

Кластер высокой готовности

Руководство оператора

Версия 2.0^{бета}

для Astra Linux Common Edition

Лаборатория 50

Кластер высокой готовности: Руководство оператора

Настоящий документ является руководством по использованию программного комплекса «Кластер высокой готовности», а также необходимых для его функционирования компонентов Astra Linux Common Edition. Данный документ не покрывает вопросы инсталляции.

© ООО «Лаборатория 50», 2014-2018

Дата публикации: 21 сентября 2018 г.

Содержание

Принятые обозначения	4
1 Назначение	5
1.1 Кластер высокой готовности	5
1.2 Система управления	5
2 Возможности	6
2.1 Кластер высокой готовности	6
2.2 Система управления	6
3 Основные понятия	7
4 Основы использования	10
4.1 Подключение к системе управления	10
4.1.1 Панель управления	10
4.1.2 Командная строка	10
4.2 Панель управления	10
4.3 Интерфейс командой строки	11
5 Узлы	12
5.1 Просмотр и изменение параметров	13
5.2 Включение и выключение	13
5.3 Обслуживание узла	14
5.4 Веб-терминал	14
6 Ресурсы	15
6.1 Свойства ресурсов	15
6.1.1 Типы ресурсов	15
6.1.2 Состояние ресурса	18
6.1.3 Режимы работы	18
6.1.4 Ресурсные агенты	19
6.1.5 Параметры агента	19
6.1.6 Операции	20
6.2 Контроль ресурсов	21
6.3 Изменение ресурса	22
6.4 Запуск и остановка ресурса	22
6.5 Сброс состояния ресурса	23
6.6 Удаление ресурсов	23
6.7 Приоритетные и запретные узлы	23
6.8 Группы	25
7 Виртуализация	26
7.1 Виртуальные машины	26
7.1.1 Свойства виртуальных машин	26
7.1.2 Загрузка образа	27
7.1.3 Создание	27
7.1.4 Подключение к рабочему столу	29

Принятые обозначения

В данном руководстве принят ряд обозначений.

В документе могут встречаться следующие типы замечаний:

Примечание

Примечание или полезная подсказка.

Важно

Существенные нюансы, которые вы должны принять во внимание, прежде чем начнёте какое-либо действие.

Предупреждение

Важная информация о проблемах безопасности или риске потери данных.

В документе используются два варианта приглашения командного интерпретатора.

\$ команда

Команды, начинающиеся с приглашения \$, может выполнять любой пользователь, в том числе root.

команда

Команды, начинающиеся с приглашения #, должен выполнять пользователь root. Но вы можете использовать утилиту **sudo** чтобы запускать их под другой учётной записью.

1 Назначение

1.1 Кластер высокой готовности

Кластер высокой готовности предназначен для обеспечения непрерывной доступности прикладных приложений и сервисов. При отказе любого сервиса или узла целиком кластер автоматически перезапустит сервис на работающем узле с минимальным прерыванием для клиентов кластера.

Программный комплекс «Кластер высокой готовности» предназначен для создания аппаратно-программных кластеров и включает следующие компоненты:

- система управления;
- кластерная файловая система (опционально);
- защищённая виртуализация (опционально).

1.2 Система управления

Система управления кластером определяет отказы программных или аппаратных средств и немедленно осуществляет перезапуск приложения или систем без вмешательства оператора. Такой процесс называется преодолением отказа. В процессе преодоления отказа кластер может производить подготовку узла для запуска на нем приложения. Например, подключать общее хранилище, настраивать сетевые интерфейсы, запускать требуемые вспомогательные службы.

В систему управления входит:

- 1) служба управления кластером;
- 2) веб-панель управления;
- 3) служба доступа к узлам веб-терминал.

2 Возможности

2.1 Кластер высокой готовности

Основные возможности ПК «Кластер высокой готовности»:

- обнаружение и восстановление после отказов аппаратного и программного обеспечения;
- работа без общего хранилища;
- управление практически любыми ресурсами;
- защита разделяемых данных от повреждения (фенсинг);
- поддержка до физических 16 узлов к кластеру;
- поддержка кворума;
- широкой диапазон схем избыточности ресурсов;
- задание различных вариантов зависимостей между ресурсами;
- управление кластером из веб-панели или командной строки;
- автоматизированная установка и настройка.

2.2 Система управления

Система управления кластером предоставляет пользователям практичный и отказоустойчивый механизм управления кластером высокой готовности. Решения в основе системы управления направлены на минимизацию вероятности ошибок пользователя.

Основные возможности:

- графический веб-интерфейс (панель управления);
- интерфейс командной строки;
- ролевая модель доступа;
- HTTP REST-интерфейс;
- управление несколькими кластерами;
- управление защищённой виртуализацией;
- управление кластерной файловой системой.

Возможности панели управления:

- контроль и управление узлами и ресурсами кластера высокой готовности;
- поддержка PAM и Kerberos аутентификации;
- работа на узлах кластера или внешнем сервере;
- веб-терминал доступа к узлам кластера.

Возможности интерфейса командной строки:

- использование на узле с любого сервера;
- управление всеми подсистемами ПО Кластер высокой готовности.

3 Основные понятия

Кластер высокой готовности — программно-аппаратный комплекс, направленный на обеспечение повышенной готовности. Коэффициент готовности — комплексный показатель надёжности, характеризующий вероятность того, что объект окажется в работоспособном состоянии в произвольный момент времени, кроме планируемых периодов, в течение которых применение объекта по назначению не предусматривается.

Готовность и отказоустойчивость достигается за счёт избыточности аппаратных средств и архитектуры программного обеспечения, которые обеспечивают непрерывное обслуживание при отказе компонентов системы.

Выбранный Координатор

Архитектура ПО «Кластер высокой готовности» построена без единой точки отказа. Все узлы являются равноправными. Конфигурация и состояние кластера хранятся в базе данных кластера (БД), непрерывно синхронизируемой между узлами.

