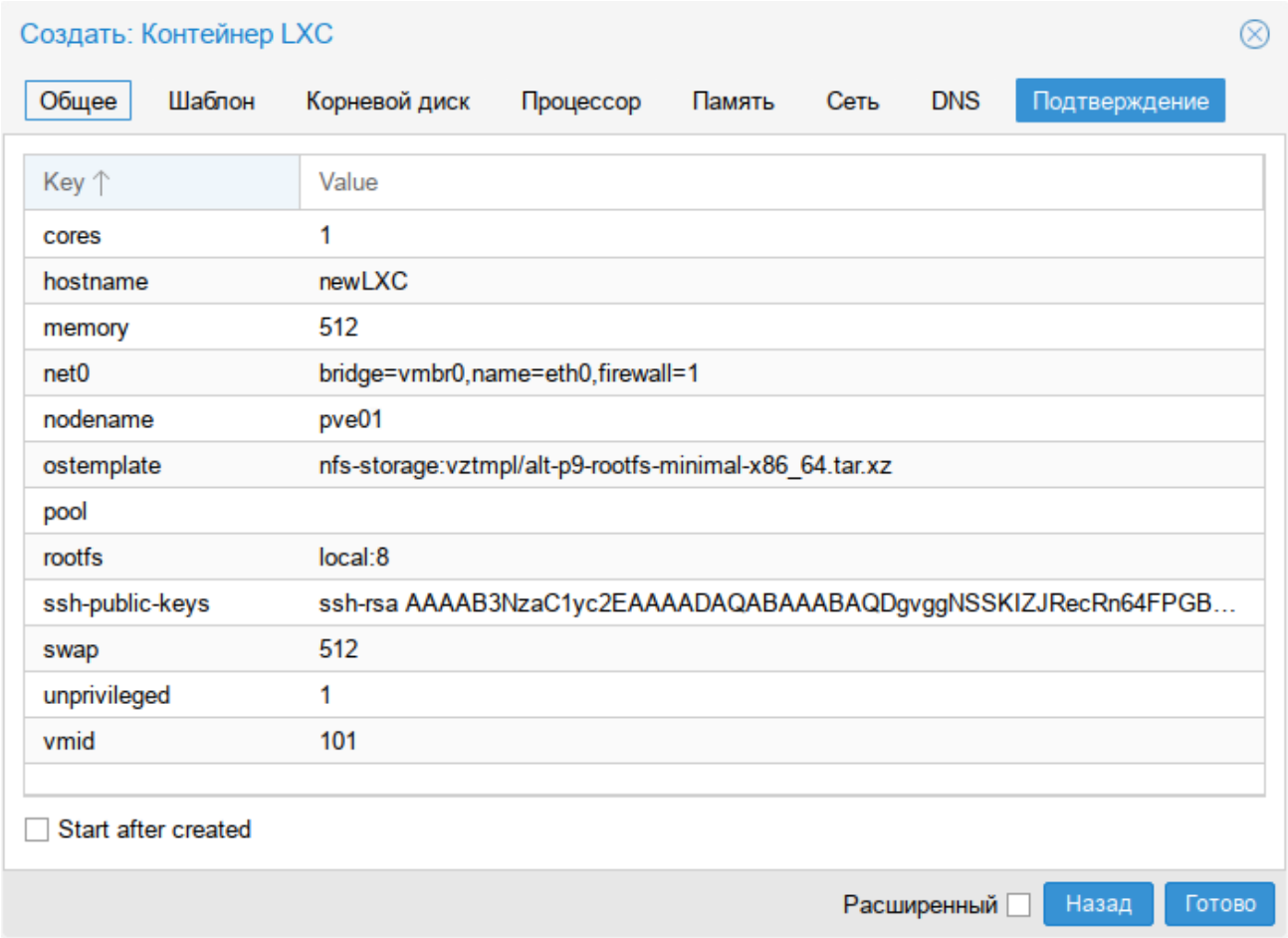

Рис. 112

Создание контейнера

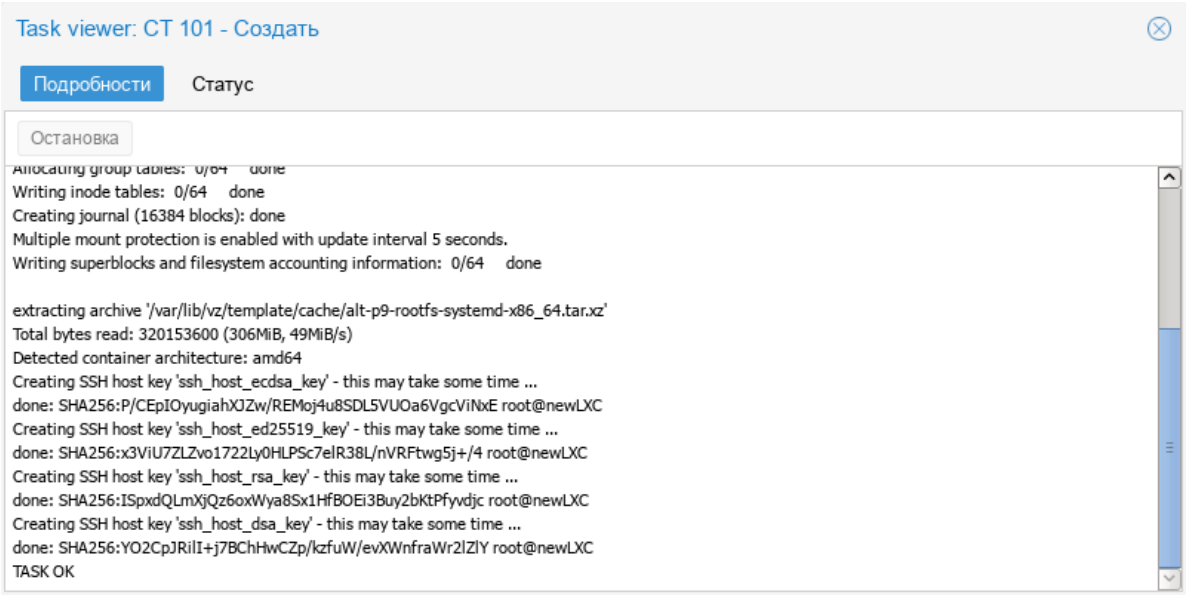

Рис. 113

4.7.2 Создание контейнера из шаблона в командной строке

Контейнер также может быть создан из шаблона при помощи командной строки хоста PVE с использованием команды `pct`.

Следующий сценарий bash иллюстрирует применение такой команды для создания контейнера:

```
#!/bin/bash
#### Set Variables ####
$hostname="pve01"
$vmid="100"
$template-path="/var/lib/vz/template/cache"
$storage="local"
$description="alt-p9"
$template="alt-p9-rootfs-systemd-x86_64.tar.xz"
$ip="192.168.0.186/24"
$nameserver="8.8.8.8"
$ram="1024"
$rootpw="changeme"
$rootfs="4"
$gateway="192.168.0.1"
$bridge="vmbr0"
$if="eth0"
#### Execute pct create using variable substitution ####
```

```
pct create $vmid \
  $template-path/$template \
  -description $description \
  -rootfs $rootfs \
  -hostname $hostname \
  -memory $ram \
  -nameserver $nameserver \
  -storage $storage \
  -password $rootpw \
  -net0 name=$if,ip=$ip,gw=$gateway,bridge=$bridge
```

4.7.3 Изменение настроек контейнера

Изменения в настройки контейнера можно вносить и после его создания. При этом изменения сразу же вступают в действие, без необходимости перезагрузки контейнера.

Есть три способа, которыми можно регулировать выделяемые контейнеру ресурсы:

- веб-интерфейс PVE;
- командная строка;
- изменение файла настройки.

4.7.3.1 Изменение настроек в веб-интерфейсе

В большинстве случаев изменения настройки контейнера и добавление виртуальных устройств может быть выполнено в веб-интерфейсе.

Для изменения ресурсов контейнера можно использовать три вкладки (Рис. 114):

- «Ресурсы» (оперативная память, подкачка, количество ядер ЦПУ, размер диска);
- «Сеть»;
- «DNS».

Изменений настроек контейнера в веб-интерфейсе PVE

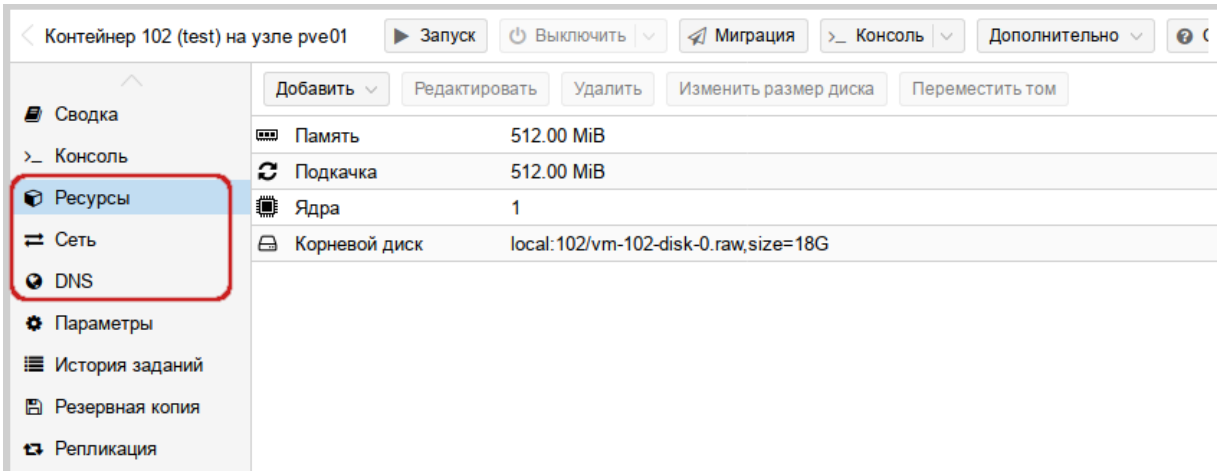

Рис. 114

Для редактирования ресурсов или добавления устройств, следует выполнить следующие действия:

1) в режиме просмотра по серверам выбрать необходимый контейнер;

2) перейти на вкладку «Ресурсы» («Resource»);

3) выбрать элемент для изменения: «Память» («Memory»), «Подкачка» («Swap»), «Ядра» («CPU units») или «Корневой диск» («Root Disk»), и нажать кнопку «Редактировать» («Edit»);

4) в открывшемся диалоговом окне ввести нужные значения и нажать кнопку «OK».

Если, например, необходимо увеличить размер диска контейнера до 18 ГБ вместо предварительно созданного 8 ГБ, нужно нажать кнопку «Изменить размер диска» («Resize disk») и в открывшемся диалоговом окне (Рис. 115) и ввести значение увеличения размера диска.

Изменение размера диска

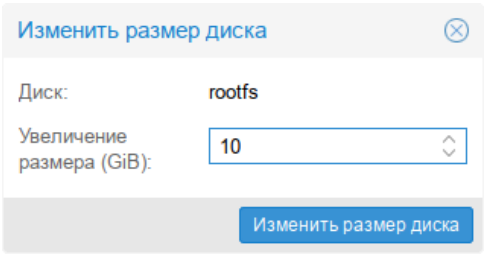

Рис. 115

Для изменения сетевых настроек контейнера необходимо:

1) в режиме просмотра по серверам выбрать необходимый контейнер;

2) перейти на вкладку «Сеть» («Network») для выбранного в левой панели навигации контейнера. На экране отобразятся все настроенные для контейнера виртуальные сетевые интерфейсы (Рис. 116);

3) выбрать нужный интерфейс, и нажать кнопку «Редактировать» («Edit») для внесения изменений (Рис. 117);

4) после выполнения необходимых изменений нажать кнопку «OK» для принятия введенных значений.

Виртуальные сетевые интерфейсы контейнера

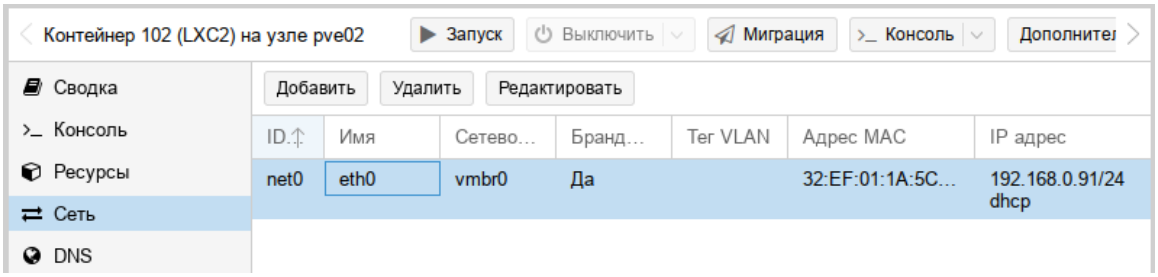

Рис. 116

Изменение сетевых настроек контейнера

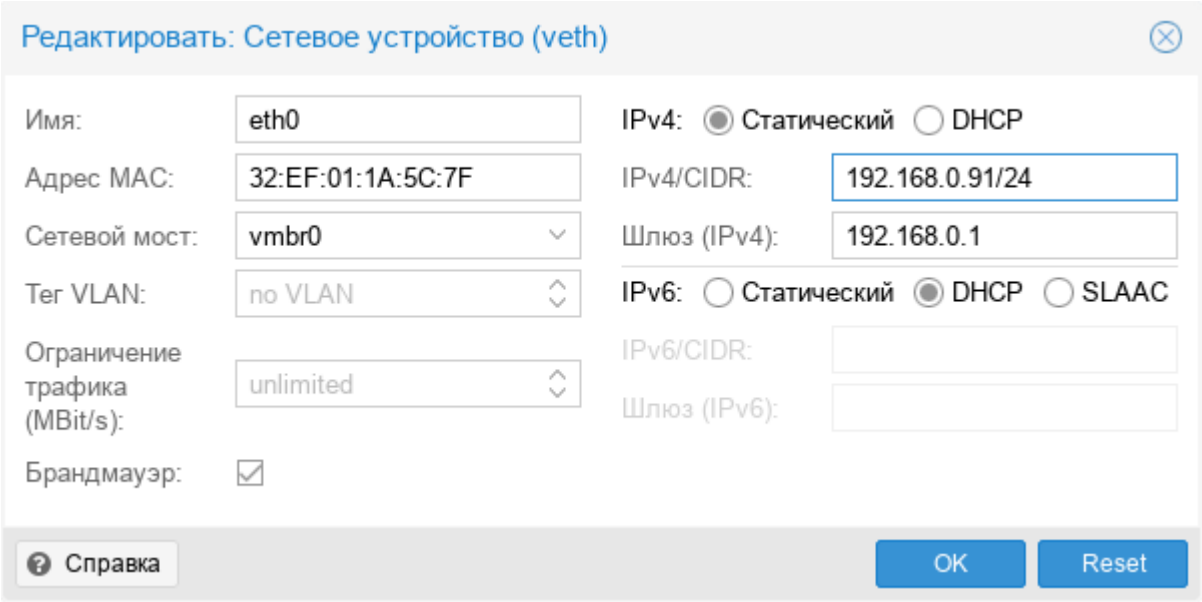

Рис. 117

4.7.3.2 Настройка ресурсов в командной строке

Если веб-интерфейс PVE недоступен, можно управлять контейнером при помощи командной строки (либо через сеанс SSH, либо из консоли noVNC, или зарегистрировавшись на физическом хосте).

`pct` – утилита управления контейнерами LXC в PVE. Чтобы просмотреть доступные для контейнеров команды PVE, можно выполнить следующую команду:

```
# pct help
```

Выполняемое командой `pct` изменение ресурсов фиксируется в контейнере немедленно, без необходимости перезапуска этого контейнера. Формат использования команды для изменения ресурсов контейнера:

```
# pct set <ct_id> [options]
```

Например, если необходимо изменить адрес IP контейнера #101, команда должна быть такой:

```
# pct set 101 –net0 name=eth0,bridge=vmbr0,ip=192.168.0.17/24,gw=192.168.0.1
```

Чтобы изменить выделение памяти контейнеру в реальном масштабе времени можно использовать следующую команду:

```
# pct set <ct_id> -memory <int_value>
```

Следующая команда изменяет имя хоста данного контейнера:

```
# pct set <ct_id> -hostname <string>
```

Команда увеличения размера диска данного контейнера:

```
# pct set <ct_id> -rootfs size=<int_value for GB>
```

Состоянием контейнера также можно управлять из командной строки PVE.

Разблокировка заблокированного контейнера в командной строке:

```
# pct set <ct_id> -unlock
```

Список контейнеров LXC данного узла:

```
# pct list
VMID       Status     Lock         Name
102        stopped                 LXC2
```

Запуск и останов контейнера LXC из командной строки:

```
# pct start <ct_id>
# pct stop <ct_id>
```

4.7.3.3 Настройка ресурсов прямым изменением

В PVE файлы конфигурации контейнеров находятся в каталоге `/etc/pve/lxc`, а файлы конфигураций ВМ – в `/etc/pve/qemu-server/`.

Примечание. Выполнять редактирование файла настройки контейнера LXC не рекомендуется, несмотря на то, что это возможно.

Любые выполненные вручную изменения в файлах не применяются, пока контейнер не будет перезапущен. Однако существуют ситуации, при которых изменение настроек вручную необходимо. Контейнеры LXC имеют большое число параметров, которые не могут быть изменены в веб-интерфейсе или с помощью утилиты `pct`. Такие параметры могут быть применены только путем изменения в файлах настройки с последующим перезапуском таких контейнеров.

Пример файла настройки `/etc/pve/lxc/102.conf`:

```
arch: amd64
cores: 1
hostname: LXC2
memory: 512
```

```
net0:
name=eth0,bridge=vmbr0,firewall=1,gw=192.168.0.1,hwaddr=32:EF:01:1A:5C
:7F,ip=192.168.0.91/24,ip6=dhcp,type=veth
ostype: altlinux
rootfs: local:102/vm-102-disk-0.raw,size=8G
swap: 512
unprivileged: 1
```

4.7.4 Запуск и останов контейнеров

4.7.4.1 Изменение состояния контейнера в веб-интерфейсе

Для запуска контейнера следует выбрать его в левой панели; его иконка должна быть серого цвета, обозначая, что контейнер не запущен (Рис. 118).

Запустить контейнер можно выбрав в контекстном меню контейнера пункт «Запуск» (Рис. 118), либо нажав на кнопку «Запуск» («Start») (Рис. 119).

Запущенный контейнер будет обозначен зеленой стрелкой на значке контейнера.

Для запущенного контейнера доступны следующие действия (Рис. 119):

– «Выключить» («Shutdown») – остановка контейнера;
– «Остановка» («Stop») – остановка контейнера, путем прерывания его работы;
– «Перезапустить» («Reboot») – перезапуск контейнера.

Контекстное меню контейнера

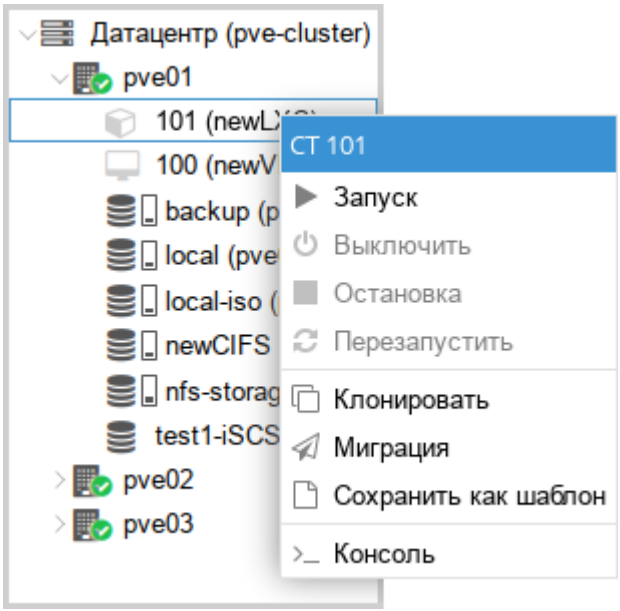

Рис. 118

Кнопки управления состоянием контейнера

Рис. 119

4.7.4.2 Изменение состояний контейнера в командной строке

Состоянием контейнера можно управлять из командной строки PVE (либо через сеанс SSH, либо из консоли noVNC, или зарегистрировавшись на физическом хосте).

Для запуска контейнера с VM ID 102 необходимо ввести команду:

```
# pct start 102
```

Этот же контейнер может быть остановлен при помощи команды:

```
# pct stop 102
```

4.7.5 Доступ к LXC контейнеру

Есть несколько вариантов при помощи, которых, может быть осуществлен доступ к LXC контейнеру:

- консоль: noVNC, SPICE или xterm.js;
- SSH;
- интерфейс командной строки PVE.

Можно получить доступ к контейнеру из веб-интерфейса при помощи консоли noVNC. Это почти визуализированный удаленный доступ к экземпляру.

Для доступа к запущенному контейнеру в консоли следует выбрать в веб-интерфейсе нужный контейнер, а затем нажать кнопку «Консоль» («Console») и в выпадающем меню выбрать нужную консоль (Рис. 120).

Кнопка «Консоль»

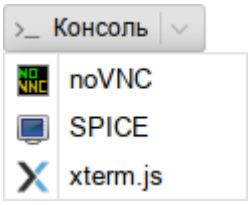

Рис. 120

Консоль также можно запустить, выбрав вкладку «Консоль» («Console») для контейнера (Рис. 121).

Консоль

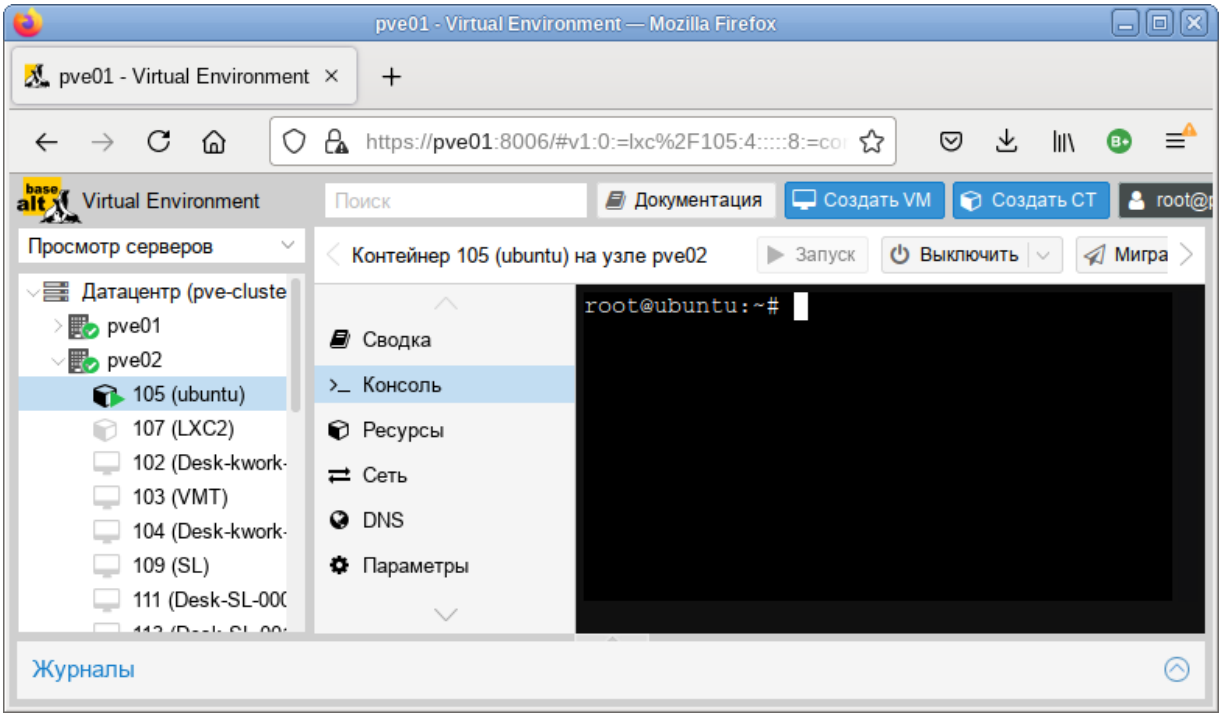

Рис. 121

Одной из функций LXC контейнера является возможность прямого доступа к оболочке контейнера через командную строку его узла хоста. Команда для доступа к оболочке контейнера LXC:

```
# pct enter <ct_id>
```

Данная команда предоставляет прямой доступ на ввод команд внутри указанного контейнера:

```
[root@pve01 ~]# pct enter 101
[root@newLXC ~]#
```

Таким образом был получен доступ к контейнеру LXC с именем newLXC на узле pve01. При этом для входа в контейнер не был запрошен пароль. Так как контейнер работает под пользователем root, можно выполнять внутри этого контейнера любые задачи. Завершив их, можно просто набрать `exit` для возвращения назад в свой узел из данного контейнера.

Примечание. При возникновении ошибки:

```
Insecure $ENV{ENV} while running with...
```

необходимо закомментировать строку: `"ENV=$HOME/.bashrc"` в файле `/root/.bashrc`.

Также можно выполнять внутри контейнера различные команды без реального входа в такой контейнер. Следующий формат команды применяется для выполнения команд внутри некоего контейнера:

```
# pct exec <ct_id> -- <command>
```

Например, если нужно создать каталог внутри контейнера и проверить что этот каталог был создан, команды и их вывод будут следующими:

```
# pct exec 101 mkdir /home/demouser
# pct exec 101 ls /home
demouser
```

Для выполнения внутри контейнера команды с параметрами необходимо изменить команду `pct`, добавив `--` после идентификатора контейнера:

```
# pct exec 101 -- df -H /
Filesystem      Size  Used Avail Use% Mounted on
/dev/loop0       19G  402M   18G   3% /
none            504k     0  504k   0% /dev
udevfs          5.3M     0  5.3M   0% /dev/tty
tmpfs           1.1G     0  1.1G   0% /dev/shm
tmpfs           1.1G   82k  1.1G   1% /run
tmpfs           5.3M     0  5.3M   0% /run/lock
tmpfs           1.1G     0  1.1G   0% /sys/fs/cgroup
tmpfs           1.1G     0  1.1G   0% /tmp
```

4.8 Миграция виртуальных машин и контейнеров

В случае, когда PVE управляет не одним физическим узлом, а кластером физических узлов, должна обеспечиваться возможность миграции ВМ с одного физического узла на другой. Миграция представляет собой заморозку состояния ВМ на одном узле, перенос данных и конфигурации на другой узел, и разморозку состояния ВМ на новом месте. Возможные сценарии, при которых может возникнуть необходимость миграции:

- отказ физического узла;
- необходимость перезагрузки узла после применения обновлений или обслуживания технических средств;
- перемещение ВМ с узла с низкой производительностью на высокопроизводительный узел.

Есть два механизма миграции:

- онлайн-миграция (Live Migration);
- автономная миграция.

Примечание. Миграция контейнеров без перезапуска в настоящее время не поддерживается. При выполнении миграции запущенного контейнера, контейнер будет выключен, перемещен, а

затем снова запущен на целевом узле. Поскольку контейнеры очень легкие, это обычно приводит к простою в несколько сотен миллисекунд.

Для возможности Live Migration необходимы следующие условия:

- ВМ не имеет локальных ресурсов;
- хосты находятся в одном кластере PVE;
- хосты имеют работающее (и надежное) сетевое соединение;
- целевой хост должен иметь одинаковые или более высокие версии пакетов PVE.

Миграция в реальном времени обеспечивает максимальное время работы, но, в то же время медленнее. Причина в том, что при миграции в реальном времени без выключения питания процесс должен скопировать все содержимое оперативной памяти ВМ на новый узел. Чем больше объем выделенной ВМ памяти, тем дольше будет происходить ее перенос.

Если образ виртуального диска ВМ хранится в локальном хранилище узла PVE миграция в реальном времени не возможна. В этом случае ВМ должна быть перед миграцией выключена. В процессе миграции ВМ, хранящейся локально, PVE будет копировать весь виртуальный диск на узел получателя с применением rsync.

Запустить процесс миграции можно как в графическом интерфейсе PVE, так в интерфейсе командной строки.

4.8.1 Миграция с применением графического интерфейса

Для миграции ВМ или контейнера необходимо выполнить следующие шаги:

1) выбрать ВМ или контейнер для миграции и нажать кнопку «Миграция» («Migrate») (Рис. 122);

2) в открывшемся диалоговом окне (Рис. 123) выбрать узел назначения, на который будет осуществляться миграция, и нажать кнопку «Миграция» («Migrate») для запуска процесса миграции.

Примечание. Режим миграции будет выбран автоматически (*Рис. 123, Рис. 124, Рис. 125*) в зависимости от состояния ВМ/контейнера (запущен/остановлен).

Выбор ВМ или контейнера для миграции

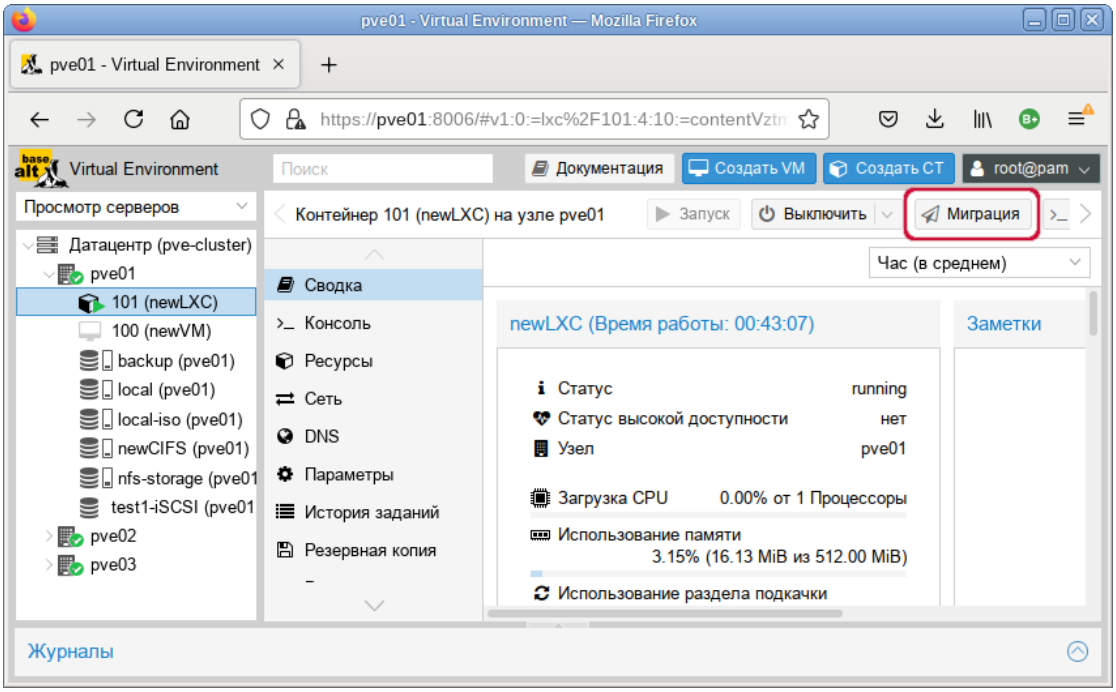

Рис. 122

Миграция контейнера с перезапуском

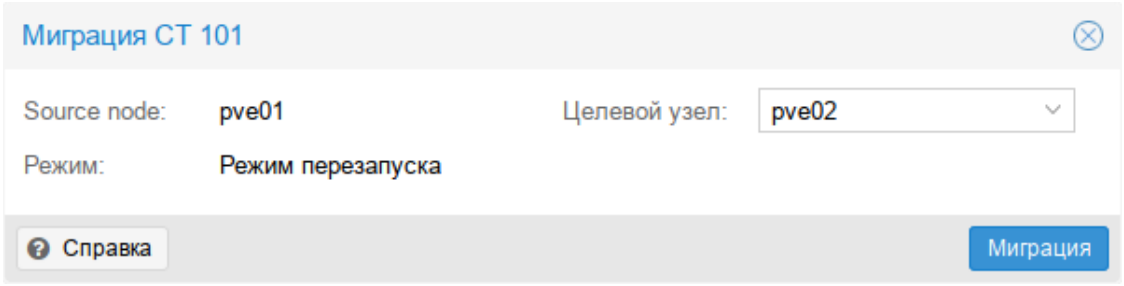

Рис. 123

Миграция ВМ Онлайн

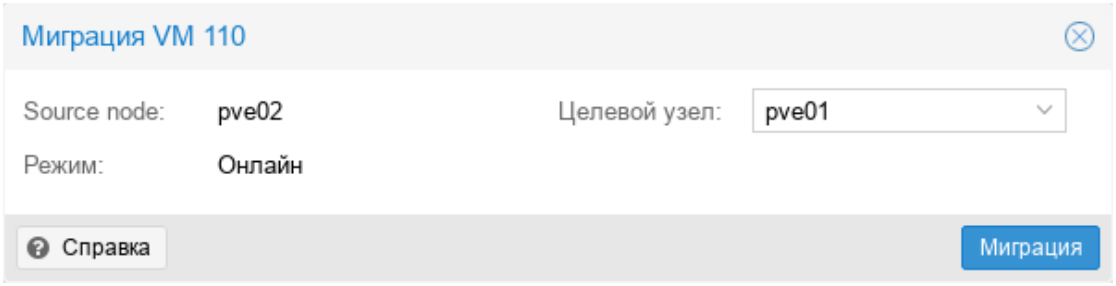

Рис. 124

Миграция ВМ Офлайн (миграция диска из локального хранилища может занять много времени)

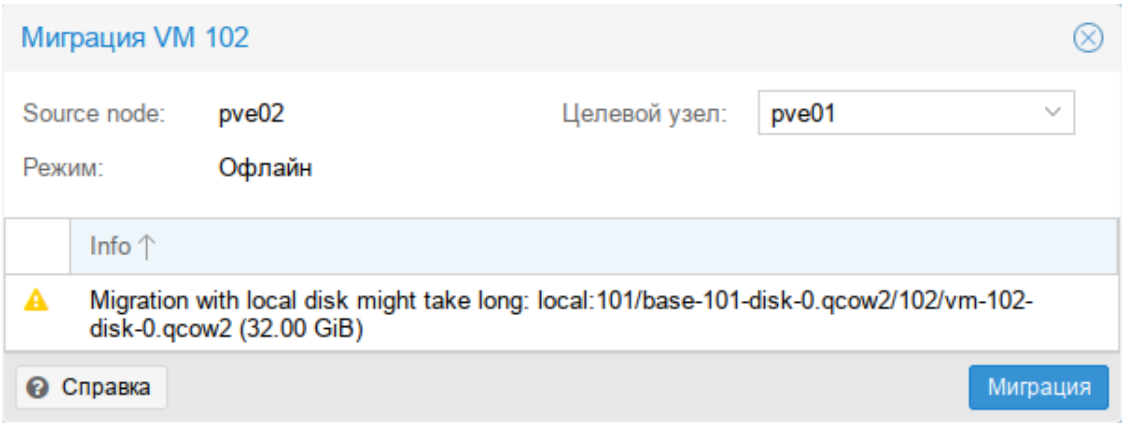

Рис. 125

4.8.2 Миграция с применением командной строки

Чтобы осуществить миграцию ВМ с применением интерфейса командной строки необходимо выполнить следующую команду:

```
# qm migrate <vmid> <target> [OPTIONS]
```

Для осуществления миграции ВМ в реальном времени необходимо использовать параметр `--online`.

Чтобы осуществить миграцию контейнера с применением интерфейса командной строки необходимо выполнить следующую команду:

```
# pct migrate <ctid> <target> [OPTIONS]
```

Поскольку миграция контейнеров в реальном времени не возможна, можно выполнить миграцию работающего контейнера с перезапуском, добавив параметр `--restart`. Например:

```
# pct migrate 101 pve02 -restart
```

4.8.3 Миграция ВМ из внешнего гипервизора

Экспорт ВМ из внешнего гипервизора обычно заключается в переносе одного или нескольких образов дисков с файлом конфигурации, описывающим настройки ВМ (ОЗУ, количество ядер). Образы дисков могут быть в формате vmdk (VMware или VirtualBox), или qcow2 (KVM).

4.8.3.1 Миграция KVM ВМ в PVE

В данном разделе рассмотрен процесс миграции ВМ из OpenNebula в PVE.

Выключить ВМ на хосте источнике. Найти путь до образа жесткого диска, который используется в ВМ (в данной команде 14 – id образа диска ВМ):

```
$ oneimage show 14
IMAGE 14 INFORMATION
ID             : 14
NAME           : ALT Linux p9
USER           : oneadmin
```

```
GROUP          : oneadmin
LOCK           : None
DATASTORE      : default
TYPE           : OS
REGISTER TIME  : 04/30 11:00:42
PERSISTENT     : Yes
SOURCE         : /var/lib/one//datastores/1/f811a893808a9d8f5bf1c029b3c7e905
FSTYPE         : save_as
SIZE           : 12G
STATE          : used
RUNNING_VMS    : 1

PERMISSIONS
OWNER          : um-
GROUP          : ---
OTHER          : ---

IMAGE TEMPLATE
DEV_PREFIX="vd"
DRIVER="qcow2"
SAVED_DISK_ID="0"
SAVED_IMAGE_ID="7"
SAVED_VM_ID="46"
SAVE_AS_HOT="YES"
```

где /var/lib/one//datastores/1/f811a893808a9d8f5bf1c029b3c7e905 – адрес образа жёсткого диска ВМ.

Скопировать данный образ на хост назначения с PVE.

Примечание. В OpenNebula любой диск ВМ можно экспортировать в новый образ (если ВМ находится в состояниях RUNNING, POWEROFF или SUSPENDED):

```
onevm disk-saveas <vmid> <diskid> <img_name>  [--type type --snapshot snapshot]
```

где --type <type> – тип нового образа (по умолчанию raw); --snapshot <snapshot_id> – снимок диска, который будет использован в качестве источника нового образа (по умолчанию текущее состояние диска).

Экспорт диска ВМ:

```
$ onevm disk-saveas 125 0 test.qcow2
Image ID: 44
```

Инфомация об образе диска ВМ:

```
$ oneimage show 44
MAGE 44 INFORMATION
ID             : 44
```

```
NAME           : test.qcow2
USER           : oneadmin
GROUP          : oneadmin
LOCK           : None
DATASTORE      : default
TYPE           : OS
REGISTER TIME  : 07/12 21:34:42
PERSISTENT     : No
SOURCE         : /var/lib/one//datastores/1/9d6336a88d6ab62ea1dce65d81e55881
FSTYPE         : save_as
SIZE           : 12G
STATE          : rdy
RUNNING_VMS    : 0

PERMISSIONS
OWNER          : um-
GROUP          : ---
OTHER          : ---

IMAGE TEMPLATE
DEV_PREFIX="vd"
DRIVER="qcow2"
SAVED_DISK_ID="0"
SAVED_IMAGE_ID="14"
SAVED_VM_ID="125"
SAVE_AS_HOT="YES"

VIRTUAL MACHINES
```

Информация о диске:

```
$ qemu-img info /var/lib/one//datastores/1/9d6336a88d6ab62ea1dce65d81e55881
image: /var/lib/one//datastores/1/9d6336a88d6ab62ea1dce65d81e55881
file format: qcow2
virtual size: 12 GiB (12884901888 bytes)
disk size: 3.52 GiB
cluster_size: 65536
Format specific information:
    compat: 1.1
    compression type: zlib
    lazy refcounts: false
    refcount bits: 16
    corrupt: false
    extended l2: false
```

На хосте нахзначения подключить образ диска к ВМ (рассмотрено подключение на основе Directory Storage):

1) Создать новую ВМ в веб-интерфейсе PVE или командой:

```
# qm create 120 --bootdisk scsi0 --net0 virtio,bridge=vmbr0 --scsihw virtio-scsi-pci
```

2) Чтобы использовать в PVE образ диска в формате qcow2 (полученный из другой системы KVM, либо преобразованный из другого формата), его необходимо импортировать. Команда импорта:

```
qm importdisk <vmid> <source> <storage> [OPTIONS]
```

Команда импорта диска f811a893808a9d8f5bf1c029b3c7e905 в хранилище local, для ВМ с ID 120 (подразумевается, что образ импортируемого диска находится в каталоге, из которого происходит выполнение команды):

```
# qm importdisk 120 f811a893808a9d8f5bf1c029b3c7e905 local --format qcow2
importing disk 'f811a893808a9d8f5bf1c029b3c7e905' to VM 120 ...
…
Successfully imported disk as 'unused0:local:120/vm-120-disk-0.qcow2'
```

3) Привязать диск к ВМ.

В веб-интерфейсе PVE: перейти на вкладку «Оборудование», созданной ВМ. В списке устройств будет показан неиспользуемый жесткий диск, выбрать его, выбрать режим «SCSI» и нажать кнопку «Добавить» (*Рис. 126*).

В командной строке:

```
# qm set 120 --scsi0 local:120/vm-120-disk-0.qcow2
update VM 120: -scsi0 local:120/vm-120-disk-0.qcow2
```

Донастроить параметры процессора, памяти, сетевых интерфейсов, порядок загрузки. Включить ВМ.

Добавление диска к ВМ

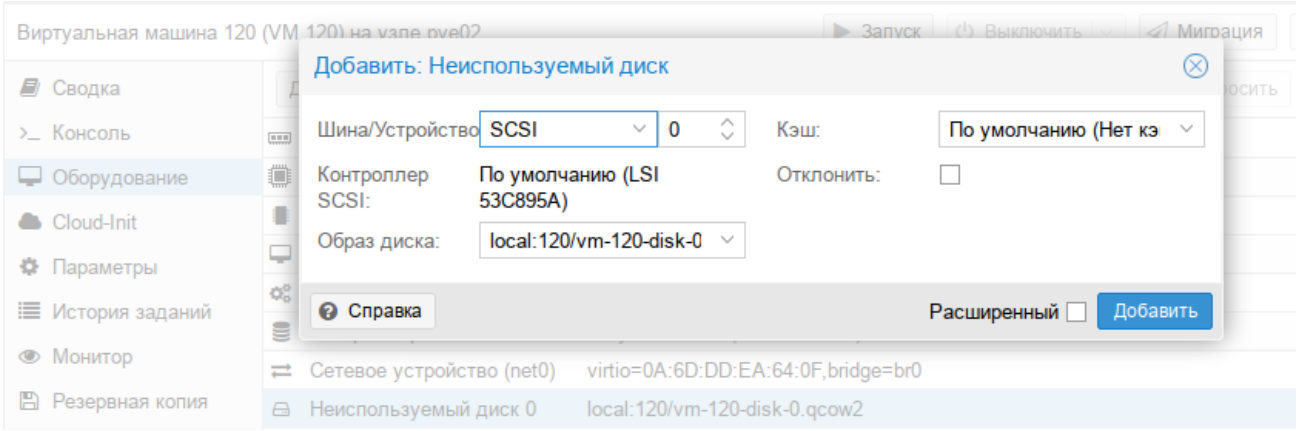

Рис. 126

4.8.3.2 Миграция ВМ из VMware в PVE

Экспорт ВМ из внешнего гипервизора обычно заключается в переносе одного или нескольких образов дисков с файлом конфигурации, описывающим настройки ВМ (ОЗУ, количество ядер). Образы дисков могут быть в формате vmdk (VMware или VirtualBox), или qcow2 (KVM).

В данном разделе рассмотрена миграция ВМ из VMware в PVE, на примере ВМ с ОС Windows 7.

Подготовить ОС Windows. ОС Windows должна загружаться с дисков в режиме IDE.

Подготовить образ диска. Необходимо преобразовать образ диска в тип `single growable virtual disk`. Сделать это можно с помощью утилиты `vmware-vdiskmanager` (поставляется в комплекте с VMWare Workstation). Для преобразования образа перейти в папку с образами дисков и выполнить команду:

```
"C:\Program Files\VMware\VMware Server\vmware-vdiskmanager"
-r win7.vmdk -t 0 win7-pve.vmdk
```

где win7.vmdk – файл с образом диска.

Подключить образ диска к ВМ одним из трёх указанных способов:

4) Подключение образа диска к ВМ на основе Directory Storage:

- в веб-интерфейсе PVE создать новую ВМ KVM;
- скопировать преобразованный образ win7-pve.vmdk в каталог с образами ВМ /var/lib/vz/images/VMID, где VMID – VMID, созданной виртуальной машины (можно воспользоваться WinSCP);
- преобразовать файл win7-pve.vmdk в qemu формат:

```
# qemu-img convert -f vmdk win7-pve.vmdk -O qcow2 win7-pve.qcow2
```

- добавить в конфигурационный файл ВМ (`/etc/pve/nodes/pve02/qemu-server/VMID.conf`) строку:

```
unused0: local:100/win7-pve.qcow2
```

где 100 – VMID, a local – хранилище в PVE.

- перейти в веб-интерфейсе PVE на вкладку «Оборудование», созданной ВМ. В списке устройств будет показан неиспользуемый жесткий диск, выбрать его, выбрать режим IDE и нажать кнопку «Добавить» (*Рис. 128*).

Добавление диска к ВМ

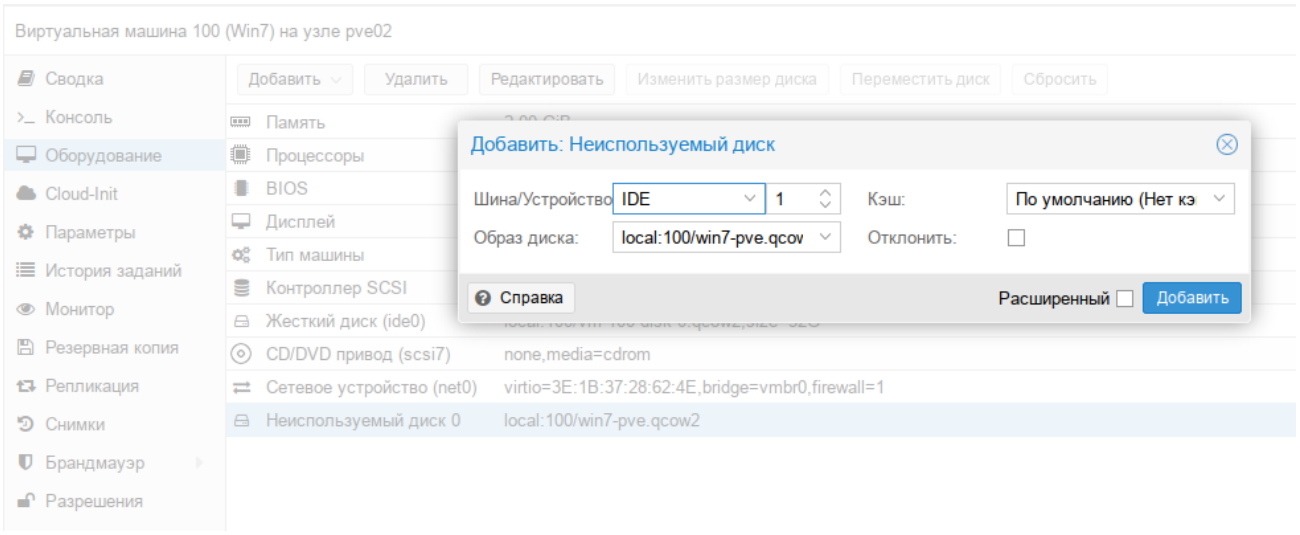

Рис. 127

5) Подключение образа диска к ВМ на основе LVM Storage:

– в веб-интерфейсе PVE создать новую ВМ с диском большего размера, чем диск в образе vmdk. Посмотреть размер диска в образе можно командой:

```
# qemu-img info win7-pve.vmdk
image: win7-pve.vmdk
file format: vmdk
virtual size: 127G (136365211648 bytes)
disk size: 20.7 GiB
cluster_size: 65536
Format specific information:
    cid: 3274246794
    parent cid: 4294967295
    create type: streamOptimized
    extents:
        [0]:
        compressed: true
        virtual size: 136365211648
        filename: win7-pve.vmdk
        cluster size: 65536
        format:
```

В данном случае необходимо создать диск в режиме IDE размером не меньше 127GB.

- скопировать преобразованный образ win7-pve.vmdk в каталог с образами ВМ /var/lib/vz/images/VMID, где VMID – VMID, созданной виртуальной машины (можно воспользоваться WinSCP);
- перейти в консоль ноды кластера и посмотреть, как называется LVM диск созданной ВМ (он должен быть в статусе ACTIVE):

```
# lvscan
ACTIVE            '/dev/sharedsv/vm-101-disk-1' [130,00 GiB] inherit
```

- сконвертировать образ vdmk в raw формат непосредственно на LVM-устройство:

```
# qemu-img convert -f vmdk win7-pve.vmdk -O raw /dev/sharedsv/vm-101-disk-1
```

6) Подключение образа диска к ВМ на основе CEPH Storage:

- в веб-интерфейсе PVE создать новую ВМ с диском большего размера, чем диск в образе vmdk. Посмотреть размер диска в образе можно командой:

```
# qemu-img info win7-pve.vmdk
```

- скопировать преобразованный образ win7-pve.vmdk в каталог с образами ВМ /var/lib/vz/images/VMID, где VMID – VMID, созданной виртуальной машины;
- перейти в консоль ноды кластера. Отобразить образ из пула CEPH в локальное блочное устройство:

```
# rbd map rbd01/vm-100-disk-1
/dev/rbd0
```

Примечание. Имя нужного пула можно посмотреть на вкладке «Датацентр» → «Хранилище» → «rbd-storage».

- сконвертировать образ vdmk в raw формат непосредственно на отображенное устройство:

```
# qemu-img convert -f vmdk win7-pve.vmdk -O raw /dev/rbd0
```

Адаптация новой ВМ:

- Проверить режим работы жесткого диска: для Windows – IDE, для Linux – SCSI.
- Режим VIRTIO жесткого диска (режим VIRTIO также доступен для Windows, но сразу загрузиться в этом режиме система не может):
  - загрузиться сначала в режиме IDE и выключить машину, добавить еще один диск в режиме VIRTIO и включить машину. Windows установит нужные драйвера;
  - выключить машину;
  - изменить режим основного диска с IDE на VIRTIO;
  - загрузить систему, которая должна применить VIRTIO драйвер и выдать сообщение, что драйвер от RedHat.

– Включить ВМ. Первое включение займет какое-то время (будут загружены необходимые драйвера).

4.9 Клонирование виртуальных машин

Простой способ развернуть множество ВМ одного типа – скопировать (создать клон) ВМ.

Существует два вида клонов:

– Полный клон – результатом такой копии является независимая ВМ. Новая ВМ не разделяет ресурсы хранения с оригинальными. При таком клонировании можно выбрать целевое хранилище, поэтому его можно использовать для переноса ВМ в совершенно другое хранилище. Также можно изменить формат образа диска, если драйвер хранилища поддерживает несколько форматов.
– Связанный клон – такой клон является записываемой копией, исходное содержимое которой совпадает с исходными данными. Создание связанного клона происходит практически мгновенно и изначально не требует дополнительного места. Клоны называются связанными, потому что новое изображение ссылается на оригинал. Немодифицированные блоки данных считываются из исходного изображения, но изменения записываются (и затем считываются) из нового местоположения. При этом требуется, чтобы исходный том работал в режиме только для чтения. С помощью PVE можно преобразовать любую ВМ в шаблон (см. ниже). Такие шаблоны впоследствии могут быть использованы для эффективного создания связанных клонов. Для связанных клонов невозможно изменить целевое хранилище, поскольку это внутренняя функция хранилища.

Примечание. Полному клону необходимо прочитать и скопировать все данные образа ВМ. Это обычно намного медленнее, чем создание связанного клона.

Весь функционал клонирования доступен из графического интерфейса PVE.

Для клонирования ВМ необходимо выполнить следующие шаги:

1) создать ВМ с необходимыми настройками (все создаваемые из такой ВМ клоны будут иметь идентичные настройки) или воспользоваться уже существующей ВМ;
2) в контекстном меню выбранной ВМ выбрать пункт «Клонировать» («Clone») (Рис. 128);
3) откроется диалоговое окно (Рис. 129), со следующими полями:

– «Целевой узел» («Target node») – узел получатель клонируемой ВМ (для создания новой ВМ на другом узле необходимо чтобы ВМ находилась в общем хранилище и это хранилище должно быть доступно на целевом узле);
– «VM ID» – идентификатор ВМ;
– «Имя» («Name») – название ВМ;
– «Пул ресурсов»(«Resource Pool») – пул, к которому будет относиться ВМ;

– «Режим» («Mode») – метод клонирования (если клонирование происходит из шаблона ВМ). Доступными параметрами являются «Полное клонирование» («Full Clone») и «Связанная копия» («Linked Clone»);
– «Снимок» («Snapshot») – снимок из которого будет создаваться клон (если снимки существуют). Если существует несколько снимков, доступных для данной виртуальной машины, из них можно выбирать на основе выпадающего меню;
– «Целевое хранилище» («Target Storage») – хранилище для клонируемых виртуальных дисков;
– «Формат» («Format») – формат образа виртуального диска.

4) для запуска процесса клонирования необходимо нажать кнопку «Клонировать».

Контекстное меню ВМ

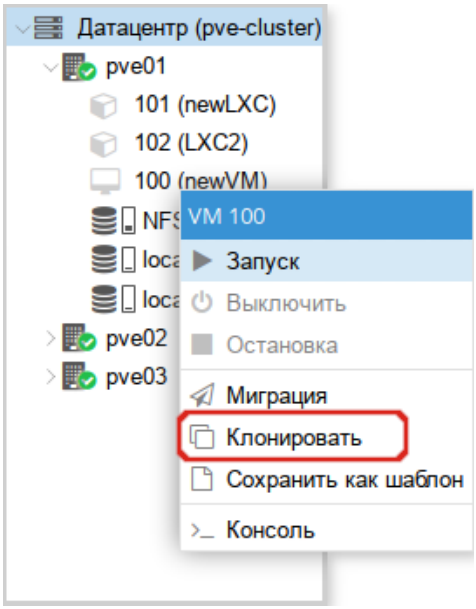

Рис. 128

Настройки клонирования

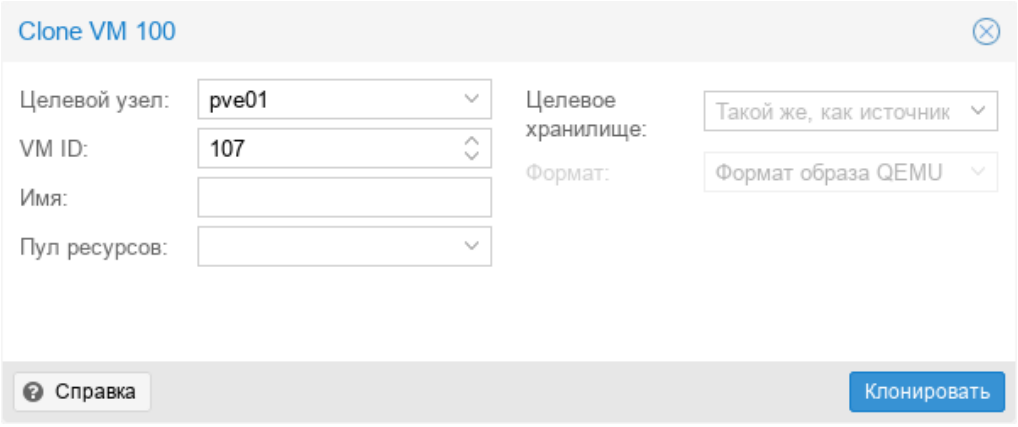

Рис. 129

Некоторые типы хранилищ позволяют копировать определенный снимок (Рис. 130), который по умолчанию соответствует текущим данным ВМ. Клон ВМ никогда не содержит дополнительных снимков оригинальной ВМ.

Выбор снимка для клонирования

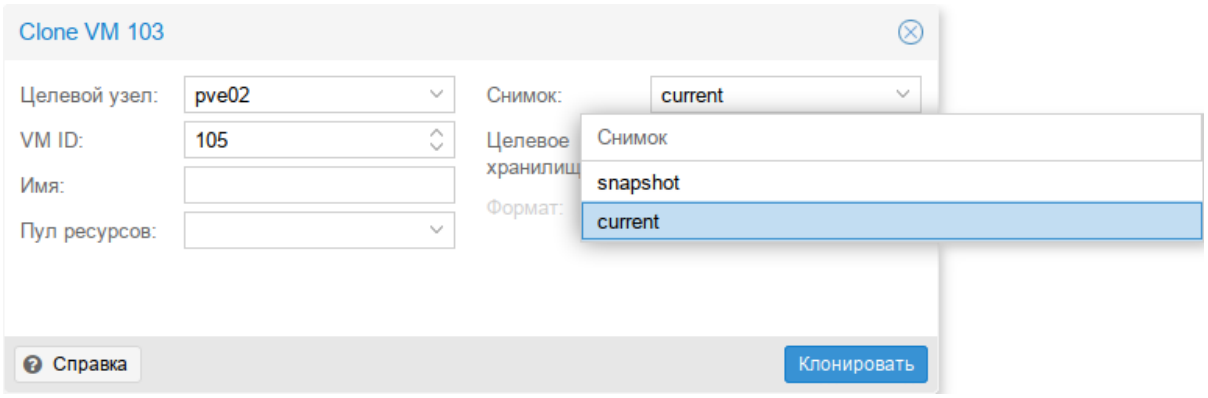

Рис. 130

ВМ можно преобразовать в шаблон. Такие шаблоны доступны только для чтения, и их можно использовать для создания связанных клонов.

Для преобразования ВМ в шаблон необходимо в контекстном меню ВМ выбрать пункт «Сохранить как шаблон» («Convert to template») и в ответ на запрос на подтверждения, нажать кнопку «Да».

Примечание. Запустить шаблоны невозможно, так как это приведет к изменению образов дисков. Если необходимо изменить шаблон, следует создать связанный клон и изменить его.