Для выполнения действий, требующих централизованного управления среди узлов избирается Выбранный Координатор (Designated Controller — DC). Этот узел может быть выбран в любое время. Если произойдёт отказ аппаратного или программного обеспечения, кластер в кратчайшее время выберёт нового координатора.

Расщепление кластера

Важнейшей особенностью ПО «Кластер высокой готовности» является способность обрабатывать состояние расщепления (split-brain), которое возникает при потере связи между частями кластера. В такой ситуации любой узел в кластере может ошибочно решить, что произошёл отказ другого узла и попытается запустить сервисы, которые по-прежнему работают. Наличие дублирующих экземпляров сервисов может привести к повреждению данных в общем хранилище.

Кворум

Кворум служит основным механизмом для предотвращения порчи данных в случае расщепления кластера.

Кворумом называется состояние кластера, когда более половины узлов работают и связаны друг с другом. Например, в кластере из 5 узлов кворум будет составлять 3 узла. Если кластер с 6 узлами расщеплён на два раздела по 3 узла, ни один из разделов не будет иметь кворума. Поэтому важно, чтобы в кластере высокой готовности было нечётное число узлов. Стандартное поведение раздела без кворума — останов.

Фенсинг

Фенсинг — это процесс защиты разделяемых ресурсов или изоляции сбойного узла кластера. При наличии разделяемых ресурсов, доступ к которым производится с нескольких узлов, необходимо обеспечить защиту системы от неисправного узла или программ-

ного обеспечения. Фенсинг может отключить узел, либо запретить доступ к разделяемому ресурсу для обеспечения целостности данных. Наряду с кворумом фенсинг является другим механизмом предотвращения потери или порчи данных в следствие сбоев на узлах или программном обеспечении.

Ресурс

С точки зрения пользователя, ресурсы – центральный элемент кластера высокой готовности. Ресурсом может быть программа, служба, внешнее или внутреннее хранилище данных, сетевой интерфейс или любой другой объект. К примеру, ресурсами могут быть веб-сервер и плавающий сетевой адрес, по которому осуществляется доступ к сайту.

Кластер высокой готовности поддерживает максимальную доступность ресурсов. ПК «Кластер высокой готовности» позволяет реализовывать различные схемы реализации отказоустойчивости, в зависимости от возможностей самого ресурса.

Управление ресурсами происходит автоматически, при этом могут учитываться потребляемые серверные ресурсы (память и процессоры). Пользователь может совершать *действия* над ресурсами: запускать, останавливать, и т.д.

Ресурсный агент

Для работы с множеством типов ресурсов предусмотрен инвентарь ресурсных агентов. Агент служит медиатором между системой управления кластера и объектом прикладного ресурса. Агенты предоставляют кластеру унифицированный интерфейс управления прикладным ресурсом.

Большинство агентов предоставляют дополнительные параметры для гибкой настройки самого прикладного ресурса или алгоритма автоматического управления.

Ресурсный агент позволяет добавлять новые виды ресурсов, не модифицируя при этом ПО кластера или ПО прикладного ресурса.

Виртуальная машина

Виртуальная машина – это один из видов ресурсов кластера высокой готовности, позволяющий запускать гостевые операционные системы.

Системный диск, с которого загружается операционная система называется *образом виртуальной машины*, или просто *образом*.

Образы располагаются в *хранилищах*.

Хранилище

Хранилище – это динамически распределяемый объем пространства хранения для использования системой защищённой виртуализации. Хранилище представляет собой совокупность однородных объектов – томов хранения.

В хранилище могут быть размещены образы дисков гостевых операционных систем, компакт-дисков, дискет и т.д.

Физически хранилище может быть каталогом на диске, удалённой файловой системой, SCSI-устройством, LVM-группой, системой хранения данных.

Образ

Образ диска — файл, содержащий в себе полную копию содержания и структуры файловой системы и данных, находящихся на диске, таком как компакт-диск, дискета, раздел жёсткого диска или весь жёсткий диск целиком.

Образ виртуальной машины — образ диска, содержащий необходимые загрузочный и системный разделы для загрузки операционной системы в виртуальной машине.

Том

Том хранилища — элемент хранилища данных, к которому может получить доступ объекты системы виртуализации.

Том (Гластера) — логический элемент хранения данных в кластерной файловой системе Гластер. Гластер предоставляет различные интерфейсы для доступа к данным, хранящимся томом. Чаще всего тома используются как файловые системы, подключаемые к серверу по протоколам GlusterFS, NFS или CIFS.

Том Гластера состоит из совокупности физических блоков хранения — *кирпичей*. В зависимости от логической конфигурации кирпичей получаются тома с зеркалированием, чередованием и другие. Аналогами конфигураций являются уровни RAID: 0, 1, 10 и т.д.

Кирпич

Кирпич (Гластера) — элементарный блок хранения данных. Это может быть каталог в файловой системе или блочное устройство, например раздел жёсткого диска. Добавлением и удалением кирпичей можно менять размер тома.

4 Основы использования

4.1 Подключение к системе управления

Управление кластером возможно двумя способами: через панель управления и через командную строку.

4.1.1 Панель управления

Веб-панель управления может быть установлена на узлах кластера или на выделенных серверах администрирования. В зависимости от способа установки вам необходимо знать:

- сетевое имя плавающего IP-адреса;
- или сетевое имя сервера администрирования.

Для подключения к панели управления используйте веб-браузер. Сетевое имя будет адресом необходимого вам сайта.

Панель управления предоставляет графический пользовательский интерфейс управления компонентами ПК «Кластер высокой готовности». Панель имеет четыре основные вкладки: «Обзор», «Узлы», «Ресурсы» и «Параметры». Так же имеется раздел «Помощь».

4.1.2 Командная строка

При использовании интерфейса командой строки вам необходимо знать сетевое имя или IP-адреса узлов кластера. В стандартной конфигурации к службе управления mgmtcd возможно подключиться *только из сети управления*.

4.2 Панель управления

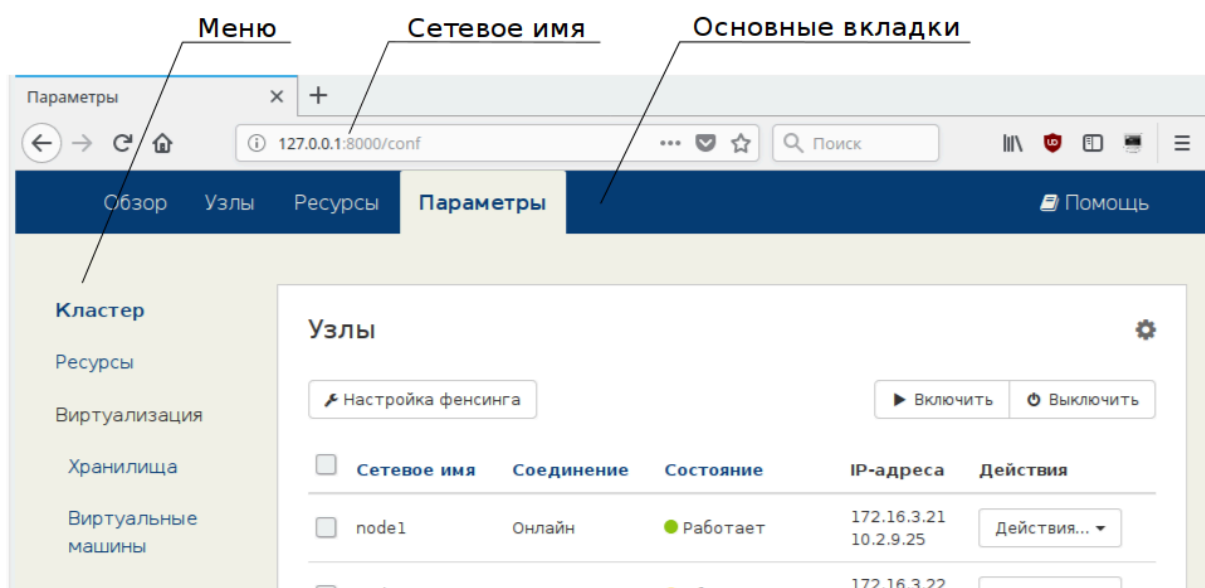


Рисунок 4.1. Панель управления

4.3 Интерфейс командой строки

Интерфейс командной строки (команда `hac`) предоставляется в пакете `python-hac-client`. После установки пакета проверьте доступность команды `hac`:

```
$ hac --help
```

Подключение к кластеру производится по сетевому имени или IP-адресу. В случае установленного кластера в варианте интегрированной панели управления следует подключаться по плавающему сетевому адресу. Если установка выполнена с выделенным сервером администрирования, следует подключаться к любому из серверов.

Сетевое имя указывается посредством параметра `--hosts`:

```
$ hac --hosts=172.16.1.15 <команда>
```

Например, проверить правильность настроек можно с помощью команды

```
$ hac --hosts=<узел> info
```

В случае успеха будет выведена версия и подключённые модули ПК «Кластер высокой готовности»:

```
Версия КВГ: 2.0.1  
Доступные модули: hac, virtualization, gluster, web_ssh
```

5 Узлы

Система управления кластером высокой готовности позволяет контролировать и управлять узлами кластера. Узлы можно включать и выключать, переводить в режим обслуживания. Основные параметры узлов:

- имя узла;
- состояние узла;
- состояние связи с узлом.

Имя узла — это сетевое имя компьютера во внутренней сети кластера. Это имя по которому узлы кластера идентифицируют друг друга.

Состояние узла — это текущее состояние кластера:

Работает

Узел работает и его состояние синхронизировано с остальными узлами кластера. В этом состоянии на узле могут запускаться ресурсы.

Обслуживание

Этот режим работы узла предназначен для обслуживания узла кластера: замена аппаратного обеспечения, обновление программного обеспечения и т.д. После перевода узла в это состояние с него будут перенесены все запущенные на узле ресурсы. В этом состоянии все системные компоненты Системы управления продолжают работать. Узел может быть выбран [Координатором](#).

Единственное состояние в котором узел можно штатно выключить.

Отказ

В состояние «Отказ» узел может перейти по ряду причин:

- сбой в программных компонентах системы высокой готовности;
- неправильная настройка компонентов Кластер высокой готовности;
- нештатное выключение узла (не из состояния «Обслуживание»).

Состояние связи с узлом — это признак включения узла в кластер:

Онлайн

Компоненты системы высокой готовности настроены и функционирует на узле. Локальная копия базы данных кластера синхронизирована с кластерной.