4.10 Резервное копирование (backup)

PVE предоставляет полностью интегрированное решение, использующее возможности всех хранилищ и всех типов гостевых систем.

Резервные копии PVE представляют собой полные резервные копии – они содержат конфигурацию ВМ/CT и все данные. Резервное копирование может быть запущено через графический интерфейс или с помощью утилиты командной строки `vzdump`.

4.10.1 Алгоритмы резервного копирования

Инструментарий для создания резервных копий PVE поддерживает следующие механизмы сжатия:

– Сжатие LZO – алгоритм сжатия данных без потерь (реализуется в PVE утилитой `lzop`). Основной особенностью этого алгоритма является скоростная распаковка. Следовательно, любая резервная копия, созданная с помощью этого алгоритма, может при необходимости быть развернута за минимальное время.

- Сжатие GZIP – при использовании этого алгоритма резервная копия будет «на лету» сжиматься утилитой GNU Zip, использующей мощный алгоритм Deflate. Основной упор делается на максимальное сжатие данных, что позволяет сократить место на диске, занимаемое резервными копиями. Главным отличием от LZO является то, что процедуры компрессии/декомпрессии занимают достаточно большое количество времени.
- Сжатие Zstandard (zstd) – алгоритм сжатия данных без потерь. В настоящее время Zstandard является самым быстрым из этих трех алгоритмов. Многопоточность – еще одно преимущество zstd перед lzo и gzip.

4.10.2 Режимы резервного копирования

Режимы резервного копирования для виртуальных машин:

- режим остановки (Stop) – обеспечивает самую высокую надежность резервного копирования, но требует полного выключения ВМ. В этом режиме ВМ отправляется команда на штатное выключение, после остановки выполняется резервное копирование и затем отдается команда на включение ВМ. Количество ошибок при таком подходе минимально и чаще всего сводится к нулю;
- режим приостановки (Suspend) – ВМ временно «замораживает» свое состояние, до окончания процесса резервного копирования. Содержимое оперативной памяти не стирается, что позволяет продолжить работу ровно с той точки, на которой работа была приостановлена. Сервер простаивает во время копирования информации, но при этом нет необходимости выключения/включения ВМ, что достаточно критично для некоторых сервисов. Особенно, если запуск части сервисов не является автоматическим. Такие резервные копии следует разворачивать в тестовой среде для проверки;
- режим снимка (Snapshot) – использование этого механизма не прерывает работу ВМ (Live backup), но имеет два очень серьезных недостатка – могут возникать проблемы из-за блокировок файлов операционной системой и самая низкая скорость создания. Резервные копии, созданные этим методом, надо всегда проверять в тестовой среде.

Режимы резервного копирования для контейнеров:

- режим остановки (Stop) – остановка контейнера на время резервного копирования. Это может привести к длительному простою;
- режим приостановки (Suspend) – этот режим использует `rsync` для копирования данных контейнера во временную папку (опция `--tmpdir`). Затем контейнер приостанавливается и `rsync` копирует измененные файлы. После этого контейнер возобновляет свою работу. Это приводит к минимальному времени простоя, но требует дополнительное пространство для хранения копии контейнера. Когда контейнер находится в локальной файловой системе

и хранилищем резервной копии является сервер NFS, необходимо установить `--tmpdir` также и в локальной файловой системе, так как это приведет к повышению производительности. Использование локального `tmpdir` также необходимо, если требуется сделать резервную копию локального контейнера с использованием списков контроля доступа (ACL) в режиме ожидания, если хранилище резервных копий – сервер NFS.

– режим снимка (Snapshot) – этот режим использует возможности мгновенных снимков основного хранилища. Сначала, контейнер будет приостановлен для обеспечения согласованности данных, будет сделан временный снимок томов контейнера, а содержимое снимка будет заархивировано в tar-файле, далее временный снимок удаляется.

4.10.3 Резервное хранилище

Перед тем, как настроить резервное копирование, необходимо определить хранилище резервных копий. Хранилище резервных копий должно быть хранилищем уровня файлов, так как резервные копии хранятся в виде обычных файлов. В большинстве случаев можно использовать сервер NFS для хранения резервных копий. Если хранилище будет использоваться только для резервных копий, следует выставить соответствующие настройки (Рис. 131).

Настройка хранилища NFS

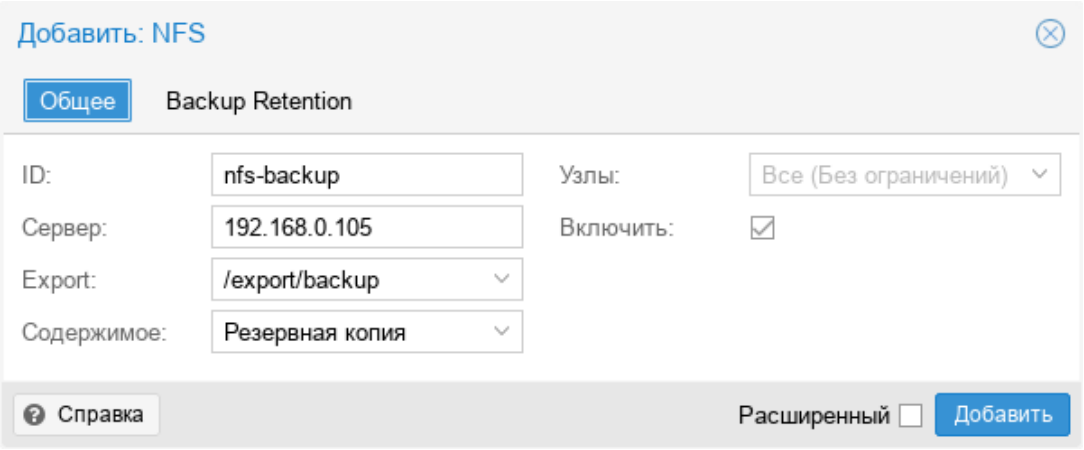

Рис. 131

На вкладке «Backup Retention» можно указать параметры хранения резервных копий (*Рис. 132*).

Параметры хранения резервных копий в хранилище NFS

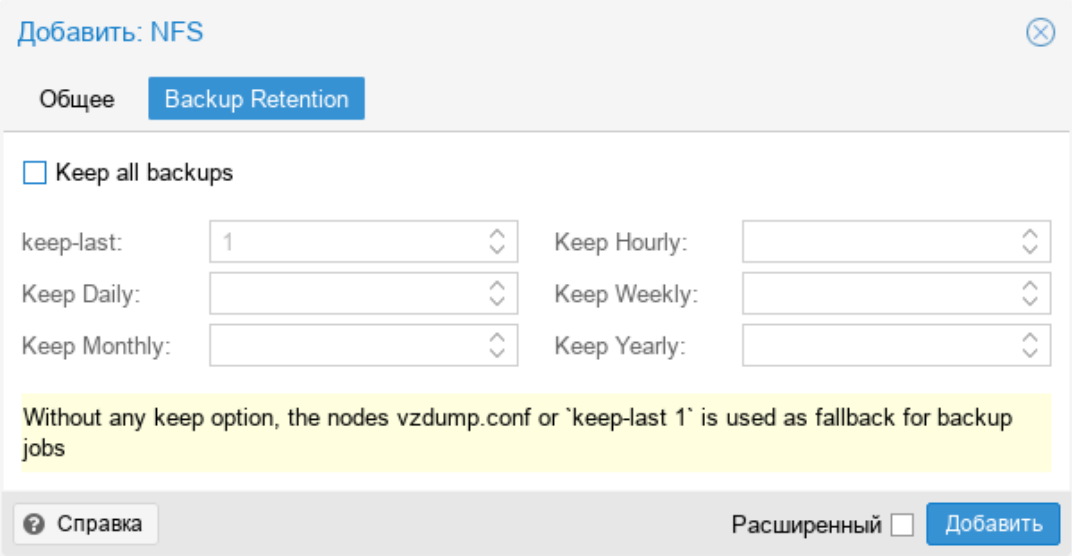

Рис. 132

Доступны следующие варианты хранения резервных копий (в скобках указаны параметры опции prune-backups команды `vzdump`):

– «Keep all backups» (`keep-all=<1|0>`) – хранить все резервные копии (если отмечен этот пункт, другие параметры не могут быть установлены);

– «keep-last» (`keep-last=<N>`) – хранить <N> последних резервных копий;

– «Keep Hourly» (`keep-hourly=<N>`) – хранить резервные копии за последние <N> часов (если за один час создается более одной резервной копии, сохраняется только последняя);

– «Keep Daily» (`keep-daily=<N>`) – хранить резервные копии за последние <N> дней (если за один день создается более одной резервной копии, сохраняется только самая последняя);

– «Keep Weekly» (`keep-weekly=<N>`) – хранить резервные копии за последние <N> недель (если за одну неделю создается более одной резервной копии, сохраняется только последняя);

– «Keep Monthly» (`keep-monthly=<N>`) – хранить резервные копии за последние <N> месяцев (если за один месяц создается более одной резервной копии, сохраняется только самая последняя);

– «Keep Yearly» (`keep-yearly <N>`) – хранить резервные копии за последние <N> лет (если за один год создается более одной резервной копии, сохраняется только самая последняя).

Варианты хранения обрабатываются в указанном выше порядке. Каждый вариант распространяется только на резервное копирование в определенный период времени.

Пример указания параметров хранения резервных копий при создании задания:

```
# vzdump 777 --prune-backups keep-last=3,keep-daily=13,keep-yearly=9
```

Хотя можно передавать параметры хранения резервных копий непосредственно в `vzdump`, часто более разумно настроить эти параметры на уровне хранилища.

4.10.4 Резервное копирование по расписанию

Задания резервного копирования можно запланировать так, чтобы они выполнялись автоматически в определенные дни и часы для конкретных узлов и гостевых систем. Конфигурирование заданий создания резервных копий выполняется на уровне центра обработки данных в веб-интерфейсе, который создает запись cron в `/etc/cron.d/vzdump`.

4.10.5 Настройка резервного копирования в графическом интерфейсе

Для того чтобы создать расписание резервного копирования, необходимо его настроить во вкладке «Резервная копия» («BackUP») датацентра (Рис. 133).

Вкладка «Резервная копия»

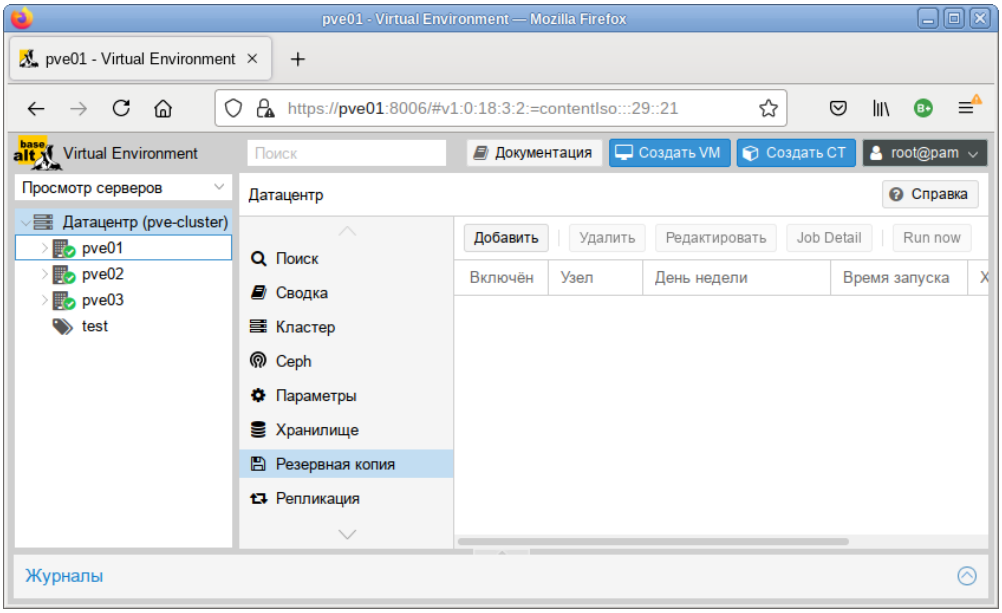

Рис. 133

При создании задания на резервирование, необходимо указать (Рис. 134):

- «Узел» («Node») – можно создавать график из одного места по разным узлам (серверам);
- «Хранилище» («Storage») – точка смонтированного накопителя, куда будет проходить копирование;
- «День недели» («Day of Week») – здесь можно указать один или несколько определенных дней, в которые будет выполняться резервное копирование (несколько дней можно выбрать при помощи зажатой клавиши <CTRL>);
- «Время запуска» («Start time») – время старта резервного копирования (указывается в формате 24);
- «Режим выбора» («Selection mode») – принимает три значения: «Учитывая выбранные VM» («Include»), «Все» («All»), «Исключить выбранные VM» («Exclude»), «Pool based»;

– «Отправить письмо» («Send email to») – адрес, на который будут приходить результаты резервного копирования;
– «Уведомление по почте» («Email notification») – принимает два значения: «Всегда» («Always») – сообщение будет приходить при любом результате резервного копирования, «Только при ошибках» («On failure only») – сообщение будет приходить только в случае неудачной попытки резервного копирования;
– «Сжатие» («Compression») – метод сжатия, принимает четыре значения: «ZSTD (быстро и хорошо)» (по умолчанию), «LZO (быстро)», «GZIP (хорошо)» и «нет»;
– «Режим» («Mode») – режим ВМ, в котором будет выполняться резервное копирование. Может принимать три значения (Рис. 135): «Снимок», «Приостановить», «Остановка».

Создание задания для резервного копирования

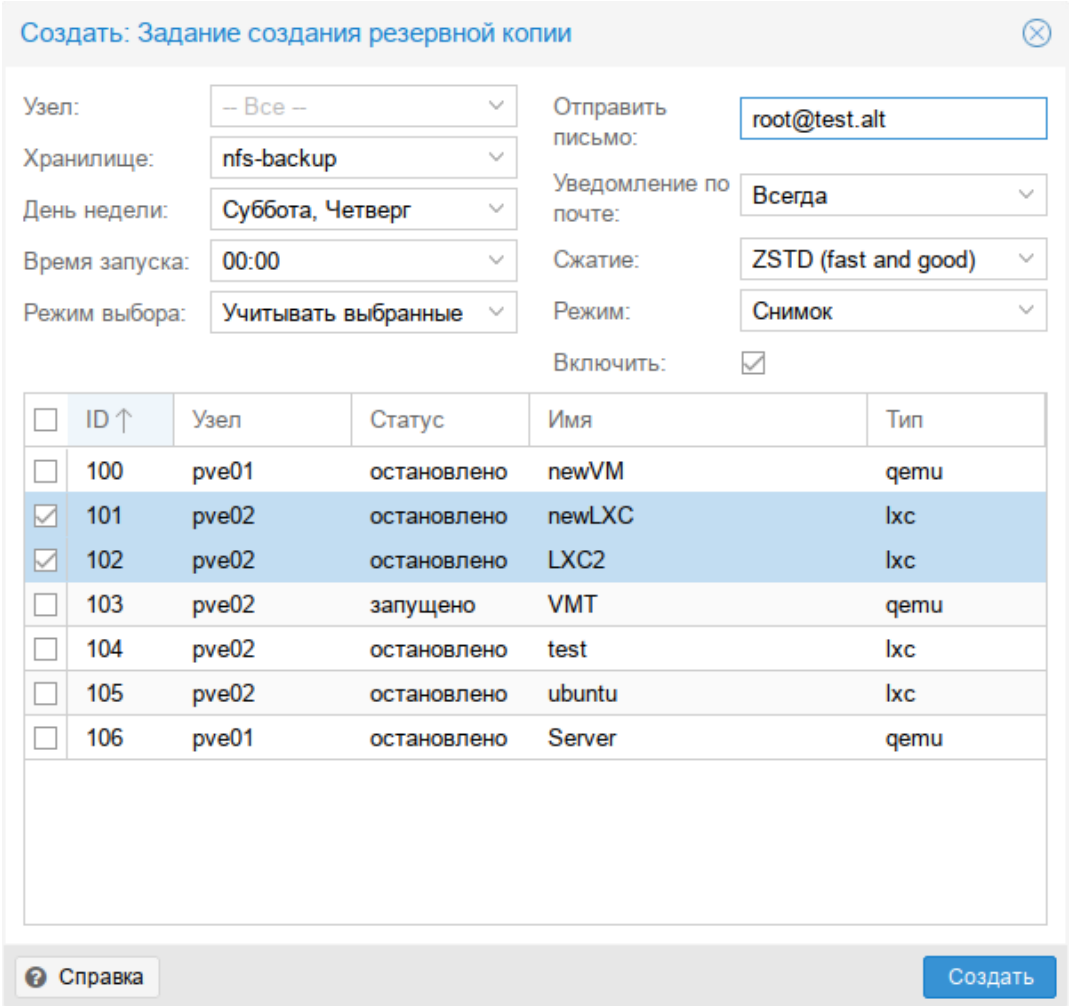

Рис. 134

Выбор режима создания резервной копии

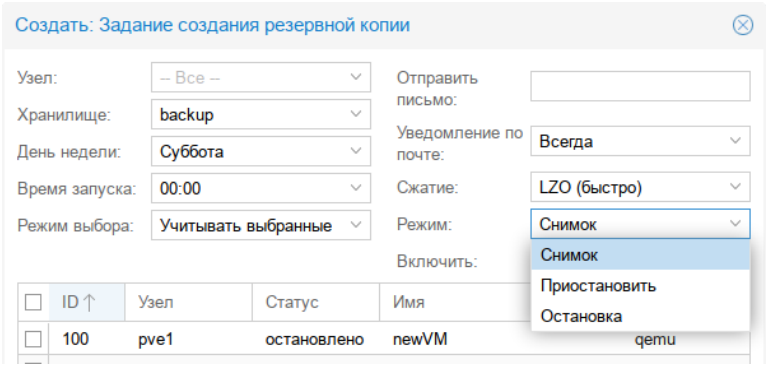

Рис. 135

После указания необходимых параметров и нажатия кнопки «Создать», задание для резервного копирования появляется в списке (Рис. 136). Оно будет запускаться в назначенное время.

Также существует возможность запустить задание по требованию – кнопка «Run now».

Задание резервного копирования

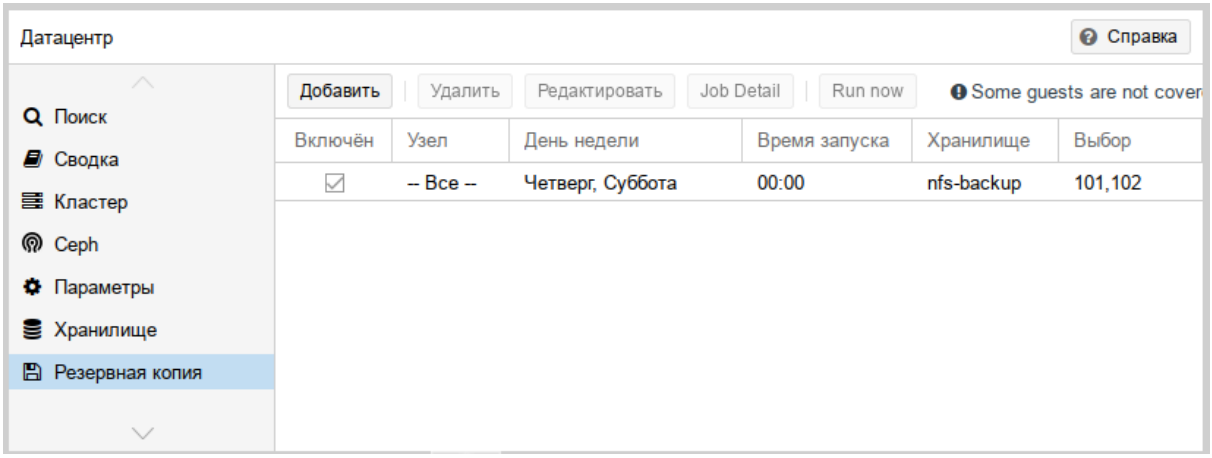

Рис. 136

Для того чтобы разово создать резервную копию конкретной ВМ, достаточно открыть ВМ, выбрать в ней раздел «Резервная копия» («Backup») и нажать кнопку «Создать резервную копию сейчас» («Backup Now») (Рис. 137).

Вкладка «Резервная копия» ВМ

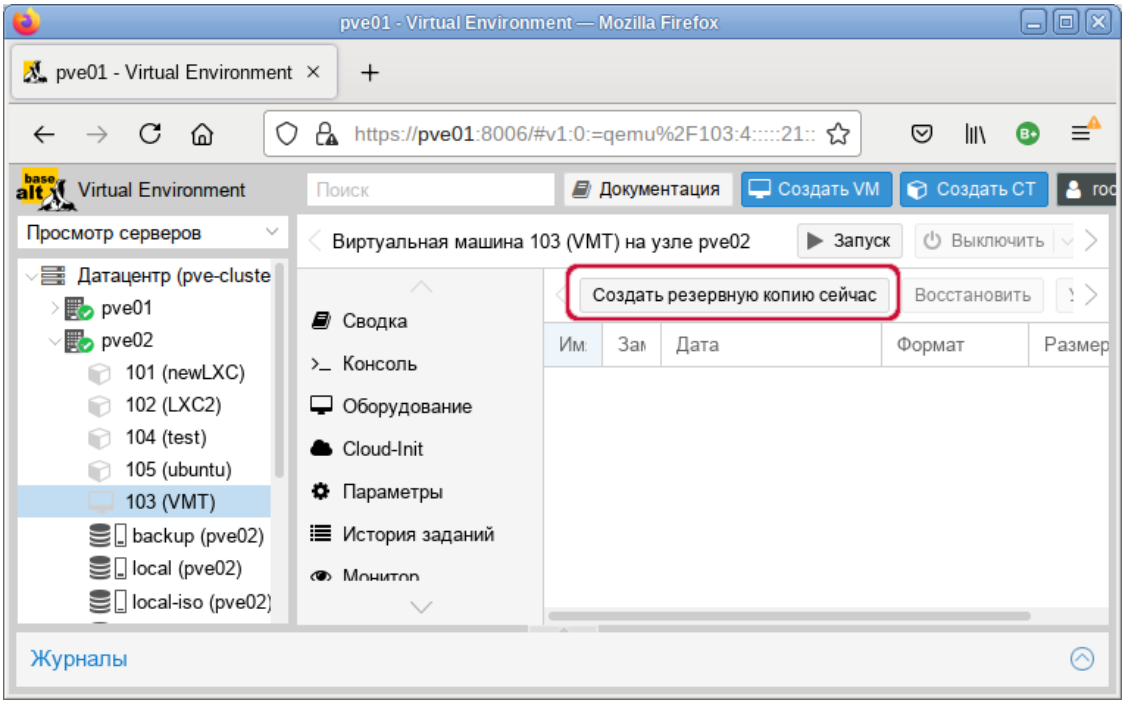

Рис. 137

Далее следует указать параметры резервного копирования (Рис. 138).

Выбор режима создания резервной копии

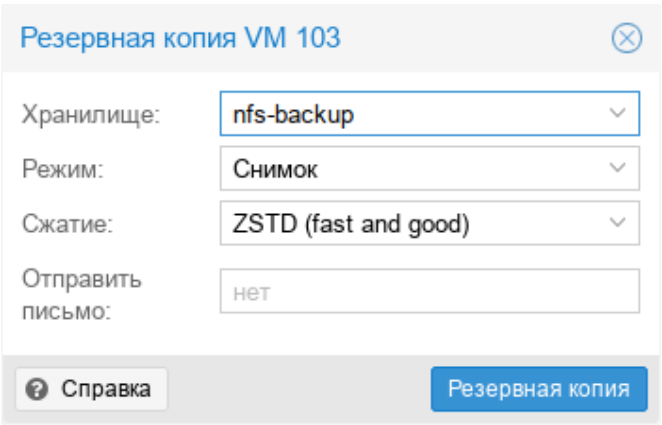

Рис. 138

После создания резервной копии рекомендуется сразу убедиться, что из нее можно восстановить ВМ. Для этого необходимо открыть хранилище с резервной копией (Рис. 139) и начать процесс восстановления (Рис. 140). При восстановлении можно указать новое имя для ВМ.

Резервная копия в хранилище nfs-backup

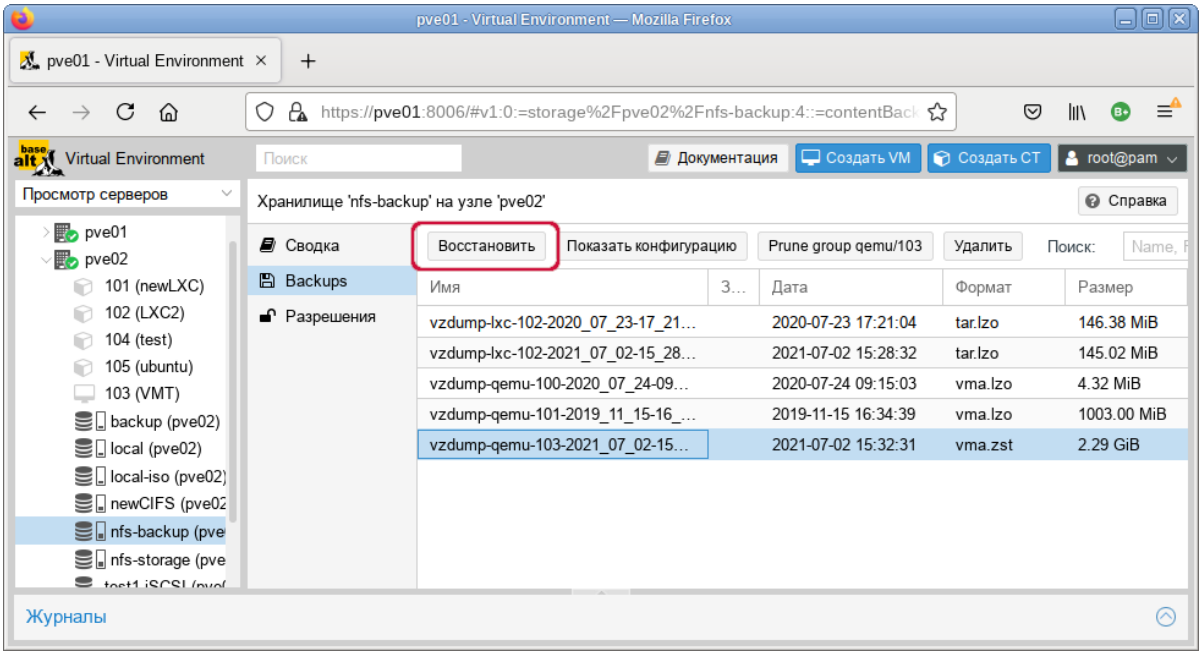

Рис. 139

Восстановить ВМ из резервной копии

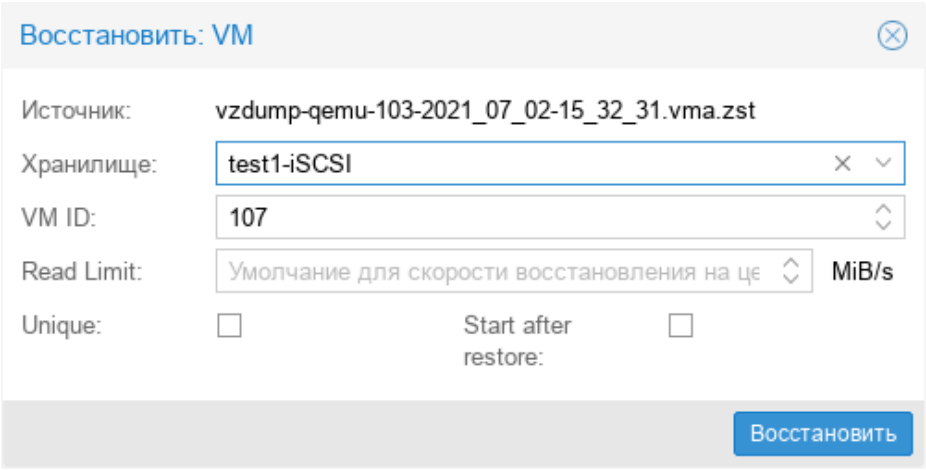

Рис. 140

Если восстанавливать из резервной копии в интерфейсе ВМ (Рис. 141), то будет предложена только замена существующей ВМ.

Восстановление из резервной копии в интерфейсе ВМ

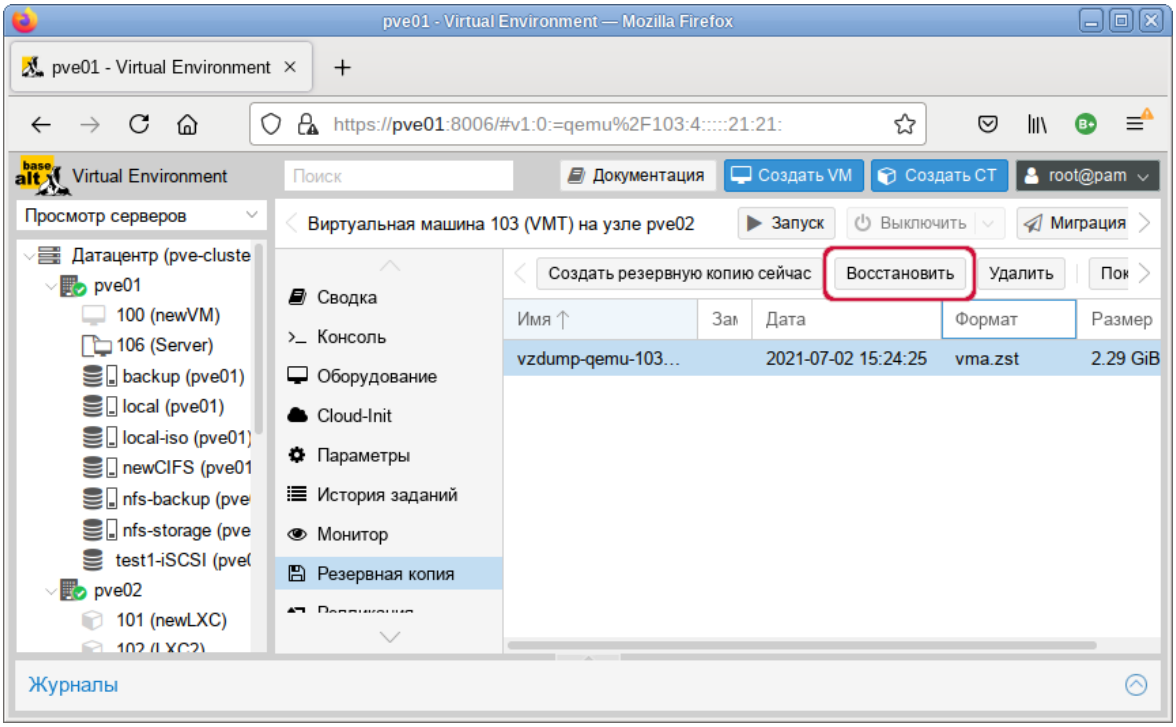

Рис. 141

4.10.6 Резервное копирование из командной строки

4.10.6.1 Файлы резервных копий

Все создаваемые резервные копии будут лежать в поддиректории «dump». Они будут иметь вид:

- vzdump-qemu-номер_машины-дата-время.vma.zst в случае выбора метода сжатия ZST;
- vzdump-qemu-номер_машины-дата-время.vma.gz в случае выбора метода сжатия GZIP;
- vzdump-qemu-номер_машины-дата-время.vma.lzo для использования метода LZO.

vzdump кодирует тип гостевой системы и время резервного копирования в имя файла, например:

```
vzdump-qemu-103-2021_07_02-15_32_31.vma.zst
```

4.10.6.2 Восстановление

Резервные копии могут быть восстановлены с помощью следующих утилит:

- pct restore – утилита восстановления контейнера;
- qmrestore – утилита восстановления QemuServer.

4.10.6.3 Ограничение пропускной способности

Для восстановления одной или нескольких больших резервных копий может потребоваться много ресурсов, особенно пропускной способности хранилища как для чтения из резервного хранилища, так и для записи в целевое хранилище. Это может негативно повлиять на другую ВМ, так как доступ к хранилищу может быть перегружен.

Чтобы избежать этого, можно установить ограничение полосы пропускания для задания резервного копирования. PVE реализует два вида ограничений для восстановления и архивирования:

- per-restore limit: обозначает максимальный объем полосы пропускания для чтения из архива резервной копии;
- предел записи для хранилища: обозначает максимальный объем полосы пропускания, используемый для записи в конкретное хранилище.

Ограничение чтения косвенно влияет на ограничение записи. Меньшее ограничение на задание перезапишет большее ограничение на хранилище. Больший лимит на каждое задание будет перезаписывать только лимит на хранилище, если у вас есть разрешения «Data.Allocate» для уязвимого хранилища.

Примечание. Чтобы отключить все ограничения для конкретного задания восстановления можно использовать значение `lim 0` для параметра `bwlimit`. Это может быть полезно, если требуется как можно быстрее восстановить ВМ.

Установить ограничение пропускной способности по умолчанию для настроенного хранилища, можно с помощью команды:

```
# pvesm set STORAGEID --bwlimit KIBs
```

4.10.6.4 Конфигурация

Глобальные настройки создания резервных копий хранятся в файле конфигурации `/etc/vzdump.conf`. Каждая строка файла имеет следующий формат (пустые строки в файле игнорируются, строки, начинающиеся с символа #, рассматриваются как комментарии и также игнорируются):

```
OPTION: value
```

Поддерживаемые опции представлены в табл. 8.

Пример `vzdump.conf`:

```
tmpdir: /mnt/fast_local_disk
storage: my_backup_storage
mode: snapshot
bwlimit: 10000
```

Т а б л и ц а 8 – Опции резервного копирования

Опция	Описание
bwlimit: integer (0 - N) (default=0)	Ограничение пропускной способности ввода/вывода (Кб/с)
compress: (0 | 1 | gzip | lzo | zstd) (default=0)	Сжатие файла резервной копии
dumpdir: string	Записать результирующие файлы в указанный каталог
exclude-path: string	Исключить определенные файлы/каталоги
ionice: integer (0 - 8) (default=7)	Установить CFQ приоритет ionice
lockwait: integer (0 - N) (default=180)	Максимальное время ожидания для глобальной блокировки (в минутах)
mailnotification: (always | failure) (default=always)	Указание, когда следует отправить отчет по электронной почте
mailto: string	Разделенный запятыми список адресов электронной почты, на которые будут приходить уведомления
maxfiles: integer (1 - N) (default=1)	Максимальное количество файлов резервных копий ВМ
mode: (snapshot | stop | suspend) (default=snapshot)	Режим резервного копирования
pigz: integer (default=0)	Использует pigz вместо gzip при N>0. N=1 использует половину ядер (uses half of cores), при N>1 N – количество потоков
prune-backups: [keep-all=<1|0>] [,keep-daily=<N>] [,keep-hourly=<N>] [,keep-last=<N>] [,keep-monthly=<N>] [,keep-weekly=<N>] [,keep-yearly=<N>]	Использовать эти параметры хранения вместо параметров из конфигурации хранилища (см.выше)
remove: boolean (default=1)	Удалить старые резервные копии, если их больше, чем установлено опцией maxfiles
script: string	Использовать указанный скрипт
stdexcludes: boolean (default=1)	Исключить временные файлы и файлы журналов
stopwait: integer (0 - N) (default=10)	Максимальное время ожидания пока гостевая система не остановится (минуты)
storage: string	Хранить полученный файл в этом хранилище
tmpdir: string	Хранить временные файлы в указанном каталоге
zstd: integer (default = 1)	Количество потоков zstd. N = 0 использовать половину доступных ядер, N> 0 использовать N как количество потоков

4.10.6.5 Файлы, не включаемые в резервную копию

П р и м е ч а н и е . Эта опция доступна только при создании резервных копий контейнеров.

vzdump пропускает следующие файлы по умолчанию (отключается с помощью опции `--stdexcludes 0`):

```
/tmp/?*
/var/tmp/?*
/var/run/?*pid
```

Кроме того, можно вручную указать какие файлы исключать (дополнительно), например:

```
# vzdump 777 --exclude-path /tmp/ --exclude-path '/var/foo*'
```

Файлы конфигурации также хранятся внутри архива резервных копий (в `/etc/vzdump/`) и будут корректно восстановлены.

4.10.6.6 Примеры

Простая резервная копия гостевой системы 777 – без снимка, только архив гостевой части и конфигурационного файла в каталог резервного копирования по умолчанию (обычно `/var/lib/vz/dump/`):

```
# vzdump 777
```

Использовать `rsync` и режим приостановки для создания снимка (минимальное время простоя):

```
# vzdump 777 --mode suspend
```

Сделать резервную копию всей гостевой системы и отправить отчет пользователю root и admin:

```
# vzdump --all --mode suspend --mailto root --mailto admin
```

Использовать режим мгновенного снимка (снапшота) (нет времени простоя) и каталог для хранения резервных копий `/mnt/backup`:

```
# vzdump 777 --dumpdir /mnt/backup --mode snapshot
```

Резервное копирование более чем одной ВМ (выборочно):

```
# vzdump 101 102 103 --mailto root
```

Резервное копирование всех ВМ, исключая 101 и 102:

```
# vzdump --mode suspend --exclude 101,102
```

Восстановить контейнер в новый контейнер 600:

```
# pct restore 600 /mnt/backup/vzdump-lxc-777.tar
```

Восстановить QemuServer VM в VM 601:

```
# qmrestore /mnt/backup/vzdump-qemu-888.vma 601
```

Клонировать существующий контейнер 101 в новый контейнер 300 с 4GB корневой файловой системы, используя pip:

```
# vzdump 101 --stdout | pct restore --rootfs 4 300 -
```

4.11 Снимки (snapshot)

Снимки ВМ – это файловые снимки состояния, данных диска и конфигурации ВМ в определенный момент времени. Можно создать несколько снимков ВМ даже во время ее работы. Затем можно возвратить ее в любое из предыдущих состояний, применив моментальный снимок к ВМ.

Чтобы создать снимок состояния системы в меню ВМ необходимо выбрать пункт «Снимки» («Snapshots») и затем нажать кнопку «Сделать снимок» (Рис. 142). В открывшемся окне (Рис. 143) следует ввести название снимка (можно также ввести его описание) и нажать кнопку «Сделать снимок».

Окно управления снимками ВМ

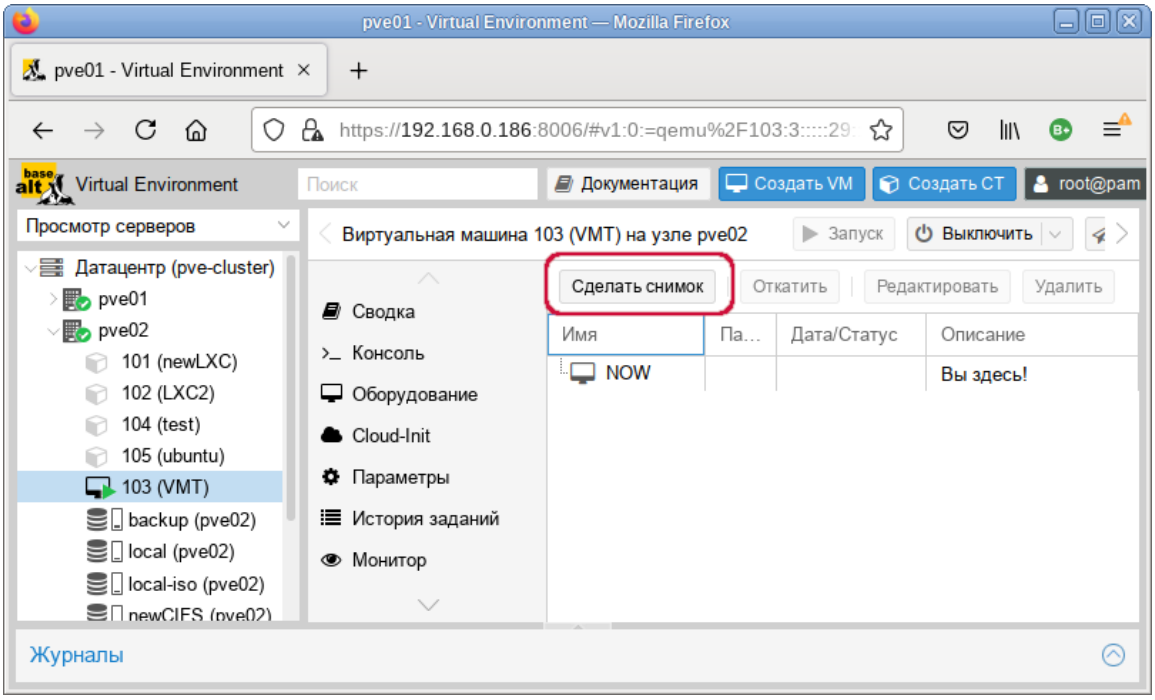

Рис. 142

Создание снимка ВМ

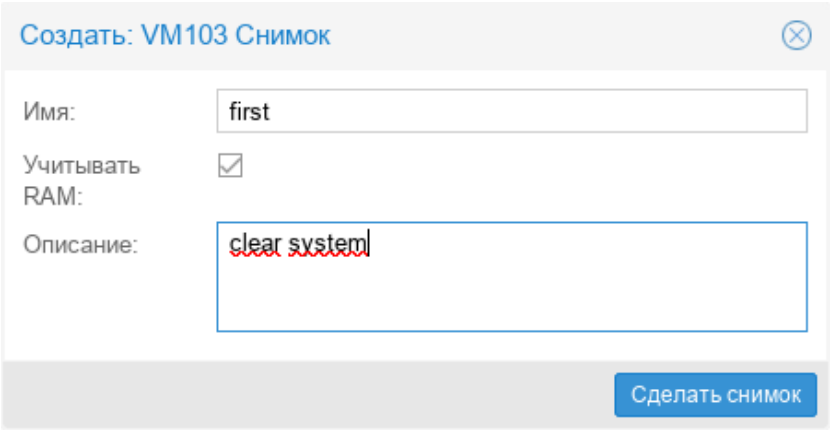

Рис. 143

Для того чтобы восстановить ВМ из снимка, необходимо в меню ВМ выбрать пункт «Снимки» («Snapshots»), выбрать снимок (Рис. 144) и нажать кнопку «Откатить».

Восстановление ОС из снимка

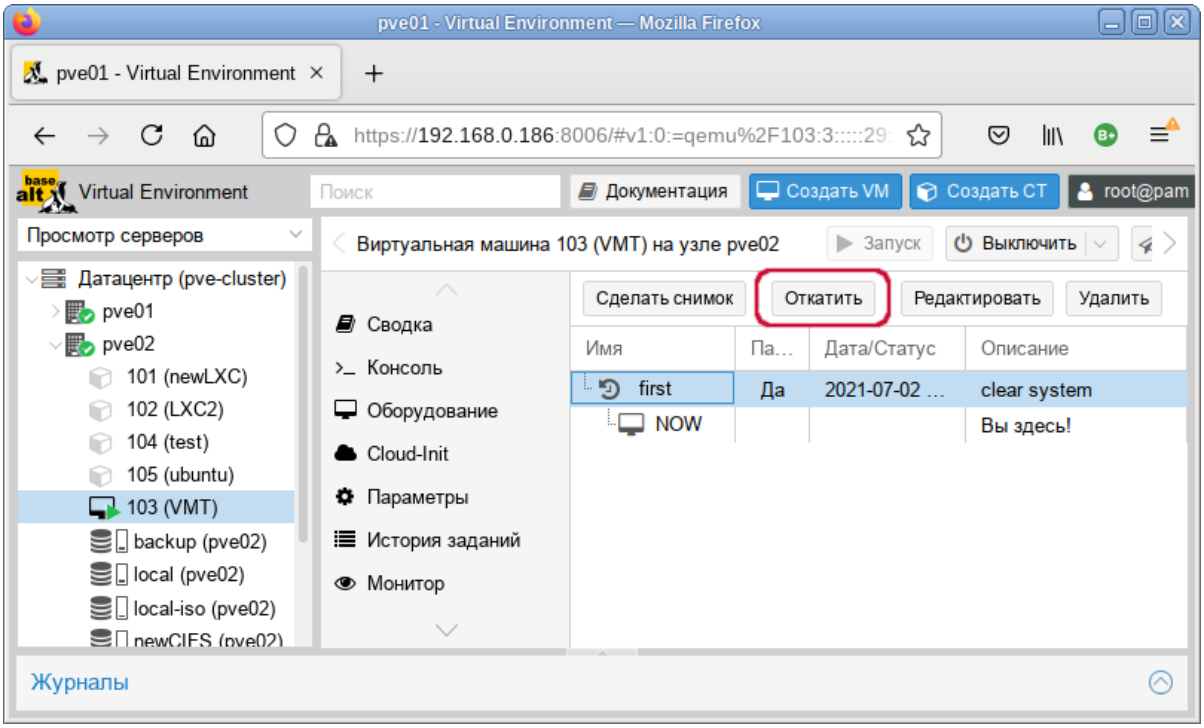

Рис. 144

При создании снимков, `qm` сохраняет конфигурацию ВМ во время снимка в отдельном разделе в файле конфигурации ВМ. Например, после создания снимка с именем `first` файл конфигурации будет выглядеть следующим образом:

```
boot: order=scsi0;scsi7;net0
cores: 1
memory: 1024
name: VMT
net0: virtio=96:53:8D:BF:FC:94,bridge=vmbr0,firewall=1
numa: 0
ostype: l26
parent: first
scsi0: local:103/vm-103-disk-0.qcow2,size=32G
scsi7: nfs-storage:iso/alt-kworkstation-9.2-x86_64.iso,media=cdrom
scsihw: virtio-scsi-pci
smbios1: uuid=e53c7a24-6fc3-4f33-8173-d36f938d1c77
sockets: 1
vmgenid: ec8a0c17-8438-439d-9ee0-13be4497f8ed
```

```
[first]
#clear system
boot: order=scsi0;scsi7;net0
cores: 1
memory: 1024
name: VMT
net0: virtio=96:53:8D:BF:FC:94,bridge=vmbr0,firewall=1
numa: 0
ostype: l26
runningcpu: kvm64,enforce,+kvm_pv_eoi,+kvm_pv_unhalt,+lahf_lm,+sep
runningmachine: pc-i440fx-5.1+pve0
scsi0: local:103/vm-103-disk-0.qcow2,size=32G
scsi7: nfs-storage:iso/alt-kworkstation-9.2-x86_64.iso,media=cdrom
scsihw: virtio-scsi-pci
smbios1: uuid=e53c7a24-6fc3-4f33-8173-d36f938d1c77
snaptime: 1625210352
sockets: 1
vmgenid: ec8a0c17-8438-439d-9ee0-13be4497f8ed
vmstate: local:103/vm-103-state-first.raw
```

Свойство `parent` используется для хранения родительских/дочерних отношений между снимками, `snaptime` – это отметка времени создания снимка (эпоха Unix).

4.12 Встроенный мониторинг PVE

Все данные о потреблении ресурсов и производительности можно найти на вкладках «Сводка» («Summary») узлов PVE и ВМ. Можно просматривать данные на основе почасового, ежедневного, еженедельного или за год периодов.

На Рис. 145 показана «Сводка» («Summary») узла pve02 со списком для выбора периода данных.

Просмотреть список всех узлов, ВМ и контейнеров в кластере можно, выбрав «Датацентр» → «Поиск» («Datacentre» → «Search») **(**Рис. 146). Этот список может быть отсортирован по полям: «Тип» («Type»), «Описание» («Description»), «Использование диска %» («Disk usage»), «Использование памяти» («Memory usage»), «Загрузка CPU» («CPU usage») и «Время работы» («Uptime»). В этом списке отображается потребление ресурсов только в реальном масштабе времени.

Выбор периода данных, для отображения отчета

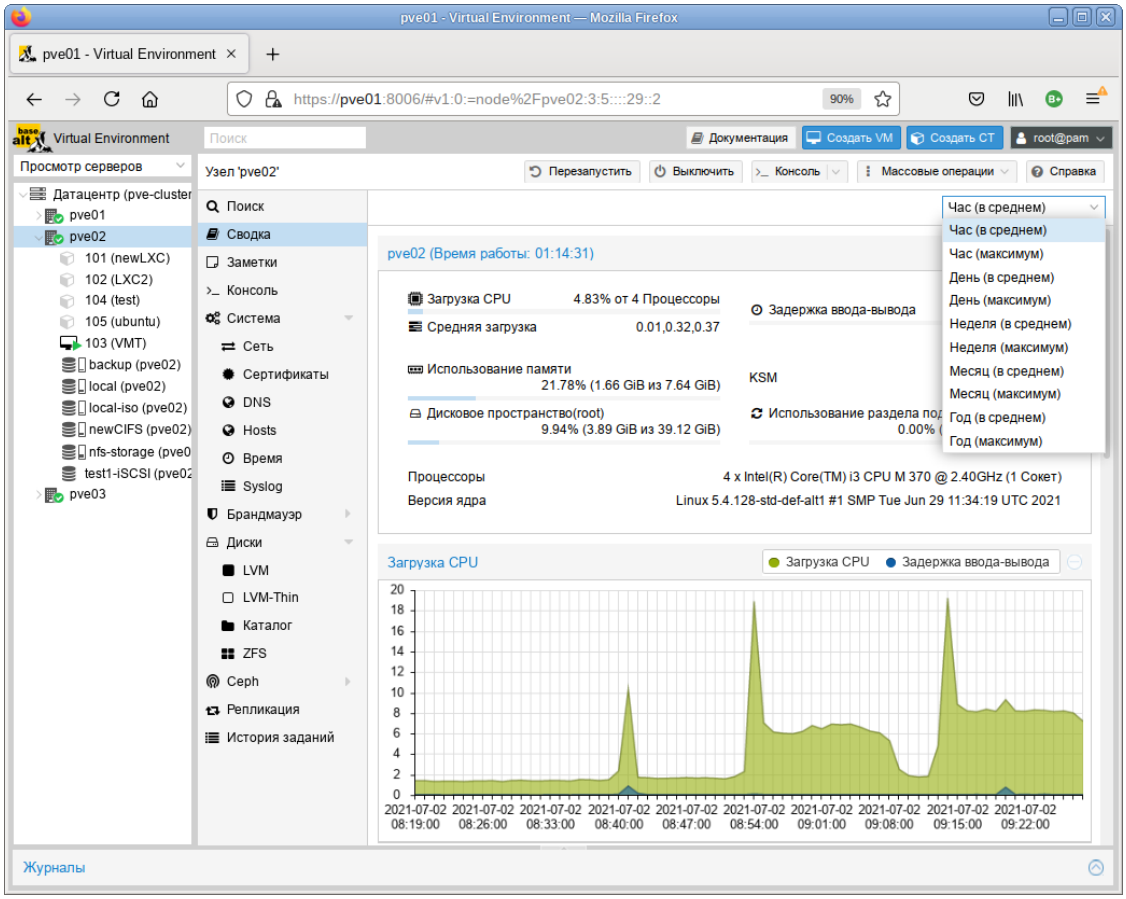

Рис. 145

Потребление ресурсов

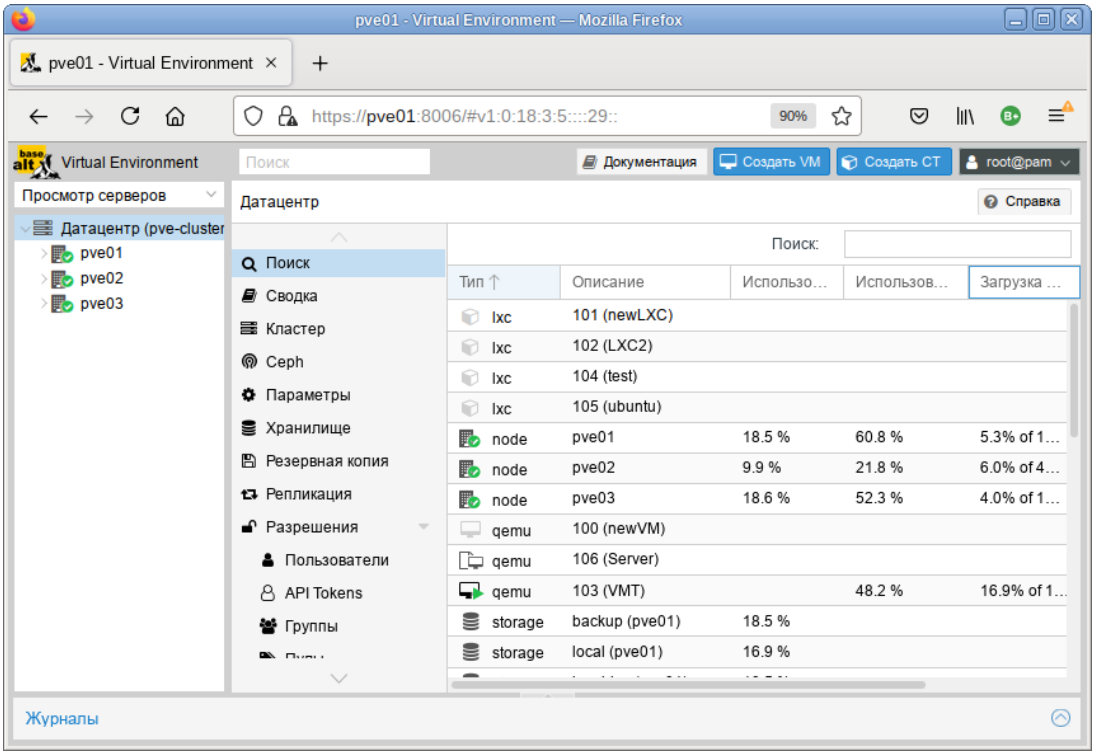

Рис. 146

Для мониторинга состояния локальных дисков используется пакет `smartmontools`. Он содержит набор инструментов для мониторинга и управления S.M.A.R.T. системой для локальных жестких дисков.

Получить статус диска можно, выполнив следующую команду:

```
# smartctl -a /dev/sdX
```

где /dev/sdX – это путь к одному из локальных дисков.

Включить поддержку SMART для диска, если она отключена:

```
# smartctl -s on /dev/sdX
```

Просмотреть S.M.A.R.T. статус диска в веб-интерфейсе можно, выбрав в разделе «Диски» нужный диск и нажав кнопку «Показать данные S.M.A.R.T.» (*Рис. 147*).

По умолчанию, `smartmontools daemon smartd` активен и включен, и сканирует диски в /dev/sdX и /dev/hdX каждые 30 минут на наличие ошибок и предупреждений, а также отправляет сообщение электронной почты пользователю root в случае обнаружения проблемы (для пользователя root в PVE должен быть введен действительный адрес электронной почты).

Кнопка «Показать данные S.M.A.R.T.»

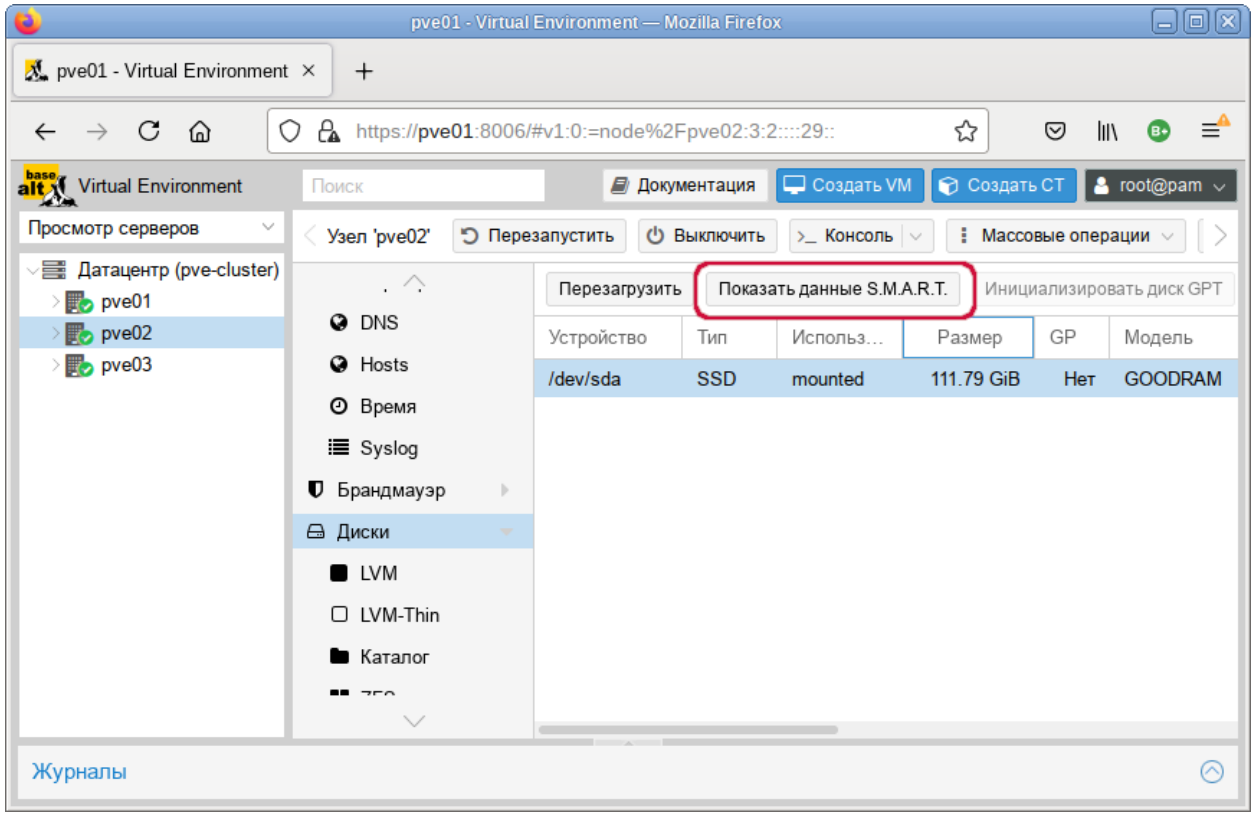

Рис. 147

Электронное сообщение будет содержать имя узла, где возникла проблема, а также параметры самого устройства, такие как серийный номер и идентификатор дискового устройства. Если та же самая ошибка продолжит возникать, узел будет отсылать электронное сообщение

каждые 24 часа. Основываясь на содержащейся в электронном сообщении информации можно определить отказавшее устройство и заменить его в случае такой необходимости.