Нет связи

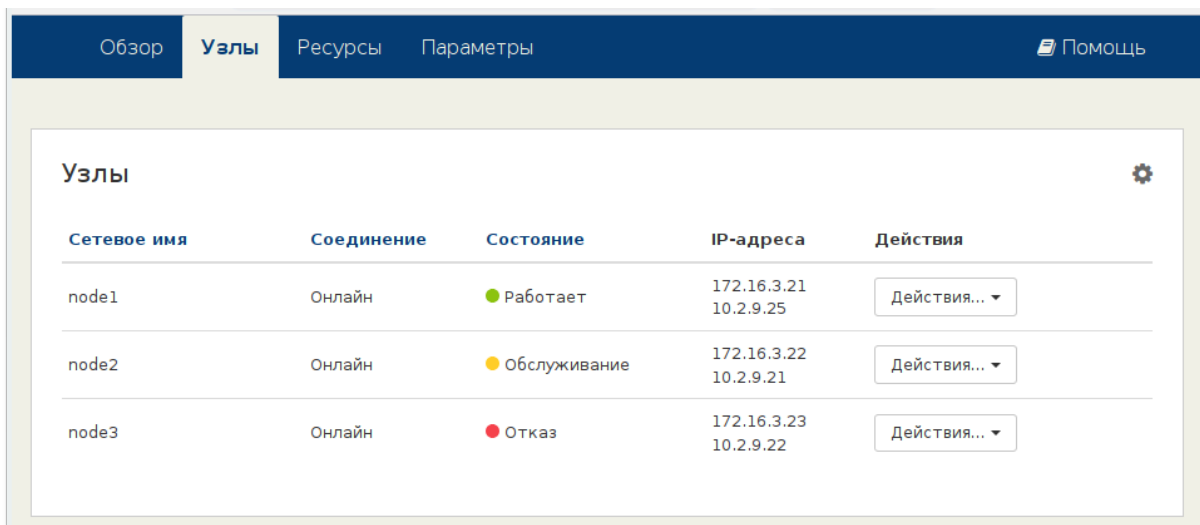
У система высокой готовности нет связи с узлом. Возможны следующие причины:

- узел выключен;
- отказ и останов компонентов системы высокой готовности;
- неправильная настройка сетевого оборудования;
- аппаратная проблема сетевых портов.

В случае, если в кластере настроен и включен фенсинг, и узел находится в состоянии отказа, то система высокой готовности предпримет попытки включить или перезагрузить узел через механизм фенсинга.

5.1 Просмотр и изменение параметров

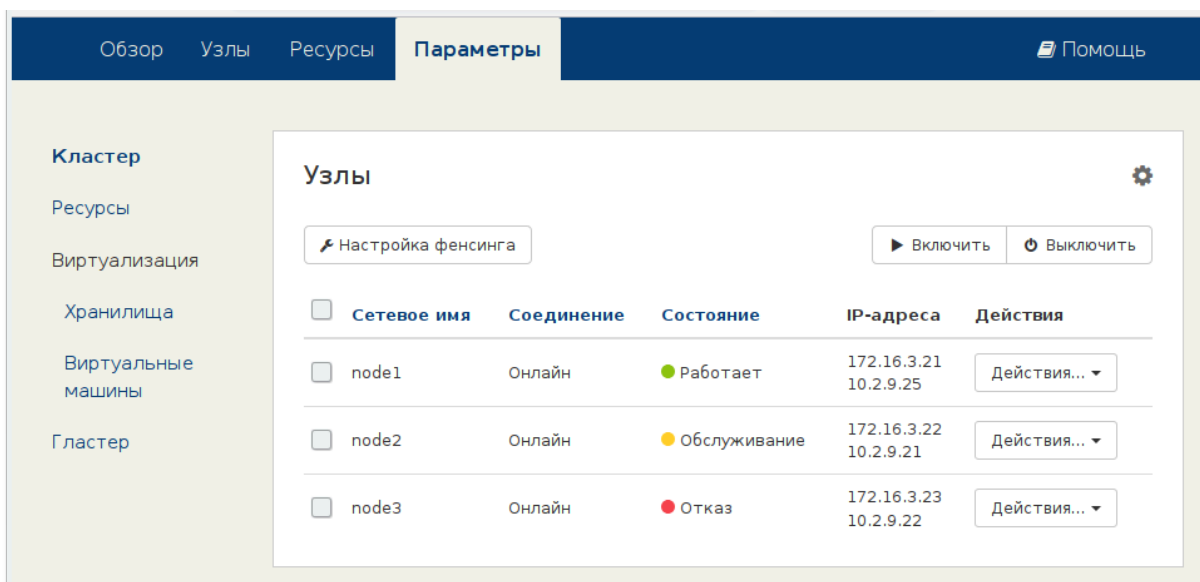
Информация об узлах отображается на основной вкладке «Узлы» (рис. 5.1).



Сетевое имя	Соединение	Состояние	IP-адреса	Действия
node1	Онлайн	● Работает	172.16.3.21 10.2.9.25	Действия... ▾
node2	Онлайн	● Обслуживание	172.16.3.22 10.2.9.21	Действия... ▾
node3	Онлайн	● Отказ	172.16.3.23 10.2.9.22	Действия... ▾

Рисунок 5.1. Информация о состоянии узлов кластера

Изменить параметры и произвести другие действия можно на основной вкладке «Параметры» в разделе меню «Кластер» (рис. 5.2).



Сетевое имя	Соединение	Состояние	IP-адреса	Действия
node1	Онлайн	● Работает	172.16.3.21 10.2.9.25	Действия... ▾
node2	Онлайн	● Обслуживание	172.16.3.22 10.2.9.21	Действия... ▾
node3	Онлайн	● Отказ	172.16.3.23 10.2.9.22	Действия... ▾

Рисунок 5.2. Параметры узлов кластера

5.2 Включение и выключение

Функция включения узла кластера доступна только при настроенном и включенном фенсинге в кластере, поскольку использует данный механизм кластера.

Выключение узла кластера выполняется штатными средствами операционной системы аналогично вызову команды poweroff на целевом узле.

Предупреждение

Перед выключением узел необходимо перевести на обслуживание. В противном случае система высокой готовности переведёт его в состояние отказа.

5.3 Обслуживание узла

Функция предназначена для предоставления администратору возможности обслуживания узла кластера: замена аппаратного обеспечения, обновление программного обеспечения и т.д.

Перевод на обслуживание доступен только для узлов, находящихся в состоянии «Работает».

После перевода узла в это состояние с него будут перенесены все запущенные на узле ресурсы. В этом состоянии все системные компоненты Системы управления продолжают работать. Узел может быть выбран **Координатором**.

Если с узлом нет связи, то перевод на обслуживание не приведёт к фактической реализации режима на узле.

В состоянии «Обслуживание» возможно штатное выключение узла.

Функции доступны через меню «Узлы» на вкладке «Параметры». Для ввода или вывода узла выберите соответствующий пункт в меню «Действия» для узла.

5.4 Веб-терминал

Функция подключения к веб-терминалу доступна при включении соответствующего модуля администратором кластера высокой готовности. Позволяет подключаться по SSH к узлам кластера.

6 Ресурсы

Ресурс кластера — это программы, данные или другие объекты, управляемые кластером.

6.1 Свойства ресурсов

Каждый ресурс обладает следующими свойствами или характеристиками:

- название;
- тип;
- ресурсный агент;
- состояние;
- режим;
- параметры агента;
- набор операций.

Название ресурса служит и его идентификатором в системе управления кластером, поэтому должно быть уникальным. Задается при создании ресурса и не может быть изменено в дальнейшем. Название должно состоять из символов английского алфавита, пробелы не допускаются.

Примечание

Мы рекомендуем следующие правила названия ресурсов:

- использование прописных букв;
- использование дефиса (-) в качестве разделителя слов.

Для каждого ресурса может быть добавлено *описание ресурса*, поясняющее его назначение или другие аспекты.

6.1.1 Типы ресурсов

Обеспечение высокой доступности ресурсов кластера может осуществляться различными приёмами. В простейшем случае это могут быть ресурсы, к примеру, виртуальные машины, которые могут работать в единственном экземпляре. Доступность для таких ресурсов обеспечивается за счёт непрерывного контроля и автоматического перезапуска в случае сбоев.

Повысить доступность в ряде случаев можно за счёт параллельного функционирования копий ресурса на нескольких узлах. В этом случае переключение на запасную копию происходит без переноса данных и запуска на другом узле кластера. К примеру, так может функционировать веб-сервер, раздающий статическую информацию.

Возможность параллельной работы нескольких экземпляров ресурса требует, как правило, выполнения одного из следующих условий:

- использование только статических данных;
- реализация синхронизации данных и состояния самим ресурсом.

Многие программы, такие как СУБД, имеют встроенные функции синхронизации между несколькими экземплярами. Но для корректной работы требуют дополнительных ограничений, таких как введение ведущих и ведомых экземпляров.

Эти и другие схемы повышения готовности (доступности) ресурсов могут быть реализованы в ПК «Кластер высокой готовности» путём назначения ресурсу одного из следующих типов:

- простой ресурс;
- группа ресурсов;
- клон (клонированный ресурс);
- несимметричные клон (ведущий-ведомый, В/В);
- уникальный клон.

Простые ресурсы

Простые ресурсы работают в единственном экземпляре на одном из узлов кластера. Для простых ресурсов доступен весь арсенал настройки взаимозависимостей: очерёдность, привязки к узлам, совместное расположение.

Группы

Один из наиболее частых видов взаимосвязи между ресурсами кластера — это когда некоторый набор ресурсов должен работать на одном узле и при этом поочерёдно запускаться и останавливаться.

Группа — тип ресурса, который объединяет несколько простых ресурсов в определённом порядке. Ресурсы в группе:

- запускаются поочерёдно в определённом порядке;
- останавливаются в определённом порядке.

Если при запуске ресурса в группе произошёл сбой, то последующие ресурсы запускаться не будут.

Как и простой ресурс, группу можно запускать и останавливать, изменять режим, устанавливать взаимозависимости с другими ресурсами. При этом ресурсы, входящие в группу, сохраняют индивидуальные свойства, такие как операции и параметры.

Клоны

Клон — набор равнозначных экземпляров ресурса или группы. Предполагается, что все экземпляры функционируют идентично на любых узлах кластера. На каждом узле одновременно может работать только один экземпляр.

Примечание

Клонировать можно только ресурсы, предоставляющие такую возможность. Например, клонировать можно ресурс подключения сетевой файловой системы, поскольку такие ФС имеют встроенные средства обработки одновременного доступа с разных узлов.

Клоны могут применяться в целях балансировки, повышения готовности ресурса, при необходимости работы служб на нескольких узлах.

Для клонов можно задать следующие специфические параметры:

- минимальное количество экземпляров;
- максимальное количество экземпляром.

Если эти параметры не заданы для клона, то минимальным количеством будет 1, а максимальным – число узлов в кластере. Таким образом стандартный клон запускает экземпляры ресурса на всех узлах кластера.

Для клонов доступны с некоторой спецификой все взаимозависимости: очерёдность, привязки к узлам, совместное расположение.

Уникальные клоны

Уникальный клон – набор экземпляров ресурса или группы, неравнозначных между собой. Поскольку каждый экземпляр уникален, то несколько экземпляров могут работать на одном узле.

Для уникальных клонов можно задать следующие специфические параметры:

- минимальное количество экземпляров;
- максимальное количество экземпляром;
- максимальное количество экземпляров на одном узле.

В остальном уникальные клоны повторяют простые.

Несимметричные клоны

В несимметричном клоне для каждого экземпляра ресурса задаётся одна из двух ролей: ведущий или ведомый.

Экземпляры ресурса могут менять роли по команде системы управления. Для несимметричных клонов можно задать следующие параметры:

- минимальное количество экземпляров;
- максимальное количество экземпляром;
- максимальное количество экземпляров на одном узле;
- максимальное количество ведущих экземпляров;
- максимальное количество ведущих экземпляров на одном узле.

Запуск несимметричного клона происходит по следующему алгоритму. Сначала все экземпляры ресурса запускаются в роли ведомых. Вторым этапом один или несколько экземпляров переводится в роль ведущих.

Типичным применением несимметричных клонов являются СУБД. При репликации баз данных, ведущий процесс базы данных устанавливается в качестве оригинального источника данных, и ведомые процессы синхронизируются с ним.

Для клонов доступны с некоторой спецификой все взаимозависимости: очерёдность, привязки к узлам, совместное расположение.