4.13 Высокая доступность PVE

Высокая доступность является комбинацией компонентов и настроек, которые делают возможной непрерывную работу вычислительной среды на протяжении длительного времени. В основном это означает, что даже если находящееся в автоматическом режиме оборудование сервера испытывает проблемы в среде реального времени, высокая доступность (HA) может управлять оставшимися серверами самостоятельно и поддерживая виртуальную среду в рабочем состоянии автоматически перемещая или выполняя миграцию виртуальных машин с одного узла на другой. Настроенная надлежащим образом HA требует очень незначительного реального вмешательства пользователей в случае отказа аппаратных средств. Без HA на своем месте, все узлы требуют постоянного мониторинга со стороны сетевых менеджеров чтобы вручную перемещать виртуальные машины на жизнеспособные узлы, когда некоторые узлы испытывают проблемы.

В небольших средах перемещение вручную ВМ не является проблемой, однако в больших средах из сотен виртуальных машин или узлов постоянный мониторинг может быть очень затратным в смысле времени. Несмотря на то, что в системе может существовать программное обеспечение мониторинга, без HA администратор будет должен вручную перемещать или выполнять миграцию любых виртуальных машин с отказавшего узла. Это может повлечь за собой значительное время простоя. Это именно то место, где вступает в действие функциональность HA PVE. HA выводит вмешательство оператора за скобки решения, просто перемещая или выполняя миграцию виртуальных машин как только возникает отказ оборудования сервера.

Для функционирования HA в PVE необходимо чтобы все ВМ были в общем хранилище. HA PVE обрабатывает только узлы PVE и виртуальные машины в пределах кластера PVE. Такую функциональность HA не следует путать с избыточностью общих хранилищ, которую PVE может применять в своем развертывании HA. Высокая доступность в общем хранилище так же важна, как и высокая доступность ВМ PVE. Общие хранилища сторонних производителей могут предоставлять свою собственную функциональность HA. Таким образом, и сам кластер PVE, и общее хранилище должны быть настроены для предоставления реальной среды с высокой доступностью.

В вычислительном узле PVE могут существовать свои уровни избыточности, такие как применение RAID, дополнительные источники питания, агрегированные сетевые связи или сцепления (bond). HA в PVE не подменяет собой ни один из этих уровней. Он просто способствует

использованию функций избыточности ВМ для сохранения их в рабочем состоянии при отказе какого-либо узла.

Перезагрузка узла PVE, вызванная необходимостью применения обновлений, вызовет выключение всех ВМ с включенной HA, перемещение их на следующий доступный узел PVE и их последующий повторный запуск. В подобной ситуации может оказаться необходимой миграция ВМ в реальном времени вручную до перезагрузки обновляемого узла.

4.13.1 Как работает высокая доступность PVE

Основной блок управления, управляемый ha-manager называется ресурсом. Ресурс (также называемый «сервис») однозначно идентифицируется идентификатором сервиса (SID), который состоит из типа ресурса и идентификатора, специфичного для данного типа, например, vm: 100. Этот пример ресурса типа vm (виртуальная машина) с идентификатором 100.

В случае, когда по какой-либо причине узел становится недоступным, HA PVE ожидает 60 секунд прежде чем выполнится ограждение (fencing) отказавшего узла. Ограждение предотвращает службы кластера от возврата в рабочее состояние в этом месте. Затем HA перемещает эти ВМ и контейнеры на следующий доступный узел в их группе участников HA. Даже если узел с ВМ все еще включен, но потерял связь с сетевой средой, HA PVE попытается переместить все ВМ с этого узла на другой узел.

При возврате отказавшего узла в рабочее состояние, HA не будет автоматически перемещать ВМ на первоначальный узел. Это необходимо выполнять вручную. При этом ВМ может быть перемещена вручную только если HA запрещен для такой ВМ. Поэтому сначала следует выключить HA, а затем переместить на первоначальный узел и включить HA на данной ВМ вновь.

4.13.2 Требования для настройки высокой доступности

Среда перед настройкой HA PVE должна отвечать следующим требованиям:

- кластер не менее чем из трех узлов (для получения надежного кворума);
- общее хранилище для ВМ и контейнеров;
- аппаратное резервирование;
- использование надежных «серверных» компонентов;
- аппаратный сторожевой таймер (если он недоступен, используется программный таймер ядра Linux);
- дополнительные устройства ограждения (fencing).

4.13.3 Настройка высокой доступности PVE

Все настройки HA PVE могут быть выполнены в веб-интерфейсе. Функциональность HA доступна при выборе «Датацентр» → «HA» («Datacenter» → «HA»). Это меню, в котором будут

выполняться все связанные с HA настройки и управление (Рис. 148). В окне отображается статус настройки HA.

Меню HA

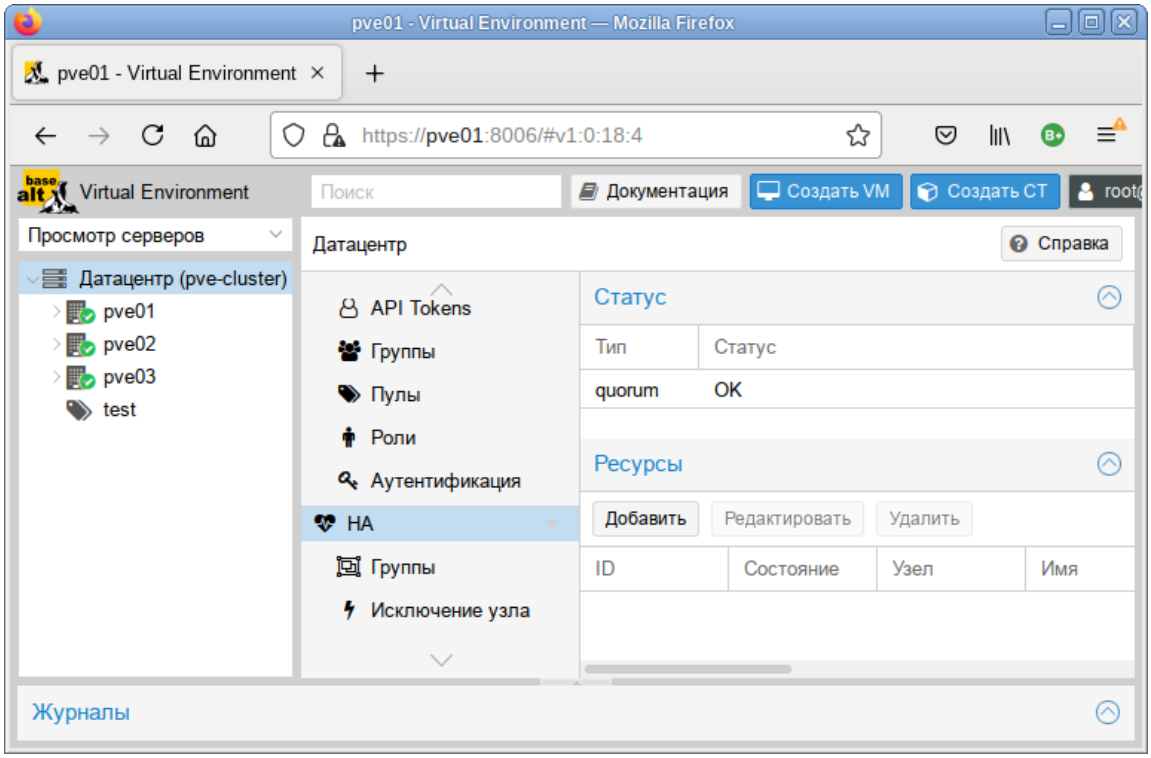

Рис. 148

4.13.3.1 Создание группы

Подменю «Группы» применяется для создания различных групп PVE для HA и управления ими. Наиболее характерным использованием групп являются некие программные решения или инфраструктура ВМ, которые должны работать совместно. Например, контроллер домена, файловый сервер и тому подобное. Назначенные в определенную группу ВМ могут перемещаться только совместно с узлами участниками этой группы. Например, шесть узлов, три из которых обладают всей полнотой ресурсов, достаточной для исполнения виртуального сервера базы данных, а другие три узла выполняют виртуальные рабочие столы или решения VDI. Можно создать две группы, для которых виртуальные серверы баз данных могут перемещаться только в пределах тех узлов, которые будут назначены для данной группы. Это гарантирует, что ВМ переместится на правильный узел, который будет способен исполнять такие ВМ.

Для включения HA для ВМ необходима как минимум одна группа.

Для создания группы необходимо нажать кнопку «Создать» («Create») в подменю «Группы» («Groups»).

Элементы, доступные в блоке диалога HA Group (Рис. 149):

– «ID» – название HA Group;

– «Узел» («Node») – назначение узлов в создаваемую группу. Чтобы создать конкретную группу, нужно выбрать, по крайней мере, один узел;
– «Restricted» – разрешение перемещения ВМ со стороны HA PVE только в рамках узлов участников данной группы HA. Если перемещаться некуда, то эти ВМ будут автоматически остановлены;
– «Nofailback» – используется для предотвращения автоматического восстановления состояния ВМ/контейнера при восстановлении узла в кластере (не рекомендуется включать эту опцию).

На Рис. 150 представлено подменю «Группы» с созданной группой.

Диалог создания группы

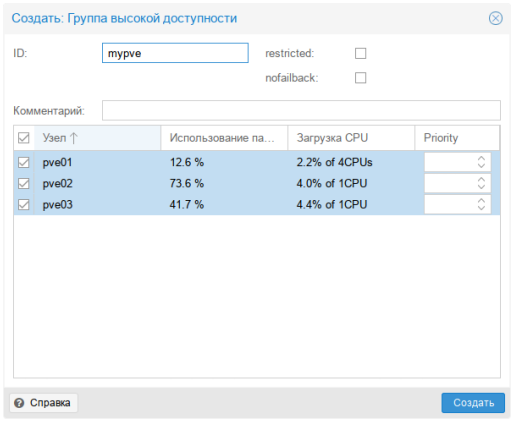

Рис. 149

Подменю «Группы» с созданной группой

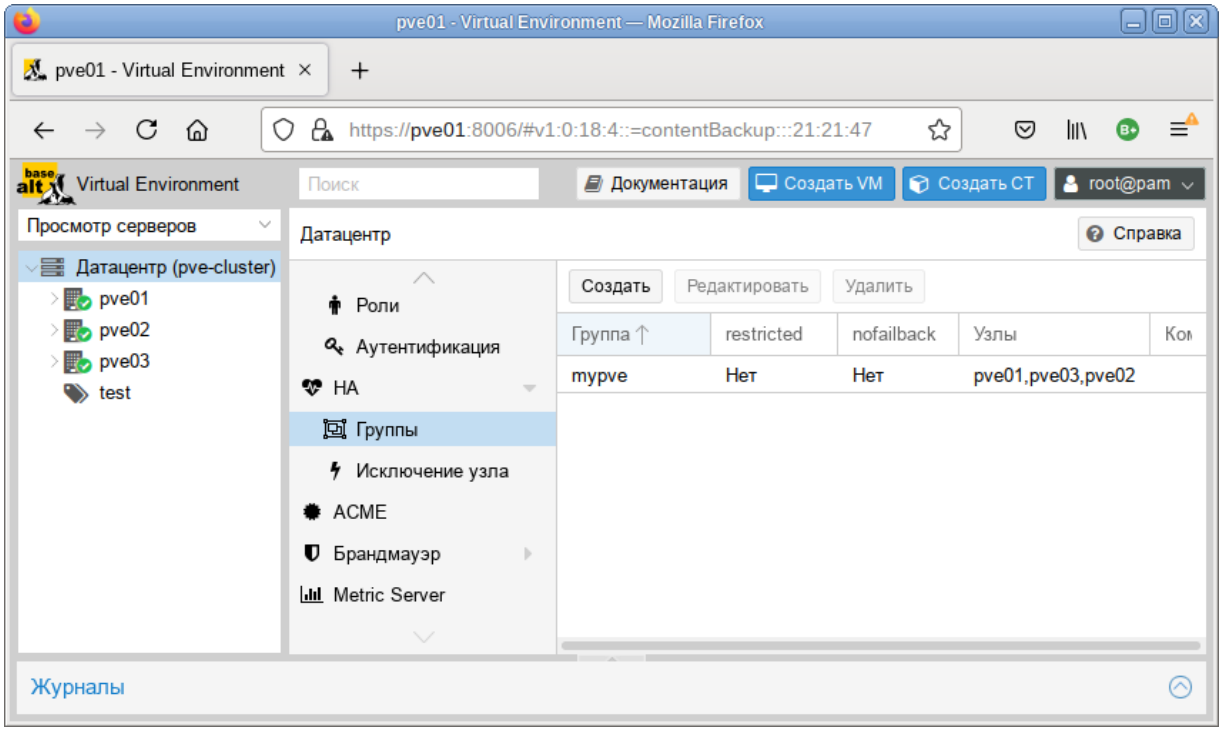

Рис. 150

4.13.3.2 Добавление ресурсов

Для включения HA для ВМ или контейнера следует нажать на кнопку «Добавить» в разделе «Ресурсы» меню «HA». В открывшемся диалоговом окне нужно выбрать ВМ и группу HA, к которой будет относиться данная ВМ (Рис. 151).

Добавление Ресурса в группу

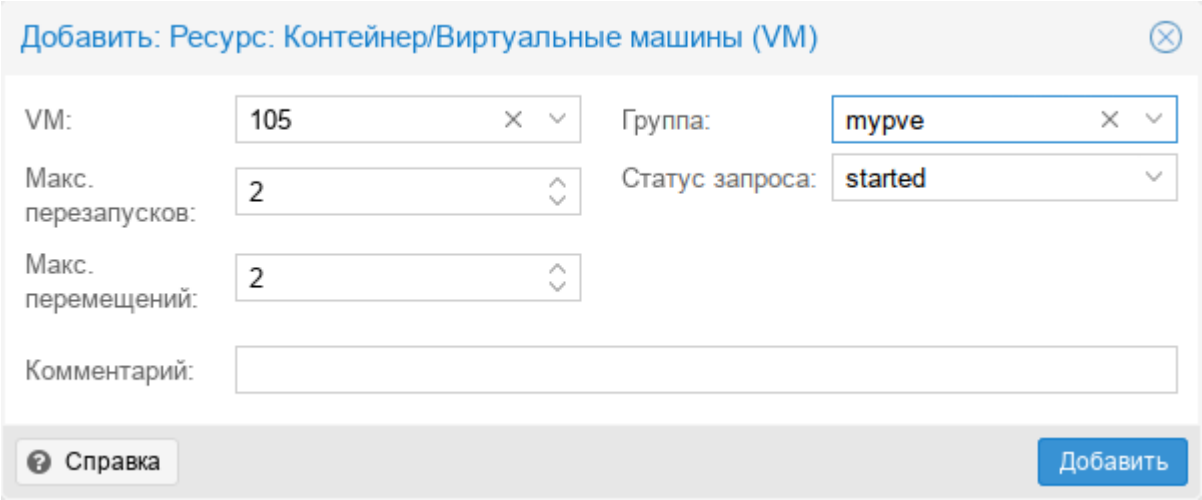

Рис. 151

В окне можно настроить следующие параметры:

- «Макс. перезапусков» («Max. Restart») – количество попыток запуска ВМ/контейнера на новом узле после перемещения;
- «Макс. перемещений» («Max. Relocate») – количество попыток перемещения ВМ/контейнера на новый узел;
- «Статус запроса» («Request State») – доступны варианты: started – кластер менеджер будет пытаться поддерживать состояние машины в запущенном состоянии; stopped – при отказе узла перемещать ресурс, но не пытаться запустить; ignored – ресурс, который не надо перемещать при отказе узла; disabled – в этот статус переходят ВМ, которые находятся в состоянии «error».

На Рис. 152 показана группа HA PVE и добавленные в нее ВМ и контейнеры, которыми будет управлять HA.

Раздел «Статус» («Status») отображает текущее состояние функциональности HA:

- кворум кластера установлен;
- главный узел pve01 группы HA активен и последний временной штамп жизнеспособности (heartbeat timestamp) проверен;
- все узлы, участвующие в группе HA активны и последний временной штамп жизнеспособности (heartbeat timestamp) проверен.

Список ресурсов

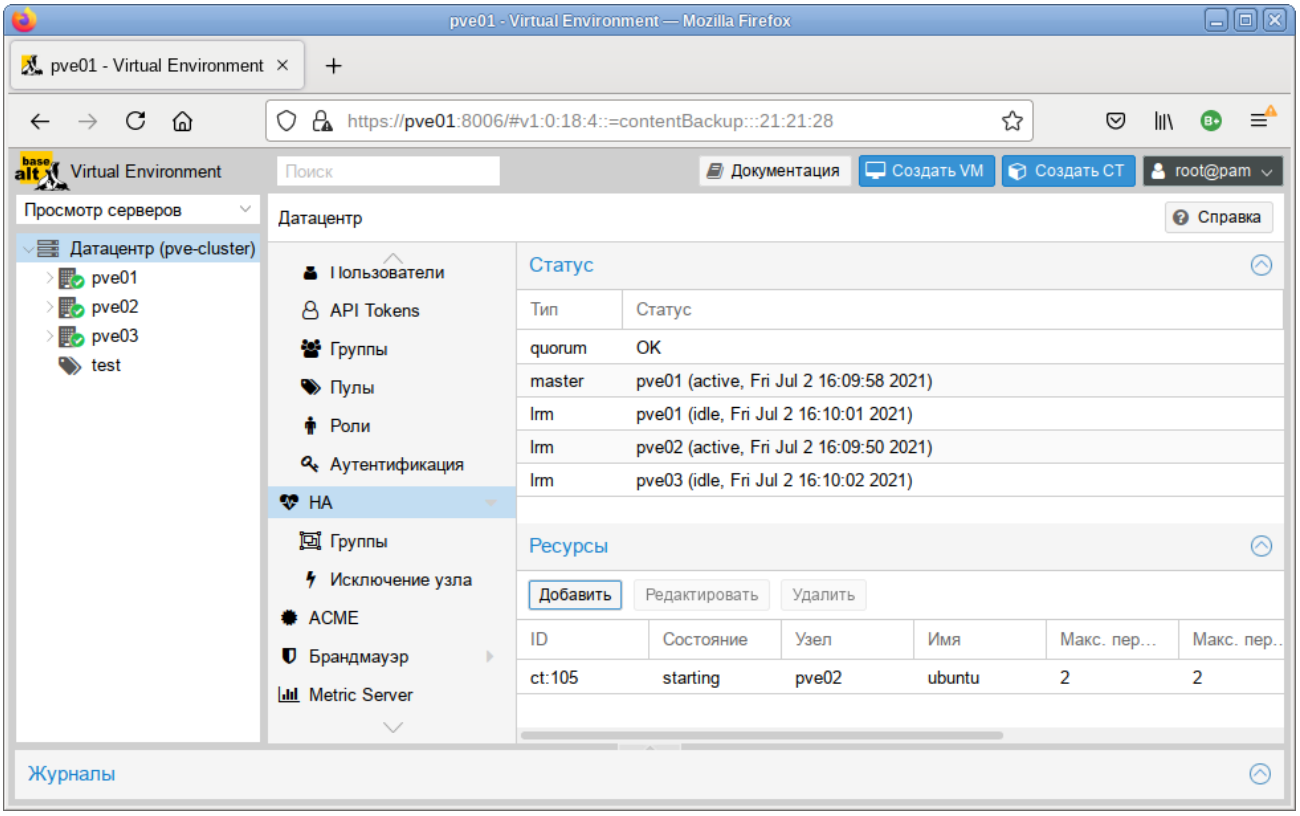

Рис. 152

Просмотреть состояние функциональности HA можно и в консоли:

```
quorum OK
master pve01 (active, Fri Jul  2 16:12:08 2021)
lrm pve01 (idle, Fri Jul  2 16:12:07 2021)
lrm pve02 (active, Fri Jul  2 16:12:05 2021)
lrm pve03 (idle, Fri Jul  2 16:12:08 2021)
service ct:105 (pve02, relocate)
```

4.13.4 Тестирование настройки высокой доступности PVE

Для того чтобы убедиться, что HA действительно работает можно отключить сетевое соединение для pve02 и понаблюдать за окном «Статус» на предмет изменений HA (Рис. 153).

После того как соединение с узлом pve02 будет потеряно, он будет помечен как не доступный (Рис. 154).

По истечению 60 секунд, HA PVE предоставит следующий доступный в группе HA узел в качестве главного (Рис. 155).

Список ресурсов

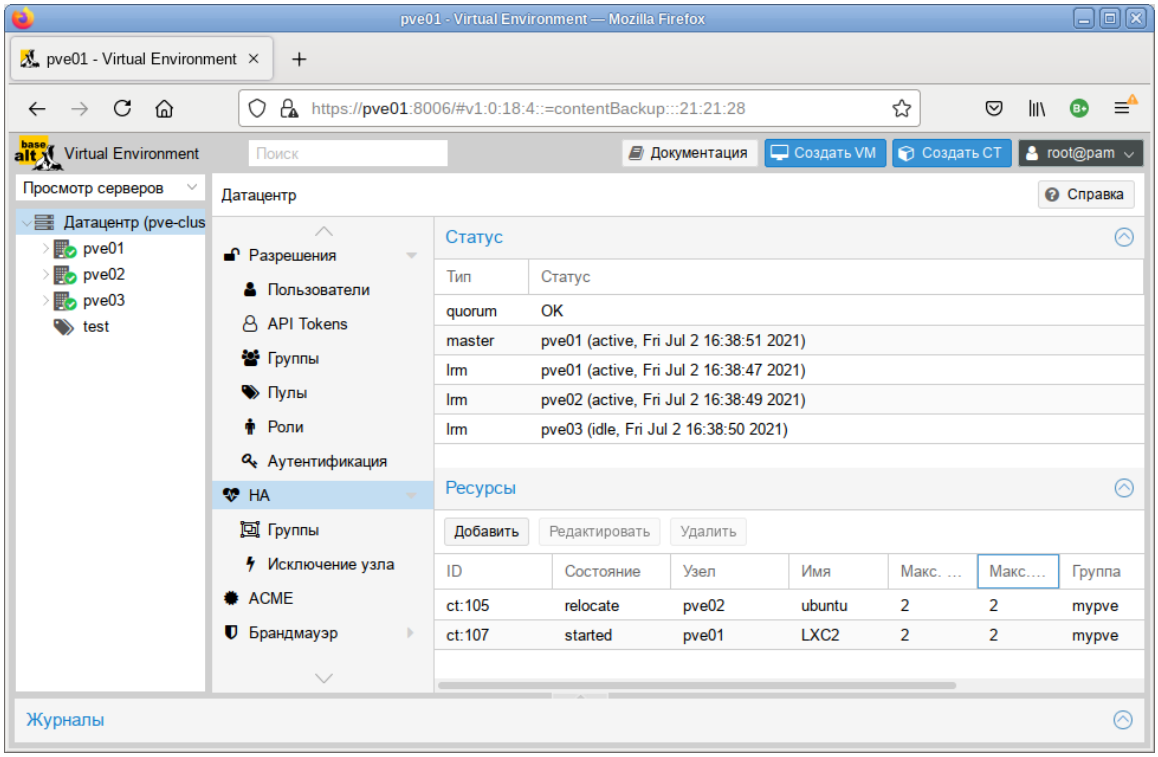

Рис. 153

Нет соединения с pve01

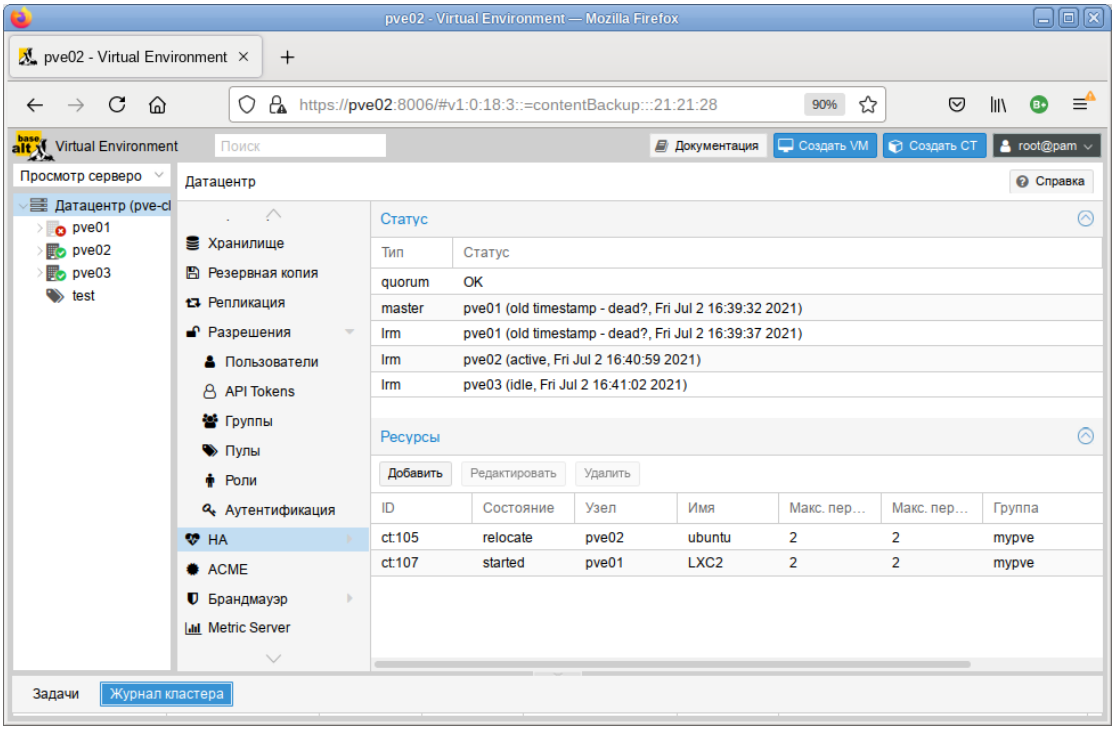

Рис. 154

Изменение главного узла на pve03

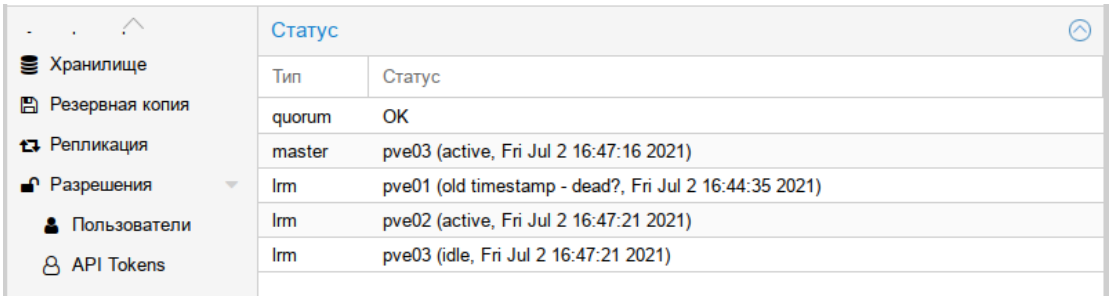

Рис. 155

После того, как HA PVE предоставит новый ведущий узел для данной группы HA, она затем запустит ограждение для ресурсов ВМ/контейнера для их подготовки к перемещению на другой узел. В процессе ограждения, все связанные с данной ВМ службы ограждаются, что означает, что даже если отказавший узел вернется в строй на этом этапе, его ВМ не будет иметь возможность восстановить свою нормальную работу.

После того, как ресурсы данной ВМ/контейнера ограждены, данная ВМ/контейнер полностью останавливается. Так как узел сам по себе отключен, ВМ/контейнер не может выполнить миграцию в реальном времени, поскольку состояние оперативной памяти исполняемой ВМ не может быть получено с отключенного узла.

После того, как отслеживаемая ВМ/контейнер остановлена, она перемещается на следующий свободный узел в группе HA и автоматически запускается. Контейнер 107 перемещен на узел pve03 и запущен (Рис. 156).

Контейнер 107 запущен на узле pve03

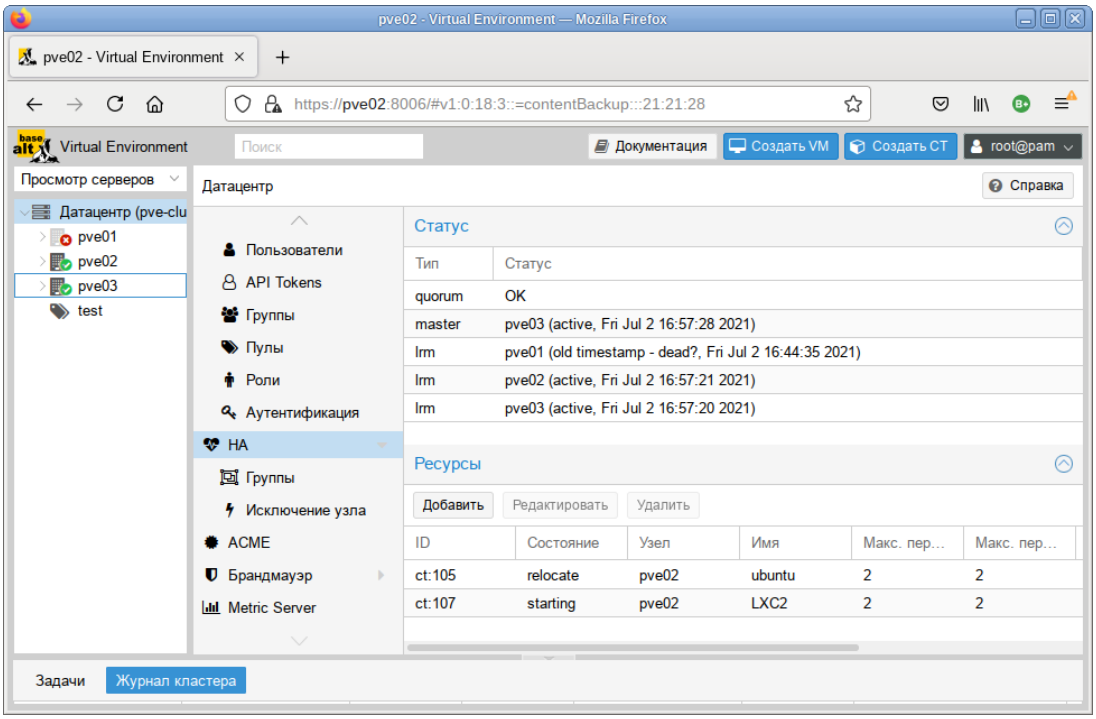

Рис. 156

В случае возникновения любой ошибки, HA PVE выполнит несколько попыток в соответствии с политиками `restart` и `relocate` для восстановления. Если все попытки окажутся неудачными, HA PVE поместит ресурсы в ошибочное состояние и не будет выполнять для них никаких задач.

4.14 Пользователи и их права

Для аутентификации пользователей в веб-интерфейсе PVE можно использовать как собственные механизмы PVE, так и PAM (Рис. 157).

Выбор типа аутентификации в веб-интерфейсе

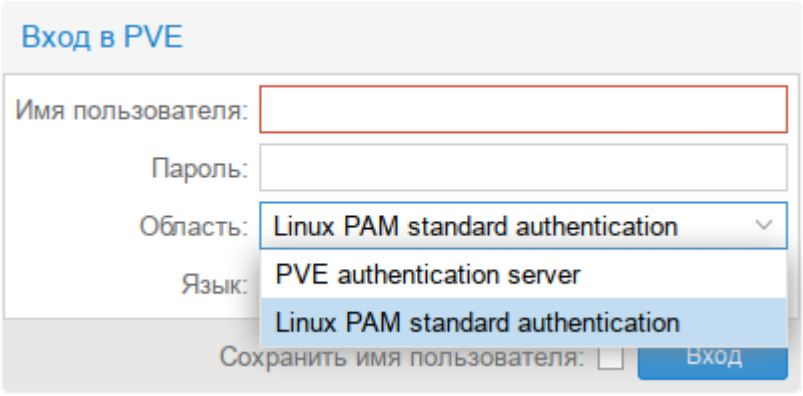

Рис. 157

Используя основанное на ролях управление пользователями и разрешениями для всех объектов (ВМ, хранилищ, узлов и т. д.), можно определить многоуровневый доступ.

Доступны следующие области (методы) аутентификации:

- Стандартная аутентификация Linux PAM – при использовании этой аутентификации системный пользователь должен существовать (должен быть создан, например, с помощью команды `adduser`) на всех узлах, на которых пользователю разрешено войти в систему. Пользователь аутентифицируется с помощью своего обычного системного пароля;
- Сервер аутентификации PVE – это хранилище паролей в стиле Unix (`/etc/pve/priv/shadow.cfg`). Пароль шифруется с использованием метода хеширования SHA-256. Этот метод аутентификации удобен для небольших (или даже средних) установок, где пользователям не нужен доступ ни к чему, кроме PVE. В этом случае пользователи полностью управляются PVE и могут менять свои пароли через графический интерфейс. PVE хранит данные пользователей в файле `/etc/pve/user.cfg`:

```
# cat /etc/pve/user.cfg
user:root@pam:1:0::::::
user:test@pve:1:0::::::
user:testuser@pve:1:0::::Just a test::
```

```
user:user@pam:1:0::::::

group:admin:user@pam::
group:testgroup:test@pve::
```

– LDAP – можно аутентифицировать пользователей через сервер LDAP (например, openldap). Сервер и дополнительный резервный сервер можно настроить, а соединение можно зашифровать с помощью SSL.

Пользователь root является неограниченным администратором и всегда может войти в систему, используя Linux PAM. Этого пользователя нельзя удалить, все системные письма будут отправляться на адрес электронной почты, назначенный этому пользователю.

Каждый пользователь может быть членом нескольких групп.

Чтобы пользователь мог выполнить какое-либо действие (например, просмотр, изменение или удаление ВМ), ему необходимо иметь соответствующие разрешения.

PVE использует систему управления разрешениями на основе ролей и путей. Запись в таблице разрешений позволяет пользователю или группе играть определенную роль при доступе к объекту или пути. Это означает, что такое правило доступа может быть представлено как тройка (путь, пользователь, роль) или (путь, группа, роль), причем роль содержит набор разрешенных действий, а путь представляет цель этих действий.

Роль – это список привилегий. В PVE предопределён ряд ролей:

– Administrator – имеет все привилегии;
– NoAccess – нет привилегий (используется для запрета доступа);
– PVEAdmin – все привилегии, кроме прав на изменение настроек системы (Sys.PowerMgmt, Sys.Modify, Realm.Allocate);
– PVEAuditor – доступ только для чтения;
– PVEDatastoreAdmin – создание и выделение места для резервного копирования и шаблонов;
– PVEDatastoreUser – выделение места для резервной копии и просмотр хранилища;
– PVEPoolAdmin – выделение пулов;
– PVESysAdmin – ACL пользователя, аудит, системная консоль и системные журналы;
– PVETemplateUser – просмотр и клонирование шаблонов;
– PVEUserAdmin – администрирование пользователей;
– PVEVMAdmin – управление ВМ;
– PVEVMUser – просмотр, резервное копирование, настройка CDROM, консоль ВМ, управление питанием ВМ.

Просмотреть список предопределенных ролей в веб-интерфейсе можно, выбрав «Датацентр» → «Разрешения»→«Роли» (*Рис. 158*).

Добавить новую роль можно как в веб-интерфейсе, так и в командной строке.

Список предопределенных ролей

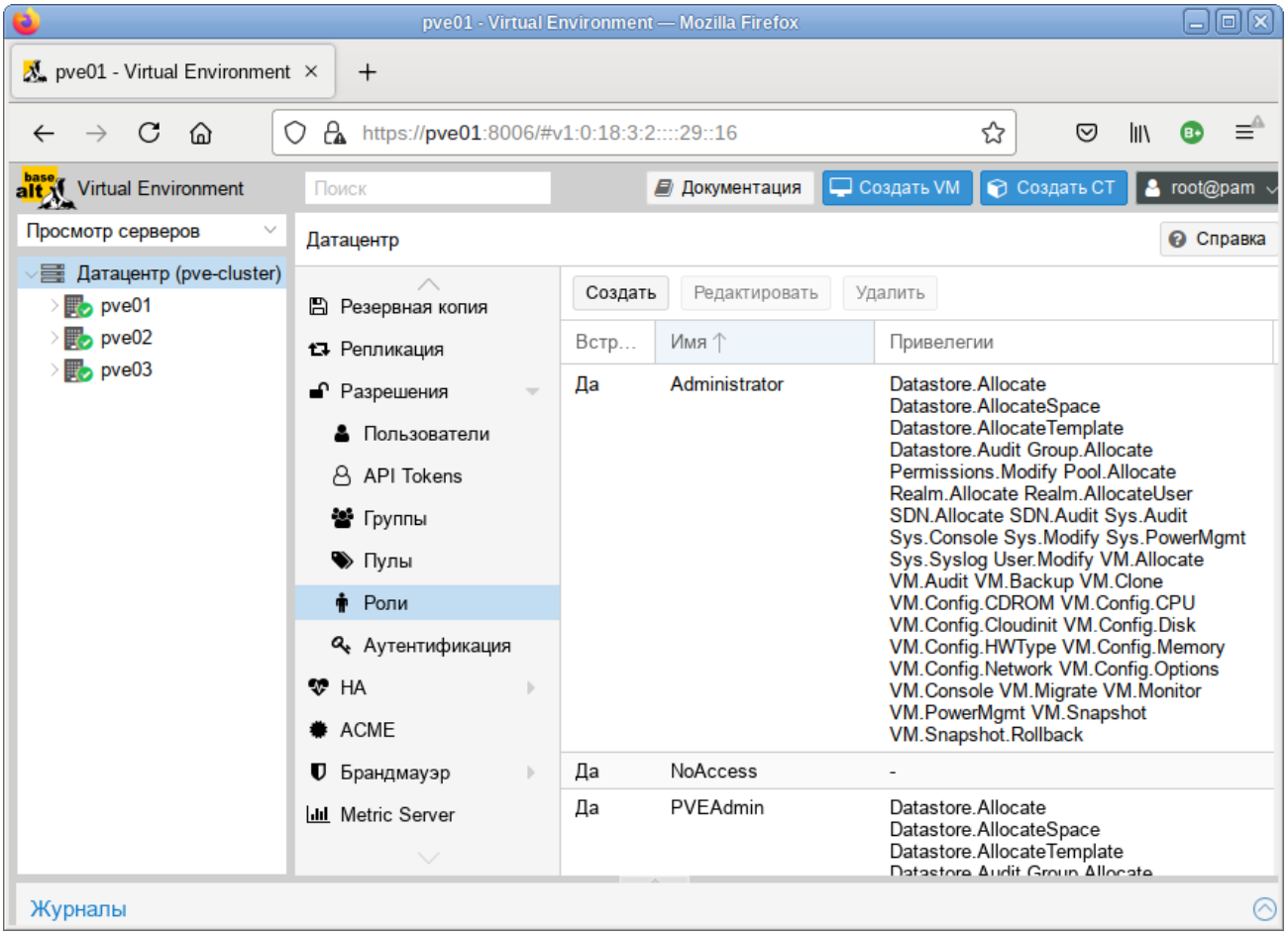

Рис. 158

Привилегия – это право на выполнение определенного действия. Для упрощения управления списки привилегий сгруппированы в роли, которые затем можно использовать в таблице разрешений. Привилегии не могут быть напрямую назначены пользователям, не будучи частью роли. Список используемых привилегий приведен в табл. 9.

Пул ресурсов – это набор ВМ, контейнеров и хранилищ. Пул ресурсов удобно использовать для обработки разрешений в случаях, когда определенные пользователи должны иметь контролируемый доступ к определенному набору ресурсов. Пулы ресурсов часто используются в тандеме с группами, чтобы члены группы имели разрешения на набор машин и хранилищ.

Т а б л и ц а 9 –Привилегии используемые в PVE

Привилегия	Описание
Привилегии узла/системы	
Permissions.Modify	Изменение прав доступа
Sys.PowerMgmt	Управление питанием узла (запуск, остановка, сброс, выключение)
Sys.Console	Консольный доступ к узлу
Sys.Syslog	Просмотр Syslog
Sys.Audit	Просмотр состояния/конфигурации узла, конфигурации кластера Corosync и конфигурации HA
Sys.Modify	Создание/удаление/изменение параметров сети узла
Group.Allocate	Создание/удаление/изменение групп
Pool.Allocate	Создание/удаление/изменение пулов
Realm.Allocate	Создание/удаление/изменение областей аутентификации
Realm.AllocateUser	Назначение пользователю области аутентификации
User.Modify	Создание/удаление/изменение пользователя
Права, связанные с ВМ	
VM.Allocate	Создание/удаление ВМ
VM.Migrate	Миграция ВМ на альтернативный сервер в кластере
VM.PowerMgmt	Управление питанием (запуск, остановка, сброс, выключение)
VM.Console	Консольный доступ к ВМ
VM.Monitor	Доступ к монитору виртуальной машины (kvm)
VM.Backup	Резервное копирование/восстановление ВМ
VM.Audit	Просмотр конфигурации ВМ
VM.Clone	Клонирование ВМ
VM.Config.Disk	Добавление/изменение/удаление дисков ВМ
VM.Config.CDROM	Извлечь/изменить CDROM
VM.Config.CPU	Изменение настроек процессора
VM.Config.Memory	Изменение настроек памяти
VM.Config.Network	Добавление/изменение/удаление сетевых устройств
VM.Config.HWType	Изменение типа эмуляции
VM.Config.Options	Изменение любой другой конфигурации ВМ
VM.Snapshot	Создание/удаление снимков ВМ
Права, связанные с хранилищем	
Datastore.Allocate	Создание/удаление/изменение хранилища данных
Datastore.AllocateSpace	Выделить место в хранилище
Datastore.AllocateTemplate	Размещение/загрузка шаблонов контейнеров и ISO-образов
Datastore.Audit	Просмотр хранилища данных

Права доступа назначаются объектам, таким как ВМ, хранилища или пулы ресурсов. PVE использует файловую систему как путь к этим объектам. Эти пути образуют естественное дерево, и права доступа более высоких уровней (более короткий путь) необязательно распространяются вниз по этой иерархии. Примеры:

– /nodes/{node} – доступ к серверам PVE;
– /vms – все ВМ;
– /vms/{vmid} – доступ к определенным ВМ;
– /storage/{storeid} – доступ к хранилищам;
– /access/groups – администрирование групп;

/access/realms/{realmid} – административный доступ.Для назначения разрешений необходимо в окне «Датацентр» → «Разрешения» нажать кнопку «Добавить» и в выпадающем меню выбрать «Разрешения группы», если разрешения назначаются группе пользователей, или «Разрешения пользователя», если разрешения назначаются пользователю. Далее в открывшемся окне (Рис. 159) выбрать путь, группу и роль и нажать кнопку «Добавить».

Добавление разрешений группе

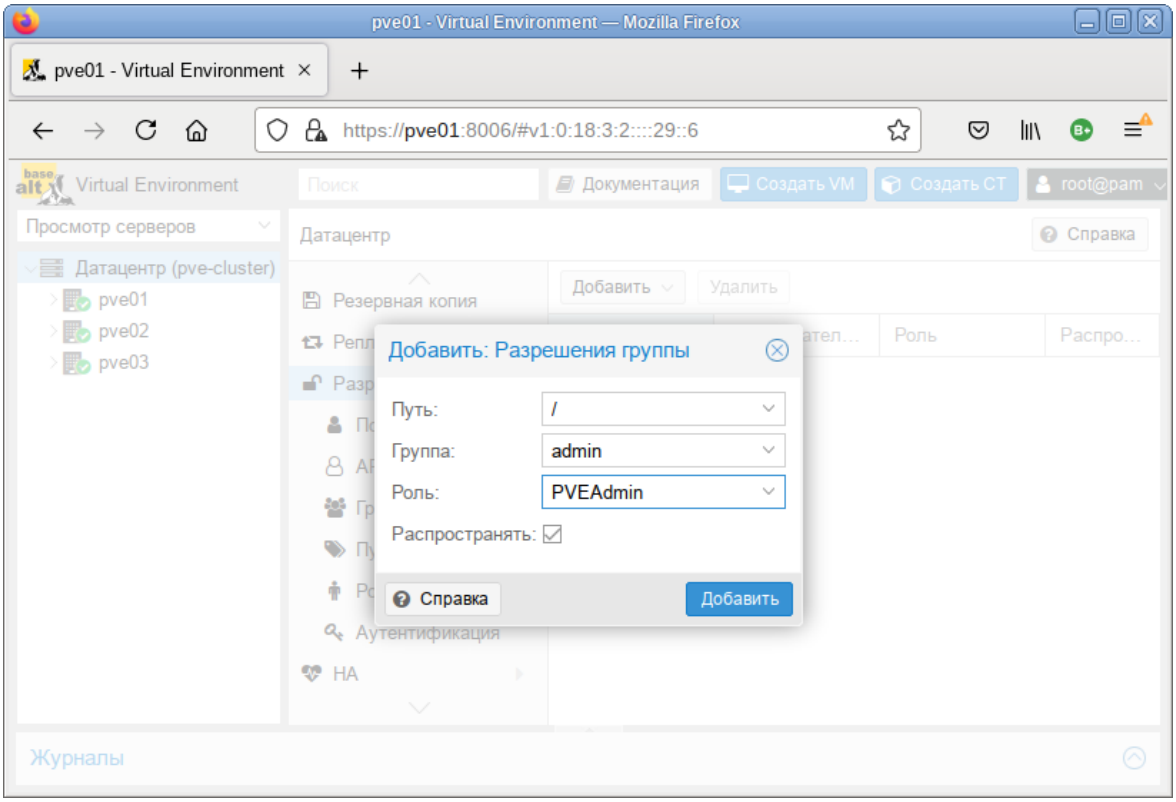

Рис. 159

Для добавления нового пользователя, необходимо в окне «Датацентр» → «Разрешения» → «Пользователи» нажать кнопку «Добавить». На Рис. 160 показано создание нового пользователя с использованием PAM аутентификации (системный пользователь user должен существовать, в ка-

честве пароля будет использоваться пароль для входа в систему). На Рис. 161 показано создание нового пользователя с использованием PVE аутентификации.

Создание нового пользователя с использованием PAM аутентификации

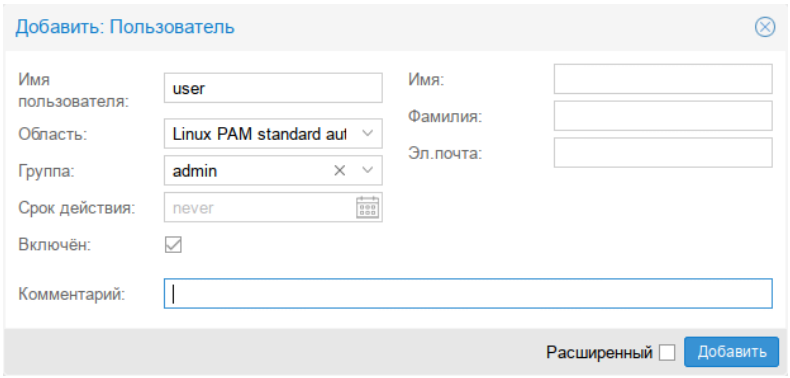

Рис. 160

Создание нового пользователя с использованием PVE аутентификации

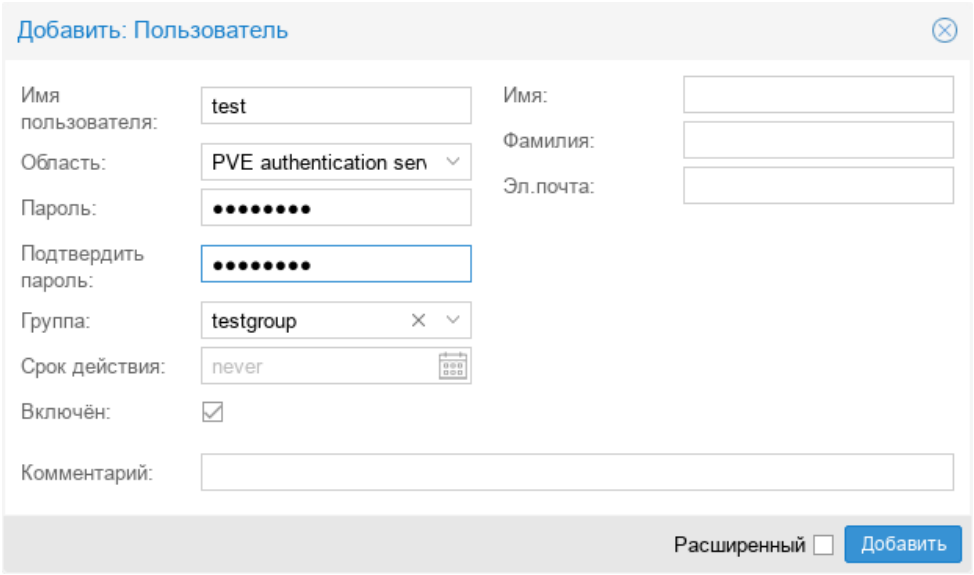

Рис. 161

Примеры использования командной строки для управления пользователями PVE:

Создать пользователя:

```
# pveum useradd testuser@pve -comment "Just a test"
```

Задать или изменить пароль:

```
# pveum passwd testuser@pve
```

Отключить пользователя:

```
# pveum usermod testuser@pve -enable 0
```

Создать новую группу:

```
# pveum groupadd testgroup
```

Создать новую роль:

```
# pveum roleadd PVE_Power-only -privs "VM.PowerMgmt VM.Console"
```

5 УПРАВЛЕНИЕ ВИРТУАЛИЗАЦИЕЙ НА ОСНОВЕ LIBVIRT

5.1 Установка и настройка libvirt

libvirt – это набор инструментов, предоставляющий единый API к множеству различных технологий виртуализации.

Кроме управления виртуальными машинами/контейнерами libvirt поддерживает управление виртуальными сетями и управление хранением образов.

Для управления из консоли разработан набор утилит virt-install, virt-clone, virsh и других. Для управления из графической оболочки можно воспользоваться virt-manager.

Любой виртуальный ресурс, необходимый для создания ВМ (compute, network, storage) представлен в виде объекта в libvirt. За процесс описания и создания этих объектов отвечает набор различных XML-файлов. Сама ВМ в терминологии libvirt называется доменом (domain). Это тоже объект внутри libvirt, который описывается отдельным XML-файлом.

При первоначальной установке и запуске libvirt по умолчанию создает мост (bridge) virbr0 и его минимальную конфигурацию. Этот мост не будет подключен ни к одному физическому интерфейсу, однако, может быть использован для связи виртуальных машин внутри одного гипервизора.

Примечание. Компоненты libvirt будут установлены в систему, если при установке дистрибутива выбрать профиль «Базовая виртуализация».

Примечание. На этапе «Подготовка диска» рекомендуется выбрать «KVM Server (large /var/lib/libvirt)».

Если же развертывание libvirt происходит в уже установленной системе на базе Девятой платформы, достаточно любым штатным способом (apt-get, aptitude, synaptic) установить пакеты:

```
# apt-get update
# apt-get install libvirt virt-install
```

Запуск службы:

```
# systemctl start libvirtd
# systemctl enable libvirtd
```

Для непривилегированного доступа (не root) к управлению libvirt, нужно добавить пользователя в группу vmusers:

```
# usermod -a -G vmusers user
```

Сервер виртуализации использует следующие каталоги хостовой файловой системы:

- /etc/libvirt/ – каталог с файлами конфигурации libvirt;
- /var/lib/libvirt/ – рабочий каталог сервера виртуализации libvirt;

– /var/log/libvirt – файлы журналов libvirt.

5.2 Утилиты управления

Основные утилиты командной строки для управления ВМ:

– qemu-img – управление образами дисков ВМ. Позволяет выполнять операции по созданию образов различных форматов, конвертировать файлы-образы между этими форматами, получать информацию об образах и объединять снимки ВМ для тех форматов, которые это поддерживают;
– virsh – консольный интерфейс управления ВМ, виртуальными дисками и виртуальными сетями;
– virt-clone –клонирование ВМ;
– virt-convert –конвертирования ВМ между различными форматами и программно-аппаратными платформами;
– virt-image – создание ВМ по их XML описанию;
– virt-install – создание ВМ с помощью опций командной строки;
– virt-xml – редактирование XML-файлов описаний ВМ.

5.2.1 Утилита Virsh

virsh – утилита для командной строки, предназначенная для управления ВМ и гипервизорами KVM.

virsh использует libvirt API и служит альтернативой графическому менеджеру виртуальных машин (virt-manager).

С помощью virsh можно сохранять состояние ВМ, переносить ВМ между гипервизорами и управлять виртуальными сетями.

В табл. 10 и табл. 11 приведены основные параметры утилиты командной строки virsh. Получить список доступных команд или параметров, можно используя команду: $ `virsh help`.

Т а б л и ц а 10 – Команды управления виртуальными машинами

Команда	Описание
`help`	Краткая справка
`list`	Просмотр всех ВМ
`dumpxml`	Вывести файл конфигурации XML для заданной ВМ
`create`	Создать ВМ из файла конфигурации XML и ее запуск
`start`	Запустить неактивную ВМ
`destroy`	Принудительно остановить работу ВМ
`define`	Определяет файл конфигурации XML для заданной ВМ
`domid`	Просмотр идентификатора ВМ

Команда	Описание
domuuid	Просмотр UUID ВМ
dominfo	Просмотр сведений о ВМ
domname	Просмотр имени ВМ
domstate	Просмотр состояния ВМ
quit	Закрыть интерактивный терминал
reboot	Перезагрузить ВМ
restore	Восстановить сохраненную в файле ВМ
resume	Возобновить работу приостановленной ВМ
save	Сохранить состояние ВМ в файл
shutdown	Корректно завершить работу ВМ
suspend	Приостановить работу ВМ
undefine	Удалить все файлы ВМ
migrate	Перенести ВМ на другой узел

Т а б л и ц а 11 – Параметры управления ресурсами ВМ и гипервизора

Команда	Описание
setmem	Определяет размер выделенной ВМ памяти
setmaxmem	Ограничивает максимально доступный гипервизору объ-ем памяти
setvcpus	Изменяет число предоставленных ВМ виртуальных про-цессоров
vcpuinfo	Просмотр информации о виртуальных процессорах
vcpupin	Настройка соответствий виртуальных процессоров
domblkstat	Просмотр статистики блочных устройств для работающей ВМ
domifstat	Просмотр статистики сетевых интерфейсов для работаю-щей ВМ
attach-device	Подключить определенное в XML-файле устройство к ВМ
attach-disk	Подключить новое дисковое устройство к ВМ
attach-interface	Подключить новый сетевой интерфейс к ВМ
detach-device	Отключить устройство от ВМ (принимает те же опреде-ления XML, что и attach-device)
detach-disk	Отключить дисковое устройство от ВМ
detach-interface	Отключить сетевой интерфейс от ВМ

5.2.2 Утилита virt-install

virt-install – это инструмент для создания ВМ, основанный на командной строке.

Далее подробно рассматриваются возможности создания ВМ при помощи утилиты командной строки virt-install. В табл. 12 приведено описание только наиболее часто используемых опций команды virt-install. Описание всех доступных опций можно получить, выполнив команду:

```
$ man virt-install
```

Утилита virt-install поддерживает как графическую установку операционных систем при помощи VNC и Spice, так и текстовую установку через последовательный порт. Гостевая система может быть настроена на использование нескольких дисков, сетевых интерфейсов, аудиоустройств и физических USB- и PCI-устройств.

Установочный носитель может располагаться как локально, так и удаленно, например, на NFS, HTTP или FTP серверах. В последнем случае virt-install получает минимальный набор файлов для запуска установки и позволяет установщику получить отдельные файлы. Также поддерживается загрузка по сети (PXE) и создание виртуальной машины/контейнера без установки операционной системы.

Утилита virt-install поддерживает большое число опции, позволяющих создать полностью независимую ВМ, готовую к работе, что хорошо подходит для автоматизации установки ВМ.

Т а б л и ц а 12 – Параметры команды virt-install

Команда	Описание
`-n NAME, --name=NAME`	Имя новой ВМ. Это имя должно быть уникально внутри одного гипервизора
`--memory MEMORY`	Определяет размер выделенной ВМ памяти (в МБ)

`--vcpus VCPUS`	Определяет количество виртуальных ЦПУ. Например:
	`--vcpus 5`
	`--vcpus 5,maxvcpus=10,cpuset=1-4,6,8`
	`--vcpus sockets=2,cores=4,threads=2`

`--cpu CPU`	Модель ЦП и его характеристики. Например:
	`--cpu coreduo,+x2apic`
	`--cpu host-passthrough`
	`--cpu host`

`--metadata METADATA`	Метаданные ВМ
Метод установки	
`--cdrom CDROM`	Установочный CD-ROM. Может указывать на файл ISO-образа или на устройство чтения CD/DVD-дисков
`-l LOCATION, --location LOCATION`	Источник установки, например, https://host/path
`--pxe`	Выполнить загрузку из сети используя протокол PXE
`--import`	Пропустить установку ОС, и создать ВМ на основе существующего образа диска

Команда	Описание

`--boot BOOT`	Параметры загрузки ВМ. Например:
	`--boot hd,cdrom,menu=on`
	`--boot init=/sbin/init (для контейнеров)`

`--os-type=DISTRO_TYPE`	Оптимизирует настройки ВМ для заданного типа ОС
`--os-variant=DISTRO_VARIANT`	Дополнительная оптимизация ВМ для конкретного варианта ОС

`--disk DISK`	Настройка пространства хранения данных. Например:
	`--disk size=10 (новый образ на 10 ГБ в выбранном по умолчанию месте)`
	`--disk /my/existing/disk,cache=none`
	`--disk device=cdrom,bus=scsi`
	`--disk=?`

`-w NETWORK, --network NETWORK`	Конфигурация сетевого интерфейса ВМ. Например:
	`--network bridge=mybr0`
	`--network network=my_libvirt_virtual_net`
	`--network network=mynet,model=virtio,mac=00:11...`
	`--network none`

`--graphics GRAPHICS`	Настройки экрана ВМ. Например:
	`--graphics spice`
	`--graphics vnc,port=5901,listen=0.0.0.0`
	`--graphics none`

`--input INPUT`	Конфигурация устройства ввода. Например:
	`--input tablet`
	`--input keyboard,bus=usb`

`--hostdev HOSTDEV`	Конфигурация физических USB/PCI и других устройств хоста для совместного использования ВМ
`-filesystem FILESYSTEM`	Передача каталога хоста гостевой системе. Например: `--filesystem /my/source/dir,/dir/in/guest`
Параметры платформы виртуализации	
`-v, --hvm`	Эта ВМ должна быть полностью виртуализированной
`-p, --paravirt`	Эта ВМ должна быть паравиртуализированной
`--container`	Тип ВМ – контейнер
`--virt-type VIRT_TYPE`	Тип гипервизора (kvm, qemu и т.п.)
`--arch ARCH`	Имитируемая архитектура процессора
`--machine MACHINE`	Имитируемый тип компьютера
Прочие параметры	

Команда	Описание
`--autostart`	Запускать домен автоматически при запуске хоста
`--transient`	Создать временный домен
`--noautoconsole`	Не подключаться к гостевой консоли автоматически
`-q, --quiet`	Подавлять вывод (за исключением ошибок)
`-d, --debug`	Вывести отладочные данные

5.2.3 Утилита qemu-img

qemu-img – инструмент для манипулирования с образами дисков машин QEMU.

Использование:

```
qemu-img command [command options]
```

Для манипуляции с образами используются следующие команды:

- create – создание нового образа диска;
- check – проверка образа диска на ошибки;
- convert – конвертация существующего образа диска в другой формат;
- info – получение информации о существующем образе диска;
- snapshot – управляет снимками состояний (snapshot) существующих образов дисков;
- commit – записывает произведенные изменения на существующий образ диска;
- rebase – создает новый базовый образ на основании существующего.

qemu-img работает со следующими форматами:

- raw – простой формат для дисковых образов, обладающий отличной переносимостью на большинство технологий виртуализации и эмуляции. Только непосредственно записанные секторы будут занимать место на диске. Действительный объем пространства, занимаемый образом, можно определить с помощью команд qemu-img info или ls -ls;
- qcow2 – формат QEMU. Этот формат рекомендуется использовать для небольших образов (в частности, если файловая система не поддерживает фрагментацию), дополнительного шифрования AES, сжатия zlib и поддержки множества снимков ВМ;
- qcow – старый формат QEMU. Используется только в целях обеспечения совместимости со старыми версиями;
- cow – формат COW (Copy On Write). Используется только в целях обеспечения совместимости со старыми версиями;
- vmdk – формат образов, совместимый с VMware 3 и 4;

– cloop – формат CLOOP (Compressed Loop). Его единственное применение состоит в обеспечении повторного использования сжатых напрямую образов CD-ROM, например, Knoppix CD-ROM.

Команда получения сведений о дисковом образе:

```
# qemu-img info /var/lib/libvirt/images/alt8.0.qcow2
image: /var/lib/libvirt/images/alt8.0.qcow2
file format: qcow2
virtual size: 12 GiB (12884901888 bytes)
disk size: 12 GiB
cluster_size: 65536
Format specific information:
    compat: 1.1
    compression type: zlib
    lazy refcounts: true
    refcount bits: 16
    corrupt: false
    extended l2: false
```

В результате будут показаны сведения о запрошенном образе, в том числе зарезервированный объем на диске, а также информация о снимках ВМ.

Команда создания образа для жесткого диска (динамически расширяемый):

```
# qemu-img create -f qcow2 /var/lib/libvirt/images/hdd.qcow2 20G
```

Команда конвертирования образ диска из формата raw в qcow2:

```
# qemu-img convert -f raw -O qcow2 disk_hd.img disk_hd.qcow2
```

5.2.4 Менеджер виртуальных машин virt-manager

Менеджер виртуальных машин virt-manager предоставляет графический интерфейс для доступа к гипервизорам и ВМ в локальной и удаленных системах. С помощью virt-manager можно создавать ВМ. Кроме того, virt-manager выполняет управляющие функции:

– выделение памяти;
– выделение виртуальных процессоров;
– мониторинг производительности;
– сохранение и восстановление, приостановка и возобновление работы, запуск и завершение работы виртуальных машин;
– доступ к текстовой и графической консоли;
– автономная и живая миграция.

Для запуска менеджера виртуальных машин, в меню приложений необходимо выбрать «Система»→ «Менеджер виртуальных машин» («Manage virtual machines»).

Примечание. На управляющей машине должен быть установлен пакет virt-manager.

В главном окне менеджера (Рис. 162), при наличии подключения к гипервизору, будут показаны все запущенные ВМ. Двойной щелчок на имени ВМ открывает ее консоль.

Главное окно менеджера виртуальных машин

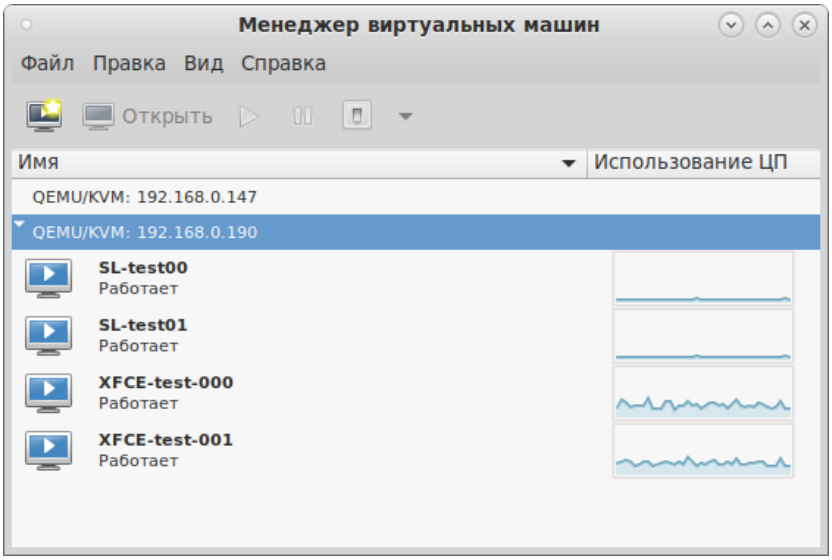

Рис. 162

5.3 Подключение к гипервизору

5.3.1 Управление доступом к libvirt через SSH

В дополнение к аутентификации SSH также необходимо определить управление доступом для службы Libvirt в хост-системе (Рис. 163).

Доступ к libvirt с удаленного узла

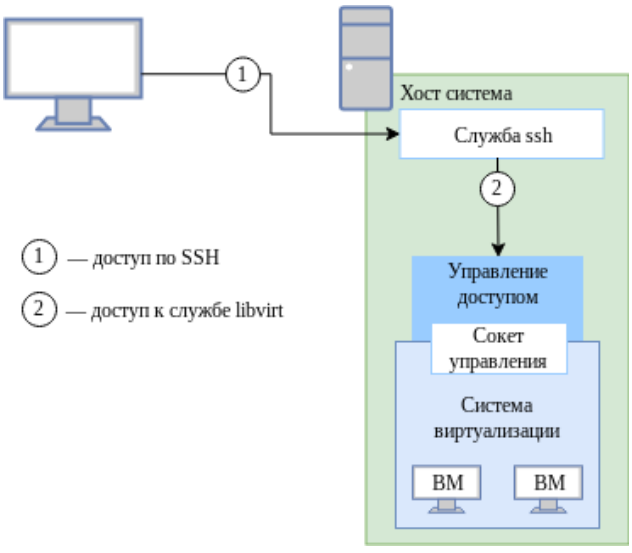

Рис. 163

Для настройки подключения к удаленному серверу виртуализации на узле, с которого будет производиться подключение, необходимо сгенерировать SSH-ключ и скопировать его публичную часть на сервер. Для этого с правами пользователя, от имени которого будет создаваться подключение, требуется выполнить в консоли следующие команды:

```
$ ssh-keygen -t rsa
$ ssh-copy-id user@192.168.0.147
```

где 192.168.0.147 – IP-адрес сервера с libvirt.

В результате получаем возможность работы с домашними каталогами пользователя user на сервере с libvirt.

Для доступа к libvirt достаточно добавить пользователя user в группу vmusers на сервере, либо скопировать публичный ключ пользователю root и подключаться к серверу по ssh от имени root – root@server

5.3.2 Подключение к сессии гипервизора с помощью virsh

Команда подключения к гипервизору:

```
virsh -c URI
```

Если параметр URI не задан, то libvirt попытается определить наиболее подходящий гипервизор.

Параметр URI может принимать следующие значения:

- qemu:///system – подключиться к службе, которая управляет KVM/QEMU-доменами и запущена под root. Этот вариант используется по умолчанию для пользователей virt-manager;
- qemu:///session – подключиться к службе, которая управляет KVM/QEMU-доменами и запущена от имени непривилегированного пользователя;
- lxc:/// – подключиться к гипервизору для создания LXC контейнеров (должен быть установлен пакет libvirt-lxc).

Чтобы установить соединение только для чтения, к приведенной выше команде следует добавить опцию `--readonly`.

Пример создания локального подключения:

```
$ virsh -c qemu:///system list --all
ID   Имя      Состояние
-------------------------
 -    alt9.1   выключен
```

Подключение к удаленному гипервизору QEMU через протокол SSH:

```
$ virsh -c qemu+ssh://user@192.168.0.147/system
Добро пожаловать в virsh — интерактивный терминал виртуализации.
```

```
Введите «help» для получения справки по командам
         «quit», чтобы завершить работу и выйти.

virsh #
```

где user – имя пользователя на удаленном хосте, который входит в группу vmusers. 192.168.0.147 – IP адрес или имя хоста виртуальных машин.

5.3.3 Настройка соединения с удаленным гипервизором в virt-manager

На управляющей системе можно запустить virt-manager, выполнив следующую команду:

```
virt-manager -c qemu+ssh://user@192.168.0.147/system
```

где user – имя пользователя на удаленном хосте, который входит в группу vmusers. 192.168.0.147 – IP адрес или имя хоста виртуальных машин.

virt-manager позволяет управлять несколькими удаленными хостами ВМ. Подключение virt-manager к удаленным хостам, также, можно настроить и в графическом интерфейсе менеджера виртуальных машин.

Для создания нового подключения необходимо в меню менеджера выбрать «Файл» → «Добавить соединение…».

В открывшемся окне необходимо выбрать сессию гипервизора, отметить пункт «Подключиться к удаленному хосту с помощью SSH», ввести имя пользователя и адрес сервера и нажать кнопку «Подключиться» (Рис. 164).

Окно соединений менеджера виртуальных машин

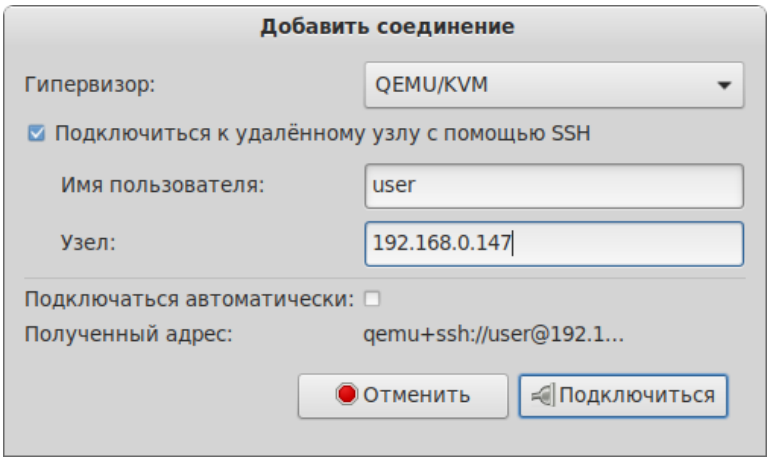

Рис. 164

5.4 Создание виртуальных машин

Наиболее важным этапом в процессе использования виртуализации является создание ВМ. Именно при создании ВМ задается используемый тип виртуализации, способы доступа к ВМ, подключение к локальной сети и другие характеристики виртуального оборудования.

Установка ВМ может быть запущена из командной строки с помощью программ virsh и virt-install или из пользовательского интерфейса программы virt-manager.

5.4.1 Создание виртуальной машины на основе файла конфигурации (утилита virsh)

ВМ могут быть созданы из файлов конфигурации. Для этого конфигурация ВМ должна быть описана в XML формате.

Команда создания ВМ из XML файла:

```
# virsh create guest.xml
```

Для получения файла конфигурации можно сделать копию существующего XML-файла ранее созданной ВМ, или использовать опцию `dumpxml`:

```
# virsh dumpxml <domain>
```

Эта команда выводит XML-файл конфигурации ВМ в стандартный вывод (stdout). Можно сохранить данные, отправив вывод в файл.

Пример передачи вывода в файл `guest.xml`:

```
# virsh dumpxml alt9.1 > guest.xml
```

Можно отредактировать этот файл конфигурации, чтобы настроить дополнительные устройства или развернуть дополнительные ВМ.

5.4.2 Создание ВМ с помощью virt-install

Минимальные требуемые опции для создания ВМ: `--name`, `--ram`, хранилище (`--disk`, `--filesystem` или `--nodisks`) и опции установки.

Чтобы использовать команду virt-install, необходимо сначала загрузить ISO-образ той ОС, которая будет устанавливаться.

Команда создания ВМ:

```
# virt-install --connect qemu:///system
--name alt-server \
--os-type=linux \
--os-variant=alt9.1 \
--cdrom /var/lib/libvirt/boot/alt-server-x86_64.iso \
--graphics vnc\
--disk pool=default,size=20,bus=virtio,format=qcow2 \
--ram 2048 \
--vcpus=2 \
--network network=default \
--hvm \
--virt-type=kvm
```

где

--name alt-server – название ВМ;

--os-type=linux – тип ОС;

--os-variant=alt9.1 – версия ОС;

--cdrom /var/lib/libvirt/boot/alt-server-x86_64.iso – путь к ISO-образу установочного диска ОС;

--graphics vnc – графическая консоль;

--disk pool=default,size=20,bus=virtio,format=qcow2 – ВМ будет создана в пространстве хранения объемом 20 ГБ, которое автоматически выделяется из пула хранилищ default. Образ диска для этой виртуальной машины будет создан в формате qcow2;

--ram 2048 – объем оперативной памяти;

--vcpus=2 – количество процессоров;

--network network=default – виртуальная сеть default;

--hvm – полностью виртуализированная система;

--virt-type=kvm – использовать модуль ядра KVM, который задействует аппаратные возможности виртуализации процессора.

Последние две опции команды virt-install оптимизируют ВМ для использования в качестве полностью виртуализированной системы (--hvm) и указывают, что KVM является базовым гипервизором (--virt-type) для поддержки новой ВМ. Обе этих опции обеспечивают определенную оптимизацию в процессе создания и установки операционной системы; если эти опции не заданы в явном виде, то вышеуказанные значения применяются по умолчанию.

Список доступных вариантов ОС можно получить, выполнив команду:

```
$ osinfo-query os
```

Запуск Live CD в ВМ без дисков:

```
# virt-install \
 --hvm \
 --name demo \
 --ram 500 \
 --nodisks \
 --livecd \
 --graphics vnc \
 --cdrom /var/lib/libvirt/boot/altlive.iso
```

Запуск /usr/bin/httpd в контейнере (LXC), с ограничением памяти в 512 МБ и двумя ядрами хост-системы:

```
# virt-install \
 --connect lxc:/// \
```

```
 --name httpd_guest \
 --ram 512 \
 --vcpus 2 \
 --init /usr/bin/httpd
```

Создать ВМ, используя существующий том хранилища:

```
# virt-install \
 --name demo \
 --ram 512 \
 --disk /home/user/VMs/mydisk.img \
 --import
```

5.4.3 Создание виртуальных машин с помощью virt-manager

Новую ВМ можно создать, нажав кнопку «Создать виртуальную машину» в главном окне virt-manager, либо выбрав в меню «Файл»→ «Создать виртуальную машину».

На первом шаге создания ВМ необходимо выбрать метод установки ОС (Рис. 165) и нажать кнопку «Вперед».

Создание ВМ. Выбор метода установки

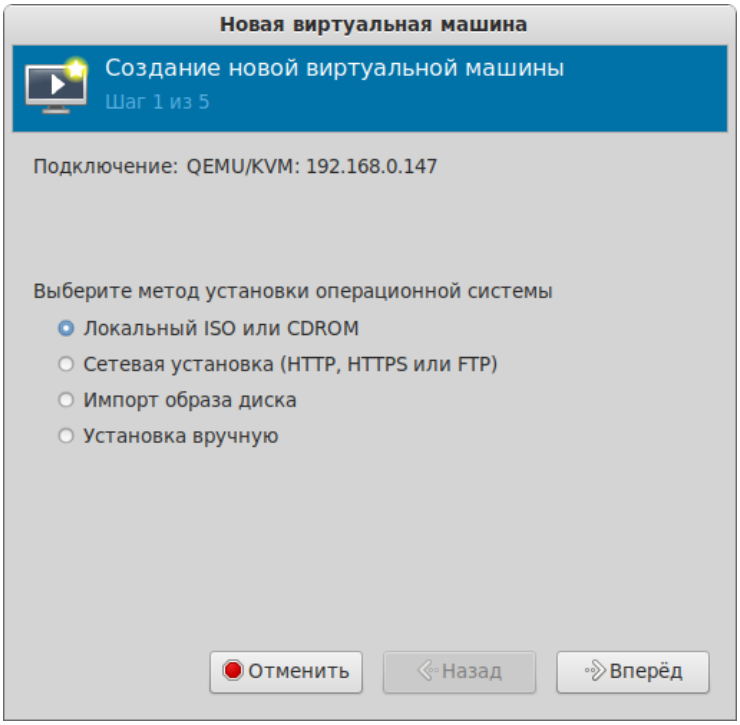

Рис. 165

В следующем окне для установки гостевой ОС требуется указать ISO-образ установочного диска ОС или CD/DVD-диск с дистрибутивом (Рис. 166). Данное окно будет выглядеть по-разному в зависимости от выбора, сделанного на предыдущем этапе. Здесь также можно указать версию устанавливаемой ОС.

Создание ВМ. Выбор ISO образа

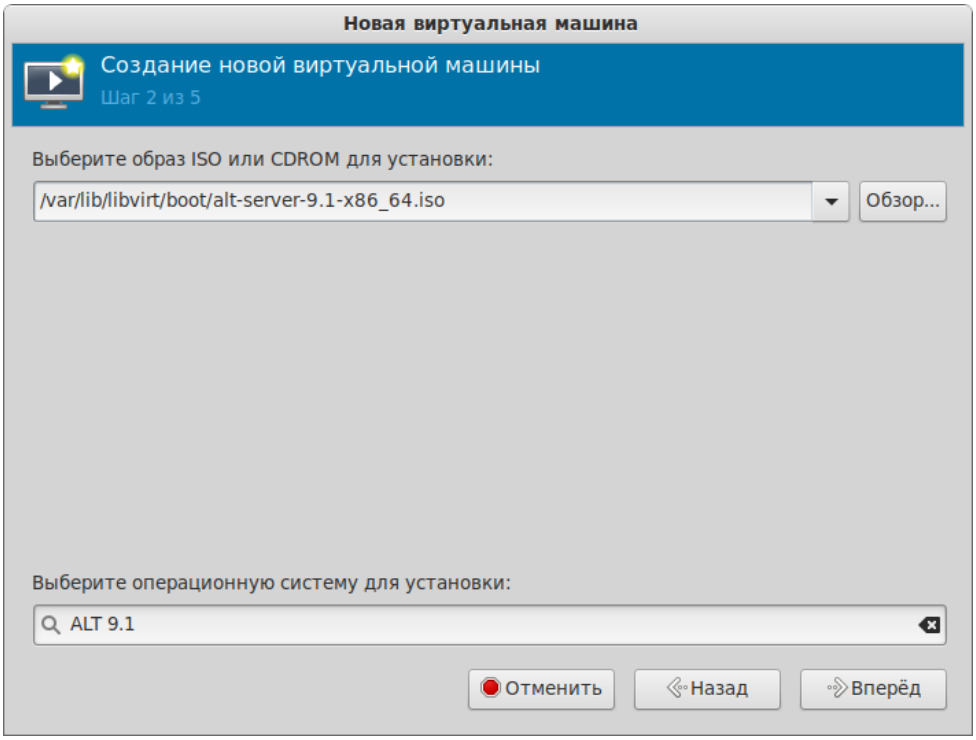

Рис. 166

На третьем шаге необходимо указать размер памяти и количество процессоров для ВМ (Рис. 167). Эти значения влияют на производительность хоста и ВМ.

Создание ВМ. Настройка ОЗУ и ЦПУ для ВМ

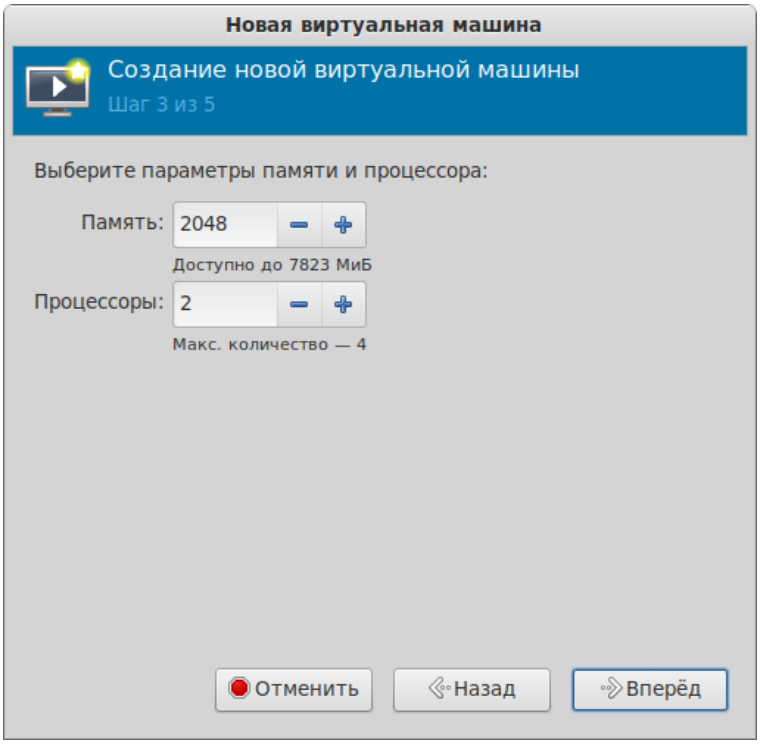

Рис. 167

На следующем этапе настраивается пространство хранения данных (Рис. 168).

Создание ВМ. Настройка пространства хранения данных

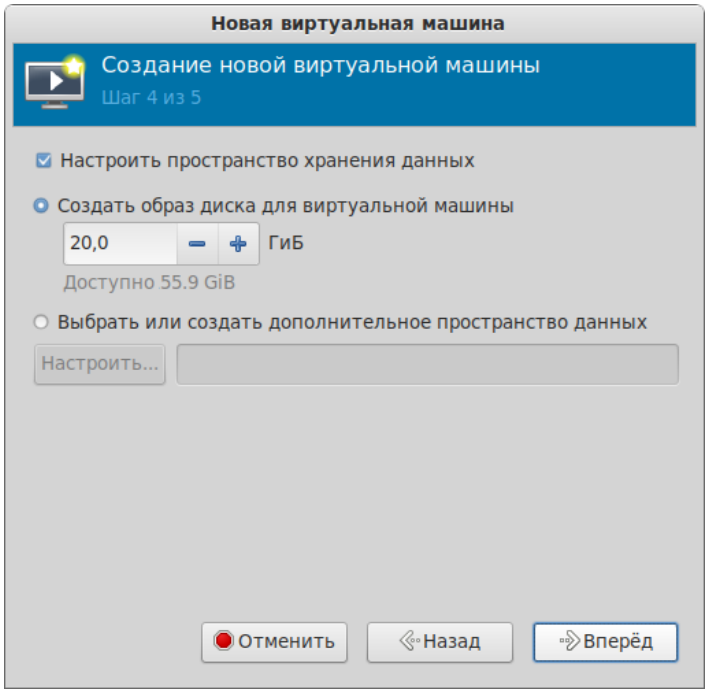

Рис. 168

На последнем этапе (Рис. 169) можно задать название ВМ, выбрать сеть и нажать кнопку «Готово».

Создание ВМ. Выбор сети

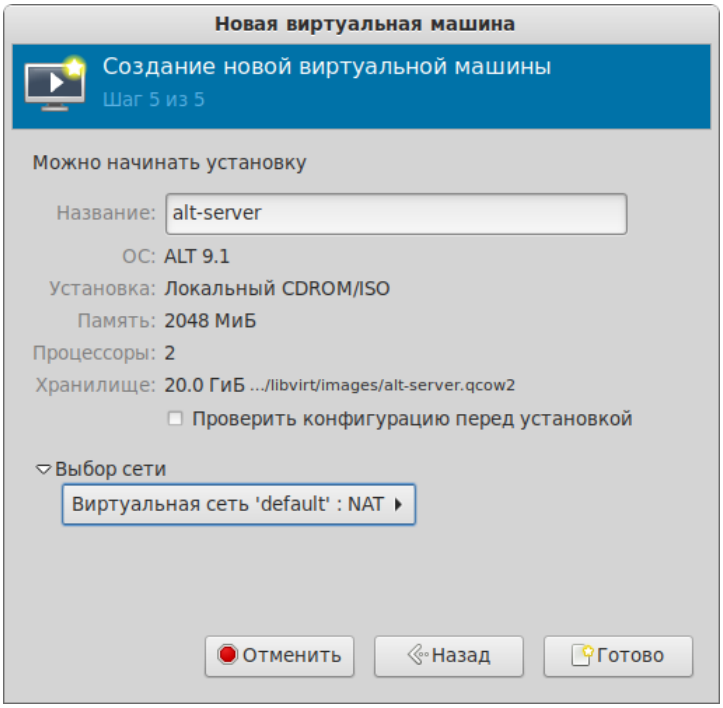

Рис. 169

В результате созданная ВМ будет запущена и после завершения исходной загрузки начнется стандартный процесс установки ОС (Рис. 170).

Установка ОС

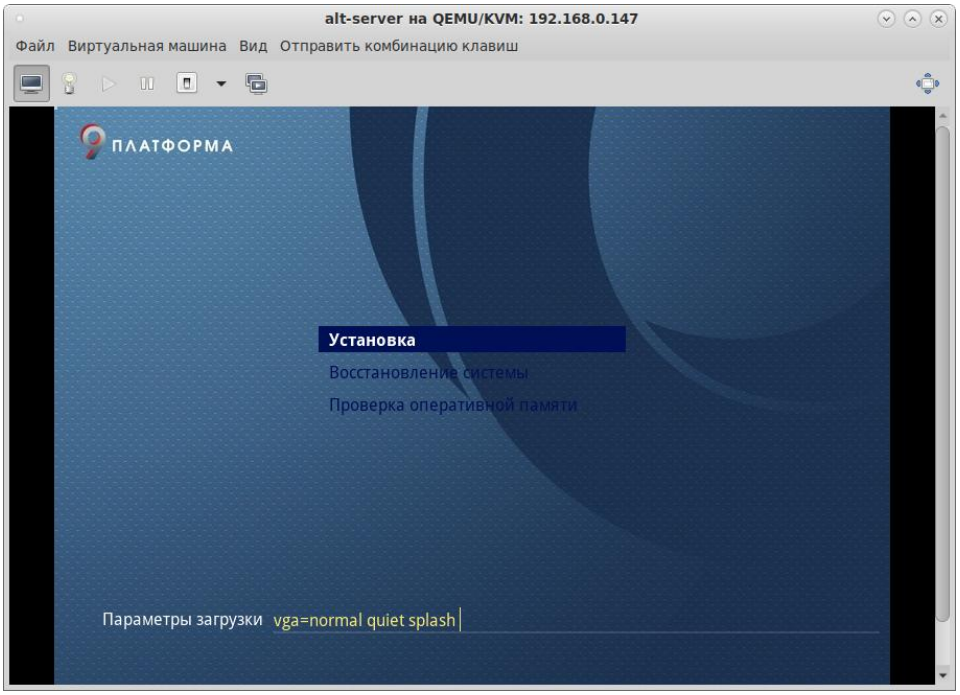

Рис. 170

Окружение локального рабочего стола способно перехватывать комбинации клавиш (например, <Ctrl>+<Alt>+<F11>) для предотвращения их отправки гостевой машине. Чтобы отправить такие последовательности, используется свойство «западания» клавиш virt-manager. Для перевода клавиши в нажатое состояние необходимо нажать клавишу модификатора (<Ctrl> или <Alt>) 3 раза. Клавиша будет считаться нажатой до тех пор, пока не будет нажата любая клавиша, отличная от модификатора. Таким образом, чтобы передать гостевой системе комбинацию <Ctrl>+<Alt>+<F11>, необходимо последовательно нажать <Ctrl>+<Ctrl>+<Ctrl>+<Alt>+<F11> или воспользоваться меню «Отправить комбинацию клавиш».

5.5 Запуск и управление функционированием ВМ

5.5.1 Управление состоянием ВМ в командной строке

Команды управления состоянием ВМ:

- start – запуск ВМ;
- shutdown – завершение работы. Поведение выключаемой ВМ можно контролировать с помощью параметра on_shutdown (в файле конфигурации);
- destroy – принудительная остановка. Использование virsh destroy может повредить гостевые файловые системы. Рекомендуется использовать опцию shutdown;
- reboot – перезагрузка ВМ. Поведение перезагружаемой ВМ можно контролировать с помощью параметра on_reboot (в файле конфигурации);

- suspend – приостановить ВМ. Когда ВМ находится в приостановленном состоянии, она потребляет системную оперативную память, но не ресурсы процессора;
- resume – возобновить работу приостановленной ВМ;
- save – сохранение текущего состояния ВМ. Эта команда останавливает ВМ, сохраняет данные в файл, что может занять некоторое время (зависит от объема ОЗУ ВМ);
- restore – восстановление ВМ, ранее сохраненной с помощью команды `virsh save`. Сохраненная машина будет восстановлена из файла и перезапущена (это может занять некоторое время). Имя и идентификатор UUID ВМ останутся неизменными, но будет предоставлен новый идентификатор домена;
- undefine – удалить ВМ (конфигурационный файл тоже удаляется);
- autostart – добавить ВМ в автозагрузку;
- autostart --disable – удалить из автозагрузки.

В результате выполнения следующих команд, ВМ alt-server будет остановлена и затем удалена:

```
# virsh destroy alt-server
# virsh undefine alt-server
```

5.5.2 Управление состоянием ВМ в менеджере виртуальных машин

Для запуска ВМ в менеджере виртуальных машин virt-manager, необходимо выбрать ВМ из списка и нажать на кнопку «Включить виртуальную машину» (Рис. 171).

Включение ВМ

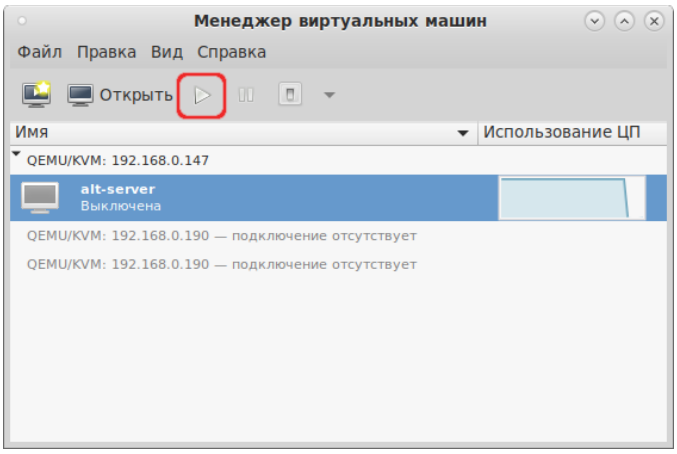

Рис. 171

Для управления запущенной ВМ используются соответствующие кнопки панели инструментов virt-manager (Рис. 172).

Управлять состоянием ВМ можно также выбрав соответствующий пункт в контекстном меню ВМ (Рис. 173).

Кнопки управления состоянием ВМ

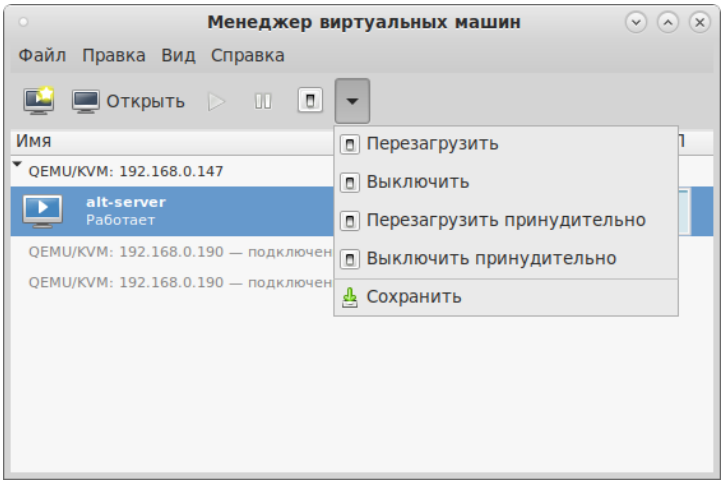

Рис. 172

Контекстное меню ВМ

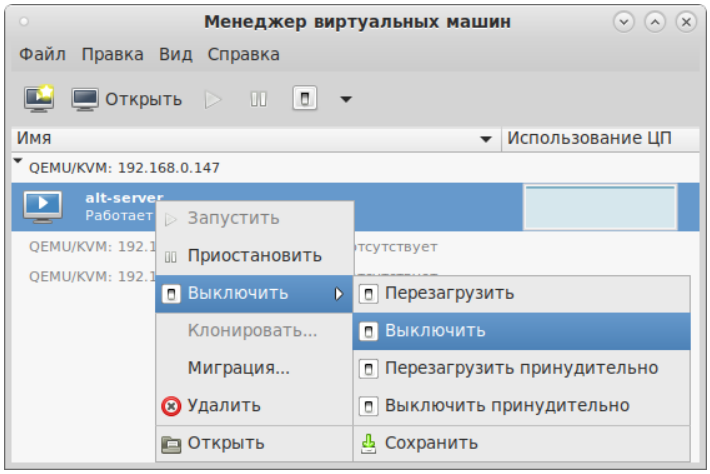

Рис. 173

5.6 Подключение к виртуальному монитору ВМ

Доступ к рабочему столу ВМ может быть организован по протоколам VNC и SPICE.

К каждой из ВМ можно подключиться, используя один IP адрес и разные порты. Порт доступа к ВМ может быть назначен вручную или автоматически. Удаленный доступ к ВМ можно защитить паролем.

5.6.1 Использование протокола SPICE

Чтобы добавить поддержку SPICE в существующую ВМ, необходимо отредактировать её конфигурацию:

```
#  virsh edit alt-server
```

Добавить графический элемент SPICE, например:

```
<graphics type='spice' port='5900' autoport='yes' listen='127.0.0.1'>
      <listen type='address' address='127.0.0.1'/>
    </graphics>
```

Добавить видеоустройство QXL:

```
<video>
    <model type='qxl'/>
</video>
```

После остановки и перезапуска ВМ она должна быть доступна через SPICE.

Проверка параметров подключения к ВМ:

```
# virsh domdisplay alt-server
spice://127.0.0.1:5900
```

В данном примере доступ к ВМ будет возможен только с локального адреса (127.0.0.1). Для удаленного подключения к ВМ SPICE-сервер должен обслуживать запросы с общедоступных сетевых интерфейсов. Для возможности подключения с других машин в конфигурации ВМ необходимо указать адрес 0.0.0.0:

```
<graphics type='spice' port='5900' autoport='yes' listen='0.0.0.0' passwd='mypasswd'>
      <listen type='address' address='0.0.0.0'/>
    </graphics>
```

Пример настроек доступа к рабочему столу по протоколу SPICE в менеджере ВМ показан на Рис. 174.

Менеджер ВМ. Вкладка «Дисплей Spice»

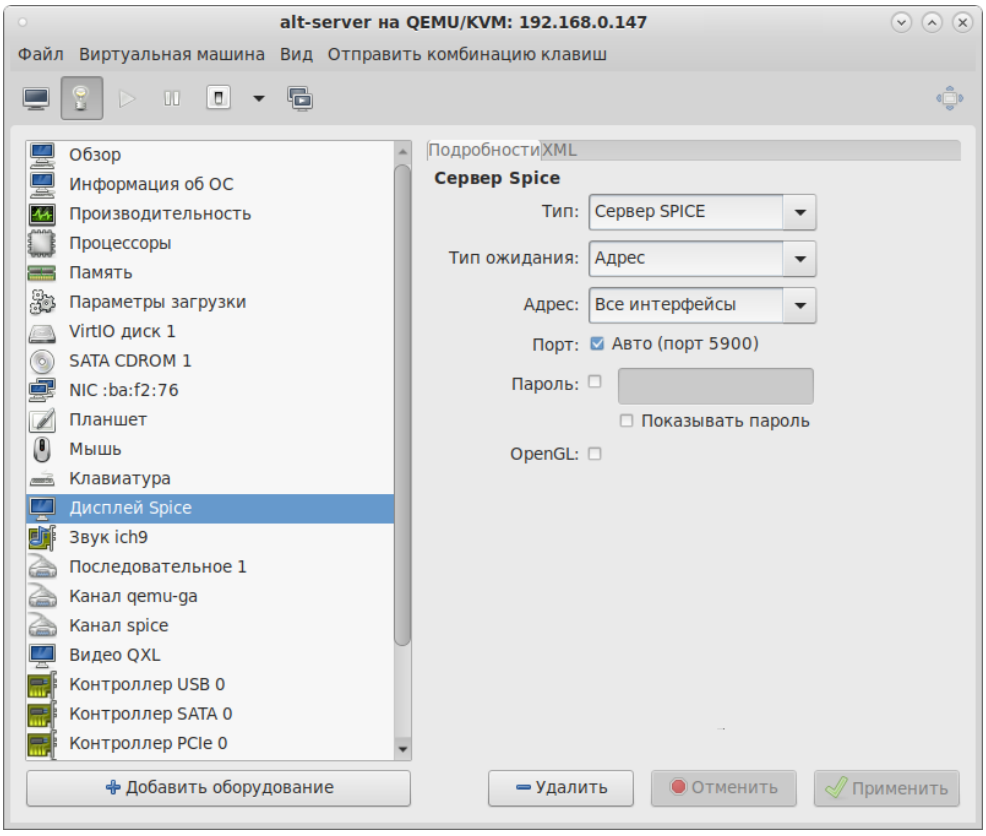

Рис. 174

Для подключения к SPICE-серверу может использоваться встроенный в virt-manager просмотрщик или любой SPICE-клиент. Примеры подключений (на хосте, с которого происходит подключение, должен быть установлен пакет virt-viewer):

```
$ virt-viewer -c qemu+ssh://user@192.168.0.147/system -d alt-server

$ remote-viewer "spice://192.168.0.147:5900"
```

Примечание. При использовании любого SPICE-клиента подключение происходит к порту и адресу хоста KVM, а не к фактическому имени/адресу ВМ.

5.6.2 Использование протокола VNC

Пример настройки доступа к рабочему столу ВМ по протоколу VNC, в файле конфигурации ВМ:

```
<graphics type='vnc' port='5900' autoport='no' listen='0.0.0.0'>
      <listen type='address' address='0.0.0.0'/>
    </graphics>
```

Пример настроек доступа к рабочему столу по протоколу VNC в менеджере ВМ показан на Рис. 175.

Менеджер ВМ. Вкладка «Дисплей VNC»

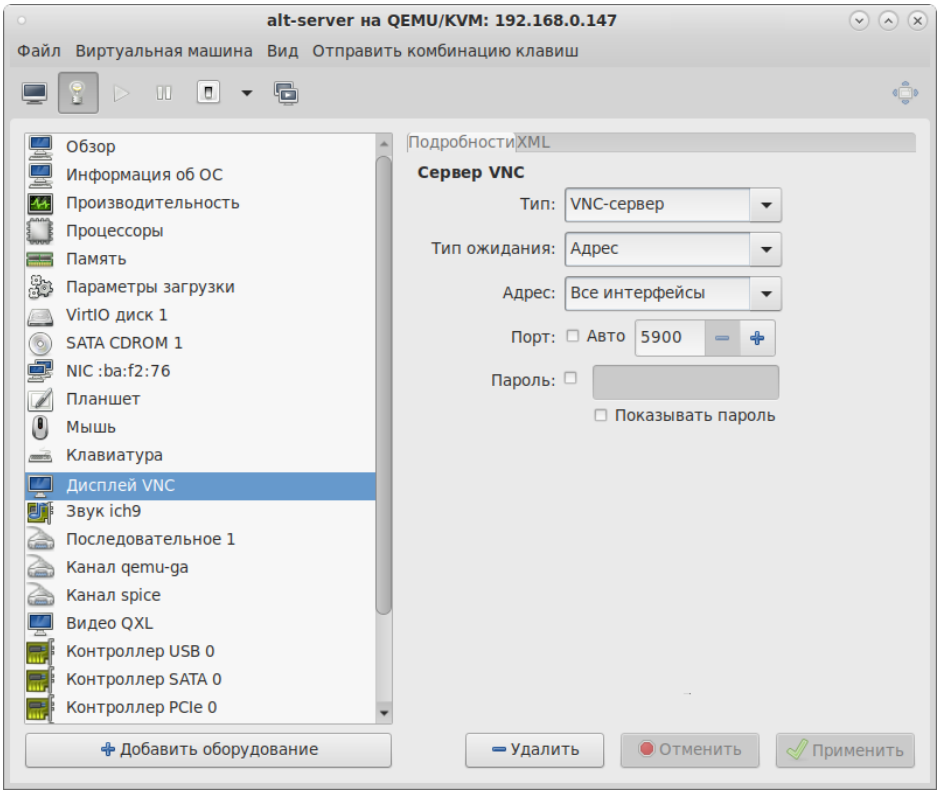

Рис. 175

Проверка параметров подключения к ВМ:

```
# virsh domdisplay alt-server
vnc://localhost:0
```

Для подключения к VNC-серверу может использоваться встроенный в virt-manager просмотрщик или любой VNC -клиент. Примеры подключений (на хосте, с которого происходит подключение, должны быть соответственно установлены пакеты virt-viewer или tigervnc):

```
$ virt-viewer -c qemu+ssh://user@192.168.0.147/system -d alt-server
$ vncviewer 192.168.0.147:5900
```

5.7 Управление ВМ

5.7.1 Управление конфигурацией ВМ

5.7.1.1 Редактирование файла конфигурации ВМ

ВМ могут редактироваться либо во время работы, либо в автономном режиме. Эту функциональность предоставляет команда `virsh edit`. Например, команда редактирования ВМ с именем alt-server:

```
# virsh edit alt-server
```

В результате выполнения этой команды откроется окно текстового редактора, заданного переменной оболочки $EDITOR.

5.7.1.2 Получение информации о ВМ

Команда для получения информации о ВМ:

```
virsh dominfo <domain>
```

где [--domain] <строка> – имя, ID или UUID домена

Пример вывода `virsh dominfo`:

```
$ virsh dominfo alt9.0
ID:             4
Имя:          alt-server
UUID:           c8170e81-92ab-44f3-bb6c-14155e11f848
Тип ОС:     hvm
Статус:   работает
CPU:            1
Время CPU: 7244,5s
Макс.память: 1048576 KiB
Занято памяти: 1048576 KiB
Постоянство: no
Автозапуск: выкл.
Управляемое сохранение: no
Модель безопасности: none
DOI безопасности: 0 0
```

Получение информации об узле:

```
virsh nodeinfo
```

Пример вывода `virsh nodeinfo`:

```
Модель процессора: x86_64
CPU:                 4
Частота процессора: 979 MHz
Сокеты:         1
Ядер на сокет: 2
Потоков на ядро: 2
Ячейки NUMA:   1
Объём памяти: 8011152 KiB
```

Просмотр списка ВМ:

```
virsh list
```

Опции команды `virsh list`:

- `--inactive` – показать список неактивных доменов;
- `--all` – показать все ВМ независимо от их состояния.

Пример вывода `virsh list`:

```
$ virsh list --all
 ID   Имя          Состояние
--------------------------
 4    alt-server   работает
```

Столбец «Статус» может содержать следующие значения:

- работает (running) – работающие ВМ, то есть те машины, которые используют ресурсы процессора в момент выполнения команды;
- blocked – заблокированные, неработающие машины. Такой статус может быть вызван ожиданием ввода/вывода или пребыванием машины в спящем режиме;
- приостановлен (paused) – приостановленные домены. В это состояние они переходят, если администратор нажал кнопку паузы в окне менеджера ВМ или выполнил команду virsh suspend. В приостановленном состоянии ВМ продолжает потреблять ресурсы, но не может занимать больше процессорных ресурсов;
- выключен (shutdown) – ВМ, завершающие свою работу. При получении ВМ сигнала завершения работы, она начнет завершать все процессы (некоторые операционные системы не отвечают на такие сигналы);
- dying – сбойные домены и домены, которые не смогли корректно завершить свою работу;

– crashed – сбойные домены, работа которых была прервана. В этом состоянии домены находятся, если не была настроена их перезагрузка в случае сбоя.

Команда получения информации о виртуальных процессорах:

```
virsh vcpuinfo <domain>
```

Пример вывода:

```
# virsh vcpuinfo alt-server
Виртуальный процессор:  0
CPU:            1
Состояние: работает
Время CPU: 845,0s
Соответствие ЦП: y
```

Команда сопоставления виртуальных процессоров физическим:

```
virsh vcpupin <domain> [--vcpu <число>] [--cpulist <строка>] [--config] [--live] [--current]
```

Здесь:

[--domain] <строка> – имя, ID или UUID домена;

--vcpu <число> – номер виртуального процессора;

--cpulist <строка> – номера физических процессоров. Если номера не указаны, команда вернет текущий список процессоров;

--config – с сохранением после перезагрузки;

--live – применить к работающему домену;

--current – применить к текущему домену.

Пример вывода:

```
 # virsh vcpupin alt-server
 Виртуальный процессор:   Соответствие ЦП
----------------------------------------
 0                        0-3
```

Команда изменения числа процессоров для домена (заданное число не может превышать значение, определенное при создании ВМ):

```
virsh setvcpus <domain> <count> [--maximum] [--config] [--live] [--current] [--guest] [--hotpluggable]
```

где

[--domain] <строка> – имя, ID или UUID домена;

[--count] <число> – число виртуальных процессоров;

--maximum – установить максимальное ограничение на количество виртуальных процессоров, которые могут быть подключены после следующей перезагрузки домена;

--config – с сохранением после перезагрузки;

--live – применить к работающему домену;

--current – применить к текущему домену;

--guest – состояние процессоров ограничивается гостевым доменом.

Команда изменения выделенного ВМ объема памяти:

```
virsh setmem <domain> <size> [--config] [--live] [--current]
```

где

[--domain] <строка> – имя, ID или UUID домена;

[--size] <число> – целое значение нового размера памяти (по умолчанию в КБ);

--config – с сохранением после перезагрузки;

--live – применить к работающему домену;

--current – применить к текущему домену.

Объем памяти, определяемый заданным числом, должен быть указан в килобайтах. Объем не может превышать значение, определенное при создании ВМ, но в то же время не должен быть меньше 64 мегабайт. Изменение максимального объема памяти может оказать влияние на функциональность ВМ только в том случае, если указанный размер меньше исходного. В таком случае использование памяти будет ограничено.

Команда изменения максимальное ограничение памяти:

```
virsh setmaxmem <domain> <size> [--config] [--live] [--current]
```

где

[--domain] <строка> – имя, ID или UUID домена;

[--size] <число> – целое значение максимально допустимого размера памяти (по умолчанию в КБ);

--config – с сохранением после перезагрузки;

--live – применить к работающему домену;

--current – применить к текущему домену.

Примеры изменения размера оперативной памяти и количества виртуальных процессоров соответственно:

```
# virsh setmaxmem --size 624000 alt-server
# virsh setmem --size 52240 alt-server
# virsh setvcpus --config alt-server 3 --maximum
```

Команда для получения информации о блочных устройствах работающей ВМ:

```
virsh domblkstat <domain> [--device <строка>] [--human]
```

где:

[--domain] <строка> – имя, ID или UUID домена;

--device <строка> – блочное устройство;

--human – форматировать вывод.

Команда для получения информации о сетевых интерфейсах работающей ВМ:

```
virsh domifstat <domain> <interface>
```

где:

[--domain] <строка> – имя, ID или UUID домена;

[--interface] <строка> – устройство интерфейса, указанное по имени или MAC-адресу.

5.7.1.3 Конфигурирование ВМ в менеджере виртуальных машин

С помощью менеджера виртуальных машин можно получить доступ к подробной информации о всех ВМ, для этого следует:

1) в главном окне менеджера выбрать ВМ;
2) нажать кнопку «Открыть» (Рис. 176);
3) в открывшемся окне нажать кнопку «Показать виртуальное оборудование» (Рис. 177);
4) появится окно просмотра сведений ВМ.

Окно менеджера виртуальных машин

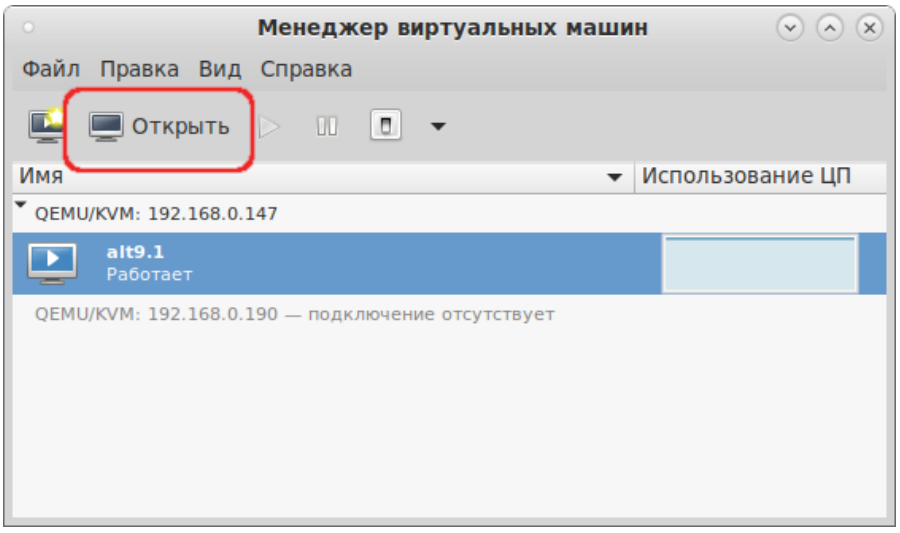

Рис. 176

Для изменения требуемого параметра необходимо перейти на нужную вкладку, внести изменения и подтвердить операцию, нажав кнопку «Применить» (Рис. 178 – Рис. 179).

Окно параметров ВМ

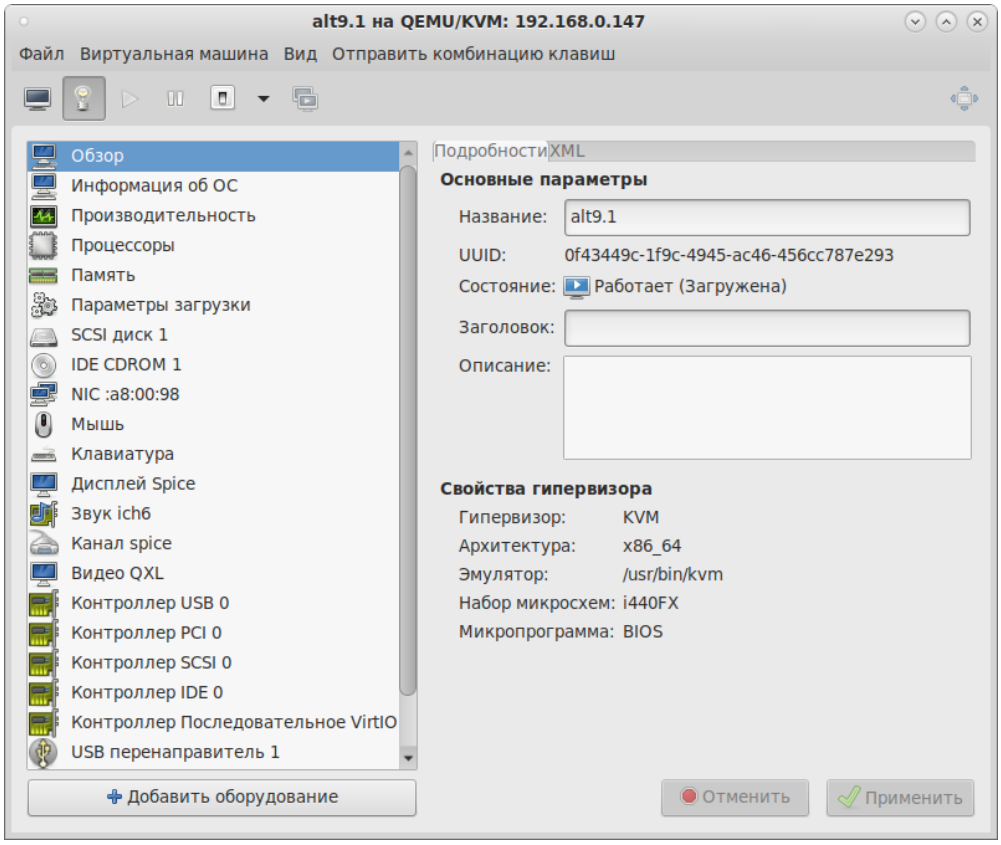

Рис. 177

Вкладка «Память»

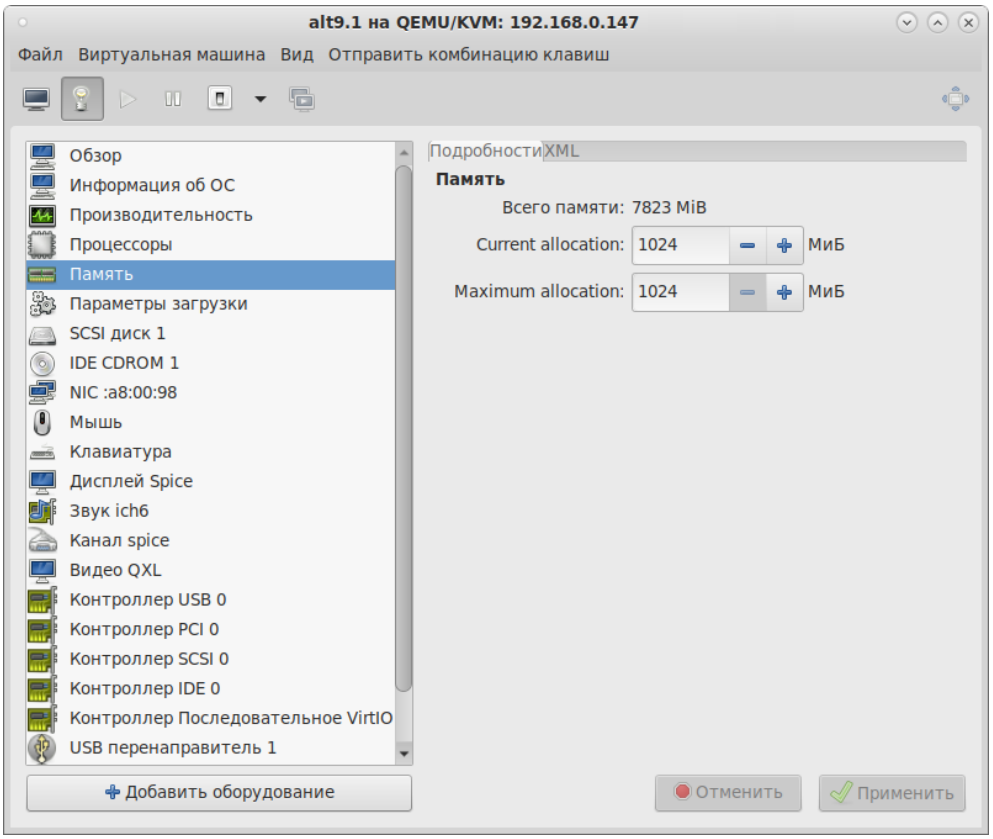

Рис. 178

Вкладка «Процессоры»

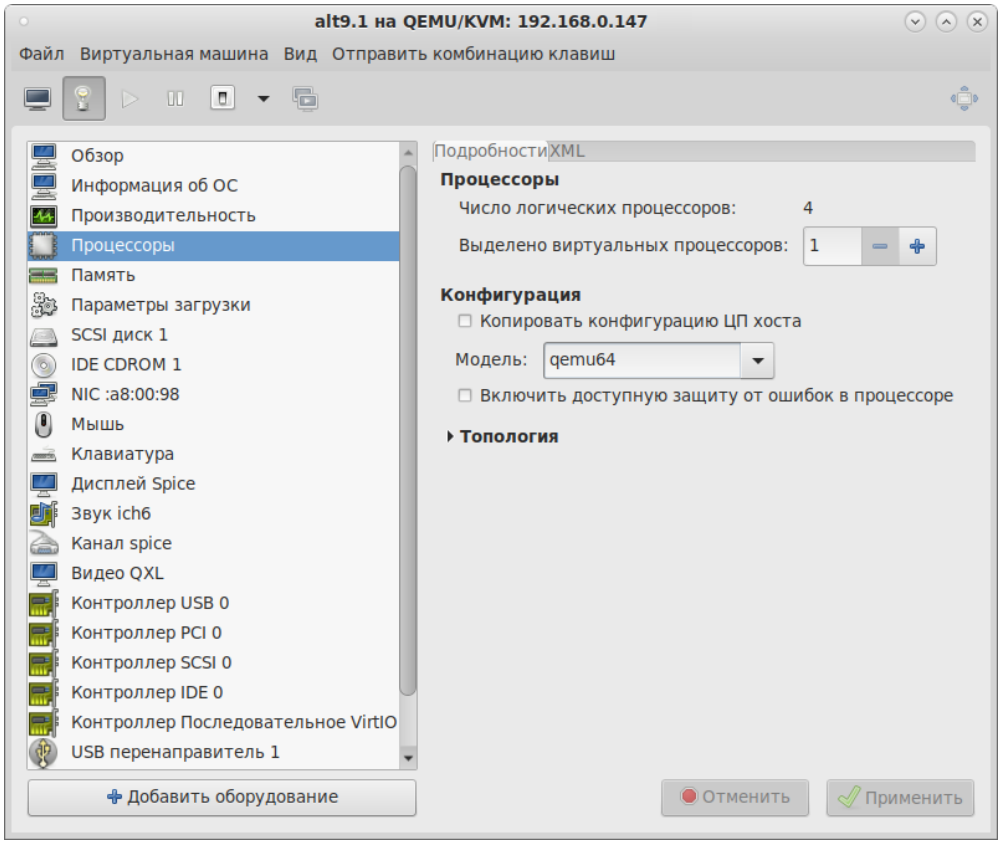

Рис. 179

5.7.1.4 *Мониторинг состояния*

С помощью менеджера виртуальных машин можно изменить настройки контроля состояния ВМ.