6.1.2 Состояние ресурса

Каждый созданный ресурс может находиться в различных состояниях, характеризующих способность работать по своему назначению:

Запущен

Ресурс находится в состоянии «запущен», если он успешно запущен и операция контроля (при наличии) подтверждает функционирование.

Ожидание

Ресурс ожидает запуска. Может быть запущен, но необходимые условия не выполнены. Возможные причины:

- ресурс зависит от ещё не запущенного ресурса;
- нет подходящих узлов для запуска;
- не хватает ресурсов для запуска.

Остановлен

Ресурс находится в состоянии «остановлен», если он целенаправленно остановлен пользователем.

Отказ

Ресурс находится с состоянием «отказ», если:

- 1) Простой: запуски на всех допустимых узлах привели к сбоям.
- 2) Группа: один из членов группы находится в состоянии отказа.
- 3) Клон: количество запущенных и ожидающих производных ресурсов меньше заданного.
- 4) В/В: условия аналогичные клону или если нет запущенных ведущих ресурсов.

6.1.3 Режимы работы

Возможны два режима работы ресурса: автоматического управления и ручного.

В режиме *автоматического управления* ресурсом система управления полностью контролирует его состояние: может запускать, останавливать, переносить, перезапускать и т.д. Вмешательство администратора в обход системы управления может привести к сбоям, отказам и порче разделяемых данных в кластере.

Режим *ручного управления* предназначен для передачи управления ресурсом администратору. В этом режиме администратор может останавливать, запускать, настраивать и производить другие действия над ресурсом. Кластер не производит операций над ресурсом, в том числе не контролирует состояние ресурса. Если остановленный ресурс запустить в обход системы управления (например, для службы – командой `service start`), то кластер не будет изменять отображаемое состояние ресурса. При возвращении ресурса на автоматическое управление кластер будет стремиться вернуть ресурс в исходное состояние.

6.1.4 Ресурсные агенты

Управление ресурсами происходит посредством **ресурсных агентов**, которые предоставляют унифицированный интерфейс управления в кластерной среде. Интерфейс основан на утверждённых отраслевых стандартах.

Название (тип) каждого ресурсного агента строится по одному из следующих правил:

- <класс>:<производитель>:<название>
- <класс>:<название>

Класс

Существует несколько классов ресурсных агентов:

service

Универсальный мета-класс для управления системными службами (демонами) Линукс, установленными на узлах кластера. На самом деле объединяет возможности следующих классов:

- `lsb` – службы, соответствующие стандарту LSB (Linux Standard Base). Такие службы располагаются в каталоге `/etc/init.d`;
- `systemd` – службы стандарта `systemd`. В соответствии с возможностями `systemd`, обобщают службу до более универсального объекта *модуль* (`unit`).

Агенты классов `service`, `lsb`, `systemd` не позволяют указывать производителя в названии.

ocf

Класс ресурсных агентов соответствующих специализированному стандарту OCF (Open Cluster Framework). Агенты данного класса наиболее полно раскрывают возможности управления ресурсами.

stonith

Класс агентов для устройств фенсинга. Свойства ресурсов фенсинга в целом аналогичны таковым ресурсам других классов, но имеют ряд второстепенных отличий.

nagios

Класс агентов для модулей инструмента мониторинга системы и инфраструктуры Nagios.

6.1.5 Параметры агента

Ресурсные агенты OCF, STONITH и Nagios позволяют задавать дополнительные параметры для ресурсов. Это могут как параметры, задающие конфигурацию целевого ресурса, или параметры, влияющие на поведение самого агента.

Список доступных параметров фиксирован и определяется ресурсным агентом. Каждый параметр строго типизирован и может быть строкой, числом или булевым значением.

Параметры задаются при создании ресурса и могут быть изменены в дальнейшем.

6.1.6 Операции

Управление ресурсами происходит путём выполнения операций ресурсного агента. Система управления запускает соответствующую операцию и ждёт её выполнения. Время и ожидания и реакция на сбои могут настраиваться пользователем.

Система управления может инициировать следующие операции:

- запуск;
- остановка;
- контроль (мониторинг).

Операций мониторинга может быть несколько, например с разными интервалами или с разной полнотой проверки.

Для **несимметричных клонов** есть дополнительные операции:

- повышение (из ведомого до ведущего);
- понижение (из ведущего до ведомого);
- уведомление (об операциях с ресурсов на других узлах).

Каждая операция обладает набором параметров, приведённых в [табл. 6.1](#).

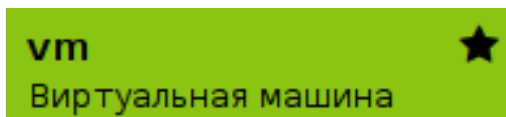
Таблица 6.1. Параметры операций

Параметр	Описание
Операция	Название операции: start, stop, monitor, notify, promote, demote
Интервал повтора	Интервал повтора операции (сек). Стандартное значение: 0 (не повторять).
Таймаут	Время ожидания выполнения операции (сек). Если операция не будет успешно выполнена за это время, она будет прервана и статус операции будет установлен в «отказ». Стандартное значение: задаётся ресурсным агентом.
Реакция	Тип реакции на произошедший сбой выполнения операции. Варианты: <ul style="list-style-type: none"> – игнорирование ошибок; – блокирование всех дальнейших операций; – остановка ресурса, в дальнейшем автоматический запуск производится не будет; – перезапуск (на этом или другом узле); – фенсинг узла на котором произошёл сбой; – вывод узла на обслуживание на котором произошёл сбой. Стандартное значение: перезапуск. Для операции stop: фенсинг (если включен), иначе блокировка.

6.2 Контроль ресурсов

Контроль ресурсов удобно осуществлять на вкладке «Ресурсы» панели управления. Страница отображает матрицу ресурсов. Информация о ресурсе представляется следующим образом.