Для этого в меню «Правка» следует выбрать пункт «Параметры», в открывшемся окне «Настройки» на вкладке «Статистика» можно задать время обновления состояния ВМ в секундах (Рис. 180).

Вкладка «Статистика»

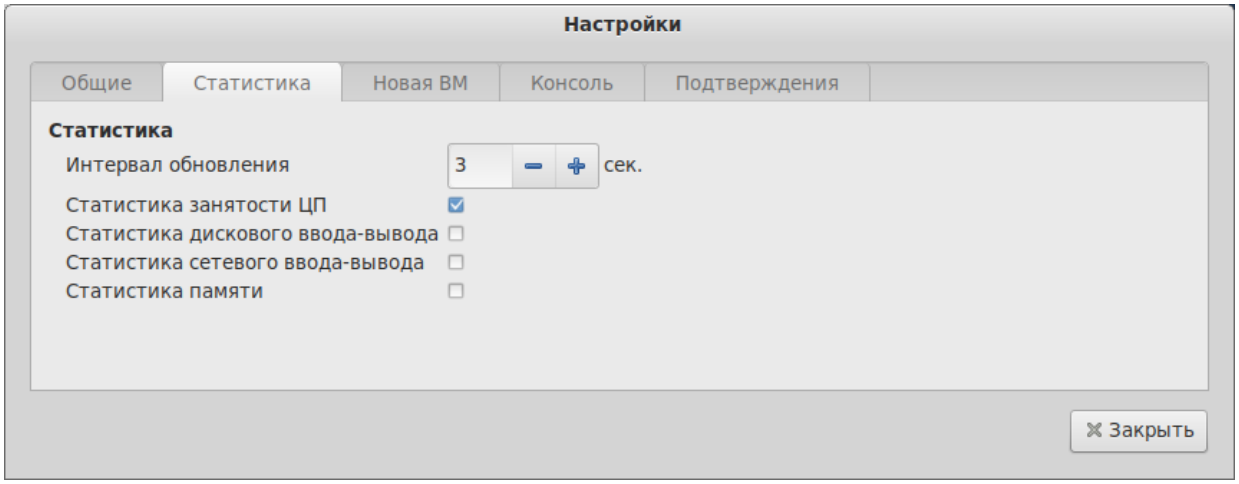

Рис. 180

Во вкладке «Консоль» (Рис. 181) можно выбрать, как открывать консоль, и указать устройство ввода.

Вкладка «Консоль»

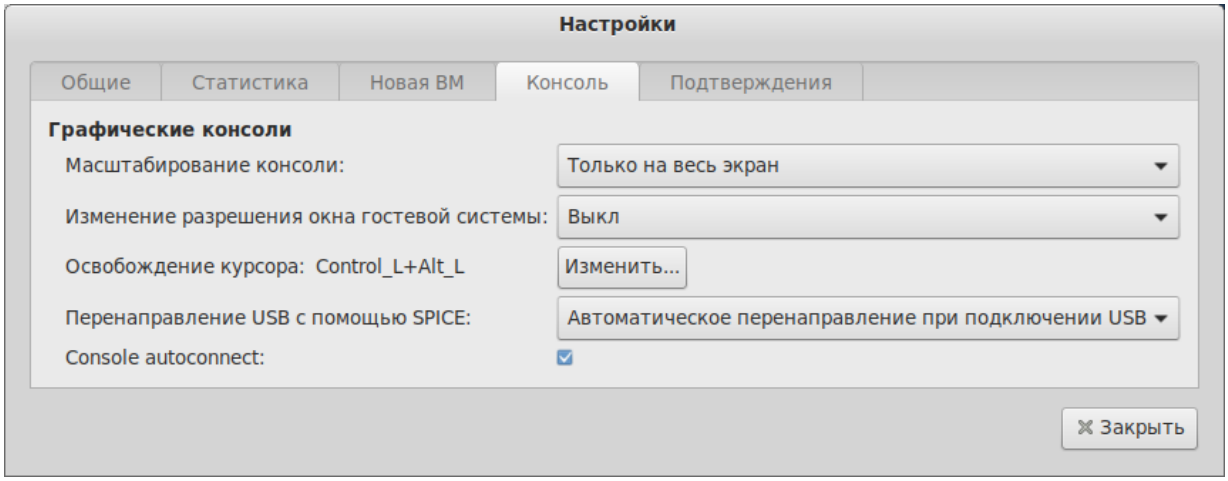

Рис. 181

5.7.2 Управление виртуальными сетевыми интерфейсами и сетями

При базовых настройках используется виртуальная сеть недоступная извне.

Доступ по IP может быть осуществлен с компьютера, на котором поднят KVM. Изнутри доступ происходит через NAT.

Возможные варианты настройки сети:

- NAT – это вариант по умолчанию. Внутренняя сеть, предоставляющая доступ к внешней сети с автоматическим применением NAT;
- Маршрутизация (Routed) – аналогично режиму NAT внутренняя сеть, предоставляющая доступ к внешней сети, но без NAT. Предполагает дополнительные настройки таблиц маршрутизации во внешней сети;
- Изолированная IPv4/IPv6 сеть (Isolated) – в этом режиме ВМ, подключенные к виртуальному коммутатору, могут общаться между собой и с хостом. При этом их трафик не будет выходить за пределы хоста;
- Bridge – подключение типа мост. Позволяет реализовать множество различных конфигураций, в том числе и назначение IP из реальной сети;
- SR-IOV pool (Single-root IOV) – перенаправление одной PCI сетевых карт хост-машины на ВМ. Технология SR-IOV повышает производительность сетевой виртуализации, избавляя гипервизор от обязанности организовывать совместное использование физического адаптера и перекладывая задачу реализации мультиплексирования на сам адаптер. В этом случае обеспечивается прямая пересылка ввода/вывода с ВМ непосредственно на адаптер.

5.7.2.1 Управление виртуальными сетями в командной строке

Команды управления виртуальными сетями:

- virsh net-autostart имя_сети – автоматический запуск заданной сети;
- virsh net-create файл_XML – создание и запуск новой сети на основе существующего XML-файла;
- virsh net-define файл_XML – создание нового сетевого устройства на основе существующего XML-файла (устройство не будет запущено);
- virsh net-list – просмотр списка виртуальных сетей;
- virsh net-dumpxml имя_сети – просмотр информации о заданной виртуальной сети;
- virsh net-destroy имя_сети – удаление заданной сети;
- virsh net-name UUID_сети – преобразование заданного идентификатора в имя сети;
- virsh net-uuid имя_сети – преобразование заданного имени в идентификатор UUID;
- virsh net-update имя_сети – обновить существующую конфигурацию сети;
- virsh net-start имя_неактивной_сети – запуск неактивной сети;
- virsh net-undefine имя_неактивной_сети — удаление определения неактивной сети.

```
# virsh net-list --all
Имя       Состояние    Автозапуск   Постоянный
----------------------------------------------
 default   не активен   no           yes

# virsh net-start default
Сеть default запущен

# virsh net-autostart default
Добавлена метка автоматического запуска сети default

# virsh net-list
 Имя       Состояние   Автозапуск   Постоянный
----------------------------------------------
 default   активен     yes          yes

# virsh net-dumpxml default
<network>
  <name>default</name>
  <uuid>0b37eff3-2234-4929-8a42-04a9cf35d3aa</uuid>
```

```
  <forward mode='nat'/>
  <bridge name='virbr0' stp='on' delay='0'/>
  <mac address='52:54:00:d2:30:b6'/>
  <ip address='192.168.122.1' netmask='255.255.255.0'>
    <dhcp>
      <range start='192.168.122.2' end='192.168.122.254'/>
    </dhcp>
  </ip>
</network>
```

Иногда бывает полезно выдавать клиенту один и тот же IP-адрес независимо от момента обращения. Пример добавления статического сопоставления MAC- и IP-адреса ВМ:

1) получить MAC-адрес ВМ (alt-server – имя ВМ):

```
# virsh dumpxml alt-server | grep 'mac address'
   <mac address='52:54:00:ba:f2:76'/>
```

2) отредактировать XML-конфигурацию сети (default – имя сети):

```
# virsh net-edit default
```

после строки:

```
<range start='192.168.122.2' end='192.168.122.254'/>
```

вставить строки с MAC-адресами виртуальных адаптеров:

```
<host mac='52:54:00:ba:f2:76' name='alt-server' ip='192.168.122.50'/>
```

3) сохранить изменения и перезапустить виртуальную сеть:

```
# virsh net-destroy default
# virsh net-start default
```

Изменения внесённые с помощью команды `virsh net-edit` не вступят в силу в силу до тех пор, пока сеть не будет перезапущена, что приведет к потере всеми ВМ сетевого подключения к хосту до тех пор, пока их сетевые интерфейсы повторно не подключаться.

Изменения в конфигурацию сети, можно внести с помощью команды `virsh net-update`, которая требует немедленного применения изменений. Например, чтобы добавить запись статического хоста, можно использовать команду:

```
# virsh net-update default add ip-dhcp-host \
   "<host           mac='52:54:00:ba:f2:76'           name='alt-server'
ip='192.168.122.50' />" \
   --live --config
```

5.7.2.2 Управление виртуальными сетями в менеджере виртуальных машин

В менеджере виртуальных машин virt-manager существует возможность настройки виртуальных сетей для обеспечения сетевого взаимодействия ВМ как между собой, так и с хостовой ОС.

Для настройки виртуальной сети с помощью virt-manager необходимо:

1) в меню «Правка» выбрать «Свойства подключения» (Рис. 182);
2) в открывшемся окне перейти на вкладку «Виртуальные сети» (Рис. 183);
3) доступные виртуальные сети будут перечислены в левой части окна. Для доступа к настройкам сети необходимо выбрать сеть.

Меню «Правка»

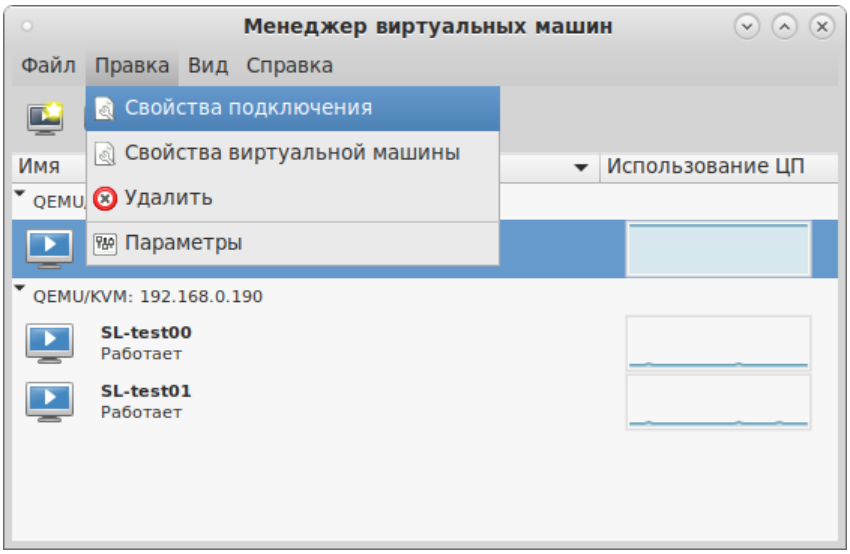

Рис. 182

Окно параметров виртуальной сети

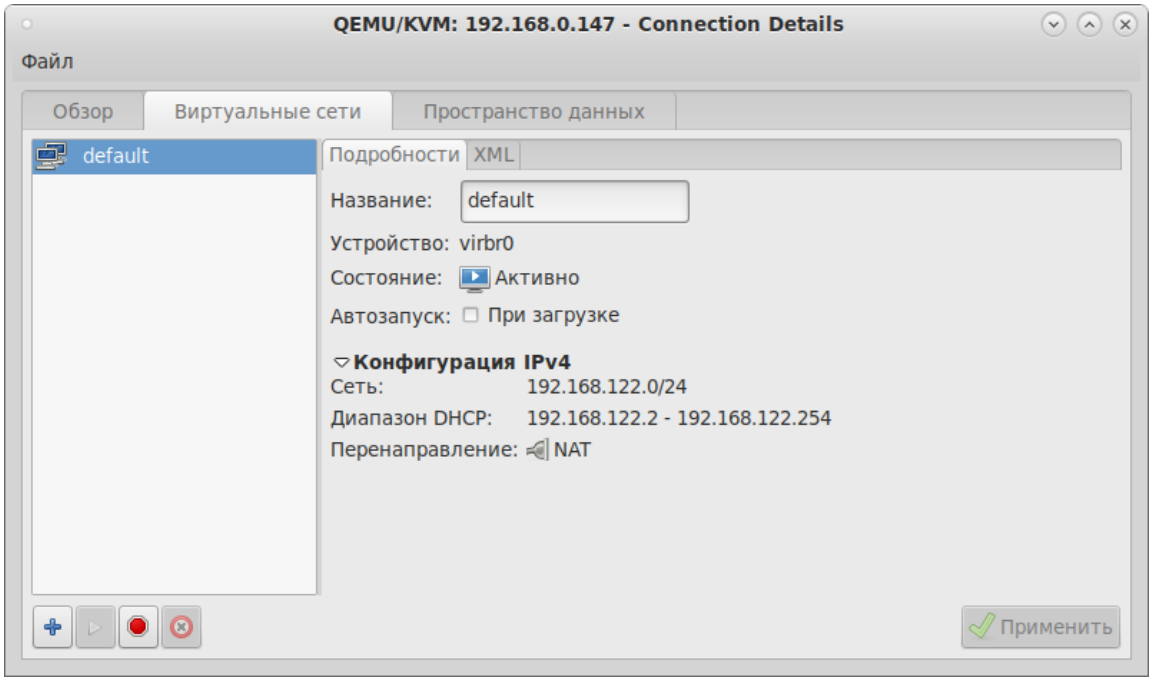

Рис. 183

Для добавления новой виртуальной сети следует нажать кнопку «Добавить сеть» («+»), расположенную в нижнем левом углу диалогового окна «Свойства соединения» (Рис. 183). В открывшемся окне (Рис. 184) следует ввести имя для новой сети и задать необходимые настройки: выбрать способ подключения виртуальной сети к физической, ввести пространство адресов IPv4 для виртуальной сети, указать диапазон DHCP, задав начальный и конечный адрес и нажать кнопку «Готово».

Создание новой виртуальной сети

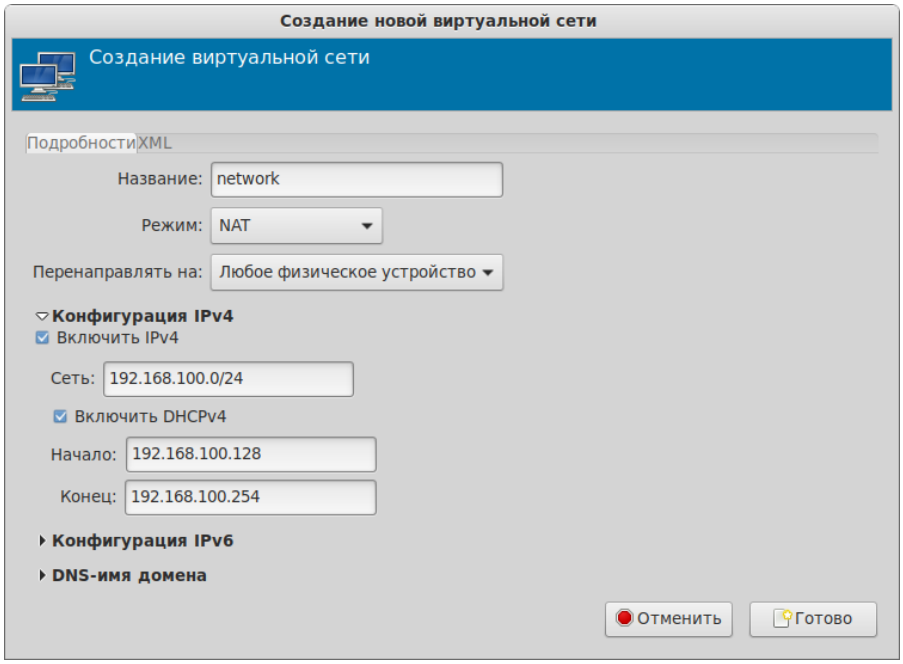

Рис. 184

5.7.3 Управление хранилищами

API-интерфейс libvirt обеспечивает удобную абстракцию для размещения образов ВМ и файловых систем, которая носит название storage pools (пул хранилищ). Пул хранилищ – это локальный каталог, локальное устройство хранения данных (физический диск, логический том или хранилище на основе хост-адаптера шины SCSI [SCSI HBA]), файловая система NFS (network file system), либо сетевое хранилище блочного уровня, управляемое посредством libvirt и позволяющее создать и хранить один или более образов виртуальных машин.

По умолчанию команды на базе libvirt используют в качестве исходного пула хранилищ для каталога файловой системы каталог `/var/lib/libvirt/images` на хосте виртуализации.

Образ диска – это снимок данных диска виртуальной машины, сохраненный в том или ином формате. libvirt понимает несколько форматов образов. Так же возможна работа с образами CD/DVD дисков. Каждый образ хранится в том или ином хранилище.

Типы хранилищ, с которыми работает libvirt:

dir – каталог в файловой системе;

disk – физический диск;

fs – отформатированное блочное устройство;

gluster – файловая система Gluster;

isci – хранилище iSCSI;

logical – группа томов LVM;

mpath – регистратор многопутевых устройств;

netfs – экспорт каталога из сети;

rbd – блочное устройство RADOS/Ceph;

scsi – хост-адаптер SCSI;

sheepdog – файловая система Sheepdog;

zfs – пул ZFS.

5.7.3.1 Управление хранилищами в командной строке

Команды управления хранилищами:

- pool-define – определить неактивный постоянный пул носителей на основе файла XML;
- pool-create – оздать пул из файла XML;
- pool-define-as – определить пул на основе набора аргументов;
- pool-create-as – создать пул на основе набора аргументов;
- pool-dumpxml – вывести файл конфигурации XML для заданного пула;
- pool-list – вывести список пулов;
- pool-build – собрать пул;

- pool-start – запустить ранее определённый неактивный пул;
- pool-autostart – автозапуск пула;
- pool-destroy – разрушить (остановить) пул;
- pool-delete – удалить пул;
- pool-edit – редактировать XML-конфигурацию пула носителей;
- pool-info – просмотр информации о пуле носителей;
- pool-refresh – обновить пул;
- pool-undefine – удалить определение неактивного пула.

Команда `virsh pool-define-as` создаст файл конфигурации для постоянного пула хранения. Позже этот пул можно запустить командой `virsh pool-start`, настроить его на автоматический запуск при загрузке хоста, остановить командой `virsh pool-destroy`.

Команда `virsh pool-create-as` создаст временный пул хранения (файл конфигурации не будет создан), который будет сразу запущен. Этот пул хранения будет удалён командой `virsh pool-destory`. Временный пул хранения нельзя запустить автоматически при загрузке. Преобразовать существующий временный пул в постоянный, можно создав файл XML-описания:

```
virsh pool-dumpxml имя_пула > имя_пула.xml && virsh pool-define имя_пула.xml
```

Пример создания пула хранения на основе NFS (netfs):

```
# virsh pool-create-as NFS-POOL netfs \
--source-host 192.168.0.105 \
--source-path /export/storage \
--target /var/lib/libvirt/images/NFS-POOL
Пул NFS-POOL создан
```

Первый аргумент (NFS-POOL) идентифицирует имя нового пула, второй аргумент идентифицирует тип создаваемого пула. Аргумент опции `--source-host` идентифицирует хост, который экспортирует каталог пула хранилищ посредством NFS. Аргумент опции `--source-path` определяет имя экспортируемого каталога на этом хосте. Аргумент опции `--target` идентифицирует локальную точку монтирования, которая будет использоваться для обращения к пулу хранилищ (этот каталог должен существовать).

Примечание. Для возможности монтирования NFS хранилища необходимо запустить службы rpcbind и nfslock:

```
# systemctl start rpcbind
# systemctl start nfslock
```

После создания нового пула хранилищ он будет указан в выходной информации команды virsh pool-list:

```
virsh pool-list --all --details
```

```
 Имя        Состояние   Автозапуск   Постоянный   Размер      Распределение   Доступно
-------------------------------------------------------------------------------------
 default    работает    yes          yes          37,15 GiB   2,94 GiB        34,21 GiB
 NFS-POOL   работает    no           no           29,40 GiB   7,26 GiB        22,14 GiB
```

В выводе команды видно, что опция «Автозапуск» («Autostart») для пула NFS-POOL имеет значение no (нет), т. е. после перезапуска системы этот пул не будет автоматически доступен для использования, и что опция «Постоянный» («Persistent») также имеет значение «no», т. е. после перезапуска системы этот пул вообще не будет определен. Пул хранилищ является постоянным только в том случае, если он сопровождается XML-описанием пула хранилищ, которое находится в каталоге `/etc/libvirt/storage`. XML-файл описания пула хранилищ имеет такое же имя, как у пула хранилищ, с которым он ассоциирован.

Чтобы создать файл XML-описания для сформированного в ручном режиме пула, следует воспользоваться командой `virsh pool-dumpxml`, указав в качестве ее заключительного аргумента имя пула, для которого нужно получить XML-описание. Эта команда осуществляет запись в стандартное устройство вывода, поэтому необходимо перенаправить ее выходную информацию в соответствующий файл.

Например, следующая команда создаст файл XML-описания для созданного ранее пула NFS-POOL и определит постоянный пул на основе этого файла:

```
# virsh pool-dumpxml NFS-POOL > NFS-POOL.xml && virsh pool-define NFS-POOL.xml
Пул NFS-POOL определён на основе NFS-POOL.xml
```

Чтобы задать для пула хранилищ опцию «Автозапуск» («Autostart»), можно воспользоваться командой `virsh pool-autostart`:

```
# virsh pool-autostart NFS-POOL
Добавлена метка автоматического запуска пула NFS-POOL
```

Маркировка пула хранилищ как автозапускаемого говорит о том, что этот пул хранилищ будет доступен после любого перезапуска хоста виртуализации (каталог `/etc/libvirt/storage/autostart` будет содержать символьную ссылку на XML-описание этого пула хранилищ).

Пример создания постоянного локального пула:

```
# virsh pool-define-as boot --type dir --target /var/lib/libvirt/boot
Пул boot определён

# virsh pool-list --all
 Имя        Состояние    Автозапуск
-------------------------------------
 default    активен      yes
 boot       не активен   no
 NFS-POOL   активен      yes
```

```
# virsh pool-build boot
Пул boot собран

# virsh pool-start boot
Пул boot запущен

# virsh pool-autostart boot
Добавлена метка автоматического запуска пула newpool

# virsh pool-list --all
 Имя         Состояние     Автозапуск
-----------------------------------------
 default     активен       yes
 boot        активен       yes
 NFS-POOL    активен       yes
```

5.7.3.2 Настройка хранилищ в менеджере виртуальных машин

Для настройки хранилищ с помощью virt-manager необходимо:

1) в меню «Правка» выбрать «Свойства подключения» (Рис. 182);
4) в открывшемся окне перейти на вкладку «Пространство данных» (Рис. 185).

Вкладка «Пространство данных»

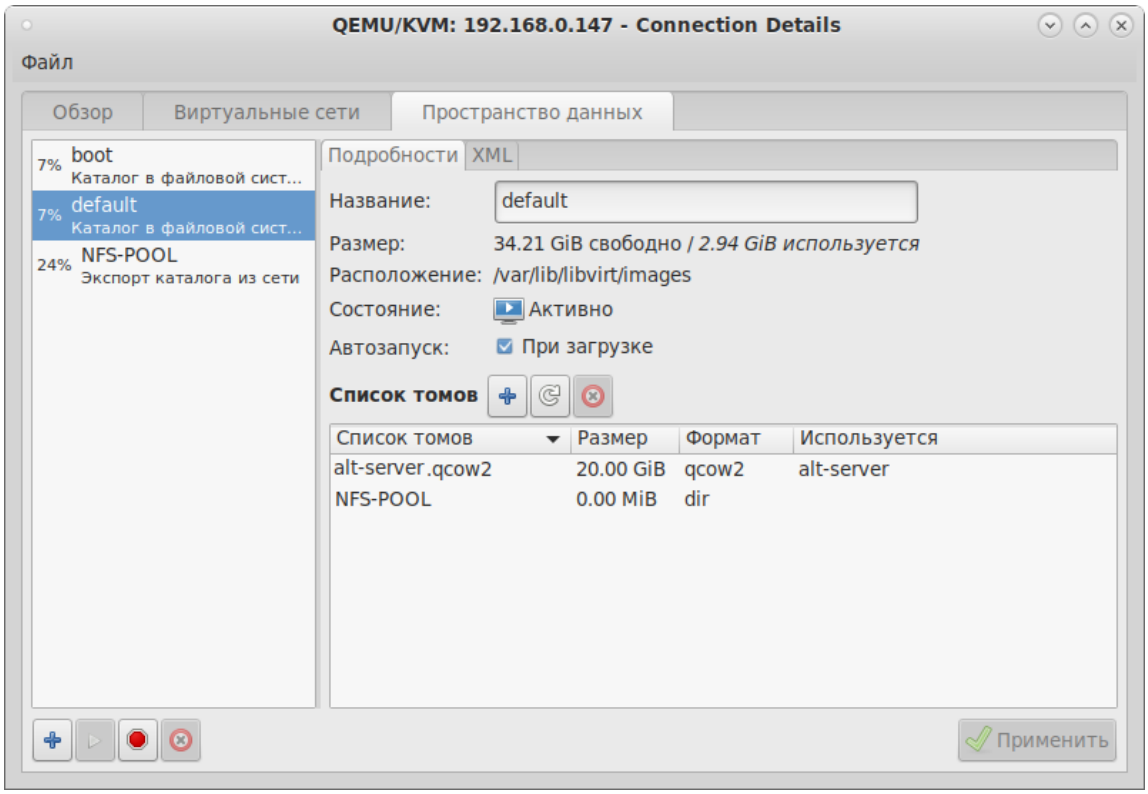

Рис. 185

Для добавления пула следует нажать кнопку «Добавить пул», расположенную в нижнем левом углу диалогового окна «Свойства соединения» (Рис. 185). В открывшемся окне (Рис. 186) следует выбрать тип пула, далее необходимо задать параметры пула (Рис. 187).

Создание пула хранения. Выбор типа пула

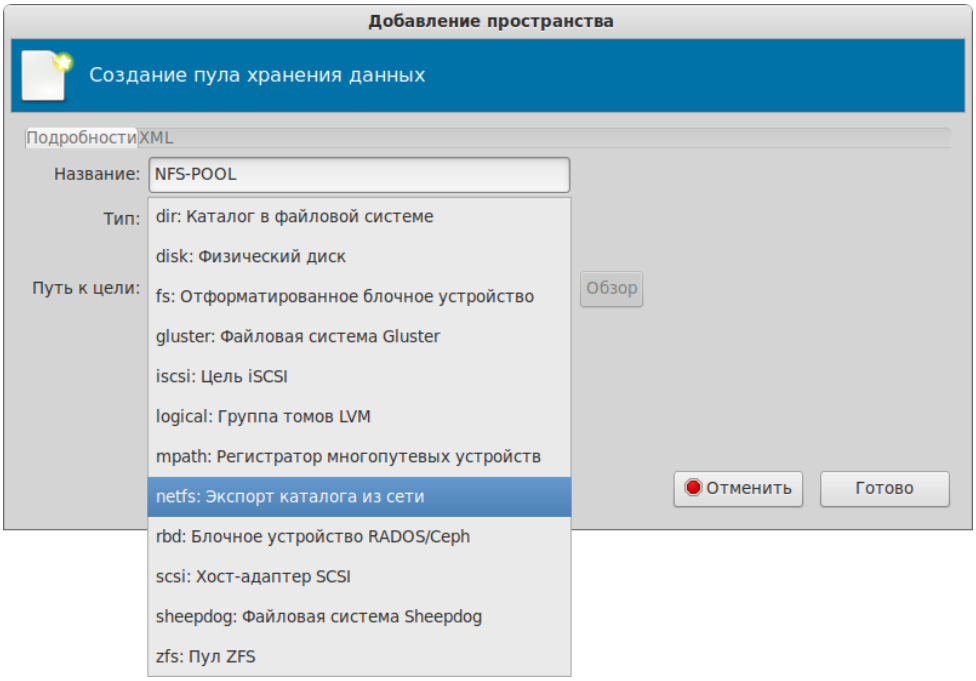

Рис. 186

Создание пула хранения. Ввод параметров

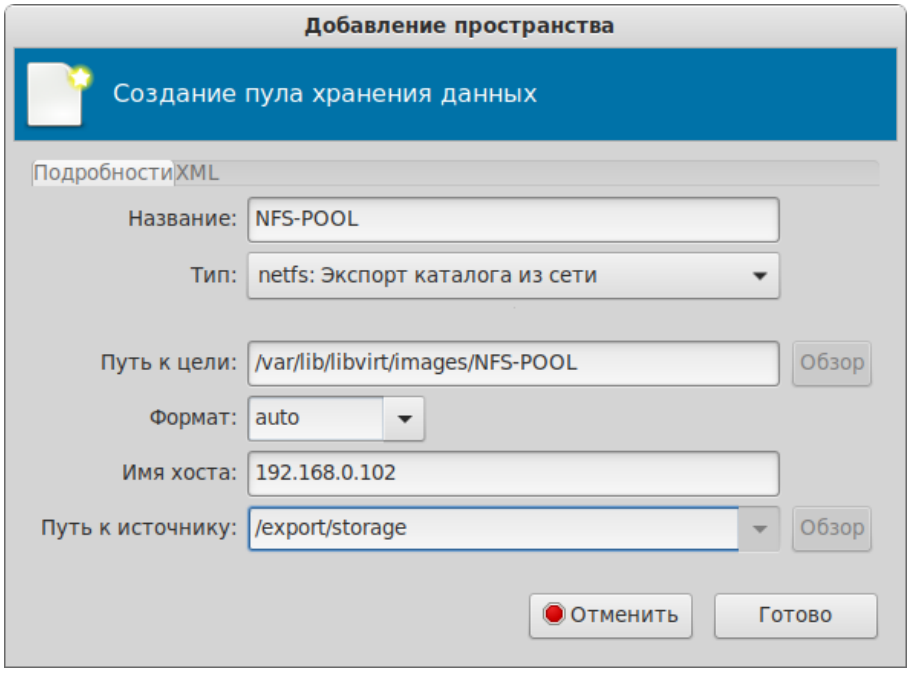

Рис. 187

5.8 Миграция ВМ

Под миграцией понимается процесс переноса ВМ с одного узла на другой.

Живая миграция позволяет перенести работу ВМ с одного физического хоста на другой без остановки ее работы.

Для возможности миграции ВМ, ВМ должна быть создана с использованием общего пула хранилищ (NFS, ISCSI, GlusterFS, CEPH).

Примечание. Живая миграция возможна даже без общего хранилища данных (с опцией `--copy-storage-all`). Но это приведет к большому трафику при копировании образа ВМ между серверами виртуализации и к заметному простою сервиса. Что бы миграция была по-настоящему «живой» с незаметным простоем необходимо использовать общее хранилище.

5.8.1 Миграция с помощью virsh

ВМ можно перенести на другой узел с помощью команды virsh. Для выполнения живой миграции нужно указать параметр `--live`. Команда переноса:

```
# virsh migrate --live VMName DestinationURL
```

где VMName – имя перемещаемой ВМ;

DestinationURL – URL или имя хоста узла назначения. Узел назначения должен использовать тот же гипервизор и служба libvirt на нем должна быть запущена.

После ввода команды будет запрошен пароль администратора узла назначения.

Для выполнения живой миграции ВМ alt-server на узел 192.168.0.190 с помощью virsh, необходимо выполнить следующие действия:

1) убедиться, что ВМ запущена:

```
# virsh list
ID   Имя          Состояние
------------------------------
 7    alt-server   работает
```

2) выполнить следующую команду, чтобы начать перенос ВМ на узел 192.168.0.190 (после ввода команды будет запрошен пароль пользователя root системы назначения):

```
# virsh migrate --live alt-server qemu+ssh://192.168.0.190/system
```

3) процесс миграции может занять некоторое время в зависимости от нагрузки и размера ВМ. virsh будет сообщать только об ошибках. ВМ будет продолжать работу на исходном узле до завершения переноса;

4) проверить результат переноса, выполнив на узле назначения команду:

```
# virsh list
```

5.8.2 Миграция с помощью virt-manager

Менеджер виртуальных машин virt-manager поддерживает возможность миграции ВМ между серверами виртуализации.

Для выполнения миграции, в virt-manager необходимо выполнить следующие действия:

1) подключить второй сервер виртуализации («Файл»→ «Добавить соединение…»);
2) в контекстном меню ВМ (она должна быть запущена) (Рис. 188) выбрать пункт «Миграция»;
3) в открывшемся окне (Рис. 189) выбрать конечный узел и нажать кнопку «Миграция».

Пункт «Миграция» в контекстном меню ВМ

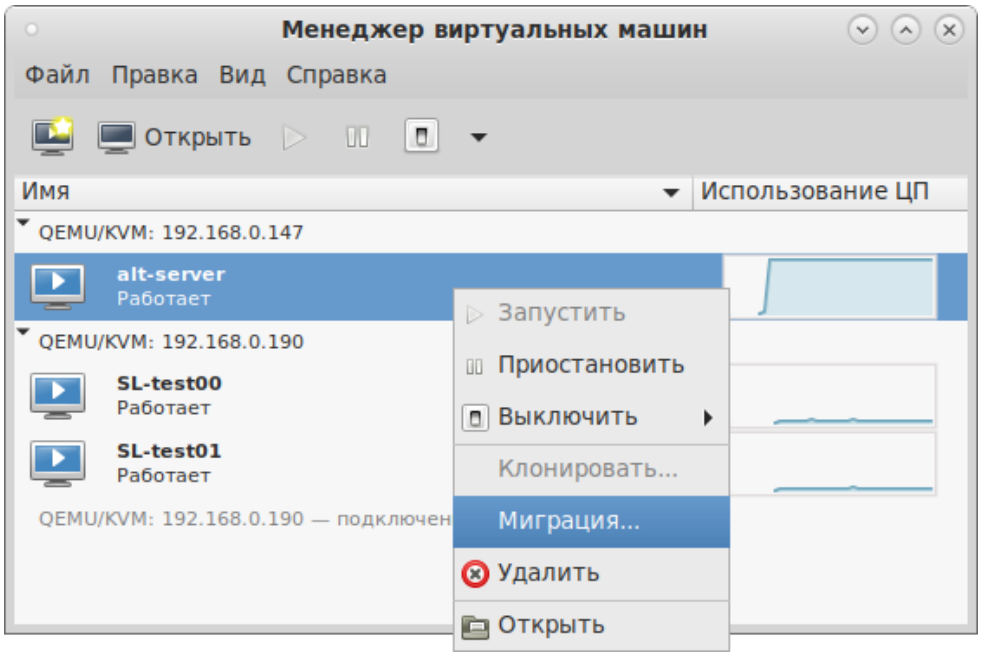

Рис. 188

Миграция ВМ

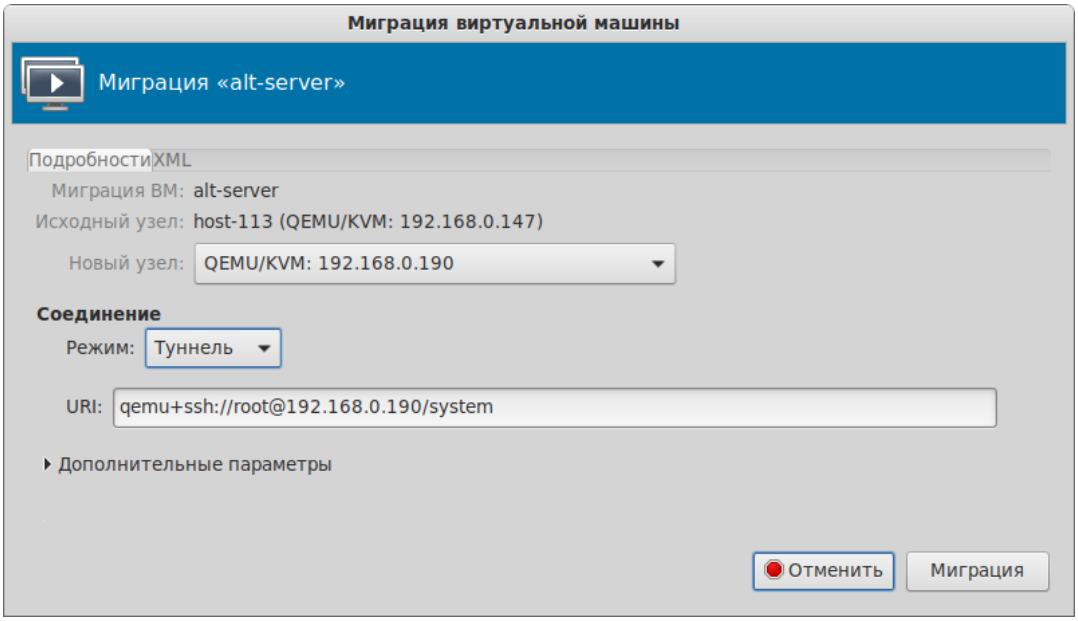

Рис. 189

При этом конфигурационный файл перемещаемой машины не переходит на новый узел, поэтому при ее выключении она вновь появится на старом хосте. В связи с этим, для совершения

полной живой миграции, при котором конфигурация ВМ будет перемещена на новый узел, необходимо воспользоваться утилитой командной строки virsh:

```
virsh migrate --live --persistent --undefinesource \
alt-server qemu+ssh://192.168.0.190/system
```

5.9 Снимки машины

Примечание. Снимок (snapshot) текущего состояния машины можно создать только если виртуальный жесткий диск в формате *.qcow2.

5.9.1 Управления снимками ВМ в консоли

Команда создания снимка (ОЗУ и диск) из файла XML:

```
virsh snapshot-create <domain> [--xmlfile <строка>] [--disk-only] [--live]...
```

Команда создания снимка (ОЗУ и диск) напрямую из набора параметров:

```
virsh snapshot-create-as <domain> [--name <строка>] [--disk-only] [--live]...
```

Пример создания снимка ВМ:

```
# virsh snapshot-create-as --domain alt-server --name alt-server-25may2021
Снимок домена alt-server-25may2021 создан
```

где

alt-server – имя ВМ;

alt-server-25may2021 – название снимка.

После того, как снимок ВМ будет сделан, резервные копии файлов конфигураций будут находиться в каталоге `/var/lib/libvirt/qemu/snapshot/`.

Пример создания снимка диска ВМ:

```
# virsh snapshot-create-as --domain alt-server --name 05may2021 -diskspec vda,file=/var/lib/libvirt/images/sn1.qcow2 --disk-only --atomic
Снимок домена 05may2021 создан
```

Просмотр существующих снимков для домена alt-server:

```
# virsh snapshot-list --domain alt-server
 Имя                    Время создания              Состояние
------------------------------------------------------------
 05may2021              2021-05-05 11:39:41 +0200   shutoff
 alt-server-25may2021   2021-05-25 16:06:50 +0200   running
```

Восстановить ВМ из снимка:

```
# virsh snapshot-revert --domain alt-server --snapshotname 05may2021 –running
```

Удалить снимок:

```
# virsh snapshot-delete --domain alt-server --snapshotname 05may2021
```

5.9.2 Управления снимками ВМ virt-manager

Для управления снимками ВМ в менеджере виртуальных машин virt-manager, необходимо:

1) в главном окне менеджера выбрать ВМ;
2) нажать кнопку «Открыть»;
3) в открывшемся окне нажать кнопку «Управление снимками» (Рис. 190). Появится окно управления снимками ВМ.

Управление снимками ВМ

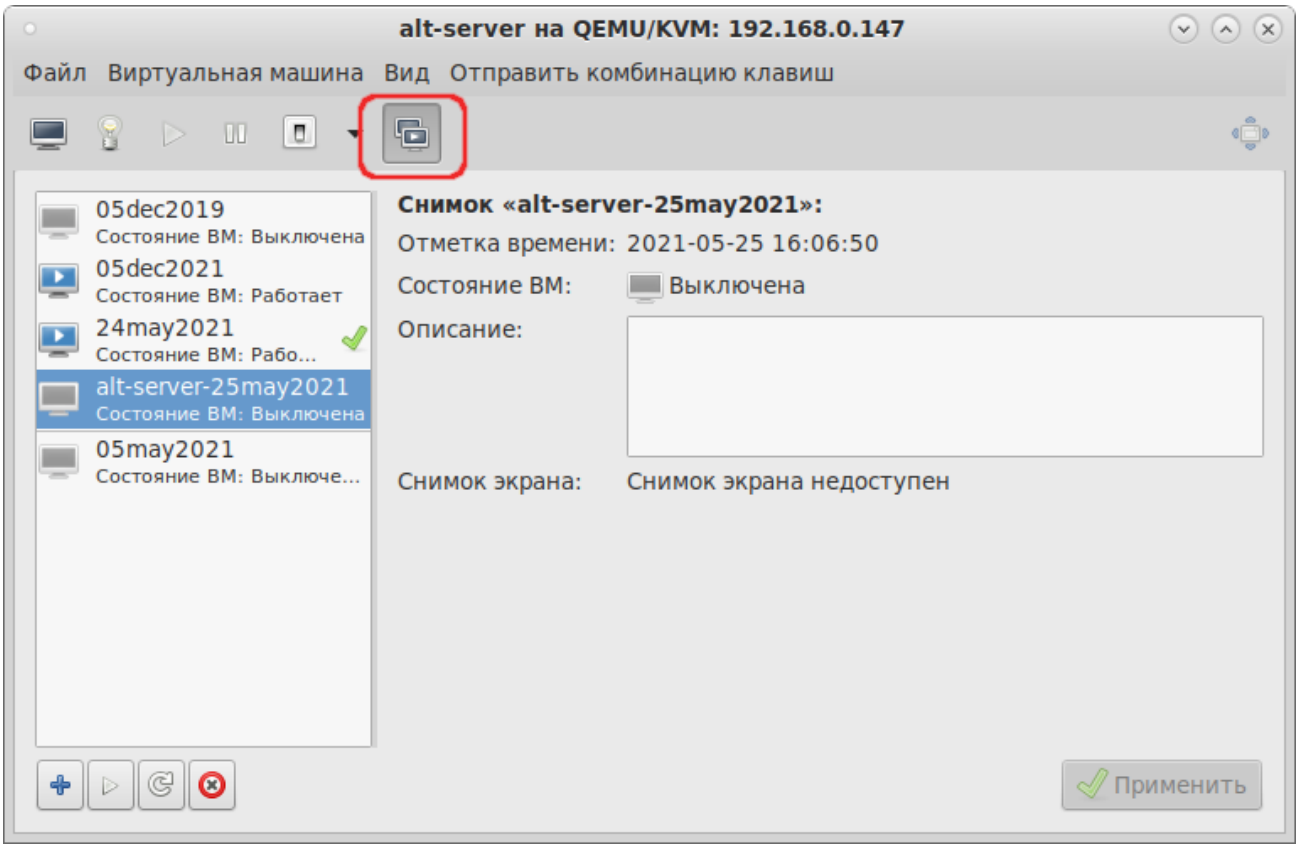

Рис. 190

Для создания нового снимка следует нажать кнопку «Создать новый снимок», расположенную в нижнем левом углу окна управления снимками ВМ. В открывшемся окне (Рис. 191) следует указать название снимка и нажать кнопку «Готово».

Для того чтобы восстановить ВМ из снимка или удалить снимок, следует воспользоваться контекстным меню снимка (Рис. 192).

Создание снимка

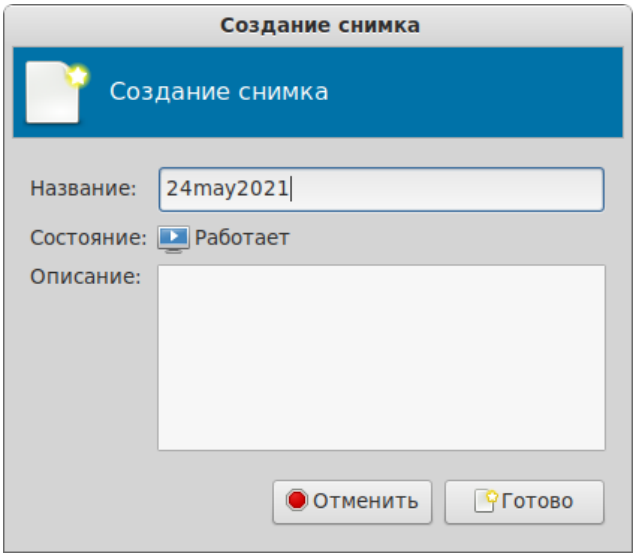

Рис. 191

Контекстное меню снимка

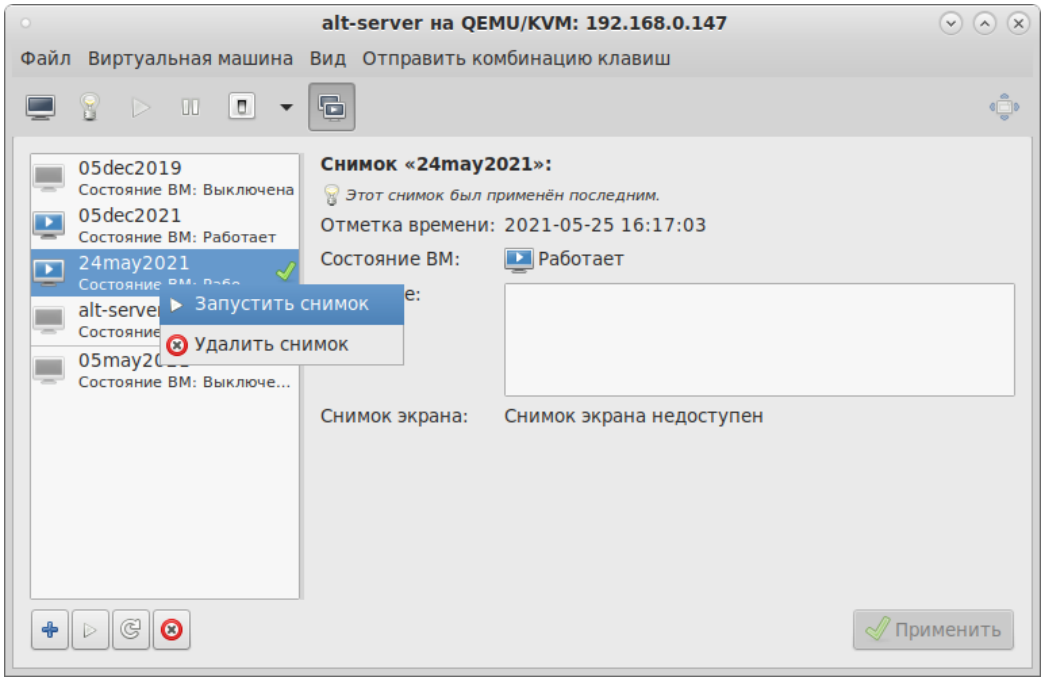

Рис. 192

5.10 Регистрация событий libvirt

Настройка регистрации событий в libvirt, осуществляется в файле `/etc/libvirt/libvirtd.conf`. Логи сохраняются в каталоге `/var/log/libvirt`.

Функция журналирования в libvirt основана на трех ключевых понятиях:

- сообщения журнала;
- фильтры;
- формат ввода.

Сообщения журнала – это информация, полученная во время работы libvirt. Каждое сообщение включает в себя уровень приоритета (отладочное сообщение – 1, информационное – 2, предупреждение – 3, ошибка – 4). По умолчанию, log_level=1, т. е. журналируются все сообщения.

Фильтры – это набор шаблонов и для записи сообщений в журнал. Если категория сообщения совпадает с фильтром, приоритет сообщения сравнивается с приоритетом фильтра, если она ниже сообщение отбрасывается, иначе сообщение записывается в журнал. Если сообщение не соответствует ни одному фильтру, то применяется общий уровень. Это позволяет, например, захватить все отладочные сообщения для QEmu, а для остальных, только сообщения об ошибках.

Формат для фильтра:

```
x:name  (log message only)
x:+name (log message + stack trace)
```

где:

- name – строка, которая сравнивается с заданной категорией, например, remote, qemu, или util.json;
- + – записывать каждое сообщение с данным именем;
- x – минимальный уровень ошибки (1, 2, 3, 4).

Пример фильтра:

```
Log_filtrers="3:remote 4:event"
```

Как только сообщение прошло через фильтрацию набора выходных данных, формат вывода определяет, куда отправить сообщение. Формат вывода также может фильтровать на основе приоритета, например, он может быть полезен для вывода всех сообщений в файл отладки.

Формат вывода может быть:

- x:stderr – вывод в STDERR;
- x:syslog:name – использовать системный журнал для вывода и использовать данное имя в качестве идентификатора;
- x:file:file_path – вывод в файл, с соответствующим filepath;
- x:journal – вывод в systemd журнал.

Пример:

```
Log_outputs=”3:syslog:libvirtd 1:file:/tmp/libvirt.log”
```

Журналы работы виртуальных машин под KVM хранятся в /var/log/libvirt/qemu/. В этом каталоге libvirt хранит журнал для каждой виртуальной машины. Например, для машины с названием alt-server журнал будет находиться по адресу: /var/log/libvirt/qemu/alt-server.log.

5.11 Управление доступом в виртуальной инфраструктуре

Права пользователя могут управляться с помощью правил polkit.

В каталоге `/usr/share/polkit-1/actions/` имеются два файла с описанием возможных действий для работы с ВМ, предоставленные разработчиками libvirt:

- файл org.libvirt.unix.policy описывает мониторинг ВМ и управление ими;
- в файле org.libvirt.api.policy перечислены конкретные действия (остановка, перезапуск и т. д.), которые возможны, если предыдущая проверка пройдена.

Перечисление конкретных свойств с комментариями доступно в файле `/usr/share/polkit-1/actions/org.libvirt.api.policy`.

В libvirt названия объектов и разрешений отображаются в имена polkit действий, по схеме:
`org.libvirt.api.$объект.$разрешение`

Например, разрешение search-storage-vols на объекте storage_pool отображено к действию polkit:
`org.libvirt.api.storage-pool.search-storage-vols`

Чтобы определить правила авторизации, polkit должен однозначно определить объект. Libvirt предоставляет ряд атрибутов для определения объектов при выполнении проверки прав доступа. Набор атрибутов изменяется в зависимости от типа объекта.

Пример тонкой настройки

Есть две виртуальные машины: alt1, alt2. Необходимо разрешить пользователю test (должен быть в группе vmusers) действия только с доменом alt1. Для этого необходимо выполнить следующие действия:

1) раскомментировать в файле `/etc/libvirt/libvirtd.conf` строку:

```
access_drivers = [ "polkit" ]
```

2) перезапустить libvirt:

```
# systemctl restart libvirtd
```

3) создать файл `/etc/polkit-1/rules.d/100-libvirt-acl.rules` (имя произвольно) следующего вида:

```
==========================
polkit.addRule(function(action, subject) {
   // разрешить пользователю test действия с доменом "alt1"
   if (action.id.indexOf("org.libvirt.api.domain.") ==0  &&
       subject.user == "test") {
         if (action.lookup("domain_name") == 'alt1') {
           return polkit.Result.YES;
```

```
        }
        else { return polkit.Result.NO; }
   }
   else {
   // разрешить пользователю test действия с
   //подключениями, хранилищем и прочим
if (action.id.indexOf("org.libvirt.api.") == 0 &&
       subject.user == "test") {
       polkit.log("org.libvirt.api.Yes");
           return polkit.Result.YES;
        }
        else { return polkit.Result.NO; }
   }})
================================
```

4) перелогиниться.

В результате выполненных действий пользователю test машина alt1 видна, а машина alt2 – нет.

Права можно настраивать более тонко, например, разрешив пользователю test запускать ВМ, но запретить ему все остальные действия с ней, для этого надо разрешить действие `org.libvirt.api.domain.start`:

```
=========================
polkit.addRule(function(action, subject) {
   // разрешить пользователю test только запускать ВМ в
   // домене "alt1"
   if (action.id. == "org.libvirt.api.domain.start") &&
       subject.user == "test") {
         if (action.lookup("domain_name") == 'alt1') {
           return polkit.Result.YES;
         }
         else { return polkit.Result.NO; }
}


   });
=========================
```

Предоставить право запускать ВМ, только пользователям группы wheel:

```
if (action.id == "org.libvirt.api.domain.start") {
        if (subject.isInGroup("wheel")) {
            return polkit.Result.YES;
        } else {
            return polkit.Result.NO;
        }
    };
```

Предоставить право останавливать ВМ, только пользователям группы wheel:

```
if (action.id == "org.libvirt.api.domain.stop") {
        if (subject.isInGroup("wheel")) {
            return polkit.Result.YES;
        } else {
            return polkit.Result.NO;
        }
    };
```

Можно также вести файл журнала, используя правила polkit. Например, делать запись в журнал при старте ВМ:

```
if (action.id.match("org.libvirt.api.domain.start") ) {
           polkit.log("action=" + action);
           polkit.log("subject=" + subject);
           return polkit.Result.YES;
         }
```

Запись в журнал при останове ВМ:

```
if (action.id.match("org.libvirt.api.domain.stop") ) {
           polkit.log("action=" + action);
           polkit.log("subject=" + subject);
           return polkit.Result.YES;
         }
```

6 KUBERNETES

6.1 Краткое описание возможностей

Kubernetes – это система для автоматизации развёртывания, масштабирования и управления контейнеризированными приложениями. Поддерживает основные технологии контейнеризации (Docker, Rocket) и аппаратную виртуализацию.

Основные задачи Kubernetes:

- развертывание контейнеров и все операции для запуска необходимой конфигурации (перезапуск остановившихся контейнеров, перемещение контейнеров для выделения ресурсов на новые контейнеры и т.д.);
- масштабирование и запуск нескольких контейнеров одновременно на большом количестве хостов;
- балансировка множества контейнеров в процессе запуска. Для этого Kubernetes использует API, задача которого заключается в логическом группировании контейнеров.

Утилиты для создания и управления кластером Kubernetes:

- `kubectl` – создание и настройка объектов в кластере;
- `kubelet` – запуск контейнеров на узлах;
- `kubeadm` – настройка компонентов, составляющие кластер.

6.2 Установка и настройка Kubernetes

Для создания управляющего или вычислительного узла, при установке дистрибутива в группе «Контейнеры» следует соответственно отметить пункт «Сервисы Kubernetes для управляющего хоста» или «Сервисы Kubernetes для вычислительного узла», пункт «Управление контейнерами Docker», при этом будет выбран автоматически (Рис. 193).

Примечание. На этапе «Подготовка диска» рекомендуется выбрать «Server KVM/Docker/LXD/Podman/CRI-O (large /var/lib/)» и не создавать раздел Swap.

Примечание. В данном руководстве рассмотрен процесс разворачивания кластера с использованием docker.

Установка Kubernetes при установке системы

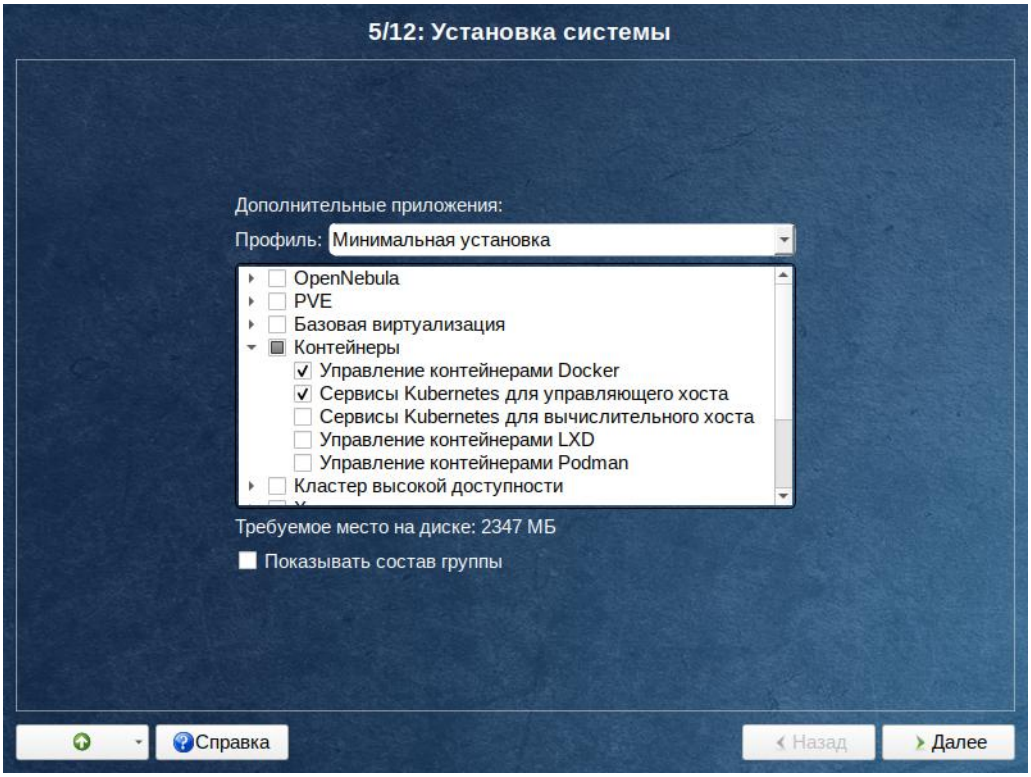

Рис. 193

6.2.1 Создание кластера Kubernetes

Для создания кластера необходимо несколько машин (nodes). Одна из которых будет мастером. Системные требования:

- 2 ГБ или больше ОЗУ на машину;
- 2 ядра процессора или больше;
- все машины должны быть доступны по сети друг для друга;
- все машины должны успешно разрешать имена hostname друг друга (через DNS или hosts);
- Swap должен быть выключен.

Примечание. Для отключения Swap нужно выполнить команду:

```
# swapoff -a
```

и удалить соответствующую строку в `/etc/fstab`.

6.2.1.1 Инициализация кластера

Для инициализации кластера запустить команду (на мастере):

```
# kubeadm init --pod-network-cidr=10.244.0.0/16 \
    --ignore-preflight-errors=SystemVerification
```

где:

- `--pod-network-cidr=10.244.0.0/16` – адрес внутренней (разворачиваемой Kubernetes) сети, данное значение рекомендуется оставить для правильной работы Flannel;

– `--ignore-preflight-errors=SystemVerification` – игнорировать ошибки проверки версии docker.

Если все сделано правильно, на экране отобразится команда, позволяющая присоединить остальные ноды кластера к мастеру:

```
…
Your Kubernetes control-plane has initialized successfully!

To start using your cluster, you need to run the following as a regular user:

  mkdir -p $HOME/.kube
  sudo cp -i /etc/kubernetes/admin.conf $HOME/.kube/config
  sudo chown $(id -u):$(id -g) $HOME/.kube/config

Alternatively, if you are the root user, you can run:

  export KUBECONFIG=/etc/kubernetes/admin.conf

You should now deploy a pod network to the cluster.
Run "kubectl apply -f [podnetwork].yaml" with one of the options listed at:
  https://kubernetes.io/docs/concepts/cluster-administration/addons/

Then you can join any number of worker nodes by running the following on each as
root:

kubeadm join 192.168.0.201:6443 --token cgmqh4.26l6pnhqagslwvae \
    --discovery-token-ca-cert-hash \
    sha256:9571e4fde1bed9ee43ed1cba98b5c2bca5184f99f54806f1a84657d161e9f0a1
```

Настроить kubernetes для работы от пользователя (на мастер-ноде):

1) создать каталог `~/.kube`:

```
$ mkdir ~/.kube
```

2) скопировать конфигурацию:

```
# cp /etc/kubernetes/admin.conf ~<пользователь>/.kube/config
```

3) изменить владельца конфигурационного файла:

```
# chown <пользователь>: ~<пользователь>/.kube/config
```

6.2.1.2 Настройка сети

Развернуть сеть (Container Network Interface), запустив команду (на мастер-ноде):

```
$ kubectl apply -f
https://raw.githubusercontent.com/coreos/flannel/master/Documentation/kube-
flannel.yml
podsecuritypolicy.policy/psp.flannel.unprivileged created
```

```
clusterrole.rbac.authorization.k8s.io/flannel created
clusterrolebinding.rbac.authorization.k8s.io/flannel created
serviceaccount/flannel created
configmap/kube-flannel-cfg created
daemonset.apps/kube-flannel-ds created
```

В выводе будут отображены имена всех созданных ресурсов. Проверить, что всё работает:

```
$ kubectl get pods --namespace kube-system
NAME                                READY   STATUS    RESTARTS   AGE
coredns-74ff55c5b-5rgmk             1/1     Running   0          38m
coredns-74ff55c5b-wjq4r             1/1     Running   0          38m
etcd-master01                       1/1     Running   0          37m
kube-apiserver-master01             1/1     Running   0          37m
kube-controller-manager-master01    1/1     Running   0          37m
kube-flannel-ds-2gl6g               1/1     Running   0          92s
kube-proxy-gknv7                    1/1     Running   0          15m
kube-scheduler-master01             1/1     Running   0          37m
```

coredns должны находиться в состоянии Running. Количество kube-flannel и kube-proxy зависит от общего числа нод.

6.2.1.3 Добавление узлов (нод) в кластер

Подключить остальные узлы (ноды) в кластер. Для этого на узле выполнить команду:

```
# kubeadm join <ip адрес>:<порт> --token <токен> \
    --discovery-token-ca-cert-hash sha256:<хэш> \
    --ignore-preflight-errors=SystemVerification
```

Данная команда была выведена при выполнении команды `kubeadm init` на мастер-ноде. В данном случае:

```
# kubeadm join 192.168.0.201:6443 --token cgmqh4.26l6pnhqagslwvae \
    --discovery-token-ca-cert-hash \
    sha256:9571e4fde1bed9ee43ed1cba98b5c2bca5184f99f54806f1a84657d161e9f0a1
[preflight] Running pre-flight checks
[preflight] Reading configuration from the cluster...
[preflight] FYI: You can look at this config file with 'kubectl -n kube-system get cm
kubeadm-config -o yaml'
[kubelet-start] Writing kubelet configuration to file "/var/lib/kubelet/config.yaml"
[kubelet-start] Writing kubelet environment file with flags to file
"/var/lib/kubelet/kubeadm-flags.env"
[kubelet-start] Starting the kubelet
[kubelet-start] Waiting for the kubelet to perform the TLS Bootstrap...

This node has joined the cluster:
* Certificate signing request was sent to apiserver and a response was received.
* The Kubelet was informed of the new secure connection details.
```

```
Run 'kubectl get nodes' on the control-plane to see this node join the cluster.
```

Примечание. Получить токен, если его нет, можно выполнив команду (на мастер-ноде):

```
$ kubeadm token list
TOKEN                     TTL       EXPIRES                     USAGES
cgmqh4.26l6pnhqagslwvae   20h       2021-06-25T11:52:05+02:00   authentication,signing
```

По умолчанию, срок действия токена – 24 часа. Если требуется добавить новый узел в кластер по окончанию этого периода, можно создать новый токен:

```
$ kubeadm token create
```

Если значение параметра `--discovery-token-ca-cert-hash` неизвестно, его можно получить, выполнив команду (на мастер-ноде):

```
$ openssl x509 -pubkey -in /etc/kubernetes/pki/ca.crt | \
   openssl rsa -pubin -outform der 2>/dev/null | \
   openssl dgst -sha256 -hex | sed 's/^.* //'
9571e4fde1bed9ee43ed1cba98b5c2bca5184f99f54806f1a84657d161e9f0a1
```

Для ввода IPv6-адреса в параметр `<control-plane-host>:<control-plane-port>`, адрес должен быть заключен в квадратные скобки:

```
[fd00::101]:2073
```

Проверить наличие нод (на мастер-ноде):

```
$ kubectl get nodes
NAME       STATUS   ROLES                  AGE    VERSION
docker03   Ready    <none>                 160m   v1.20.2
master01   Ready    control-plane,master   3h3m   v1.20.2
```

или:

```
$ kubectl get nodes -o wide
```

Информация о кластере:

```
$ kubectl cluster-info
Kubernetes control plane is running at https://192.168.0.201:6443
KubeDNS is running at https://192.168.0.201:6443/api/v1/namespaces/kube-
system/services/kube-dns:dns/proxy

To further debug and diagnose cluster problems, use 'kubectl cluster-info dump'.
```

Посмотреть подробную информацию о ноде:

```
$ kubectl describe node docker03
```

6.2.2 Тестовый запуск nginx

Deployment – это объект Kubernetes, представляющий работающее приложение в кластере.

Создать Deployment с nginx:

```
$ kubectl apply -f https://k8s.io/examples/application/deployment.yaml
deployment.apps/nginx-deployment created
```

Список подов:

```
$ kubectl get pods
NAME                                READY   STATUS    RESTARTS   AGE
nginx-deployment-66b6c48dd5-89lq4   1/1     Running   1          9s
nginx-deployment-66b6c48dd5-bt7zp   1/1     Running   1          9s
```

Создать сервис, с помощью которого можно получить доступ к приложению из внешней сети. Для этого создать файл `nginx-service.yaml`, со следующим содержимым:

```
apiVersion: v1
    kind: Service
    metadata:
      name: nginx
      labels:
        app: nginx
    spec:
      type: NodePort
      ports:
      - port: 80
        targetPort: 80
      selector:
        app: nginx
```

Запустить новый сервис:

```
$ kubectl apply -f nginx-service.yaml
service/nginx created
```

Просмотреть порт сервиса nginx:

```
$ kubectl get svc nginx
NAME    TYPE       CLUSTER-IP      EXTERNAL-IP   PORT(S)        AGE
nginx   NodePort   10.105.18.148   <none>        80:30723/TCP   7s
```

Проверить работу nginx, выполнив команду (сервер должен вернуть код 200):

```
curl -I <ip адрес>:<порт>
```

где <ip адрес> – это адрес любой из нод, а <порт> – это порт сервиса, полученный с помощью предыдущей команды. В данном кластере возможна команда:

```
$ curl -I 192.168.0.204:30723
HTTP/1.1 200 OK
Server: nginx/1.14.2
```

7 НАСТРОЙКА СИСТЕМЫ

7.1 Центр управления системой

Для управления настройками установленной системы можно использовать Центр управления системой. Центр управления системой (ЦУС) представляет собой удобный интерфейс для выполнения наиболее востребованных административных задач: добавление и удаление пользователей, настройка сетевых подключений, просмотр информации о состоянии системы и т.п. ЦУС имеет веб-ориентированный интерфейс, позволяющий управлять сервером с любого компьютера сети.

Центр управления системой состоит из нескольких независимых диалогов-модулей. Каждый модуль отвечает за настройку определённой функции или свойства системы.

7.1.1 Применение ЦУС

ЦУС можно использовать для разных целей, например:

- настройка даты и времени;
- управление системными службами;
- просмотр системных журналов;
- управление выключением удаленного компьютера;
- настройка ограничений выделяемых ресурсов памяти пользователям (квоты);
- настройка ограничений на использование внешних носителей;
- конфигурирование сетевых интерфейсов;
- настройка межсетевого экрана;
- изменения пароля администратора системы (root);
- создание, удаление и редактирование учётных записей пользователей.

Все модули ЦУС имеют справочную информацию.

7.1.2 Использование веб-ориентированного ЦУС

ЦУС имеет веб-ориентированный интерфейс, позволяющий управлять данным компьютером с любого другого компьютера сети.

Для запуска веб-ориентированного интерфейса должен быть запущен сервис ahttpd и alteratord:

```
# systemctl enable --now ahttpd
# systemctl enable --now alteratord
```

Работа с ЦУС может происходить из любого веб-браузера. Для начала работы необходимо перейти по адресу https://ip-адрес:8080/.

Если для сервера задан IP-адрес 192.168.0.122, то интерфейс управления будет доступен по адресу: https://192.168.0.122:8080/.

Примечание. IP-адрес сервера можно узнать, введя на сервере команду:

```
$ ip addr
```

При запуске ЦУС необходимо ввести в соответствующие поля имя пользователя (root) и пароль пользователя root (Рис. 194).

Запуск веб-ориентированного центра управления системой

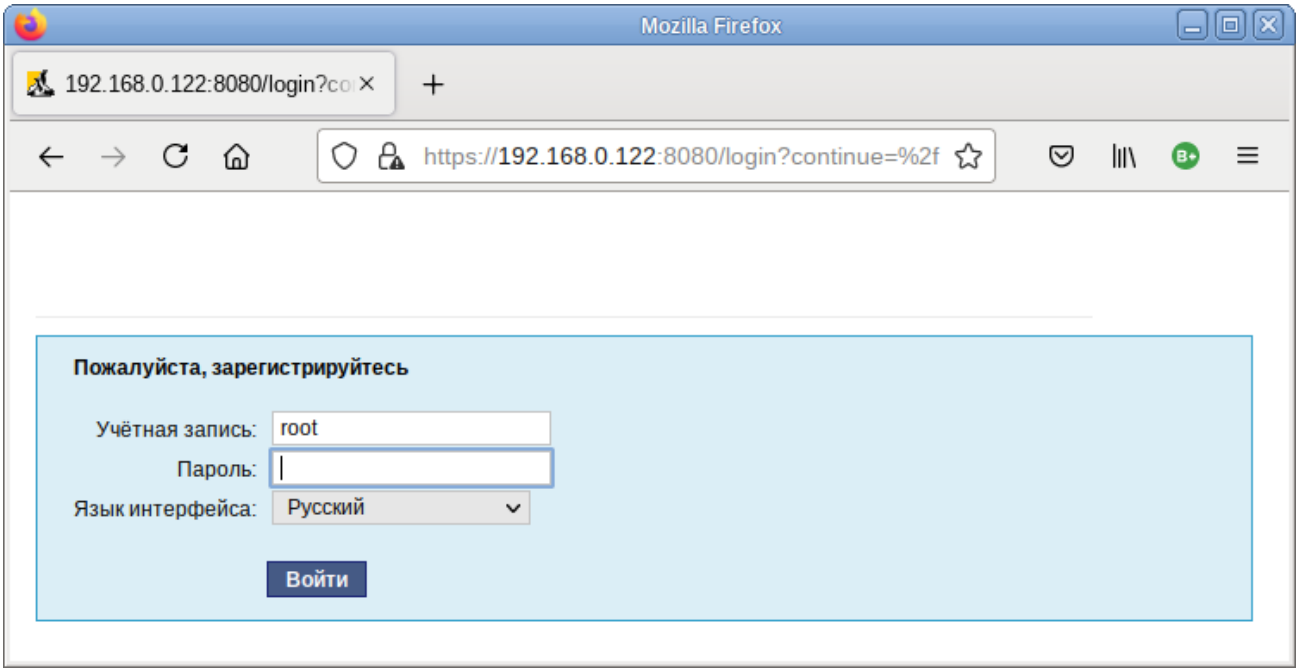

Рис. 194

После этого будут доступны все возможности ЦУС на той машине, к которой было произведено подключение через веб-интерфейс (Рис. 195).

Веб-интерфейс ЦУС можно настроить (кнопка «Режим эксперта»), выбрав один из режимов:

- основной режим;
- режим эксперта.

Выбор режима влияет на количество отображаемых модулей. В режиме эксперта отображаются все модули, а в основном режиме только наиболее используемые.

ЦУС содержит справочную информацию по всем включённым в него модулям. Об использовании самого интерфейса системы управления можно прочитать, нажав, на кнопку «Справка» на начальной странице центра управления системой (Рис. 195).

После работы с ЦУС, в целях безопасности, не следует оставлять открытым браузер. Необходимо обязательно выйти из сеанса работы с ЦУС, нажав на кнопку «Выйти».

Подробнее об использовании ЦУС можно узнать в главе «Работа с центром управления системой».

Веб-ориентированный центр управления системой

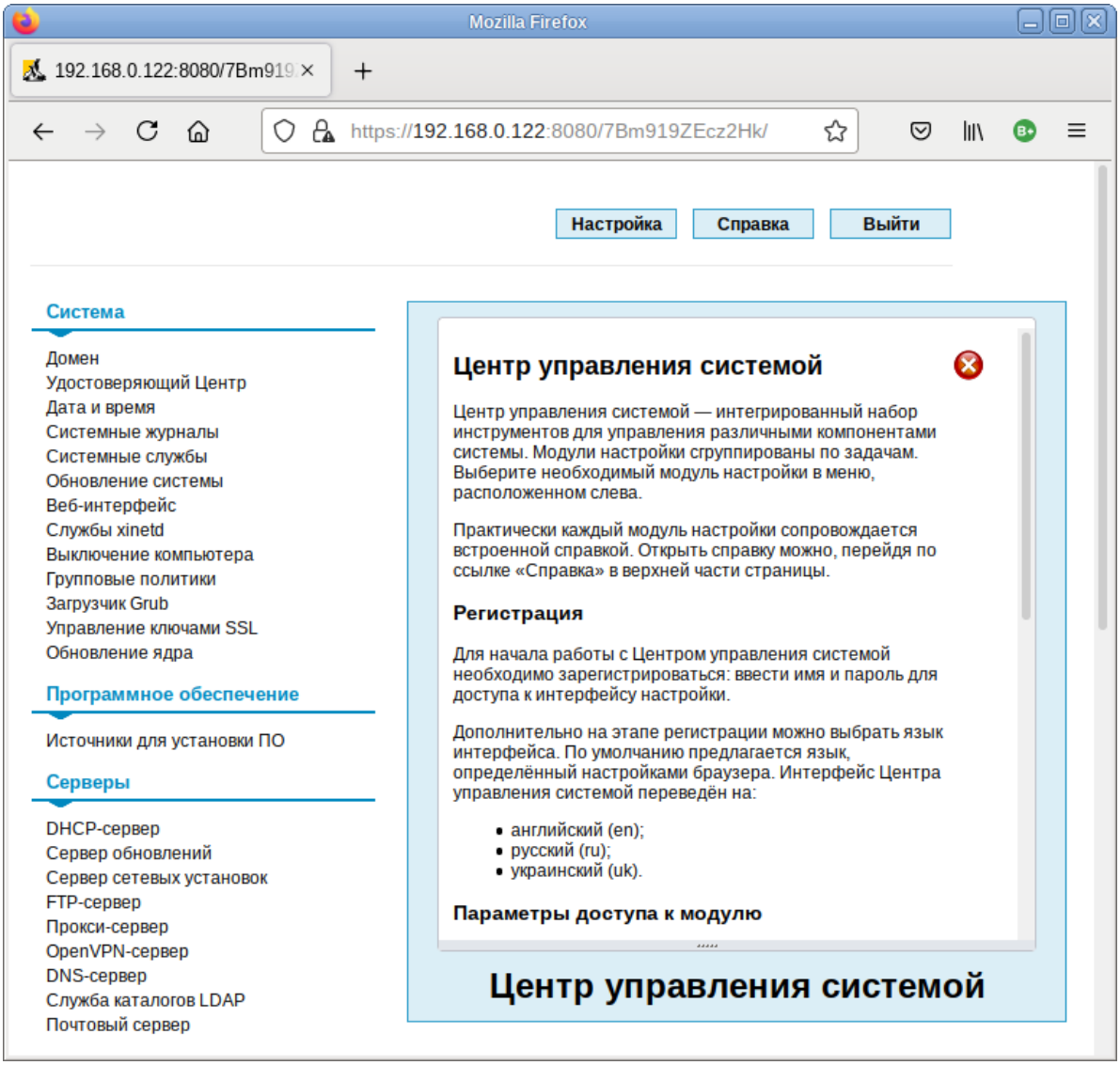

Рис. 195

8 РАБОТА С ЦЕНТРОМ УПРАВЛЕНИЯ СИСТЕМОЙ

Дальнейшие разделы описывают некоторые возможности использования ОС «Альт Сервер Виртуализации», настраиваемые в ЦУС.

8.1 Настройка подключения к Интернету

Помимо множества различных служб, которые ОС «Альт Сервер Виртуализации» может предоставлять компьютерам сети, важно определить, будет ли сервер предоставлять общий доступ в Интернет для компьютеров домена или нет. В зависимости от этого сервер можно рассматривать как:

- Сервер без подключения к сети Интернет – это сервер с одним сетевым интерфейсом (одной сетевой картой), который и связывает его с компьютерами локальной сети. Такой сервер называется также сервер рабочей группы.
- Шлюз – в этом случае сервер обычно имеет два сетевых интерфейса (например, две сетевые карты), одна из которых служит для подключения к локальной сети, а другая – для подключения к сети Интернет.

Как для обеспечения доступа в сеть Интернет самого сервера, так и для настройки общего выхода в Интернет для компьютеров сети необходимо настроить подключение к Интернету на самом сервере. ОС «Альт Сервер Виртуализации» поддерживает самые разные способы подключения к сети Интернет:

- Ethernet;
- PPTP;
- PPPoE;
- и т.д.

Для настройки подключения можно воспользоваться одним из разделов ЦУС «Сеть»:

- Ethernet-интерфейсы;
- PPTP-соединения;
- PPPoE-соединения;
- OpenVPN-соединения.

8.1.1 Конфигурирование сетевых интерфейсов

Конфигурирование сетевых интерфейсов осуществляется в модуле ЦУС «Ethernet-интерфейсы» (пакет alterator-net-eth) из раздела раздел «Сеть» (Рис. 196).

Настройка модуля «Ethernet-интерфейсы»

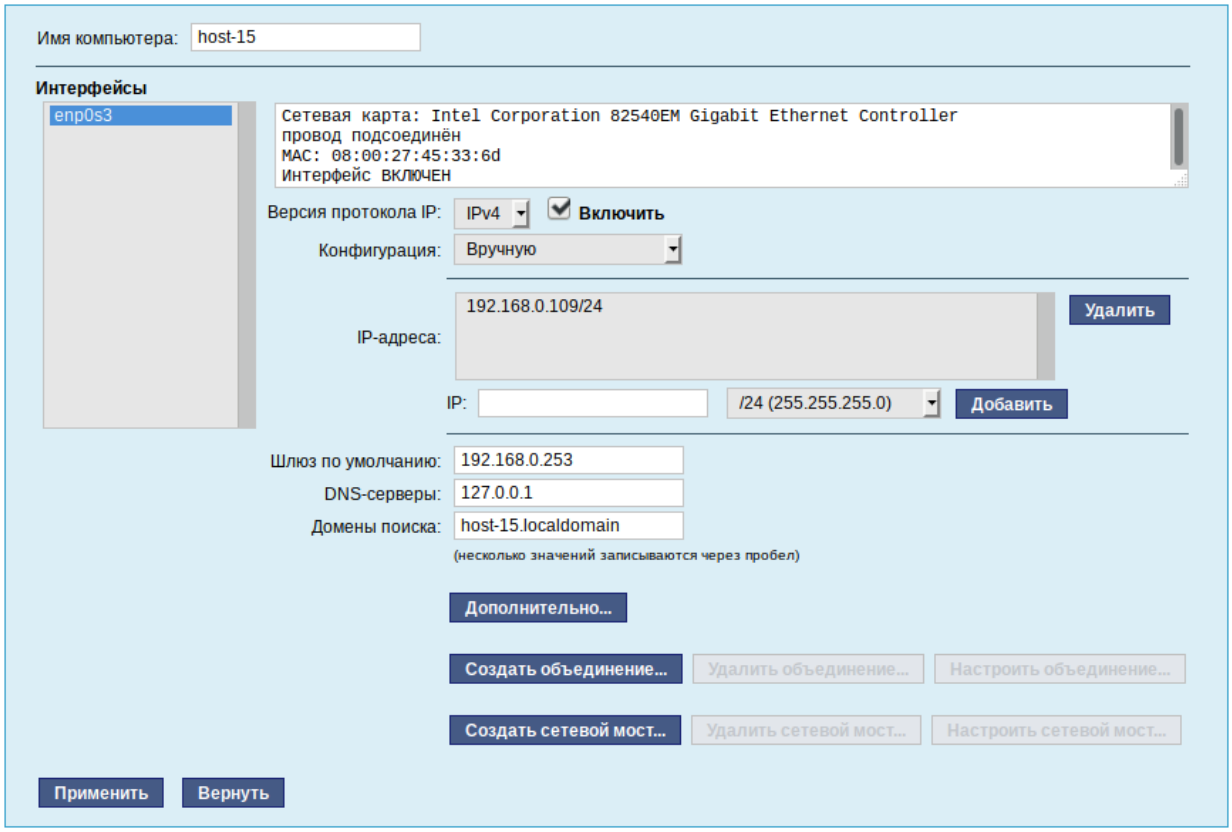

Рис. 196

В модуле «Ethernet-интерфейсы» можно заполнить следующие поля:

– «Имя компьютера» – указать сетевое имя ПЭВМ в поле для ввода имени компьютера (это общий сетевой параметр, не привязанный к какому либо конкретному интерфейсу). Имя компьютера, в отличие от традиционного имени хоста в Unix (hostname), не содержит названия сетевого домена;

– «Интерфейсы» – выбрать доступный сетевой интерфейс, для которого будут выполняться настройки;

– «Версия протокола IP» – указать в выпадающем списке версию используемого протокола IP (IPv4, IPv6) и убедиться, что пункт «Включить», обеспечивающий поддержку работы протокола, отмечен;

– «Конфигурация» – выбрать способ назначения IP-адресов (службы DHCP, Zeroconf, вручную);

– «IP-адреса» – пул назначенных IP-адресов из поля «IP», выбранные адреса можно удалить нажатием кнопки «Удалить»;

– «IP» – ввести IP-адрес вручную и выбрать в выпадающем поле предпочтительную маску сети, затем нажать кнопку «Добавить» для переноса адреса в пул поля «IP-адреса»;

– «Шлюз по умолчанию» – в поле для ввода необходимо ввести адрес шлюза, который будет использоваться сетью по умолчанию;
– «DNS-серверы» – в поле для ввода необходимо ввести список предпочтительных DNS-серверов, которые будут получать информацию о доменах, выполнять маршрутизацию почты и управлять обслуживающими узлами для протоколов в домене;
– «Домены поиска» – в поле для ввода необходимо ввести список предпочтительных доменов, по которым будет выполняться поиск.

«IP-адрес» и «Маска сети» – обязательные параметры каждого узла IP-сети. Первый параметр – уникальный идентификатор машины, от второго напрямую зависит, к каким машинам локальной сети данная машина будет иметь доступ. Если требуется выход во внешнюю сеть, то необходимо указать параметр «Шлюз по умолчанию».

В случае наличия DHCP-сервера можно все вышеперечисленные параметры получить автоматически – выбрав в списке «Конфигурация» пункт «Использовать DHCP» (Рис. 197).

Автоматическое получение настроек от DHCP-сервера

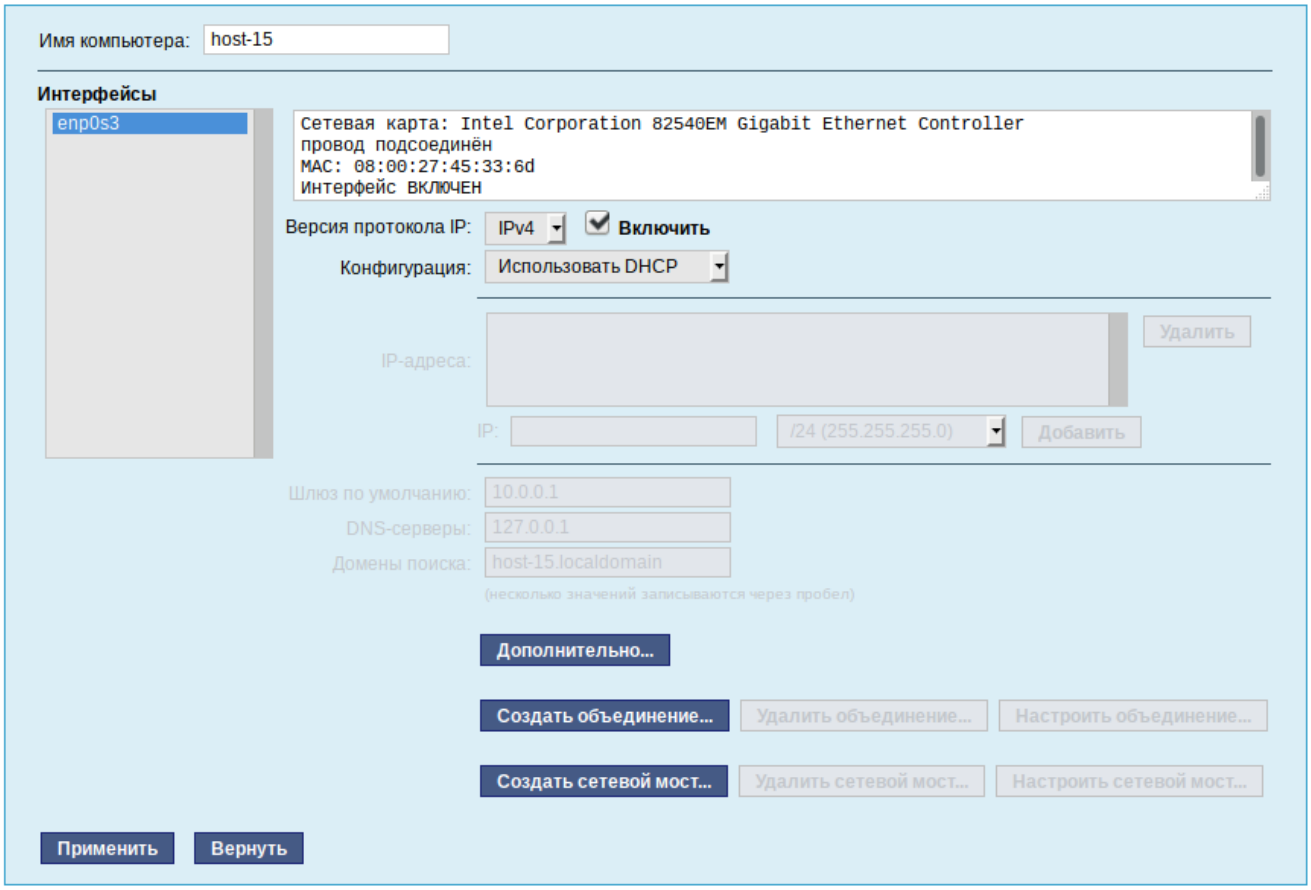

Рис. 197

Если в компьютере имеется несколько сетевых карт, то возможна ситуация, когда при очередной загрузке ядро присвоит имена интерфейсов (enp0s3, enp0s8) в другом порядке. В результате интерфейсы получат не свои настройки. Чтобы этого не происходило, можно привязать

интерфейс к имени по его аппаратному адресу (MAC) или по местоположению на системной шине.

Дополнительно для каждого интерфейса можно настроить сетевую подсистему (NetworkManager, Etcnet), а также должен ли запускаться данный интерфейс при загрузке системы (Рис. 198).

Выбор сетевой подсистемы

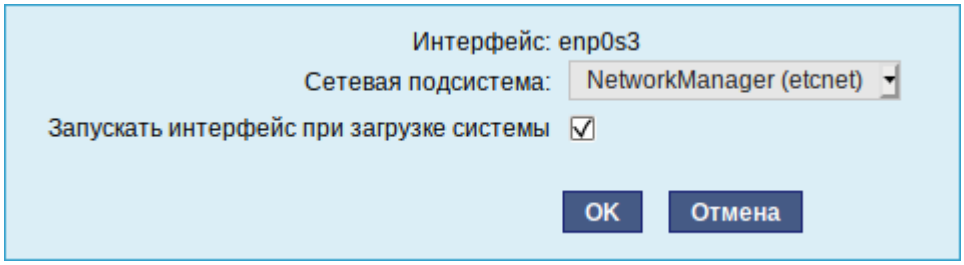

Рис. 198

В списке «Сетевая подсистема» можно выбрать следующие режимы:

– «Etcnet» – в этом режиме настройки берутся исключительно из файлов находящихся в каталоге настраиваемого интерфейса `/etc/net/ifaces/<интерфейс>`. Настройки сети могут изменяться либо в ЦУС в данном модуле, либо напрямую через редактирование файлов `/etc/net/ifaces/<интерфейс>`;
– «NetworkManager (etcnet)» – в этом режиме NetworkManager сам инициирует сеть, используя в качестве параметров – настройки из файлов Etcnet. Настройки сети могут изменяться либо в ЦУС в данном модуле, либо напрямую через редактирование файлов `/etc/net/ifaces/<интерфейс>`. В этом режиме можно просмотреть настройки сети, например полученный по DHCP IP-адрес, через графический интерфейс NetworkManager;
– «NetworkManager (native)» – в данном режиме управление настройками интерфейса передаётся NetworkManager и не зависит от файлов Etcnet. Управлять настройками можно через графический интерфейс NetworkManager. Файлы с настройками находятся в каталоге `/etc/NetworkManager/system-connections`. Этот режим особенно актуален для задач настройки сети на клиенте, когда IP-адрес необходимо получать динамически с помощью DHCP, а DNS-сервер указать явно. Через ЦУС так настроить невозможно, так как при включении DHCP отключаются настройки, которые можно задавать вручную;
– «Не контролируется» – в этом режиме интерфейс находится в состоянии DOWN (выключен).

8.1.2 Настройка общего подключения к сети Интернет

Пользователи корпоративных сетей обычно подключаются к сети Интернет через один общий канал. Для организации совместного доступа к сети Интернет стандартными средствами поддерживаются две технологии, которые могут использоваться как по отдельности, так и совместно:

- использование прокси-сервера;
- использование NAT.

Оба способа предполагают, что соединение с Интернет компьютера, через который предполагается настроить общий выход, предварительно сконфигурировано.

8.1.2.1 Прокси-сервер

Отличительной особенностью использования прокси-сервера является то, что, помимо предоставления доступа к веб-сайтам, прокси-сервер кэширует загруженные страницы, а при повторном обращении к ним – отдаёт их из своего кэша. Это может существенно снизить потребление трафика.

У прокси-сервера есть два основных режима работы:

- прозрачный;
- обычный.

Для работы с прокси-сервером в прозрачном режиме специальная настройка рабочих станций не потребуется. Они лишь должны использовать сервер в качестве шлюза по умолчанию. Этого можно добиться, сделав соответствующие настройки на DHCP-сервере.

Для использования прокси-сервера в обычном режиме потребуется на каждом клиенте в настройках браузера указать данные прокси-сервера (IP-адрес и порт).

Преимуществом обычного режима работы, требующего перенастройки программ локальной сети, является возможность производить аутентификацию пользователей и контролировать их доступ во внешнюю сеть.

В различных браузерах местоположение формы настройки на прокси-сервер различное.

По умолчанию прокси-сервер не предоставляет доступ в Интернет никому кроме себя самого. Список сетей, обслуживаемых прокси-сервером можно изменить, нажав на кнопку «Разрешённые сети…» в модуле ЦУС «Прокси-сервер» (пакет alterator-squid) из раздела «Серверы» (Рис. 199).

Модуль «Прокси-сервер»

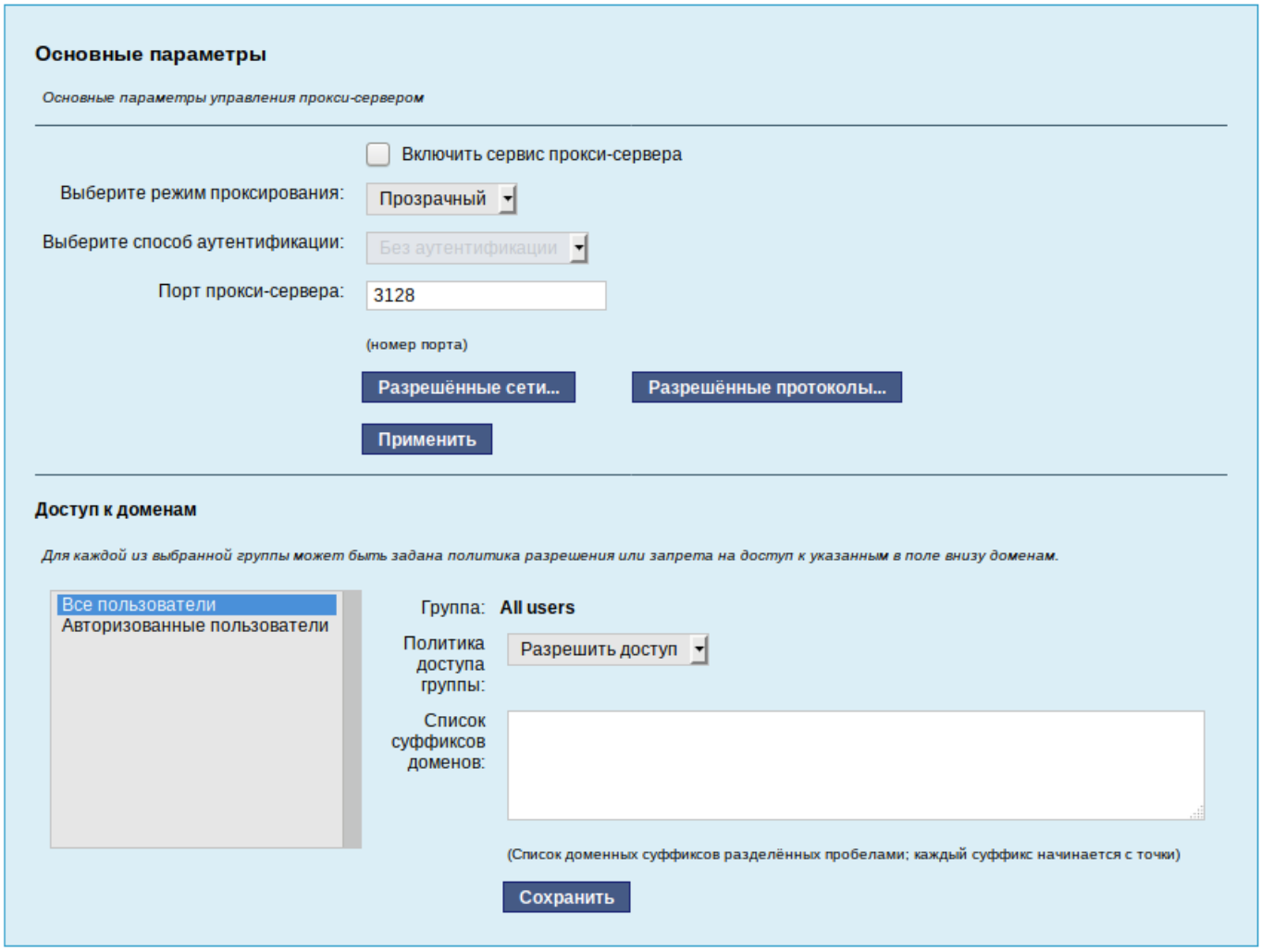

Рис. 199

Для того чтобы включить аутентификацию пользователей и контролировать их доступ во внешнюю сеть, необходимо выбрать обычный режим проксирования и способ аутентификации, отличный от «Без аутентификации» (Рис. 200).

Настройка аутентификации пользователей

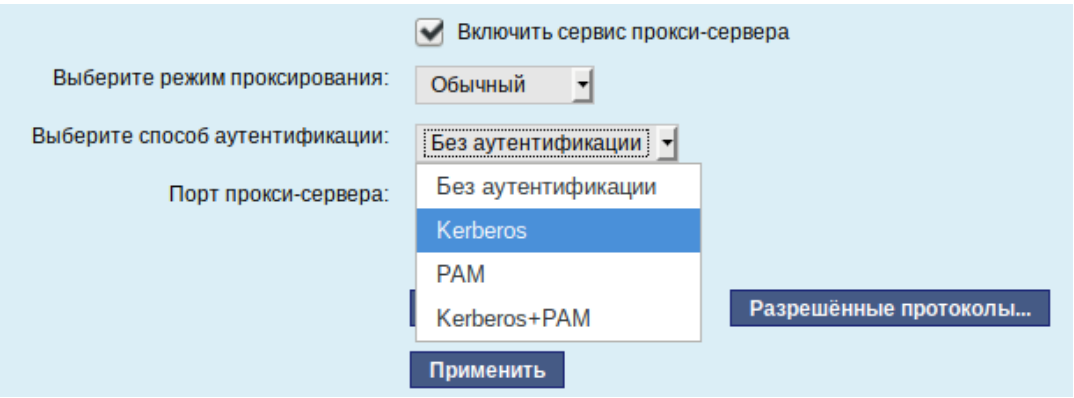

Рис. 200

Прокси-сервер принимает запросы из локальной сети и, по мере необходимости, передаёт их во внешнюю сеть. Поступление запроса ожидается на определённом порту, который по умолчанию имеет стандартный номер 3128.

Перед тем как выполнить перенаправление запроса, прокси-сервер проверяет принадлежность сетевого адрес узла, с которого запрос был отправлен к группе внутренних сетевых адресов. Для того чтобы запросы, отправленные из локальной сети, обрабатывались прокси-сервером, необходимо добавить соответствующую группу адресов (адрес подсети и адресную маску) в список внутренних сетей в разделе «Разрешённые сети» (Рис. 201).

Настройка списка внутренних сетей

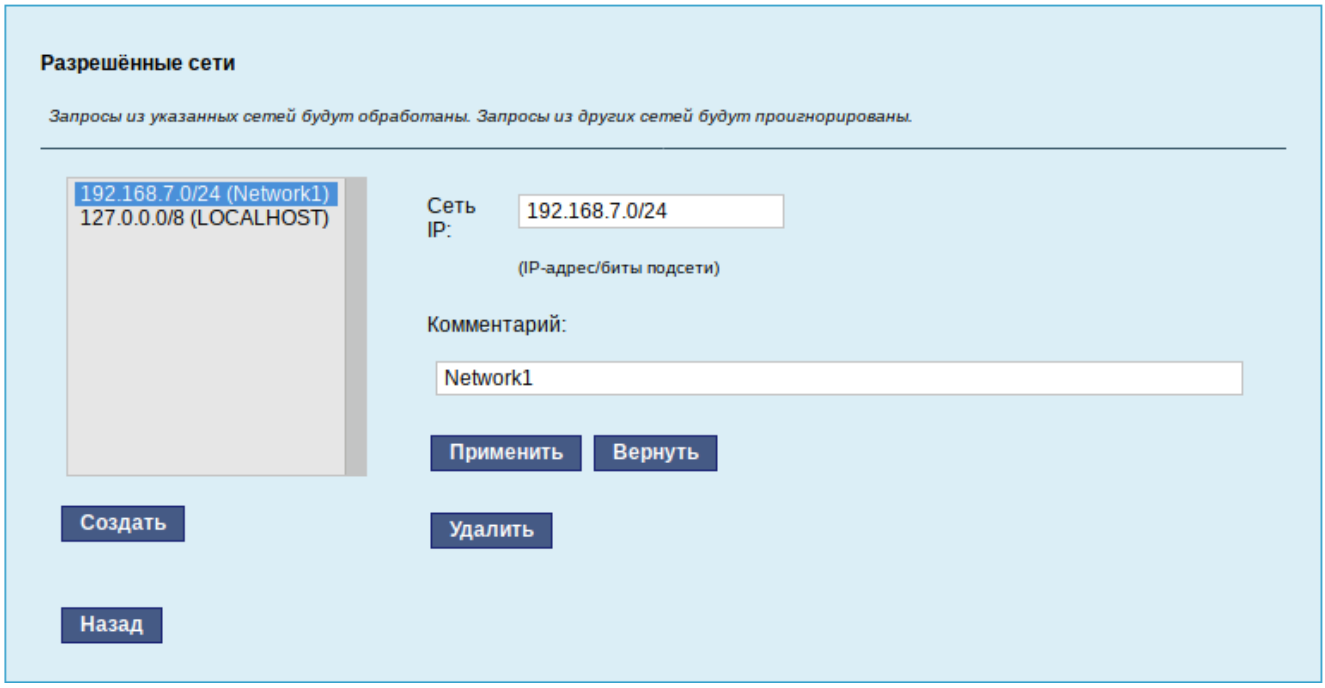

Рис. 201

Вторым условием передачи запроса является принадлежность целевого порта к разрешённому диапазону. Посмотреть и отредактировать список разрешённых целевых портов можно в разделе «Разрешённые протоколы» (Рис. 202).

Настройка списка разрешённых целевых портов

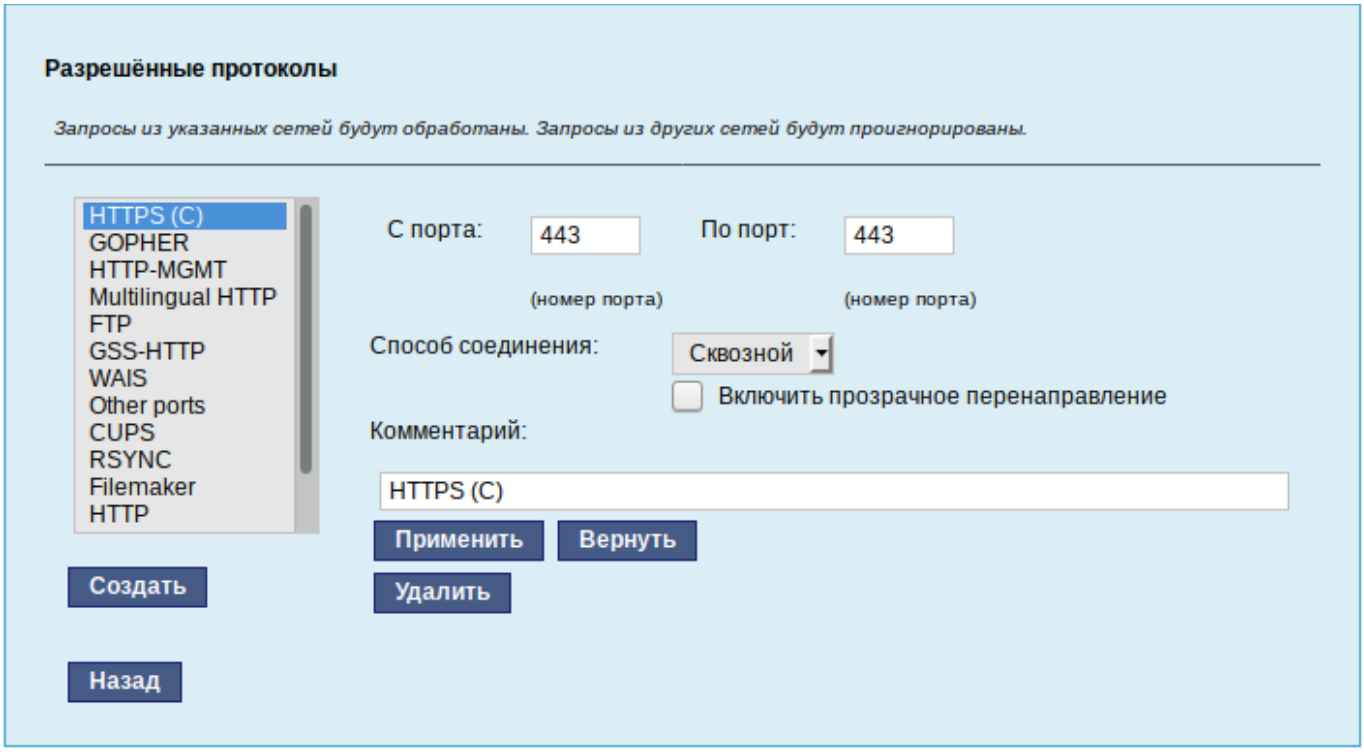

Рис. 202

Прокси-сервер позволяет вести статистику посещений страниц в Интернете. Она доступна в модуле ЦУС «Прокси-сервер» (пакет alterator-squidmill) в разделе «Статистика». Основное предназначение статистики – просмотр отчёта об объёме полученных из Интернета данных в привязке к пользователям (если включена аутентификация) или к IP-адресам клиентов.

8.1.2.2 NAT

NAT (Network Address Translation, преобразование сетевых адресов) – это механизм в сетях TCP/IP, позволяющий преобразовывать IP-адреса транзитных пакетов. Таким образом, компьютеры локальной сети, имеющие IP-адреса, зарезервированные для использования исключительно в локальных сетях, могут использовать общий канал доступа к сети Интернет (общий внешний IP-адрес). При этом на компьютере-шлюзе, непосредственно подключённом к сети Интернет, выполняется преобразование адресов.

Настройка NAT осуществляется в модуле ЦУС «Внешние сети» (пакет alterator-net-iptables) из раздела «Брандмауэр». Для минимальной настройки достаточно выбрать режим работы «Шлюз» (NAT), отметить правильный внешний сетевой интерфейс (Рис. 203) и нажать на кнопку «Применить».

Настройка NAT в модуле «Внешние сети»

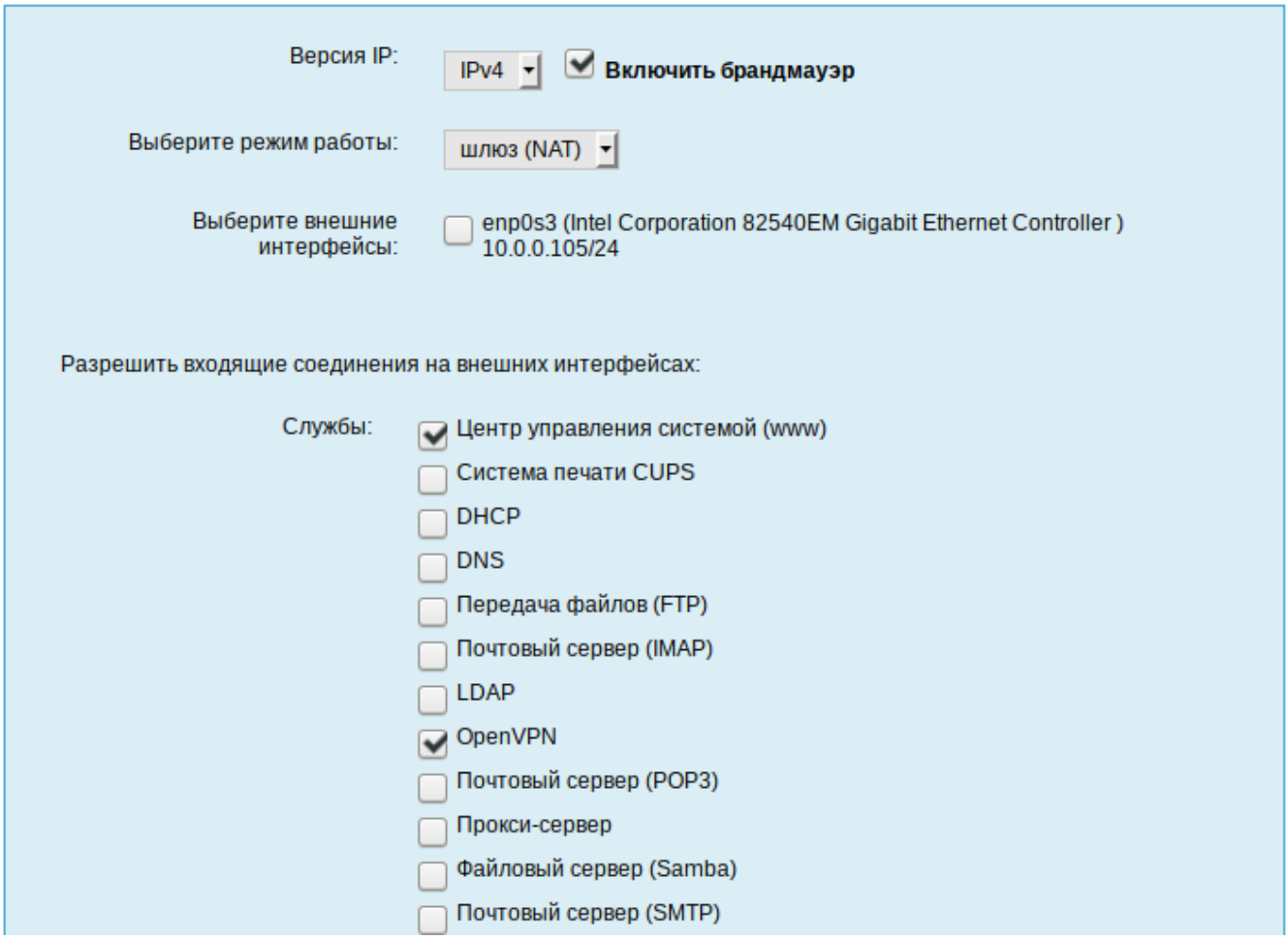

Рис. 203

8.1.3 Автоматическое присвоение IP-адресов (DHCP-сервер)

DHCP (Dynamic Host Configuration Protocol) – протокол, позволяющий клиенту самостоятельно получить IP-адрес из зарезервированного диапазона адресов, а также дополнительную информацию о локальной сети (DNS-сервер сети, домен поиска, шлюз по умолчанию).

Чтобы настраивать DHCP-сервер, на машине должен быть хотя бы один статически сконфигурированный Ethernet-интерфейс.

Настройка DHCP-сервера осуществляется в модуле ЦУС «DHCP-сервер» (пакет alterator-dhcp) из раздела «Серверы».

Для включения DHCP-сервера необходимо установить флажок «Включить службу DHCP» (Рис. 204), указать начальный и конечный IP-адрес, а также шлюз по умолчанию (обычно, это IP-адрес сервера на сетевом интерфейсе, обслуживающем локальную сеть).

Теперь при включении любой клиентской машины с настройкой «получение IP и DNS автоматически» будет присваиваться шлюз 192.168.8.250, DNS 192.168.8.251 и адреса начиная с 192.168.8.50 по порядку включения до 192.168.8.60.

Настройка модуля DHCP-сервер

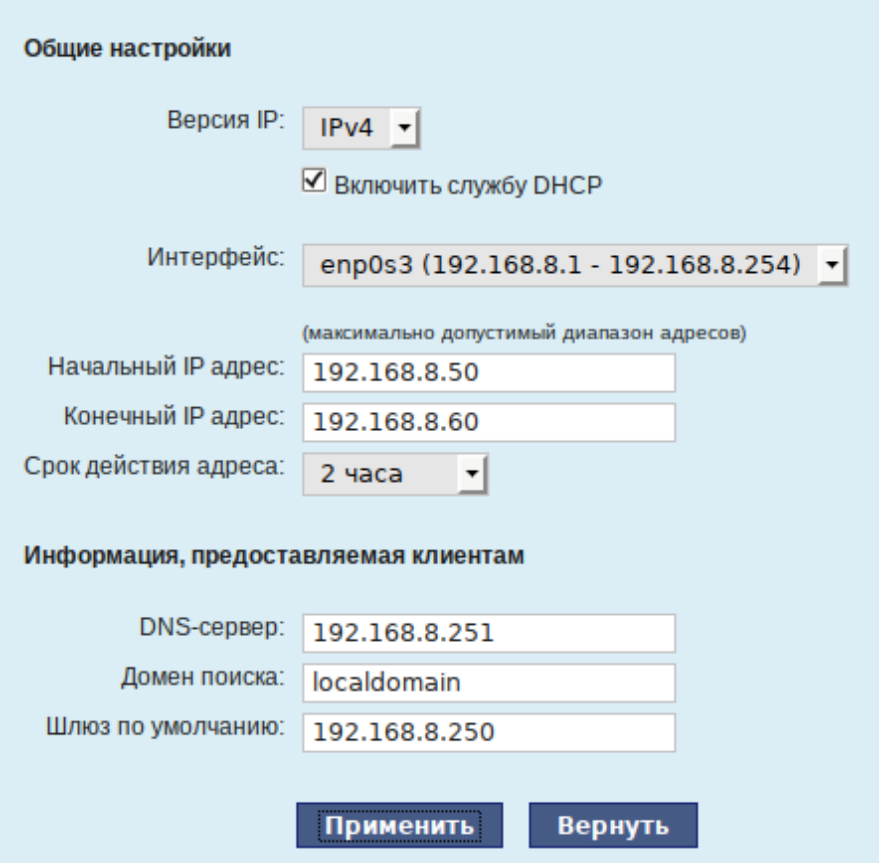

Рис. 204

Иногда бывает полезно выдавать клиенту один и тот же IP-адрес независимо от момента обращения. В этом случае он определяется по аппаратному адресу (MAC-адресу) сетевой карты клиента. Для добавления своих значений в таблицу соответствия статических адресов следует ввести IP-адрес и соответствующий ему MAC-адрес и нажать кнопку «Добавить» (Рис. 205).