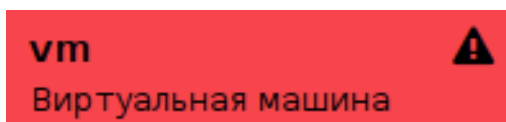
Основная информация: название, ресурсный агент, наличие сбоев, привязка к узлу — отображаются в ячейке строки соответствующего ресурса. Описание ресурса появляется при наведении мыши на ячейку.



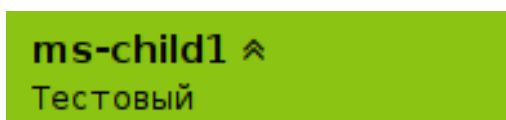
приоритетный узел



запретный узел

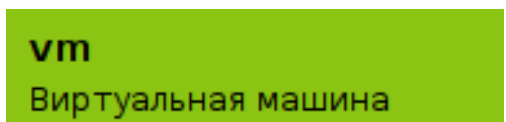


сбой

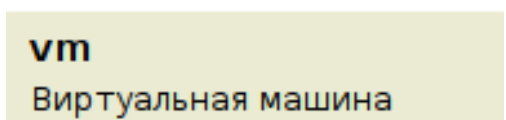


ведущий

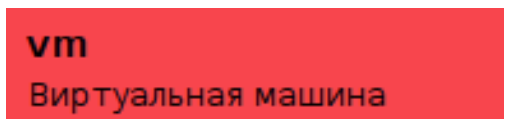
Состояние и режим работы кодируются цветом:



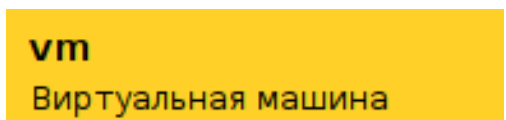
запущен



ожидание



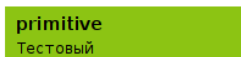
отказ



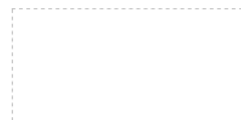
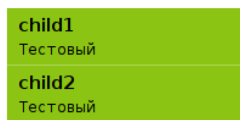
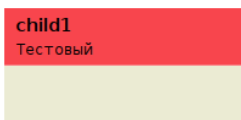
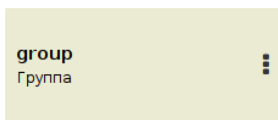
ручное управление

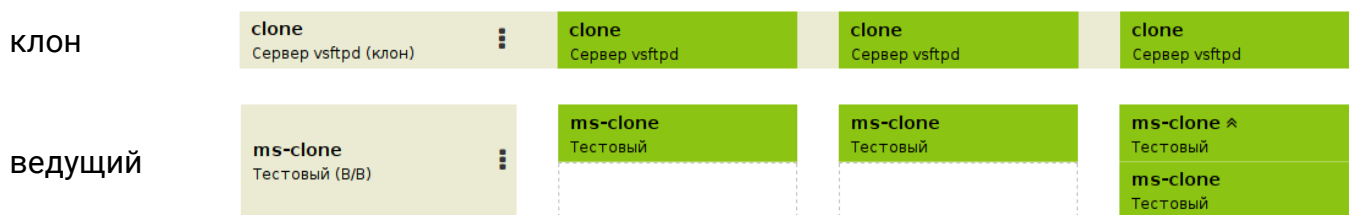
Фон строки ресурса несёт информацию о **типе ресурса**:

простой



группа





Также на данной странице возможно выполнить ряд основных функций для каждого ресурса. Номенклатура операций зависит от прав доступа пользователя и состояния ресурса.

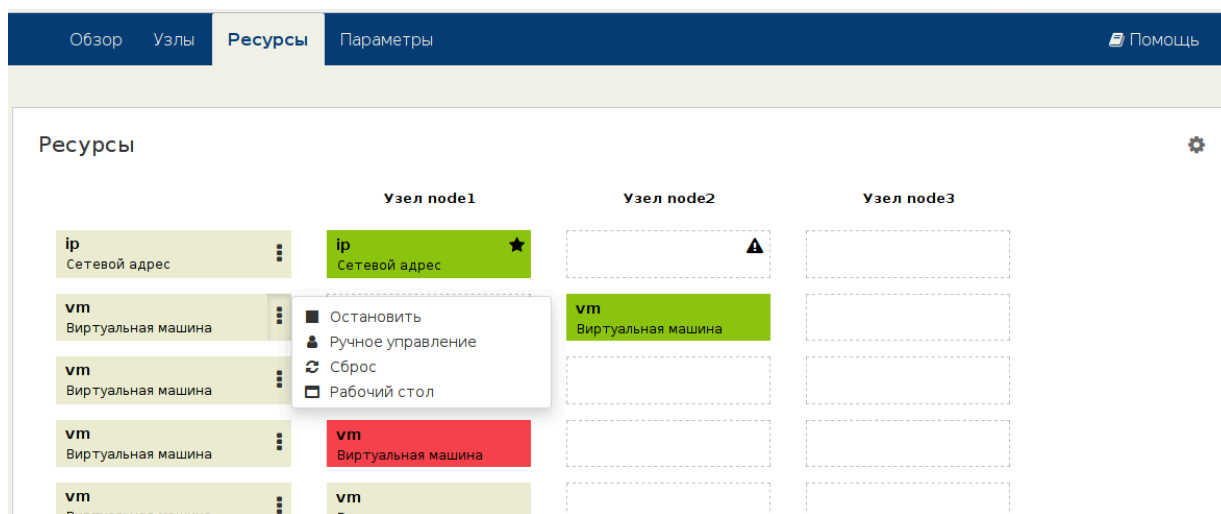


Рисунок 6.1. Основные функции над ресурсом

6.3 Изменение ресурса

Изменение параметров ресурсов возможно на вкладке «Параметры» в меню «Ресурсы». В выпадающем меню «Действия» для каждого ресурса доступны функции запуска, удаления, редактирования и др. Номенклатура операций зависит от прав доступа пользователя и состояния ресурса.