Привязка IP-адреса к MAC-адресу

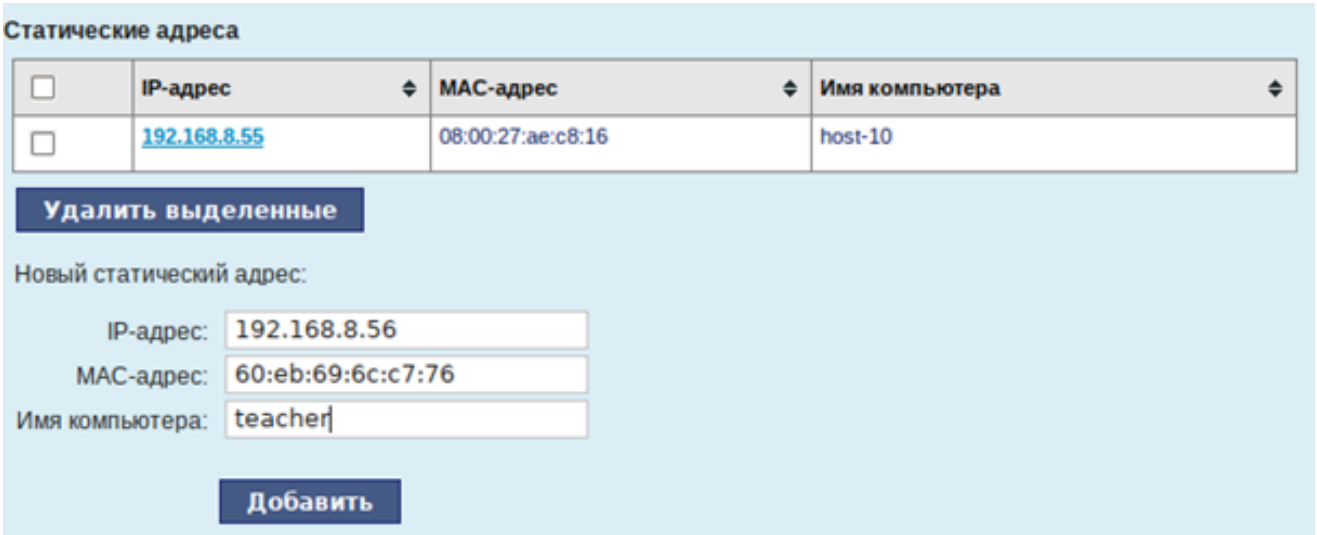

Рис. 205

Выданные IP-адреса можно увидеть в списке «Текущие динамически выданные адреса» (Рис. 206). Также имеется возможность зафиксировать выданные адреса, за данными компьютерами. Для этого необходимо отметить хост, за которым нужно закрепить IP-адрес и нажать кнопку «Зафиксировать адрес для выбранных компьютеров».

Список динамически выданных адресов

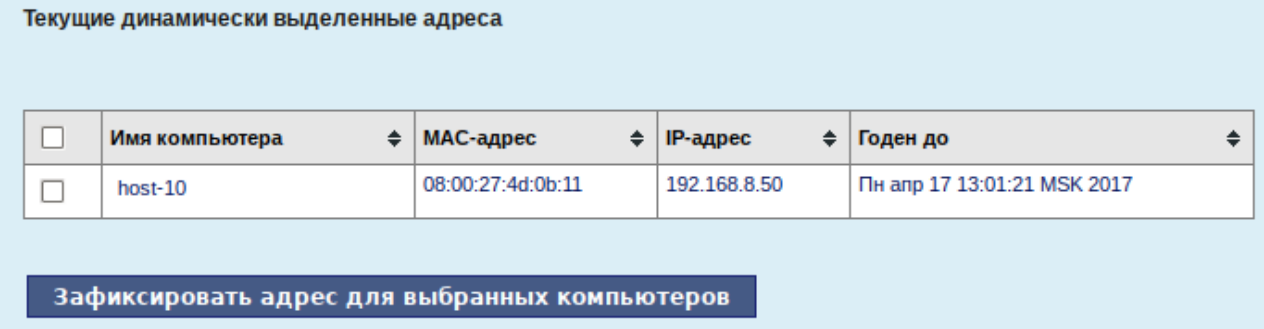

Рис. 206

8.2 Доступ к службам сервера из сети Интернет

8.2.1 Внешние сети

Сервер предоставляет возможность организовать доступ к своим службам извне. Например, можно предоставить доступ к корпоративному веб-сайту из сети Интернет. Для обеспечения такой возможности необходимо разрешить входящие соединения на внешних интерфейсах. По умолчанию такие соединения блокируются.

Для разрешения внешних и внутренних входящих соединений предусмотрен раздел ЦУС «Брандмауэр». В списке «Разрешить входящие соединения на внешних интерфейсах» модуля «Внешние сети» (пакет alterator-net-iptables) перечислены наиболее часто используемые службы, отметив которые, можно сделать их доступными для соединений на внешних сетевых интерфейсах (Рис. 207). Если необходимо предоставить доступ к службе, отсутствующей в списке, то нужно задать используемые этой службой порты в соответствующих полях.

Можно выбрать один из двух режимов работы:

- роутер – перенаправление пакетов между сетевыми интерфейсами происходит без трансляции сетевых адресов;
- шлюз (NAT) – в этом режиме будет настроена трансляция сетевых адресов (NAT) при перенаправлении пакетов на внешние интерфейсы. Использование этого режима имеет смысл, если на компьютере настроен, по крайней мере, один внешний и один внутренний интерфейс.

Модуль «Внешние сети»

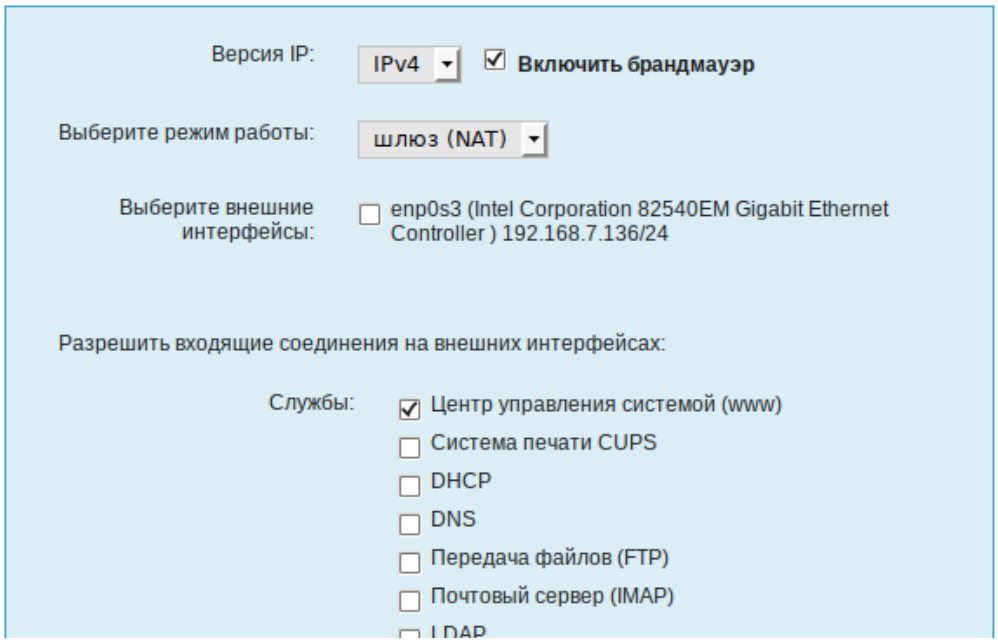

Рис. 207

В любом режиме включено только перенаправление пакетов с внутренних интерфейсов. Перенаправление пакетов с внешних интерфейсов всегда выключено. Все внутренние интерфейсы открыты для любых входящих соединений.

8.2.2 Список блокируемых хостов

Модуль «Список блокируемых хостов» (пакет alterator-net-bl) позволяет настроить блокировку любого сетевого трафика с указанных в списке узлов (входящий, исходящий и пересылаемый).

Блокирование трафика с указанных в списке узлов начинается после установки флажка «Использовать чёрный список» (Рис. 208).

Модуль «Список блокируемых хостов»

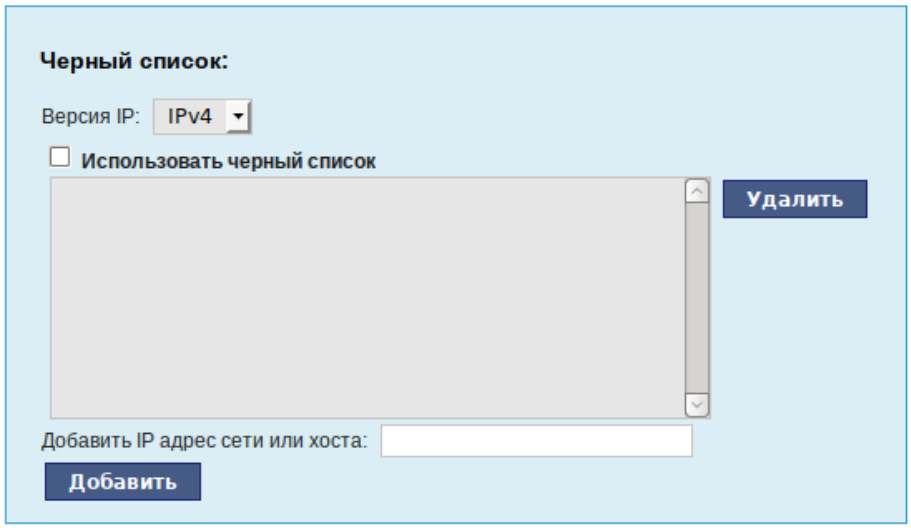

Рис. 208

Для добавления блокируемого узла необходимо ввести IP-адрес в поле «Добавить IP адрес сети или хоста» и нажать кнопку «Добавить».

Для удаления узла необходимо выбрать его из списка и нажать кнопку «Удалить».

8.3 Статистика

8.3.1 Сетевой трафик

Все входящие и исходящие с сервера сетевые пакеты могут подсчитываться, и выводится по запросу для анализа.

Модуль «Сетевой трафик» (пакет alterator-ulogd) из раздела «Статистика» предназначен для просмотра статистики входящих и исходящих с сервера сетевых пакетов. Данный модуль позволяет оценить итоговый объём полученных и переданных данных за всё время работы сервера, за определённый период времени и по каждой службе отдельно.

Для включения сбора данных необходимо установить флажок «Включить сбор данных», и нажать кнопку «Применить» (Рис. 209).

Просмотр статистики входящих и исходящих пакетов

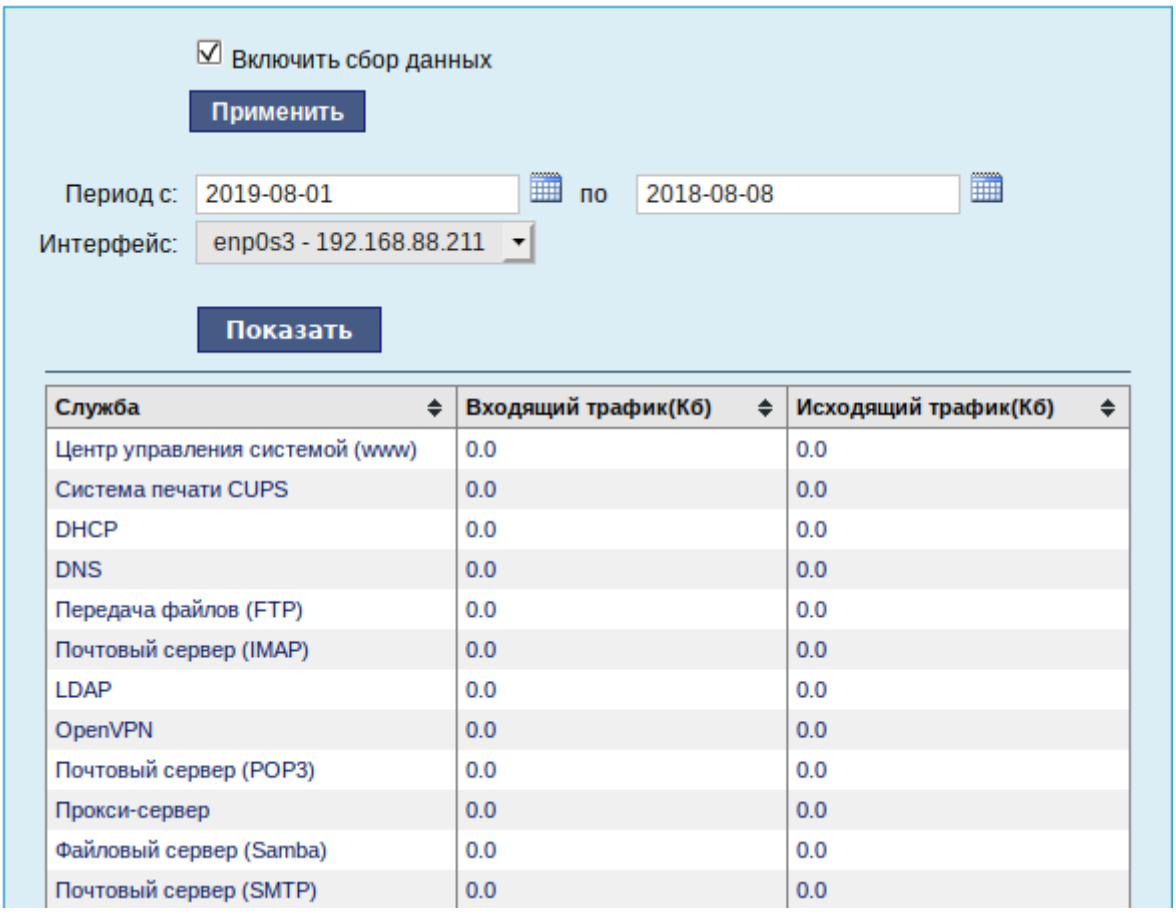

Рис. 209

Для просмотра статистики указывается период (в виде начальной и конечной дат). Дата указывается в формате YYYY-MM-DD (год-месяц-день) или выбирается из календаря справа от

поля ввода даты. Из списка доступных сетевых интерфейсов необходимо выбрать интересующий и нажать на кнопку «Показать» (Рис. 209).

Трафик на указанном интерфейсе за заданный период показывается в виде:

- служба (название протокола);
- входящий трафик в килобайтах;
- исходящий трафик в килобайтах.

8.3.2 Прокси-сервер

Пересылка каждого запроса во внешнюю сеть фиксируется прокси-сервером в специальном журнале. На основании этих данных автоматически формируются отчёты о статистике использования ресурсов сети, в том числе потраченного времени и количества переданных данных (трафика).

Статистика не собирается по умолчанию. Включить её сбор следует в модуле ЦУС «Прокси-сервер» (пакет alterator-squidmill) из раздела «Статистика». Для включения сбора статистики прокси-сервера необходимо установить флажок «Включить сбор данных прокси-сервера» (Рис. 210).

Настройка сбора статистики прокси-сервера

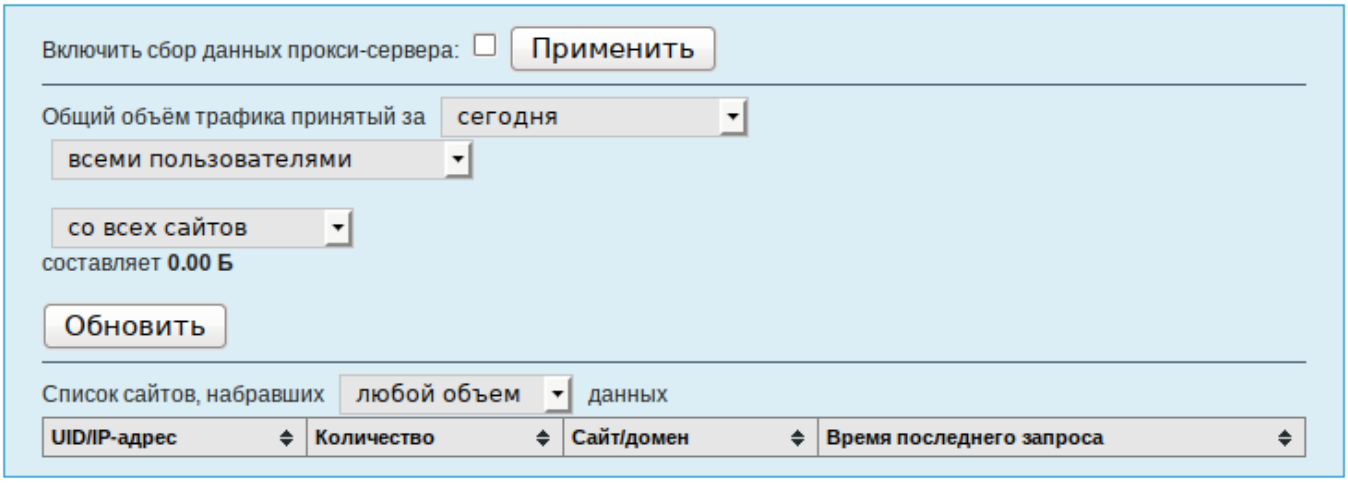

Рис. 210

В том случае, если на прокси-сервере производилась аутентификация пользователей, отчёты будут содержать данные об обращениях каждого пользователя. Иначе отчёты будут формироваться только на основании адресов локальной сети.

Для показа отчёта необходимо задать условия фильтра и нажать кнопку «Показать». Данные в таблице отсортированы по объёму трафика в порядке убывания.

Для учёта пользователей в статистике необходимо добавить хотя бы одно правило. Самое очевидное правило – запрет неаутентифиуцированных пользователей. Только после этого в статистике начнут показываться пользователи.

8.4 Обслуживание сервера

Регулярный мониторинг состояния сервера, своевременное резервное копирование, обновление установленного ПО, являются важной частью комплекса работ по обслуживанию сервера.

8.4.1 Мониторинг состояния системы

Для обеспечения бесперебойной работы сервера крайне важно производить постоянный мониторинг его состояния. Все события, происходящие с сервером, записываются в журналы, анализ которых помогает избежать сбоев в работе сервера и предоставляет возможность разобраться в причинах некорректной работы сервера.

Для просмотра журналов предназначен модуль ЦУС «Системные журналы» (пакет alterator-logs) из раздела «Система». Интерфейс позволяет просмотреть различные типы журналов с возможностью перехода к более старым или более новым записям.

Различные журналы могут быть выбраны из списка «Журналы» (Рис. 211).

Модуль «Системные журналы»

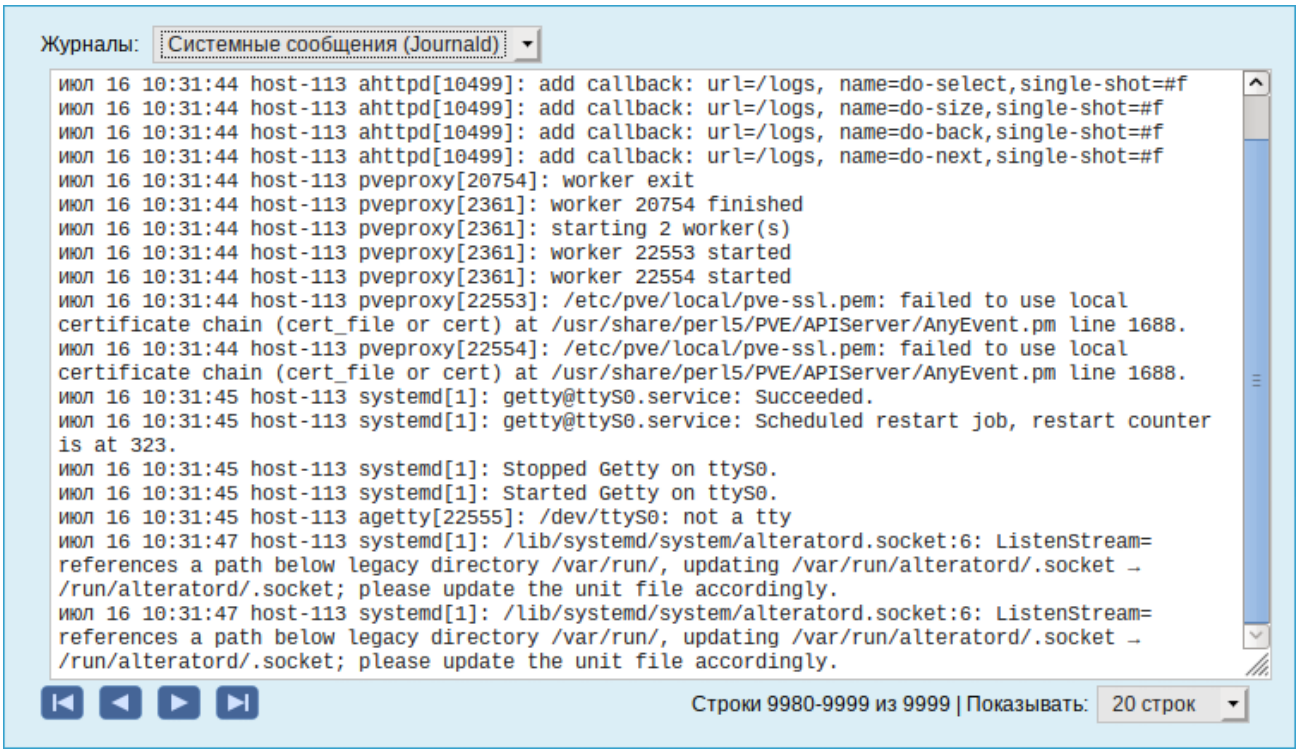

Рис. 211

Доступны следующие виды журналов:

- Брандмауэр – отображаются события безопасности, связанные с работой межсетевого экрана ОС;
- Системные сообщения – сообщения от системных служб (сообщения с типом DAEMON).

Каждый журнал может содержать довольно большое количество сообщений. Уменьшить либо увеличить количество выводимых строк можно, выбрав нужное значение в списке «Показывать».

8.4.2 Системные службы

Для изменения состояния служб можно использовать модуль ЦУС «Системные службы» » (пакет alterator-services) из раздела «Система». Интерфейс позволяет изменять текущее состояние службы и, если необходимо, применить опцию запуска службы при загрузке системы (Рис. 212).

Модуль «Системные службы»

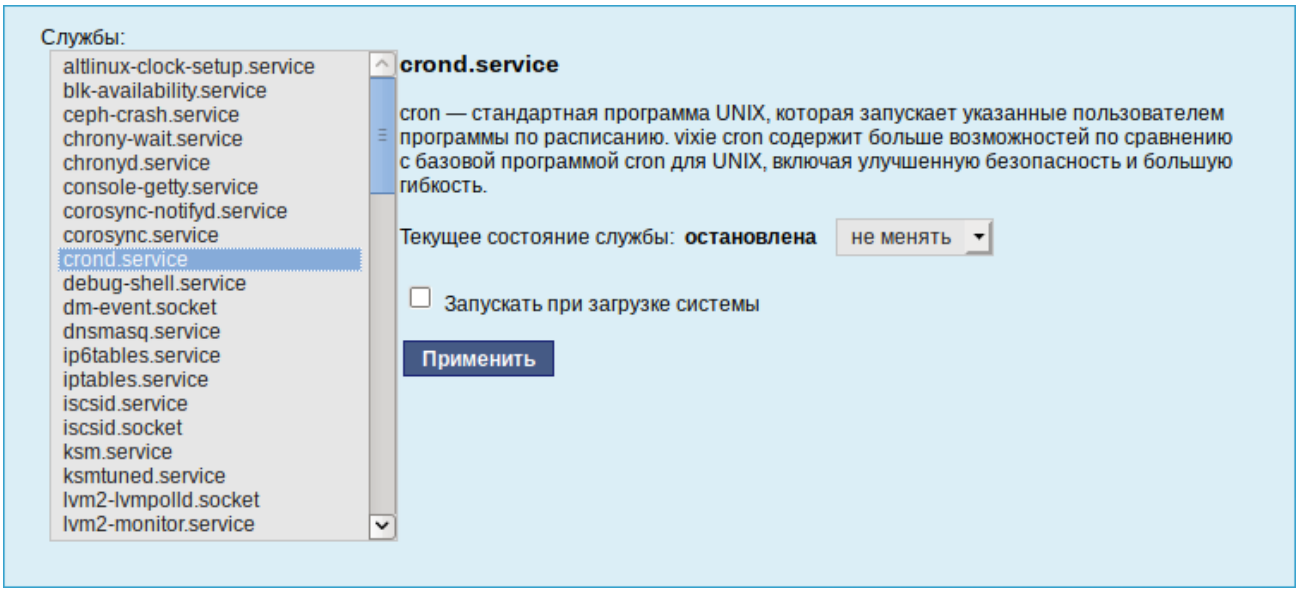

Рис. 212

После выбора названия службы из списка отображается описание данной службы, а также текущее состояние: Работает/Остановлена/Неизвестно.

8.4.3 Обновление системы

После установки системы крайне важно следить за обновлениями ПО. Обновления для ОС «Альт Сервер Виртуализации» могут содержать как исправления, связанные с безопасностью, так и новый функционал или просто улучшение и ускорение алгоритмов. В любом случае настоятельно рекомендуется регулярно обновлять систему для повышения надёжности работы сервера.

Для автоматизации процесса установки обновлений предусмотрен модуль ЦУС «Обновление системы» (пакет alterator-updates) из раздела «Система». Здесь можно включить автоматическое обновление через Интернет с одного из предлагаемых серверов или задать собственные настройки (Рис. 213).

Источник обновлений указывается явно (при выбранном режиме «Обновлять систему автоматически из сети Интернет») или вычисляется автоматически (при выбранном режиме

«Обновление системы управляемое сервером» и наличии в локальной сети настроенного сервера обновлений).

Процесс обновления системы будет запускаться автоматически согласно заданному расписанию.

Модуль «Обновление системы»

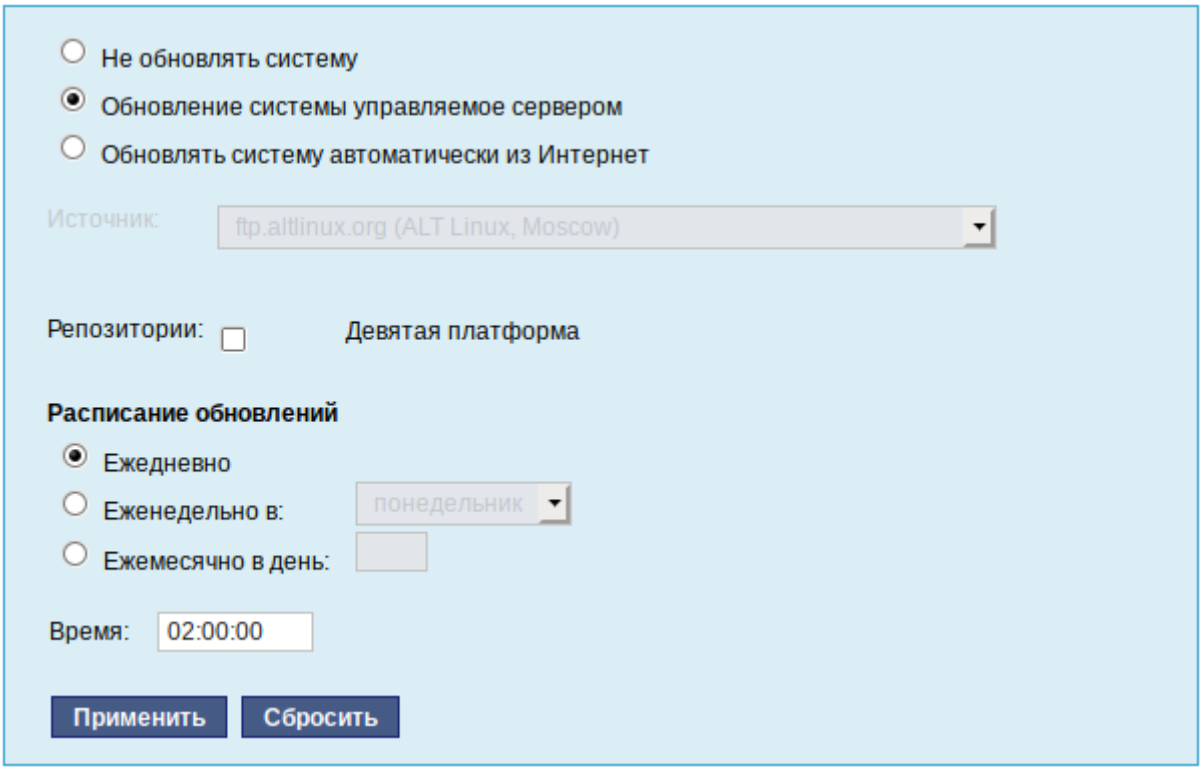

Рис. 213

8.4.4 Обновление систем, не имеющих выхода в Интернет

Для систем, не имеющих прямого выхода в Интернет, рекомендуется установка отдельного сервера обновлений на базе ОС «Альт Сервер Виртуализации», находящегося вне защищенного контура и организация ограниченного доступа к этому серверу.

Модуль ЦУС «Сервер обновлений» (пакет alterator-mirror) из раздела «Серверы» предназначен для зеркалирования репозиториев и публикации их для обновлений рабочих станций и серверов.

По умолчанию локальное зеркало репозитория находится в `/srv/public/mirror`. Для того чтобы зеркалирование происходило в другую папку необходимо эту папку примонтировать в папку `/srv/public/mirror`. Для этого в файл `/etc/fstab` следует вписать следующую строку

```
/media/disk/localrepo /srv/public/mirror none rw,bind,auto 0 0
```

где `/media/disk/localrepo` – папка-хранилище локального репозитория.

На странице модуля можно выбрать, как часто выполнять закачку пакетов, можно выставить время, когда начинать зеркалирование (Рис. 214).

Модуль «Сервер обновлений»

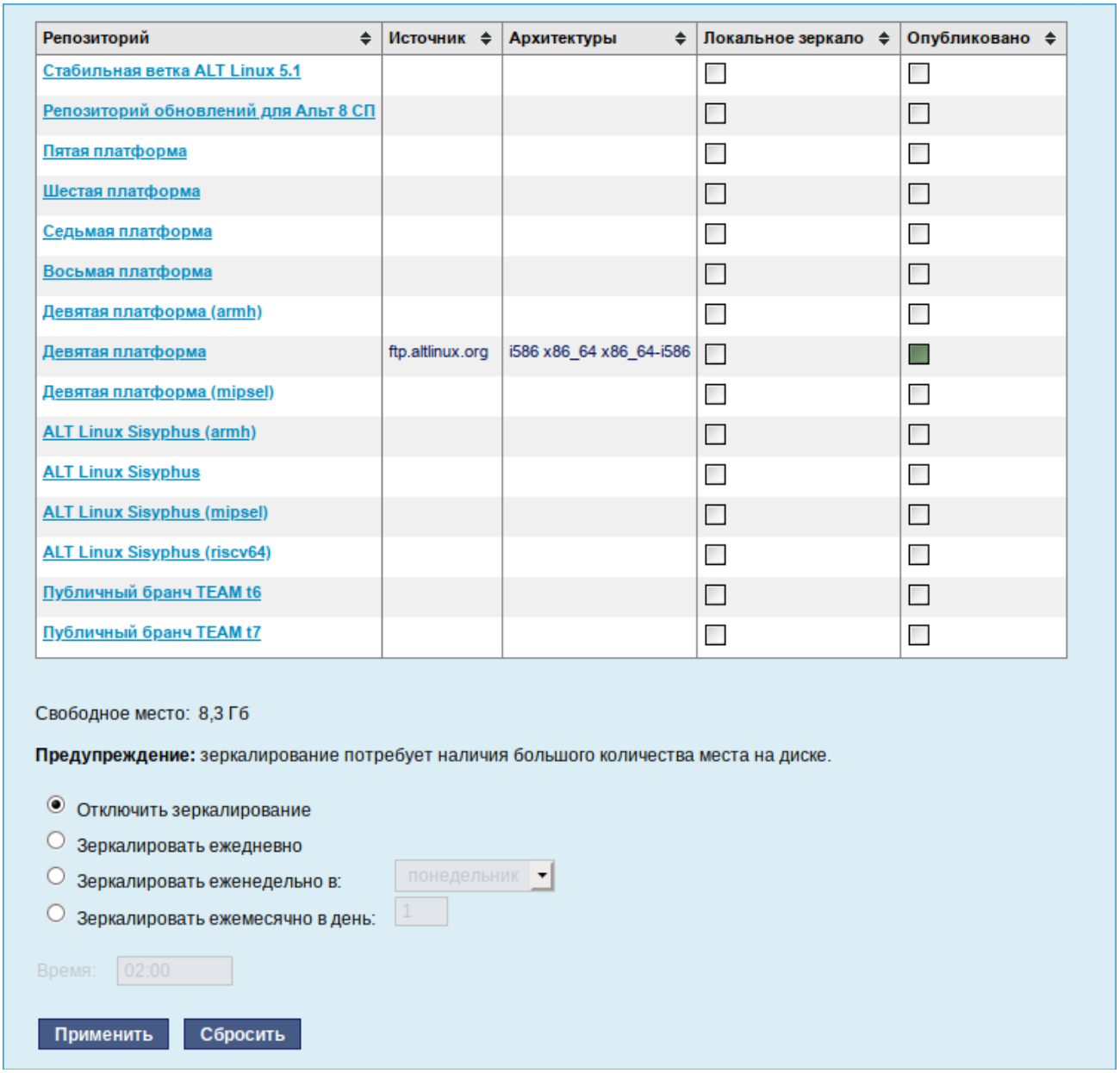

Рис. 214

Здесь также можно выбрать репозитории, локальные срезы которых необходимы. При нажатии на название репозитория, появляются настройки этого репозитория (Рис. 215). Необходимо выбрать источник, архитектуру процессора (если их несколько, то стоит выбрать соответствующие).

Сервер обновлений предоставляет возможность автоматически настроить обновление клиентских машин в нужном режиме:

– Локальное зеркало репозитория – в этом режиме на сервере создаётся копия удалённого репозитория, доступная клиентским машинам по протоколу FTP. Загрузка ПО клиентскими

машинами производится с локального сервера. Наличие на локальном сервере зеркала репозитория при большом количестве машин в сети позволяет существенно сэкономить на трафике.

– Публикация репозитория – в этом случае реального зеркалирования (загрузки пакетов) не происходит. Публикуется URL внешнего сервера, содержащего репозиторий. Такая публикация позволяет клиентским машинам автоматически настроить свои менеджеры пакетов на использование внешнего сервера. Загрузка ПО клиентскими машинами производится с внешнего сервера.

Настройки репозитория

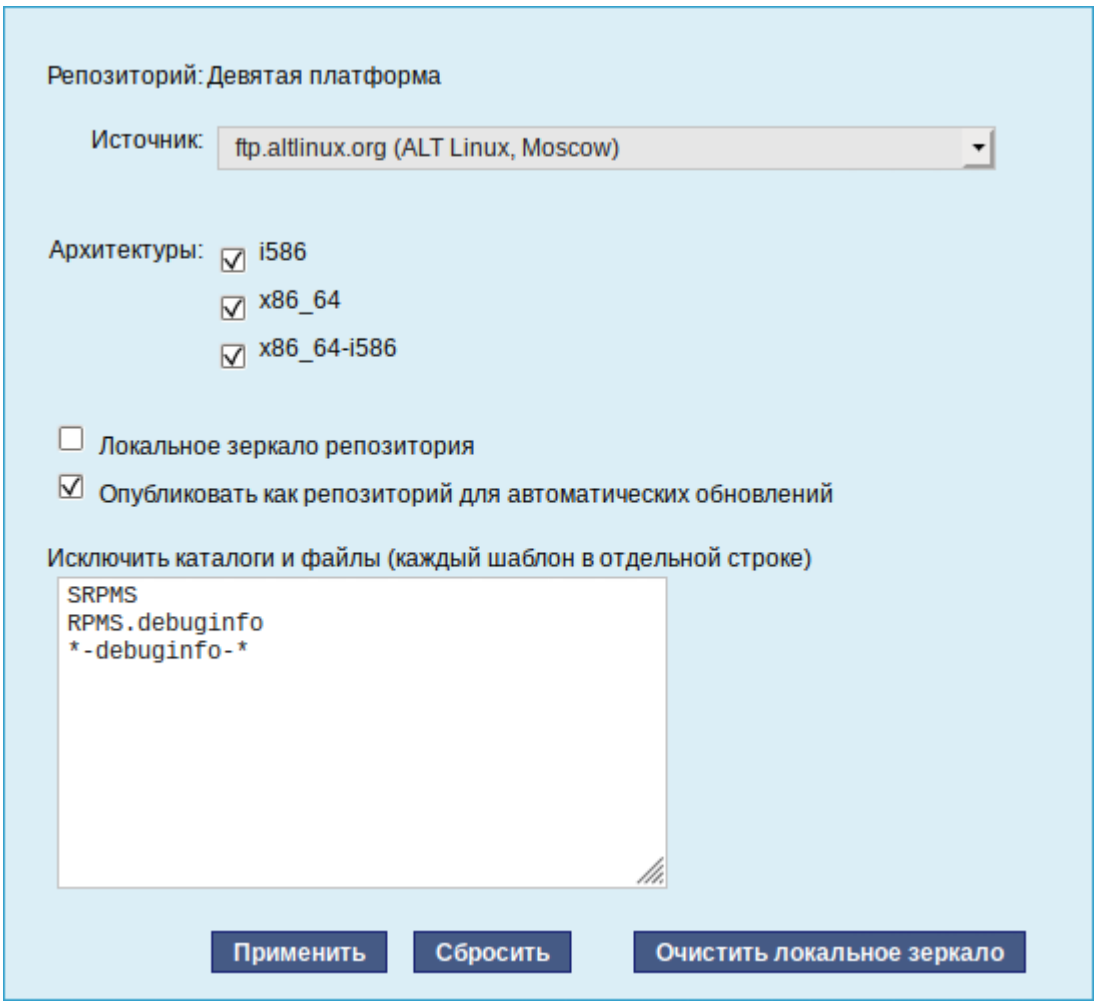

Рис. 215

Здесь также можно указать имена каталогов и файлов, которые будут исключены из синхронизации, что позволит уменьшить размер скачиваемых файлов и занимаемое репозиторием место на диске. Например, не скачивать пакеты с исходным кодом и пакеты с отладочной информацией:

```
SRPMS
*-debuginfo-*
```

Шаблоны указываются по одному в отдельной строке. Символ «*» используется для подстановки любого количества символов.

Настройка локального репозитория заканчивается нажатием на кнопку «Применить».

Далее необходимо отредактировать файл `/etc/httpd2/conf/extra-available/Directory_html_default.conf`, изменив следующие строки:

```
Options Indexes FollowSymLinks MultiViews
AllowOverride None
Order allow,deny
allow from all
```

Этим серверу apache будет разрешено обрабатывать символические ссылки.

Перезапустить apache:

```
# service httpd2 restart
```

Перейти в папку веб-сервера:

```
cd /var/www/html
```

Создать здесь символическую ссылку на репозиторий:

```
ln -s /srv/public/mirror mirror
```

На клиентских машинах необходимо настроить репозитории. Для этого необходимо запустить Synaptic, в параметрах выбрать репозитории. И далее настроить URL доступных репозиториев:

```
http://<IP-адрес>/mirror/
```

Так же со стороны клиентских машин на них необходимо настроить модуль «Обновление системы» отметив в нем пункт «Обновление системы, управляемое сервером».

8.4.5 Локальные учётные записи

Модуль «Локальные учётные записи» (пакет alterator-users) из раздела «Пользователи» предназначен для администрирования системных пользователей (Рис. 216).

Веб-интерфейс модуля alterator-users

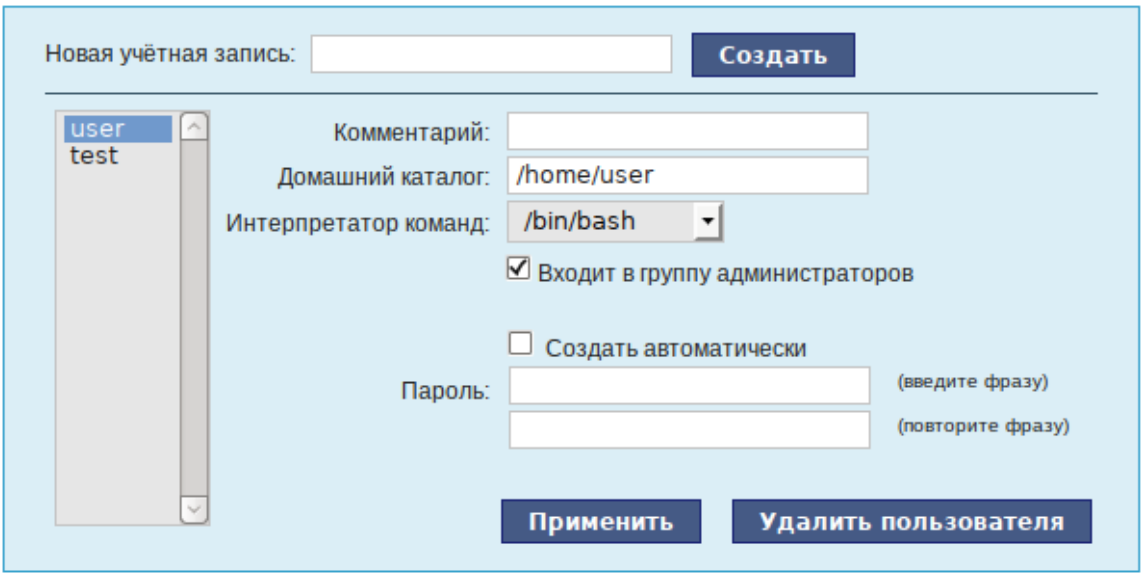

Рис. 216

Для создания новой учётной записи необходимо ввести имя новой учётной записи и нажать кнопку «Создать», после чего имя отобразится в списке слева.

Для дополнительных настроек необходимо выделить добавленное имя, либо, если необходимо изменить существующую учётную запись, выбрать её из списка.

8.4.6 Администратор системы

В модуле «Администратор системы» (пакет alterator-root) из раздела «Пользователи» можно изменить пароль суперпользователя (root), заданный при начальной настройке системы (Рис. 217).

Модуль «Администратор системы»

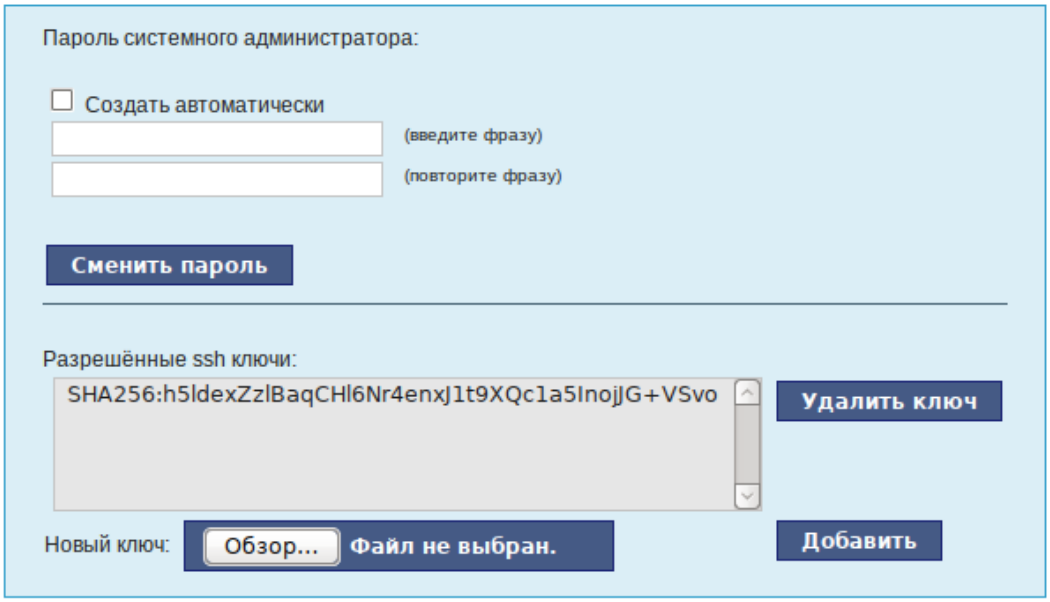

Рис. 217

В данном модуле (только в веб-интерфейсе) можно добавить публичную часть ключа RSA или DSA для доступа к серверу по протоколу SSH.

8.4.7 Дата и время

В модуле «Дата и время» (пакет alterator-datetime) из раздела «Система» можно изменить дату и время на сервере, сменить часовой пояс, а также настроить автоматическую синхронизацию часов на самом сервере по протоколу NTP и предоставление точного времени по этому протоколу для рабочих станций локальной сети (Рис. 218).

Системное время зависит от следующих факторов:

- часы в BIOS – часы, встроенные в компьютер; они работают, даже если он выключен;
- системное время – часы в ядре операционной системы. Во время работы системы все процессы пользуются именно этими часами;
- часовые пояса – регионы Земли, в каждом из которых принято единое местное время.

Модуль «Дата и время»

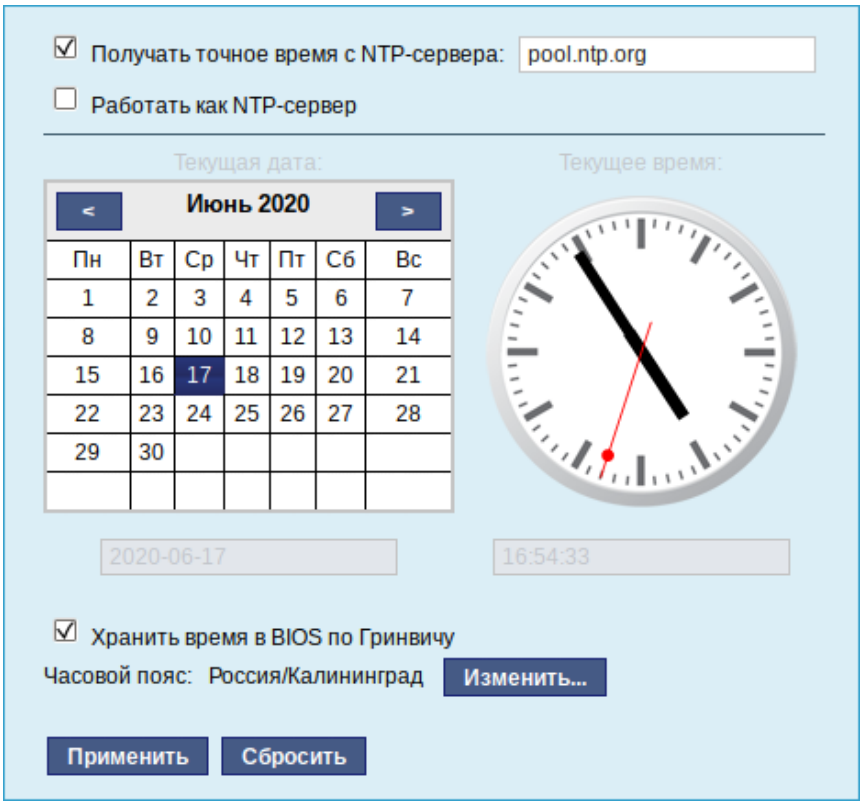

Рис. 218

При запуске системы происходит активация системных часов и их синхронизация с аппаратными, кроме того, в определённых случаях учитывается значение часового пояса. При завершении работы системы происходит обратный процесс.

Если настроена синхронизация времени с NTP-сервером, то сервер сможет сам работать как сервер точного времени. Для этого достаточно отметить соответствующий пункт «Работать как NTP-сервер».

8.4.8 Ограничение использования диска

Модуль «Использование диска» (пакет alterator-quota) из раздела «Пользователи» позволяет ограничить использование дискового пространства пользователями, заведёнными на сервере в модуле «Пользователи».

Модуль позволяет задать ограничения (квоты) для пользователя при использовании определённого раздела диска. Ограничить можно как суммарное количество килобайт, занятых файлами пользователя, так и количество этих файлов (Рис. 219).

Модуль «Использование диска»

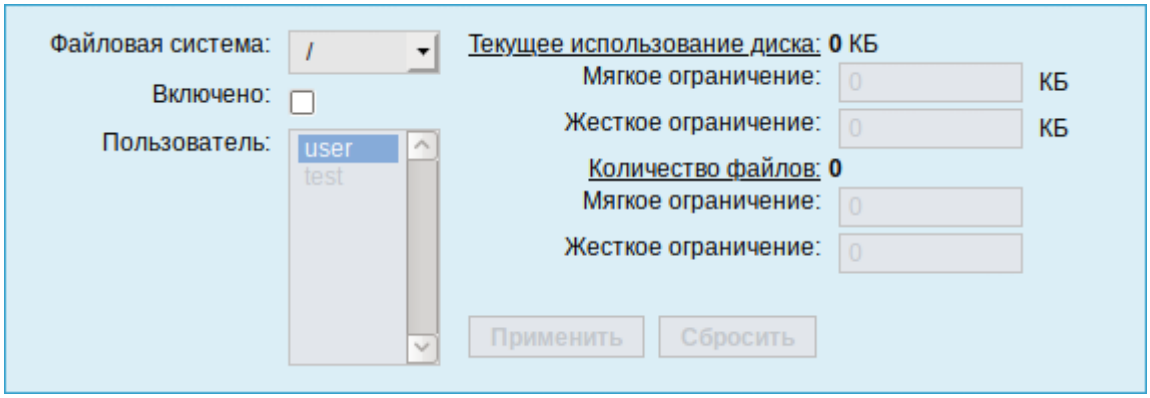

Рис. 219

Для управления квотами файловая система должна быть подключена с параметрами usrquota, grpquota. Для этого следует выбрать нужный раздел в списке «Файловая система» и установить отметку в поле «Включено» (Рис. 220).

Модуль «Использование диска»

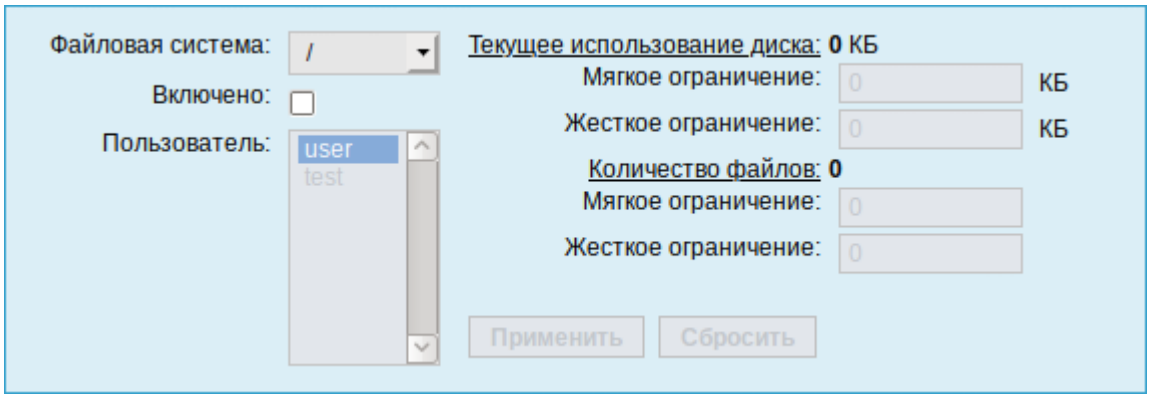

Рис. 220

Для того чтобы задать ограничения для пользователя, необходимо выбрать пользователя в списке «Пользователь», установить ограничения и нажать кнопку «Применить».

При задании ограничений различают жёсткие и мягкие ограничения:

– мягкое ограничение: нижняя граница ограничения, которая может быть временно превышена. Временное ограничение – одна неделя;

– жёсткое ограничение: использование диска, которое не может быть превышено ни при каких условиях.

Значение 0 при задании ограничений означает отсутствие ограничений.

8.4.9 Выключение и перезагрузка компьютера

Модуль ЦУС «Выключение компьютера» (пакет alterator-ahttpd-power) в разделе «Система» позволяет выполнить:

– выключить компьютер;

– перезагрузить компьютер;

– приостановить работу компьютера;

– погрузить компьютер в сон.

Возможна настройка ежедневного применения данных действий в заданное время.

Так как выключение и перезагрузка – критичные для функционирования компьютера операции, то по умолчанию настройка выставлена в значение «Продолжить работу» (Рис. 221). Для выключения, перезагрузки или перехода в энергосберегающие режимы нужно отметить соответствующий пункт и нажать «Применить».

Модуль «Выключение компьютера»

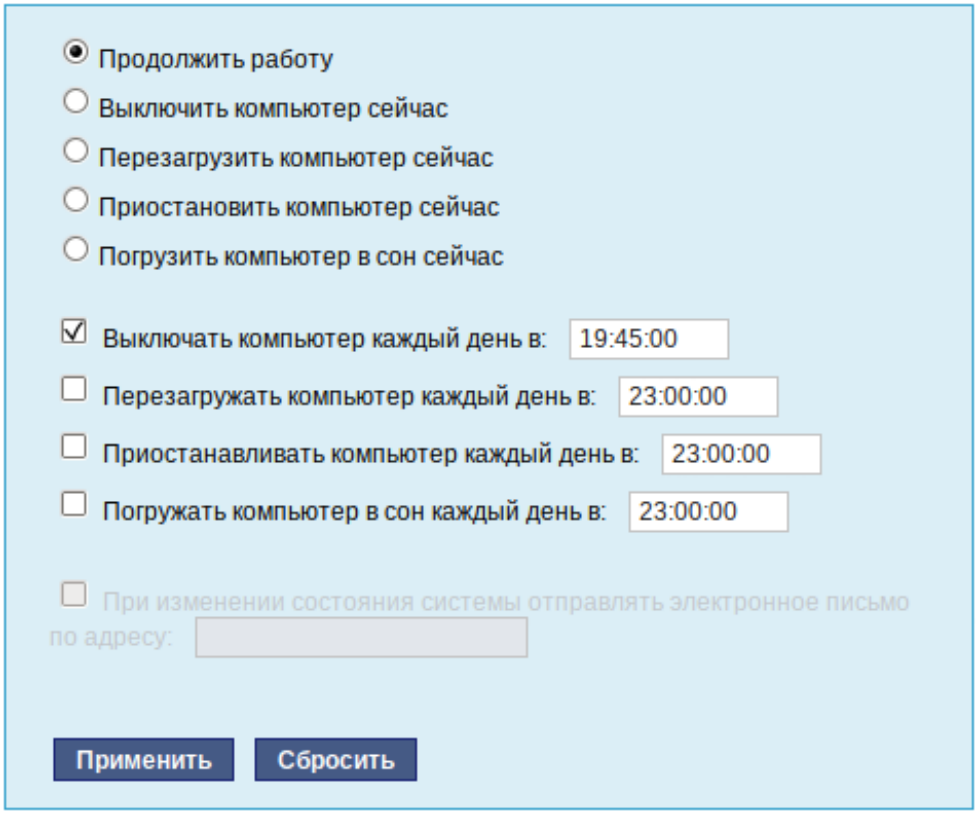

Рис. 221

Для ежедневного автоматического выключения компьютера, перезагрузки, а также перехода в энергосберегающие режимы необходимо отметить соответствующий пункт и задать желаемое

время. Например, для выключения компьютера следует отметить пункт «Выключать компьютер каждый день в», задать время выключения в поле ввода слева от этого флажка и нажать кнопку «Применить».

Примечание. Для возможности настройки оповещений на e-mail, должен быть установлен пакет state-change-notify-postfix:

```
# apt-get install state-change-notify-postfix
```

Для настройки оповещений необходимо отметить пункт «При изменении состояния системы отправлять электронное письмо по адресу», ввести e-mail адрес и нажать кнопку «Применить» (Рис. 222).

Модуль «Выключение компьютера». Настройка оповещений

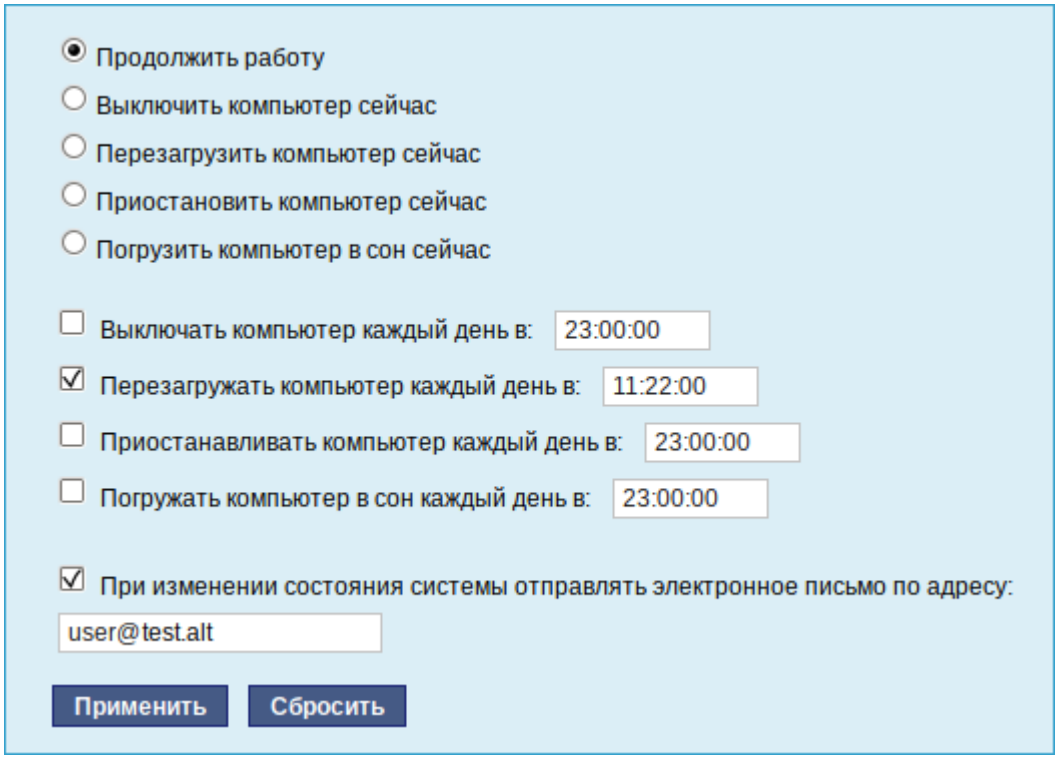

Рис. 222

По указанному адресу, при изменении состоянии системы будут приходить электронные письма. Например, при включении компьютера, содержание письма будет следующее:

```
Tue Jun 16 11:46:59 EET 2020: The server.test.alt is about to start.
```

При выключении:

```
Tue Jun 16 12:27:02 EET 2020: The server.test.alt is about to
shutdown.
```

Кнопка «Сбросить» возвращает сделанный выбор к безопасному значению по умолчанию: «Продолжить работу», перечитывает расписания и выставляет отметки для ежедневного автоматического действия в соответствии с прочитанным.

8.5 Прочие возможности ЦУС

Возможности ЦУС ОС «Альт Сервер Виртуализации» не ограничиваются только теми, что были описаны выше.

Установленные пакеты, которые относятся к ЦУС, можно посмотреть, выполнив команду:

```
rpm -qa | grep alterator*
```

Прочие пакеты для ЦУС можно найти, выполнив команду:

```
apt-cache search alterator*
```

Модули можно дополнительно загружать и удалять как обычные программы:

```
# apt-get install alterator-net-openvpn
# apt-get remove alterator-net-openvpn
```

8.6 Права доступа к модулям ЦУС

Администратор системы (root) имеет доступ ко всем модулям, установленным в системе, и может назначать права доступа для пользователей к определенным модулям.

Для разрешения доступа пользователю к конкретному модулю, администратору в веб-интерфейсе ЦУС необходимо выбрать нужный модуль и нажать ссылку «Параметры доступа к модулю», расположенную в нижней части окна модуля (Рис. 223).

Ссылка «Параметры доступа к модулю»

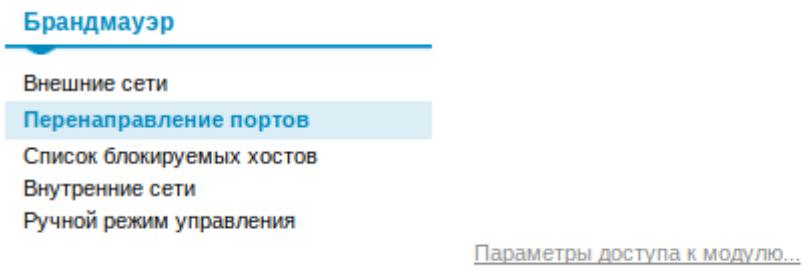

Рис. 223

В открывшемся окне, в списке «Новый пользователь» необходимо выбрать пользователя, который получит доступ к данному модулю, и нажать кнопку «Добавить» (Рис. 224). Для сохранения настроек необходимо перезапустить HTTP-сервер, для этого достаточно нажать кнопку «Перезапустить HTTP-сервер».

Параметры доступа к модулю

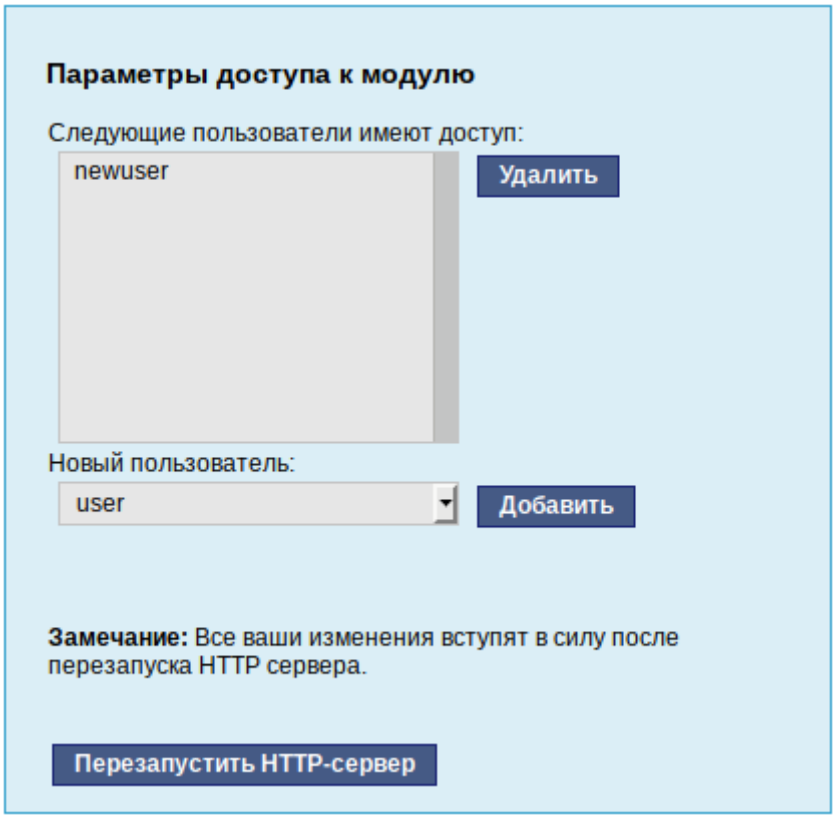

Рис. 224

Для удаления доступа пользователя к определенному модулю, администратору, в окне этого модуля необходимо нажать ссылку «Параметры доступа к модулю», в открывшемся окне в списке пользователей которым разрешен доступ, должен выбрать пользователя, нажать кнопку «Удалить» (Рис. 224) и перезапустить HTTP-сервер.

Системный пользователь, пройдя процедуру аутентификации, может просматривать и вызывать модули, к которым он имеет доступ.

9 УСТАНОВКА ДОПОЛНИТЕЛЬНОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

После установки ОС «Альт Сервер Виртуализации», при первом запуске, доступен тот или иной набор программного обеспечения. Количество предустановленных программ зависит от выбора, сделанного при установке системы. Имеется возможность доустановить программы, которых не хватает в системе, из разных источников.

Дополнительное программное обеспечение может находиться на установочном диске и/или в специальных банках программ (репозиториях), расположенных в сети Интернет и/или в локальной сети. Программы, размещённые в указанных источниках, имеют вид подготовленных для установки пакетов.

Для установки, удаления и обновления программ и поддержания целостности системы в ОС семейства Linux используются менеджеры пакетов типа «rpm». Для автоматизации этого процесса и применяется Усовершенствованная система управления программными пакетами APT (Advanced Packaging Tool).

Автоматизация достигается созданием одного или нескольких внешних репозиториев, в которых хранятся пакеты программ и относительно которых производится сверка пакетов, установленных в системе. Репозитории могут содержать как официальную версию дистрибутива, обновляемую его разработчиками по мере выхода новых версий программ, так и локальные наработки, например, пакеты, разработанные внутри компании.

Таким образом, в распоряжении APT находятся две базы данных: одна описывает установленные в системе пакеты, вторая – внешний репозиторий. APT отслеживает целостность установленной системы и, в случае обнаружения противоречий в зависимостях пакетов, руководствуется сведениями о внешнем репозитории для разрешения конфликтов и поиска корректного пути их устранения.

Система APT состоит из нескольких утилит. Чаще всего используется утилита управления пакетами apt-get, которая автоматически определяет зависимости между пакетами и строго следит за их соблюдением при выполнении любой из следующих операций: установка, удаление или обновление пакетов.

9.1 Источники программ (репозитории)

Репозитории, с которыми работает APT, отличаются от обычного набора пакетов наличием мета информации – индексов пакетов, содержащихся в репозитории, и сведений о них. Поэтому, чтобы получить всю информацию о репозитории, APT достаточно получить его индексы.

APT может работать с любым количеством репозиториев одновременно, формируя единую

информационную базу обо всех содержащихся в них пакетах. При установке пакетов APT обращает внимание только на название пакета, его версию и зависимости, а расположение в том или ином репозитории не имеет значения. Если потребуется, APT в рамках одной операции установки группы пакетов может пользоваться несколькими репозиториями.

Подключая одновременно несколько репозиториев, нужно следить за тем, чтобы они были совместимы друг с другом по пакетной базе – отражали один определенный этап разработки. Совместимыми являются основной репозиторий дистрибутива и репозиторий обновлений по безопасности к данному дистрибутиву. В то же время смешение среди источников APT репозиториев, относящихся к разным дистрибутивам, или смешение стабильного репозитория с нестабильной веткой разработки (Sisyphus) чревато различными неожиданными трудностями при обновлении пакетов.

APT позволяет взаимодействовать с репозиторием с помощью различных протоколов доступа. Наиболее популярные – HTTP и FTP, однако существуют и некоторые дополнительные методы.

Для того чтобы APT мог использовать тот или иной репозиторий, информацию о нем необходимо поместить в файл `/etc/apt/sources.list`, либо в любой файл `.list` (например, `mysources.list`) в каталоге `/etc/apt/sources.list.d/`. Описания репозиториев заносятся в эти файлы в следующем виде:

```
rpm [подпись] метод: путь база название
rpm-src [подпись] метод: путь база название
```

где:

- rpm или rpm-src – тип репозитория (скомпилированные программы или исходные тексты);
- [подпись] – необязательная строка-указатель на электронную подпись разработчиков. Наличие этого поля подразумевает, что каждый пакет из данного репозитория должен быть подписан соответствующей электронной подписью. Подписи описываются в файле `/etc/apt/vendor.list`;
- метод – способ доступа к репозиторию: ftp, http, file, cdrom, copy;
- путь – путь к репозиторию в терминах выбранного метода;
- база – относительный путь к базе данных репозитория;
- название – название репозитория.

При выборе пакетов для установки APT руководствуется всеми доступными репозиториями вне зависимости от способа доступа к ним. Таким образом, если в репозитории, доступном по сети Интернет, обнаружена более новая версия программы, чем на CD (DVD)-носителе информации, APT начнет загружать данный пакет по сети.

9.1.1 Добавление репозиториев

Непосредственно после установки дистрибутива «Альт Сервер Виртуализации» в `/etc/apt/sources.list`, а также в файлах `/etc/apt/sources.list.d/*.list` обычно указывается несколько репозиториев:

- репозиторий с установочного диска дистрибутива;
- интернет-репозиторий, совместимый с установленным дистрибутивом.

9.1.1.1 Скрипт apt-repo для работы с репозиториями

Для добавления репозиториев можно воспользоваться скриптом `apt-repo`.

Примечание. Для выполнения большинства команд необходимы права администратора.

Просмотреть список активных репозиториев можно, выполнив команду:

```
$ apt-repo list
```

Команда добавления репозитория в список активных репозиториев:

```
apt-repo add <репозиторий>
```

Команда удаления или выключения репозитория:

```
apt-repo rm <репозиторий>
```

Команда удаления всех репозиториев:

```
apt-repo clean
```

Обновление информации о репозиториях:

```
apt-repo update
```

Вывод справки:

```
man apt-repo
```

или

```
apt-repo –help
```

9.1.1.2 Добавление репозитория на CD/DVD-носителе

Для добавления в `sources.list` репозитория на CD/DVD-носителе информации в APT предусмотрена специальная утилита – `apt-cdrom`. Чтобы добавить запись о репозитории на носителе, достаточно вставить его в привод для чтения (записи) CD (DVD)-носителей информации и выполнить следующую команду:

```
# apt-cdrom add
```

После этого в `sources.list` появится запись о подключенном диске примерно такого вида:

```
rpm cdrom:[ALT Server 9.1 v x86_64 build 2019-07-22]/ ALTLinux main
```

Примечание. В случае если записи для cdrom в файле `/etc/fstab` нет, потребуется примонтировать носитель информации вручную (каталог `/media/ALTLinux` должен существо-

вать):

```
# mount /dev/cdrom /media/ALTLinux
```

Затем использовать команду добавления носителя с дополнительным ключом:

```
# apt-cdrom add -m
```

9.1.1.3 Добавление репозиториев вручную

Для редактирования списка репозиториев можно отредактировать в любом текстовом редакторе файлы из папки `/etc/apt/sources.list.d/`. Для изменения этих файлов необходимы права администратора. В файле `alt.list` может содержаться такая информация:

```
rpm [alt] http://ftp.altlinux.org/pub/distributions/ALTLinux p9/x86_64
classic
rpm [alt] http://ftp.altlinux.org/pub/distributions/ALTLinux
p9/x86_64-i586 classic
rpm [alt] http://ftp.altlinux.org/pub/distributions/ALTLinux p9/noarch
classic
```

По сути, каждая строчка соответствует некому репозиторию. Для выключения репозитория достаточно закомментировать соответствующую строку (дописать символ решётки перед строкой). Для добавления нового репозитория необходимо дописать его вниз этого или любого другого файла.

9.1.2 Обновление информации о репозиториях

Практически любое действие с системой apt начинается с обновления данных от активированных источников. Список источников необходимо обновлять при поиске новой версии пакета, установке пакетов или обновлении установленных пакетов новыми версиями.

Обновление данных осуществляется командой:

```
# apt-get update
```

После выполнения этой команды, apt обновит свой кэш новой информацией.

9.2 Поиск пакетов

Утилита `apt-cache` предназначена для поиска программных пакетов, в репозитории, и позволяет искать не только по имени пакета, но и по его описанию.

Команда `apt-cache search <подстрока>` позволяет найти все пакеты, в именах или описании которых присутствует указанная подстрока. Пример поиска может выглядеть следующим образом:

```
$ apt-cache search ^telegraf
ceph-mgr-telegraf - Telegraf module for Ceph Manager Daemon
telegraf - The plugin-driven server agent for collecting and reporting
```

```
metrics
```

Символ «^» в поисковом выражении, указывает на то, что необходимо найти совпадения только в начале строки (в данном случае – в начале имени пакета).

Для того чтобы подробнее узнать о каждом из найденных пакетов и прочитать его описание, можно воспользоваться командой `apt-cache show`, которая покажет информацию о пакете из репозитория:

```
$ apt-cache show telegraf
Package: telegraf
Section: Development/Other
Installed Size: 72171378
Maintainer: Alexey Gladkov <legion@altlinux.ru>
Version: 1.12.1-alt1:p9+237461.200.1.1@1568252449
Pre-Depends: /bin/sh, /usr/sbin/groupadd, /usr/sbin/useradd, /usr/sbin/usermod, /usr/sbin/post_service, /usr/sbin/preun_service, rpmlib(PayloadIsLzma)
Depends: /bin/sh, /etc/logrotate.d, /etc/rc.d/init.d, /etc/rc.d/init.d(SourceIfNotEmpty), /etc/rc.d/init.d(msg_reloading), /etc/rc.d/init.d(msg_usage), /etc/rc.d/init.d(start_daemon), /etc/rc.d/init.d(status), /etc/rc.d/init.d(stop_daemon), /etc/rc.d/init.d/functions
Provides: telegraf (= 1.12.1-alt1:p9+237461.200.1.1)
Architecture: x86_64
Size: 14742951
MD5Sum: 1f97475f9494e7e14111541d8967cc7d
Filename: telegraf-1.12.1-alt1.x86_64.rpm
Description: The plugin-driven server agent for collecting and reporting metrics
 Telegraf is an agent written in Go for collecting, processing, aggregating, and writing metrics.
```

При поиске с помощью `apt-cache` можно использовать русскую подстроку. В этом случае будут найдены пакеты, имеющие описание на русском языке.

9.3 Установка или обновление пакета

Установка пакета с помощью APT выполняется командой:

```
# apt-get install <имя_пакета>
```

Примечание. Перед установкой и обновлением пакетов необходимо выполнить команду обновления индексов пакетов:

```
# apt-get update
```

Если пакет уже установлен и в подключенном репозитории нет обновлений для данного пакета, система сообщит об уже установленном пакете последней версии. Если в репозитории присутствует более новая версия или новое обновление – программа начнет процесс установки.

`apt-get` позволяет устанавливать в систему пакеты, требующие для работы другие, пока еще не установленные. В этом случае он определяет, какие пакеты необходимо установить, и устанавливает их, пользуясь всеми доступными репозиториями.

Установка пакета `telegraf` командой `apt-get install telegraf` приведет к следующему диалогу с APT:

```
# apt-get install telegraf
Чтение списков пакетов... Завершено
Построение дерева зависимостей... Завершено
Следующие НОВЫЕ пакеты будут установлены:
  telegraf
0 будет обновлено, 1 новых установлено, 0 пакетов будет удалено и 6 не будет обновлено.
Необходимо получить 14,7MB архивов.
После распаковки потребуется дополнительно 72,2MB дискового пространства.
Получено: 1 http://mirror.yandex.ru p9/branch/x86_64/classic telegraf 1.12.1-alt1:p9+237461.200.1.1@1568252449 [14,7MB]
Получено 14,7MB за 5s (2787kB/s).
Совершаем изменения...
Подготовка...
################################# [100%]
Обновление / установка...
1:                                                telegraf-1.12.1-alt1
################################# [100%]
Завершено.
```

Команда `apt-get install <имя_пакета>` используется и для обновления уже установленного пакета или группы пакетов. В этом случае `apt-get` дополнительно проверяет, не обновилась ли версия пакета в репозитории по сравнению с установленным в системе.

При помощи APT можно установить и отдельный бинарный rpm-пакет, не входящий ни в

один из репозиториев. Для этого достаточно выполнить команду `apt-get install путь_к_файлу.rpm`. При этом APT проведет стандартную процедуру проверки зависимостей и конфликтов с уже установленными пакетами.

В результате операций с пакетами без использования APT может нарушиться целостность ОС «Альт Сервер Виртуализации», и `apt-get` в таком случае откажется выполнять операции установки, удаления или обновления.

Для восстановления целостности ОС «Альт Сервер Виртуализации» необходимо повторить операцию, задав опцию `-f`, заставляющую `apt-get` исправить нарушенные зависимости, удалить или заменить конфликтующие пакеты. Любые действия в этом режиме обязательно требуют подтверждения со стороны пользователя.

При установке пакетов происходит запись в системный журнал вида:

```
apt-get: имя-пакета installed
```

9.4 Удаление установленного пакета

Для удаления пакета используется команда `apt-get remove <имя_пакета>`. Удаление пакета с сохранением его файлов настройки производится при помощи следующей команды:

```
# apt-get remove <значимая_часть_имени_пакета>
```

В случае, если при этом необходимо полностью очистить систему от всех компонент удаляемого пакета, то применяется команда:

```
# apt-get remove --purge <значимая_часть_имени_пакета>
```

Для того чтобы не нарушать целостность системы, будут удалены и все пакеты, зависящие от удаляемого.

В случае удаления с помощью `apt-get` базового компонента системы появится запрос на подтверждение операции:

```
# apt-get remove filesystem
Обработка файловых зависимостей... Завершено
Чтение списков пакетов... Завершено
Построение дерева зависимостей... Завершено
Следующие пакеты будут УДАЛЕНЫ:
basesystem filesystem ppp sudo
Внимание: следующие базовые пакеты будут удалены:
В обычных условиях этого не должно было произойти, надеемся, вы точно
представляете, чего требуете!
basesystem filesystem (по причине basesystem)
0 пакетов будет обновлено, 0 будет добавлено новых, 4 будет
```

```
удалено(заменено) и 0 не будет обновлено.
Необходимо получить 0B архивов. После распаковки 588kБ будет
освобождено.
Вы делаете нечто потенциально опасное!
Введите фразу 'Yes, do as I say!' чтобы продолжить.
```

Каждую ситуацию, в которой APT выдает такое сообщение, необходимо рассматривать отдельно. Однако, вероятность того, что после выполнения этой команды система окажется неработоспособной, очень велика.

При удалении пакетов происходит запись в системный журнал вида:

```
apt-get: имя-пакета removed
```

9.5 Обновление всех установленных пакетов

Полное обновление всех установленных в системе пакетов производится при помощи команд:

```
# apt-get update
# apt-get dist-upgrade
```

Первая команда (`apt-get update`) обновит индексы пакетов. Вторая команда (`apt-get dist-upgrade`) позволяет обновить только те установленные пакеты, для которых в репозиториях, перечисленных в `/etc/apt/sources.list`, имеются новые версии.

В случае обновления всего дистрибутива APT проведёт сравнение системы с репозиторием и удалит устаревшие пакеты, установит новые версии присутствующих в системе пакетов, отследит ситуации с переименованиями пакетов или изменения зависимостей между старыми и новыми версиями программ. Всё, что потребуется поставить (или удалить) дополнительно к уже имеющемуся в системе, будет указано в отчете apt-get, которым APT предварит само обновление.

Примечание. Команда `apt-get dist-upgrade` обновит систему, но ядро ОС не будет обновлено.

9.6 Обновление ядра

Для обновления ядра ОС необходимо выполнить команду:

```
# update-kernel
```

Примечание. Если индексы сегодня еще не обновлялись перед выполнением команды update-kernel необходимо выполнить команду `apt-get update`.

Команда `update-kernel` обновляет и модули ядра, если в репозитории обновилось что-то из модулей без обновления ядра.

Новое ядро загрузится только после перезагрузки системы.

10 КОРПОРАТИВНАЯ ИНФРАСТРУКТУРА

10.1 Zabbix

Zabbix – система мониторинга и отслеживания статусов разнообразных сервисов компьютерной сети, серверов и сетевого оборудования.

Для управления системой мониторинга и чтения данных используется веб-интерфейс.

Перед установкой должен быть установлен и запущен сервер PostgreSQL, с созданным пользователем zabbix и созданной базой zabbix.

10.1.1 Установка клиента Zabbix

Установить необходимый пакет:

```
# apt-get install zabbix-agent
```

Добавить Zabbix agent в автозапуск и запустить его:

```
# systemctl enable zabbix_agentd
# systemctl start zabbix_agentd
```

Адрес сервера, которому разрешено обращаться к агенту задается в конфигурационном файле `/etc/zabbix/zabbix_agentd.conf` параметрами:

```
Server=127.0.0.1
ServerActive=127.0.0.1
```

11 ОБЩИЕ ПРИНЦИПЫ РАБОТЫ ОС

Работа с операционной средой заключается во вводе определенных команд (запросов) к операционной среде и получению на них ответов в виде текстового отображения.

Основой операционной среды является операционная система.

Операционная система (ОС) – совокупность программных средств, организующих согласованную работу операционной среды с аппаратными устройствами компьютера (процессор, память, устройства ввода-вывода и т. д.).

Диалог с ОС осуществляется посредством командных интерпретаторов и системных библиотек.

Каждая системная библиотека представляет собой набор программ, динамически вызываемых операционной системой.

Командные интерпретаторы – особый род специализированных программ, позволяющих осуществлять диалог с ОС посредством команд.

Для удобства пользователей при работе с командными интерпретаторами используются интерактивные рабочие среды (далее – ИРС), предоставляющие пользователю удобный интерфейс для работы с ОС.

В самом центре ОС изделия находится управляющая программа, называемая ядром. В ОС изделия используется новейшая модификация «устойчивого» ядра Linux – версия 5.4.

Ядро взаимодействует с компьютером и периферией (дисками, принтерами и т. д.), распределяет ресурсы и выполняет фоновое планирование заданий.

Другими словами, ядро ОС изолирует вас от сложностей аппаратуры компьютера, командный интерпретатор от ядра, а ИРС от командного интерпретатора.

ОС «Альт Сервер Виртуализации» является многопользовательской интегрированной системой. Это значит, что она разработана в расчете на одновременную работу нескольких пользователей.

Пользователь может либо сам работать в системе, выполняя некоторую последовательность команд, либо от его имени могут выполняться прикладные процессы.

Пользователь взаимодействует с системой через командный интерпретатор, который представляет собой, как было сказано выше, прикладную программу, которая принимает от пользователя команды или набор команд и транслирует их в системные вызовы к ядру системы. Интерпретатор позволяет пользователю просматривать файлы, передвигаться по дереву файловой системы, запускать прикладные процессы. Все командные интерпретаторы UNIX имеют развитый командный язык и позволяют писать достаточно сложные программы, упрощающие процесс администрирования системы и работы с ней.

11.1 Процессы функционирования ОС

Все программы, которые выполняются в текущий момент времени, называются процессами. Процессы можно разделить на два основных класса: системные процессы и пользовательские процессы. Системные процессы – программы, решающие внутренние задачи ОС, например, организацию виртуальной памяти на диске или предоставляющие пользователям те или иные сервисы (процессы-службы).

Пользовательские процессы – процессы, запускаемые пользователем из командного интерпретатора для решения задач пользователя или управления системными процессами. Linux изначально разрабатывался как многозадачная система. Он использует технологии, опробованные и отработанные другими реализациями UNIX, которые существовали ранее.

Фоновый режим работы процесса – режим, когда программа может работать без взаимодействия с пользователем. В случае необходимости интерактивной работы с пользователем (в общем случае) процесс будет «остановлен» ядром, и работа его продолжится только после переведения его в «нормальный» режим работы.

11.2 Файловая система ОС

В ОС использована файловая система Linux, которая в отличие от файловых систем DOS и Windows(™) является единым деревом. Корень этого дерева – каталог, называемый root (рут), и обозначаемый «/». Части дерева файловой системы могут физически располагаться в разных разделах разных дисков или вообще на других компьютерах, – для пользователя это прозрачно. Процесс присоединения файловой системы раздела к дереву называется монтированием, удаление – размонтированием. Например, файловая система CD-ROM в изделии монтируется по умолчанию в каталог /media/cdrom (путь в изделии обозначается с использованием «/», а не «\», как в DOS/Windows). Текущий каталог обозначается «./».

Файловая система изделия содержит каталоги первого уровня:

- /bin (командные оболочки (shell), основные утилиты);
- /boot (содержит ядро системы);
- /dev (псевдофайлы устройств, позволяющие работать с ними напрямую);
- /etc (файлы конфигурации);
- /home (личные каталоги пользователей);
- /lib (системные библиотеки, модули ядра);
- /lib64 (64-битные системные библиотеки);
- /media (каталоги для монтирования файловых систем сменных устройств);

- /mnt (каталоги для монтирования файловых систем сменных устройств и внешних файловых систем);
- /proc (файловая система на виртуальном устройстве, ее файлы содержат информацию о текущем состоянии системы);
- /root (личный каталог администратора системы);
- /sbin (системные утилиты);
- /sys (файловая система, содержащая информацию о текущем состоянии системы);
- /usr (программы и библиотеки, доступные пользователю);
- /var (рабочие файлы программ, очереди, журналы);
- /tmp (временные файлы).

11.3 Организация файловой структуры

Система домашних каталогов пользователей помогает организовывать безопасную работу пользователей в многопользовательской системе. Вне своего домашнего каталога пользователь обладает минимальными правами (обычно чтение и выполнение файлов) и не может нанести ущерб системе, например, удалив или изменив файл.

Кроме файлов, созданных пользователем, в его домашнем каталоге обычно содержатся персональные конфигурационные файлы некоторых программ.

Маршрут (путь) – это последовательность имён каталогов, представляющий собой путь в файловой системе к данному файлу, где каждое следующее имя отделяется от предыдущего наклонной чертой (слэшем). Если название маршрута начинается со слэша, то путь в искомый файл начинается от корневого каталога всего дерева системы. В обратном случае, если название маршрута начинается непосредственно с имени файла, то путь к искомому файлу должен начаться от текущего каталога (рабочего каталога).

Имя файла может содержать любые символы за исключением косой черты (/). Однако следует избегать применения в именах файлов большинства знаков препинания и непечатаемых символов. При выборе имен файлов рекомендуется ограничиться следующими символами:

- строчные и ПРОПИСНЫЕ буквы. Следует обратить внимание на то, что регистр всегда имеет значение;
- цифры;
- символ подчеркивания (_);
- точка (.).

Для удобства работы можно использовать точку (.) для отделения имени файла от расширения файла. Данная возможность может быть необходима пользователям или некоторым программам, но не имеет значение для shell.

11.3.1 Иерархическая организация файловой системы

Каталог /:

/boot – место, где хранятся файлы необходимые для загрузки ядра системы;

/lib – здесь располагаются файлы динамических библиотек, необходимых для работы большей части приложений и подгружаемые модули ядра;

/lib64 – здесь располагаются файлы 64-битных динамических библиотек, необходимых для работы большей части приложений;

/bin – минимальный набор программ необходимых для работы в системе;

/sbin – набор программ для административной работы с системой (программы необходимые только суперпользователю);

/home – здесь располагаются домашние каталоги пользователей;

/etc – в данном каталоге обычно хранятся общесистемные конфигурационные файлы для большинства программ в системе;

/etc/rc?.d,/etc/init.d,/etc/rc.boot,/etc/rc.d – директории, где расположены командные файлы системы инициализации SysVinit;

/etc/passwd – база данных пользователей, в которой содержится информация об имени пользователя, его настоящем имени, личном каталоге, закодированный пароль и другие данные;

/etc/shadow – теневая база данных пользователей. При этом информация из файла /etc/passwd перемещается в /etc/shadow, который недоступен по чтению всем, кроме пользователя root. В случае использования альтернативной схемы управления теневыми паролями (TCB) все теневые пароли для каждого пользователя располагаются в директории /etc/tcb/<имя пользователя>/shadow;

/dev – в этом каталоге находятся файлы устройств. Файлы в /dev создаются сервисом udev;

/usr – обычно файловая система /usr достаточно большая по объему, так как все программы установлены именно здесь. Вся информация в каталоге /usr помещается туда во время установки системы. Отдельно устанавливаемые пакеты программ и другие файлы размещаются в каталоге /usr/local. Некоторые подкаталоги системы /usr рассмотрены ниже;

/usr/bin – практически все команды, хотя некоторые находятся в /bin или в /usr/local/bin;

/usr/sbin – команды, используемые при администрировании системы и не предназначенные для размещения в файловой системе root;

/usr/local – здесь рекомендуется размещать файлы, установленные без использования пакетных менеджеров, внутренняя организация каталогов практически такая же, как и корневого каталога;

/usr/man – каталог, где хранятся файлы справочного руководства man;

/usr/share – каталог для размещения общедоступных файлов большей части приложений.

Каталог /var:

/var/log – место, где хранятся файлы аудита работы системы и приложений;

/var/spool – каталог для хранения файлов находящих в очереди на обработку для того или иного процесса (очередь на печать, отправку почты и т. д.);

/tmp – временный каталог необходимый некоторым приложениям;

/proc – файловая система /proc является виртуальной и в действительности она не существует на диске. Ядро создает её в памяти компьютера. Система /proc предоставляет информацию о системе.

11.3.2 Имена дисков и разделов

Все физические устройства вашего компьютера отображаются в каталог /dev файловой системы изделия (об этом – ниже). Диски (в том числе IDE/SATA/SCSI жёсткие диски, USB-диски) имеют имена:

/dev/sda – первый диск;

/dev/sdb – второй диск;

и т. д.

Диски обозначаются /dev/sdX, где X – a,b,c,d,e,... в порядке обнаружения системой.

Раздел диска обозначается числом после его имени. Например, /dev/sdb4 – четвертый раздел второго диска.

11.4 Разделы, необходимые для работы ОС

Для работы ОС необходимо создать на жестком диске (дисках) по крайней мере, два раздела: корневой (то есть тот, который будет содержать каталог /) и раздел подкачки (swap). Размер последнего, как правило, составляет от однократной до двукратной величины оперативной памяти компьютера. Если свободного места на диске много, то можно создать отдельные разделы для каталогов /usr, /home, /var.

11.5 Управление системными сервисами и командами

11.5.1 Сервисы

Сервисы – это программы, которые запускаются и останавливаются через инициализированные скрипты, расположенные в каталоге /etc/init.d. Многие из этих сервисов запускаются на этапе старта ОС «Альт Сервер Виртуализации». В ОС существует шесть системных уровней выполнения, каждый из которых определяет список служб (сервисов), запускаемых на данном уровне. Уровни 0 и 6 соответствуют выключению и перезагрузке системы.

При загрузке системы процесс init запускает все сервисы, указанные в каталоге /etc/rc (0-6).d/ для уровня по умолчанию. Поменять его можно в конфигурационном файле /etc/inittab. Следующая строка соответствует второму уровню выполнения:

```
id:2:initdefault:
```

Для тестирования изменений, внесенных в файл inittab, применяется команда telinit. При указании аргумента -q процесс init повторно читает inittab.

Для перехода ОС «Альт Сервер Виртуализации» на нужный уровень выполнения можно воспользоваться командой init, например:

```
init 3
```

Данная команда переведет систему на третий уровень выполнения, запустив все сервисы, указанные в каталоге /etc/rc3.d/.

11.5.2 Команды

Далее приведены основные команды, использующиеся в ОС «Альт Сервер Виртуализации»:

- ar – создание и работа с библиотечными архивами;
- at – формирование или удаление отложенного задания;
- awk – язык обработки строковых шаблонов;
- batch – планирование команд в очереди загрузки;
- bc – строковый калькулятор;
- chfn – управление информацией учетной записи (имя, описание);
- chsh – управление выбором командного интерпретатора (по умолчанию – для учетной записи);
- cut – разбивка файла на секции, задаваемые контекстными разделителями;
- df – вывод отчета об использовании дискового пространства;
- dmesg – вывод содержимого системного буфера сообщений;
- du – вычисление количества использованного пространства элементов ФС;
- echo – вывод содержимого аргументов на стандартный вывод;
- egrep – поиск в файлах содержимого согласно регулярным выражениям;
- fgrep – поиск в файлах содержимого согласно фиксированным шаблонам;
- file – определение типа файла;
- find – поиск файла по различным признакам в иерархии каталогов;
- gettext – получение строки интернационализации из каталогов перевода;
- grep – вывод строки, содержащей шаблон поиска;
- groupadd – создание новой учетной записи группы;

- groupdel – удаление учетной записи группы;
- groupmod – изменение учетной записи группы;
- groups – вывод списка групп;
- gunzip – распаковка файла;
- gzip – упаковка файла;
- hostname – вывод и задание имени хоста;
- install – копирование файла с установкой атрибутов;
- ipcrm – удаление ресурса IPC;
- ipcs – вывод характеристик ресурса IPC;
- kill – прекращение выполнения процесса;
- killall – удаление процессов по имени;
- lpr – система печати;
- ls – вывод содержимого каталога;
- lsb_release – вывод информации о дистрибутиве;
- m4 – запуск макропроцессора;
- md5sum – генерация и проверка MD5-сообщения;
- mknod – создание файла специального типа;
- mktemp – генерация уникального имени файла;
- more – постраничный вывод содержимого файла;
- mount – монтирование ФС;
- msgfmt – создание объектного файла сообщений из файла сообщений;
- newgrp – смена идентификатора группы;
- nice – изменение приоритета процесса перед его запуском;
- nohup – работа процесса после выхода из системы;
- od – вывод содержимого файла в восьмеричном и других видах;
- passwd – смена пароля учетной записи;
- patch – применение файла описания изменений к оригинальному файлу;
- pidof – вывод идентификатора процесса по его имени;
- ps – вывод информации о процессах;
- renice – изменение уровня приоритета процесса;
- sed – строковый редактор;
- sendmail – транспорт системы электронных сообщений;
- sh – командный интерпретатор;

- shutdown – команда останова системы;
- su – изменение идентификатора запускаемого процесса;
- sync – сброс системных буферов на носители;
- tar – файловый архиватор;
- umount – размонтирование ФС;
- useradd – создание новой учетной записи или обновление существующей;
- userdel – удаление учетной записи и соответствующих файлов окружения;
- usermod – модификация информации об учетной записи;
- w – список пользователей, кто в настоящий момент работает в системе и с какими файлами;
- who – вывод списка пользователей системы.

Узнать об опциях команд можно с помощью команды man.

12 РАБОТА С НАИБОЛЕЕ ЧАСТО ИСПОЛЬЗУЕМЫМИ КОМПОНЕНТАМИ

12.1 Командные оболочки (интерпретаторы)

Для управления ОС используется командные интерпретаторы (shell).

Зайдя в систему, можно увидеть приглашение – строку, содержащую символ «$» (далее, этот символ будет обозначать командную строку). Программа ожидает ввода команд. Роль командного интерпретатора – передавать команды пользователя операционной системе. При помощи командных интерпретаторов можно писать небольшие программы – сценарии (скрипты). В Linux доступны следующие командные оболочки:

bash – самая распространённая оболочка под linux. Она ведет историю команд и предоставляет возможность их редактирования.

pdksh – клон korn shell, хорошо известной оболочки в UNIX(™) системах.

Оболочкой по умолчанию является «Bash» (Bourne Again Shell) Проверить, какая оболочка используется, можно выполнив команду:

```
$ echo $SHELL
```

У каждой оболочки свой синтаксис. Все примеры в дальнейшем построены с использованием оболочки Bash.

12.1.1 Командная оболочка Bash

В Bash имеется несколько приемов для работы со строкой команд. Например, используя клавиатуру, можно:

<Ctrl> + <A> – перейти на начало строки;

<Ctrl> + <U> – удалить текущую строку;

<Ctrl> + <C> – остановить текущую задачу.

Для ввода нескольких команд одной строкой можно использовать разделитель «;». По истории команд можно перемещаться с помощью клавиш <↑> и <↓>. Чтобы найти конкретную команду в списке набранных, не пролистывая всю историю, необходимо набрать <Ctrl> + <R> и начать вводить символы ранее введенной команды.

Для просмотра истории команд можно воспользоваться командой history. Команды, присутствующие в истории, отображаются в списке пронумерованными. Чтобы запустить конкретную команду необходимо набрать:

```
!номер команды
```

Если ввести:

```
!!
```

запустится последняя, из набранных команд.

В Bash имеется возможность самостоятельного завершения имен команд из общего списка команд, что облегчает работу при вводе команд, в случае, если имена программ и команд слишком длинны. При нажатии клавиши <Tab> Bash завершает имя команды, программы или каталога, если не существует нескольких альтернативных вариантов. Например, чтобы использовать программу декомпрессии bunzip2, можно набрать следующую команду:

```
$ bu
```

Затем нажать <Tab>. Так как в данном случае существует несколько возможных вариантов завершения команды, то необходимо повторно нажать клавишу <Tab>, чтобы получить список имен, начинающихся с bu.

В предложенном примере можно получить следующий список:

```
$ bu
buildhash builtin bunzip2
```

Если набрать: n (bunzip – это единственное имя, третьей буквой которого является «n»), а затем нажать клавишу <Tab>, то оболочка самостоятельно дополнит имя. Чтобы запустить команду нужно нажать <Enter>.

Программы, вызываемые из командной строки, Bash ищет в каталогах, определяемых в системной переменной PATH. По умолчанию в этот перечень каталогов не входит текущий каталог, обозначаемый ./ (точка слеш) (если только не выбран один из двух самых слабых уровней защиты). Поэтому, для запуска программы из текущего каталога, необходимо использовать команду (в примере запускается команда prog):

```
./prog
```

12.1.2 Базовые команды оболочки Bash

Все команды, приведенные ниже, могут быть запущены в режиме консоли. Для получения более подробной информации следует использовать команду `man`. Пример:

```
$ man ls
```

12.1.2.1 Учетные записи пользователей

Команда su

Команда su позволяет получить права администратора. При вводе команды su, будет запрошен пароль суперпользователя (root). И в случае ввода корректного пароля, оператор получит привилегии суперпользователя. Чтобы вернуться к правам оператора, необходимо ввести команду:

```
# exit
```

Команда id

Команда id выводит информацию о пользователе и группах, в которых он состоит для заданного пользователя или о текущем пользователе (если ничего не указано).

Синтаксис:

```
id [параметры] [ПОЛЬЗОВАТЕЛЬ]
```

Команда passwd

Команда passwd меняет (или устанавливает) пароль, связанный с входным_именем пользователя.

Обычный пользователь может менять только пароль, связанный с его собственным входным_именем.

Команда запрашивает у обычных пользователей старый пароль (если он был), а затем дважды запрашивает новый. Новый пароль должен соответствовать техническим требованиям к паролям, заданным администратором системы.

12.1.2.2 Основные операции с файлами и каталогами

Команда ls

Команда ls (list) выдает список файлов каталога.

Синтаксис:

```
ls [-CFRacdilqrtu1] [[-H] | [-L]] [-fgmnopsx] [файл...]
```

Основные опции:

-a – просмотр всех файлов, включая скрытые;

-l – отображение более подробной информации;

-R – выводить рекурсивно информацию о подкаталогах.

Команда cd

Команда cd предназначена для смены каталога. Команда работает как с абсолютными, так и с относительными путями. Если каталог не указан, используется значение переменной окружения HOME (домашний каталог пользователя). Если каталог задан полным маршрутным именем, он становится текущим. По отношению к новому каталогу нужно иметь право на выполнение, которое в данном случае трактуется как разрешение на поиск.

Синтаксис:

```
cd [-L|-P] [каталог]
```

Если в качестве аргумента задано -, то это эквивалентно $OLDPWD. Если переход был осуществлен по переменной окружения CDPATH или в качестве аргумента был задан - и смена каталога была успешной, то абсолютный путь нового рабочего каталога будет выведен на стандартный вывод.

Пример. Находясь в домашнем каталоге перейти в его подкаталог docs/ (относительный путь):

```
$ cd docs/
```

Сделать текущим каталог /usr/bin (абсолютный путь):

```
$ cd /usr/bin/
```

Сделать текущим родительский каталог:

```
$ cd ..
```

Вернуться в предыдущий каталог:

```
$ cd -
```

Сделать текущим домашний каталог:

```
$ cd
```

Команда pwd

Команда pwd выводит абсолютный путь текущего (рабочего) каталога.

Синтаксис:

```
pwd [-L|-P]
```

Опции:

-P – не выводить символические ссылки;

-L – выводить символические ссылки.

Команда rm

Команда rm используется для удаления файлов.

Синтаксис:

```
rm [-fiRr] имя_файла
```

Основные опции:

-f – не запрашивать подтверждения;

-i – запрашивать подтверждение;

-r, -R – рекурсивно удалять содержимое указанных каталогов.

Пример. Удалить все файлы html в каталоге ~/html:

```
$ rm -i ~/html/*.html
```

Команда mkdir

Команда mkdir позволяет создать каталог.

Синтаксис:

```
mkdir [-p] [-m права] [каталог...]
```

Команда rmdir

Команда rmdir удаляет записи, соответствующие указанным пустым каталогам.

Синтаксис:

```
rmdir [-p] [каталог...]
```

Команда rmdir часто заменяется командой rm -rf, которая позволяет удалять каталоги, даже если они не пусты.

Команда cp

Команда cp предназначена для копирования файлов.

Синтаксис:

`cp [-fip] [исх_файл] [цел_файл]`

`cp [-fip] [исх_файл...] [каталог]`

`cp [-R] [[-H] | [-L] | [-P]] [-fip] [исх_файл...] [каталог]`

Основные опции:

-p – сохранять по возможности времена изменения и доступа к файлу, владельца и группу, права доступа;

-i – запрашивать подтверждение перед копированием в существующие файлы;

-r, -R – рекурсивно копировать содержимое каталогов.

Команда mv

Команда mv предназначена для перемещения файлов.

Синтаксис:

`mv [-fi] [исх_файл...] [цел_файл]`

`mv [-fi] [исх_файл...] [каталог]`

В первой синтаксической форме, характеризующейся тем, что последний операнд не является ни каталогом, ни символической ссылкой на каталог, mv перемещает исх_файл в цел_файл.

Во второй синтаксической форме mv перемещает исходные файлы в указанный каталог под именами, совпадающими с краткими именами исходных файлов.

Основные опции:

-f – не запрашивать подтверждения перезаписи существующих файлов;

-i – запрашивать подтверждение перезаписи существующих файлов.

Команда cat

Команда cat последовательно выводит содержимое файлов.

Синтаксис:

`cat [параметры] [файл...]`

Основные опции:

-n, --number – нумеровать все строки при выводе;

-E, --show-ends – показывать $ в конце каждой строки.

Если файл не указан, читается стандартный ввод. Если в списке файлов присутствует имя -, вместо этого файла читается стандартный ввод.

Команда less

Команда less позволяет постранично просматривать текст (для выхода необходимо нажать <q>).

Синтаксис:

`less имя_файла`

Команда grep

Команда grep имеет много опций и предоставляет возможности поиска символьной строки в файле.

Синтаксис:

```
grep шаблон_поиска файл
```

Команда chmod

Команда chmod изменяет права доступа к файлу.

Синтаксис:

```
chmod ОПЦИЯ]... РЕЖИМ[,РЕЖИМ]... [Файл...]
chmod ОПЦИЯ]... --reference=ИФАЙЛ ФАЙЛ...
```

Основные опции:

-R – рекурсивно изменять режим доступа к файлам, расположенным в указанных каталогах;

--reference=ИФАЙЛ – использовать режим файла ИФАЙЛ.

Команда chmod изменяет права доступа каждого указанного файла в соответствии с правами доступа, указанными в параметре режим, который может быть представлен как в символьном виде, так и в виде восьмеричного, представляющего битовую маску новых прав доступа.

Формат символьного режима следующий:

```
[ugoa...][[+-=][разрешения...]...]
```

Здесь разрешения – это ноль или более букв из набора «rwxXst» или одна из букв из набора «ugo».

Каждый аргумент – это список символьных команд изменения прав доступа, разделены запятыми. Каждая такая команда начинается с нуля или более букв «ugoa», комбинация которых указывает, чьи права доступа к файлу будут изменены: пользователя, владеющего файлом (u), пользователей, входящих в группу, к которой принадлежит файл (g), остальных пользователей (o) или всех пользователей (a). Если не задана ни одна буква, то автоматически будет использована буква «a», но биты, установленные в umask, не будут затронуты.

Оператор «+» добавляет выбранные права доступа к уже имеющимся у каждого файла, «-» удаляет эти права, «=» присваивает только эти права каждому указанному файлу.

Буквы «rwxXst» задают биты доступа для пользователей: «r» – чтение, «w» – запись, «x» – выполнение (или поиск для каталогов), «X» – выполнение/поиск, только если это каталог или же файл с уже установленным битом выполнения, «s» – задать ID пользователя и группы при выполнении, «t» – запрет удаления.

Примеры. Позволить всем выполнять файл f2:

```
$ chmod +x f2
```

Запретить удаление файла f3:

```
$ chmod+t f3
```

Команда chown

Команда chown изменяет владельца и/или группу для каждого заданного файла.

Синтаксис:

```
chown [КЛЮЧ]…[ВЛАДЕЛЕЦ][:[ГРУППА]] ФАЙЛ ...
```

Изменить владельца может только владелец файла или суперпользователь. Владелец не изменяется, если он не задан в аргументе. Группа также не изменяется, если не задана, но если после символьного ВЛАДЕЛЬЦА стоит символ «:», подразумевается изменение группы на основную группу текущего пользователя. Поля ВЛАДЕЛЕЦ и ГРУППА могут быть как числовыми, так и символьными.

Примеры. Поменять владельца /u на пользователя test:

```
$ chown test /u
```

Поменять владельца и группу /u:

```
$ chown test:staff /u
```

Поменять владельца /u и вложенных файлов на test:

```
$ chown -hR test /u
```

12.1.2.3 Поиск файлов

Команда find

Команда find предназначена для поиска всех файлов, начиная с корневой директории. Поиск может осуществляться по имени, типу или владельцу файла.

Синтаксис:

```
find [-H] [-L] [-P] [-Оуровень] [-D
help|tree|search|stat|rates|opt|exec] [путь…] [выражение]
```

Ключи для поиска:

-name – поиск по имени файла;

-type – поиск по типу f=файл, d=каталог, l=ссылка(lnk);

-user – поиск по владельцу (имя или UID).

Когда выполняется команда find, можно выполнять различные действия над найденными файлами. Основные действия:

-exec команда \; – выполнить команду. Запись команды должна заканчиваться экранированной точкой с запятой. Строка «{}» заменяется текущим маршрутным именем файла;

-execdir команда \; – то же самое что и exec, но команда вызывается из подкаталога, содержащего текущий файл;

-ok команда – эквивалентно -exec за исключением того, что перед выполнением команды запрашивается подтверждение (в виде сгенерированной командной строки со знаком вопроса в конце) и она выполняется только при ответе: y;

-print – вывод имени файла на экран.

Путем по умолчанию является текущий подкаталог. Выражение по умолчанию -print.

Примеры. Найти в текущем каталоге обычные файлы (не каталоги), имя которых начинается с символа «~»:

```
$ find . -type f -name "~*" -print
```

Найти в текущем каталоге файлы, измененные позже, чем файл file.bak:

```
$ find . -newer file.bak -type f -print
```

Удалить все файлы с именами a.out или *.o, доступ к которым не производился в течение недели:

```
$ find / \( -name a.out -o -name '*.o' \) \ -atime +7 -exec rm {} \;
```

Удалить из текущего каталога и его подкаталогов все файлы нулевого размера, запрашивая подтверждение:

```
$ find . -size 0c -ok rm {} \;
```

Команда whereis

Команда whereis сообщает путь к исполняемому файлу программы, ее исходным файлам (если есть) и соответствующим страницам справочного руководства.

Синтаксис:

```
whereis [options] <name>
```

Опции:

-b – вывод информации только об исполняемых файлах;

-m – вывод информации только о страницах справочного руководства;

-s – вывод информации только об исходных файлах.

12.1.2.4 Мониторинг и управление процессами

Команда ps

Команда ps отображает список текущих процессов.

Синтаксис:

```
ps [-aA] [-defl] [-G список] [-o формат...] [-p список] [-t список] [-
U список] [-g список] [-n список] [-u список]
```

По умолчанию выводится информация о процессах с теми же действующим UID и управляющим терминалом, что и у подающего команду пользователя.

Основные опции:

-a – вывести информацию о процессах, ассоциированных с терминалами;

-f – вывести «полный» список;

-l – вывести «длинный» список;

-p список – вывести информацию о процессах с перечисленными в списке PID;

-u список – вывести информацию о процессах с перечисленными идентификаторами или именами пользователей.

Команда kill

Команда kill позволяет прекратить исполнение процесса или передать ему сигнал.

Синтаксис:

```
kill [-s] [сигнал] [идентификатор] [...]
kill [-l] [статус_завершения]
kill [-номер_сигнала] [идентификатор] [...]
```

Идентификатор – PID ведущего процесса задания или номер задания, предварённый знаком «%».

Основные опции:

-l – вывести список поддерживаемых сигналов;

-s сигнал, -сигнал – послать сигнал с указанным именем.

Если обычная команда kill не дает желательного эффекта, необходимо использовать команду kill с параметром -9:

```
$ kill -9 PID_номер
```

Команда df

Команда df показывает количество доступного дискового пространства в файловой системе, в которой содержится файл, переданный как аргумент. Если ни один файл не указан, показывается доступное место на всех смонтированных файловых системах. Размеры по умолчанию указаны в блоках по 1КБ по умолчанию.

Синтаксис:

```
df [опция]... [файл]...
```

Основные опции:

-total – подсчитать общий объем в конце;

-h, --human-readable – печатать размеры в удобочитаемом формате (например, 1K 234M 2G);

-h, --human-readable – печатать размеры в удобочитаемом формате (например, 1K 234M 2G).

Команда du

Команда du подсчитывает использование диска каждым файлом, для каталогов подсчет происходит рекурсивно.

Синтаксис:

```
du [опции][файл...]
```

Основные опции:

-a, --all – выводить общую сумму для каждого заданного файла, а не только для каталогов;

-c, --total – подсчитать общий объем в конце. Может быть использовано для выяснения суммарного использования дискового пространства для всего списка заданных файлов;

-d, --max-depth=N – выводить объем для каталога (или файлов, если указано --all) только если она на N или менее уровней ниже аргументов командной строки;

-S, --separate-dirs – выдавать отдельно размер каждого каталога, не включая размеры подкаталогов;

-s, --summarize – отобразить только сумму для каждого аргумента.

Команда which

Команда which – отображает полный путь к указанным командам или сценариям.

Синтаксис:

```
which [опции] [--] имя_программы [...]
```

Основные опции:

-a, --all – выводит все совпавшие исполняемые файлы по содержимому в переменной окружения PATH, а не только первый из них;

-c, --total – подсчитать общий объем в конце. Может быть использовано для выяснения суммарного использования дискового пространства для всего списка заданных файлов;

-d, --max-depth=N – выводить объем для каталога (или файлов, если указано --all) только если она на N или менее уровней ниже аргументов командной строки;

-S, --separate-dirs – выдавать отдельно размер каждого каталога, не включая размеры подкаталогов;

--skip-dot – пропускает все каталоги из переменной окружения PATH, которые начинаются с точки.

12.1.2.5 Использование многозадачности

ОС «Альт Сервер Виртуализации» – многозадачная система.

Для того чтобы запустить программу в фоновом режиме необходимо набрать «&» после имени программы. После этого оболочка дает возможность запускать другие приложения.

Так как некоторые программы интерактивны – их запуск в фоновом режиме бессмысленен. Подобные программы просто остановятся, если их запустить в фоновом режиме.

Можно также запускать нескольких независимых сеансов. Для этого в консоли необходимо набрать <Alt> и одну из клавиш, находящихся в интервале от <F1> до <F6>. На экране появится новое приглашение системы, и можно открыть новый сеанс.

Команда bg

Команда bg используется для того, чтобы перевести задание на задний план.

Синтаксис:

```
bg [идентификатор ...]
```

Идентификатор – PID ведущего процесса задания или номер задания, предварённый знаком «%».

Команда fg

Команда fg позволяет перевести задание на передний план.

Синтаксис:

```
fg [идентификатор ...]
```

Идентификатор – PID ведущего процесса задания или номер задания, предварённый знаком «%».

12.1.2.6 Сжатие и упаковка файлов

Команда tar

Сжатие и упаковка файлов выполняется с помощью команды tar, которая преобразует файл или группу файлов в архив без сжатия (tarfile).

Упаковка файлов в архив чаще всего выполняется следующей командой:

```
$ tar -cf [имя создаваемого файла архива] [упаковываемые файлы и/или директории]
```

Пример использования команды упаковки архива:

```
$ tar -cf moi_dokumenti.tar Docs project.tex
```

Распаковка содержимого архива в текущий каталог выполняется командой:

```
$ tar -xf [имя файла архива]
```

Для сжатия файлов используются специальные программы сжатия: gzip, bzip2 и 7z.

12.2 Стыкование команд в системе

12.2.1 Стандартный ввод и стандартный вывод

Многие команды системы имеют так называемые стандартный ввод (standard input) и стандартный вывод (standard output), часто сокращаемые до stdin и stdout. Ввод и вывод здесь – это входная и выходная информация для данной команды. Программная оболочка делает так, что стандартным вводом является клавиатура, а стандартным выводом – экран монитора.

Пример с использованием команды cat. По умолчанию команда cat читает данные из всех файлов, которые указаны в командной строке, и посылает эту информацию непосредственно в стандартный вывод (stdout). Следовательно, команда:

```
$ cat history-final masters-thesis
```

выведет на экран сначала содержимое файла history-final, а затем – файла masters-thesis.

Если имя файла не указано, программа cat читает входные данные из stdin и возвращает их в stdout. Пример:

```
$ cat
Hello there.
```

```
Hello there.
Bye.
Bye.
<Ctrl>-<D>
```

Каждую строку, вводимую с клавиатуры, программа cat немедленно возвращает на экран. При вводе информации со стандартного ввода конец текста сигнализируется вводом специальной комбинации клавиш, как правило, <Ctrl>-<D>. Сокращённое название сигнала конца текста – EOT (end of text).

12.2.2 Перенаправление ввода и вывода

При необходимости можно перенаправить стандартный вывод, используя символ > и стандартный ввод, используя символ <.

Фильтр (filter) – программа, которая читает данные из стандартного ввода, некоторым образом их обрабатывает и результат направляет на стандартный вывод. Когда применяется перенаправление, в качестве стандартного ввода и вывода могут выступать файлы. Как указывалось выше, по умолчанию, stdin и stdout относятся к клавиатуре и к экрану соответственно. Программа sort является простым фильтром – она сортирует входные данные и посылает результат на стандартный вывод. Совсем простым фильтром является программа cat – она ничего не делает с входными данными, а просто пересылает их на выход.

12.2.3 Использование состыкованных команд

Стыковку команд (pipelines) осуществляет командная оболочка, которая stdout первой команды направляет на stdin второй команды. Для стыковки используется символ |. Направить stdout команды ls на stdin команды sort:

```
$ ls | sort -r
notes
masters-thesis
history-final
english-list
```

Вывод списка файлов частями:

```
$ ls /usr/bin | more
```

Пример стыкования нескольких команд. Команда head – является фильтром следующего свойства: она выводит первые строки из входного потока (в примере на вход будет подан выход от нескольких состыкованных команд). Если необходимо вывести на экран последнее по алфавиту имя файла в текущем каталоге, можно использовать следующую команду:

```
$ ls | sort -r | head -1 notes
```

где команда head -1 выводит на экран первую строку получаемого ей входного потока строк (в примере поток состоит из данных от команды ls), отсортированных в обратном алфавитном порядке.

12.2.4 Не деструктивное перенаправление вывода

Эффект от использования символа > для перенаправления вывода файла является деструктивным; то есть, команда

```
$ ls > file-list
```

уничтожит содержимое файла file-list, если этот файл ранее существовал, и создаст на его месте новый файл. Если вместо этого перенаправление будет сделано с помощью символов >>, то вывод будет приписан в конец указанного файла, при этом исходное содержимое файла не будет уничтожено.

Примечание. Перенаправление ввода и вывода и стыкование команд осуществляется командными оболочками, которые поддерживают использование символов >, >> и |. Сами команды не способны воспринимать и интерпретировать эти символы.

13 ОБЩИЕ ПРАВИЛА ЭКСПЛУАТАЦИИ

13.1 Включение компьютера

Для включения компьютера необходимо:

- включить стабилизатор напряжения, если компьютер подключен через стабилизатор напряжения;
- включить принтер, если он нужен;
- включить монитор компьютера, если он не подключен к системному блоку кабелем питания;
- включить компьютер (переключателем на корпусе компьютера либо клавишей с клавиатуры).

После этого на экране компьютера появятся сообщения о ходе работы программ проверки и начальной загрузки компьютера.

13.2 Выключение компьютера

Для выключения компьютера надо:

- закончить работающие программы;
- выбрать функцию завершения работы и выключения компьютера, после чего ОС самостоятельно выключит компьютер, имеющий системный блок формата ATX;
- выключить компьютер (переключателем на корпусе AT системного блока);
- выключить принтер;
- выключить монитор компьютера (если питание монитора не от системного блока);
- выключить стабилизатор, если компьютер подключен через стабилизатор напряжения.