Набор параметров, доступных для редактирования зависит от типа ресурсов.

6.4 Запуск и остановка ресурса

Запуск возможен только остановленного ресурса, а остановка — запущенного. Если ресурс находится в состоянии «отказ», то попробуйте сбросить состояние ресурса и повторить попытку.

Запуск ресурса не означает того, что он будет немедленно запущен. Ресурс будет переведён в состояние ожидания, если условия для запуска не выполнены. Наиболее распространённые причины такого поведения:

- не запущен ресурс, от которого зависит запускаемый;
- нет подходящего узла для запуска ресурса.

В случае, если при запуске на одном из узлов произойдёт сбой, то система управления, как правило, попытается запустить ресурс на другом узле. Произошедший сбой при операции остановки приведёт либо к блокированию ресурса либо к фенсингу (перезагрузке)

узла, в зависимости от доступности механизма фенсинга. Это стандартные реакции на сбой системы управления, конкретные зависят от настройки соответствующей операции данного ресурса.

Ресурсы в группе запускаются последовательно друг за другом. В случае сбоя запуска одного из ресурсов, последующие запускаться не будут.

6.5 Сброс состояния ресурса

Сброс состояние ресурсов необходим в случае отказа или блокировки ресурса. Напомним, что состояние отказа фиксируется при переполнении счётчика сбоев на всех подходящих для ресурса узлах. Типичная ситуация отказа — исчерпание попыток запуска ресурса на подходящих узлах. После выяснения и устранения причины перед повторным запуском необходимо сбросить состояние ресурса.

При сбросе выполняются следующие операции:

- очистка истории операций;
- обнуление счётчика сбоев (ошибок) на всех узлах;
- вычисление действительного состояние ресурса.

6.6 Удаление ресурсов

Удаление ресурса действительно только в рамках системы управления и не обязательно очищает или удаляет задействованные прикладные ресурсы. К примеру, при удалении виртуальной машины будет удалена машина из кластера, но при этом используемые тома удалены не будут.

Помимо самого ресурса будут удалены и его взаимозависимости с другими ресурсами.

Примечание

Перед удалением ресурса его необходимо остановить.

Предупреждение

При удалении группы или клона будут удалены все вложенные ресурсы.

6.7 Приоритетные и запретные узлы

Ресурс кластера можно привязать к определённому узлу или, наоборот, запретить работу на нём. Система управления будет располагать ресурс на его приоритетном узле, когда тот находится в рабочем состоянии. Напротив, если узел отмечен как запретный, то ресурс не будет там запущен, даже в случае отсутствия подходящих узлов.

Задать приоритетный или запретный узел можно и для группы ресурсов. В этом случае каждому члену группы будет задан соответствующий узел. При этом невозможно задавать приоритетный и запретный узел для члена группы в отдельности.

Примечание

Мы рекомендуем задавать приоритетные узел только в случае наличия на этом узле уникальных аппаратных ресурсов. Использование данной возможности, к примеру, для балансировки, может ухудшить работу кластера при сбоях.

Назначение узла

1. Выберите пункт меню «Ресурсы» на вкладке «Параметры».
2. Нажмите кнопку «Действия» и выберите «Привязка» для выбранного ресурса.
3. В появившемся окне (рис. 6.2) выберите узел, который необходимо назначить приоритетным/запретным, выберите соответствующий пункт из выпадающего меню и нажмите кнопку «Сохранить».
4. После задания настройки система управления при необходимости произведёт перенос ресурса.

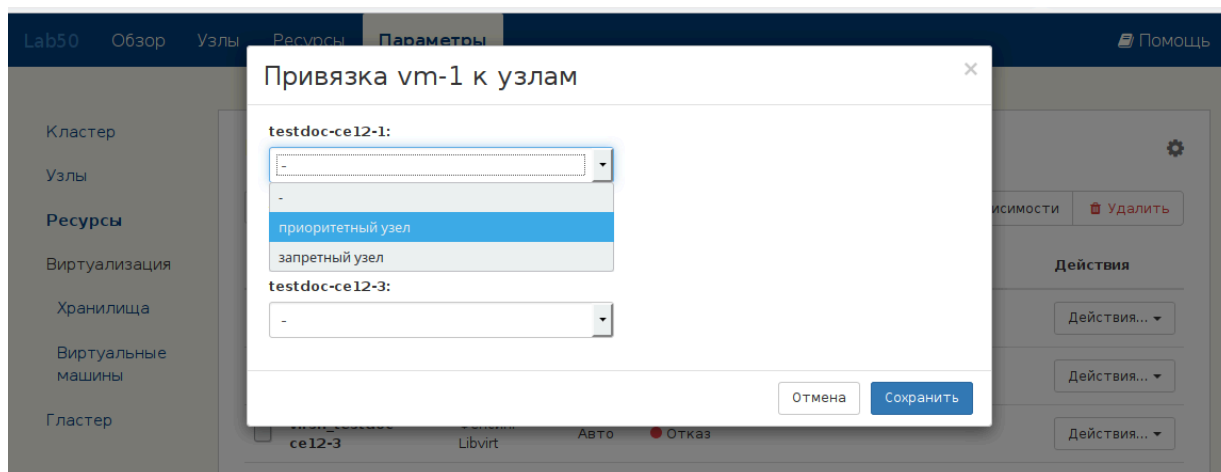


Рисунок 6.2. Назначение приоритетного/запретного узла для ресурса

Удаление узла

1. Выберите пункт меню «Ресурсы» на вкладке «Параметры».
2. Нажмите кнопку «Действия» и выберите «Привязка» для выбранного ресурса.
3. В появившемся окне (рис. 6.3) выберите узел, для которого необходимо сбросить статус приоритетный/запретный, выберите соответствующий пункт из выпадающего меню и нажмите кнопку «Сохранить».

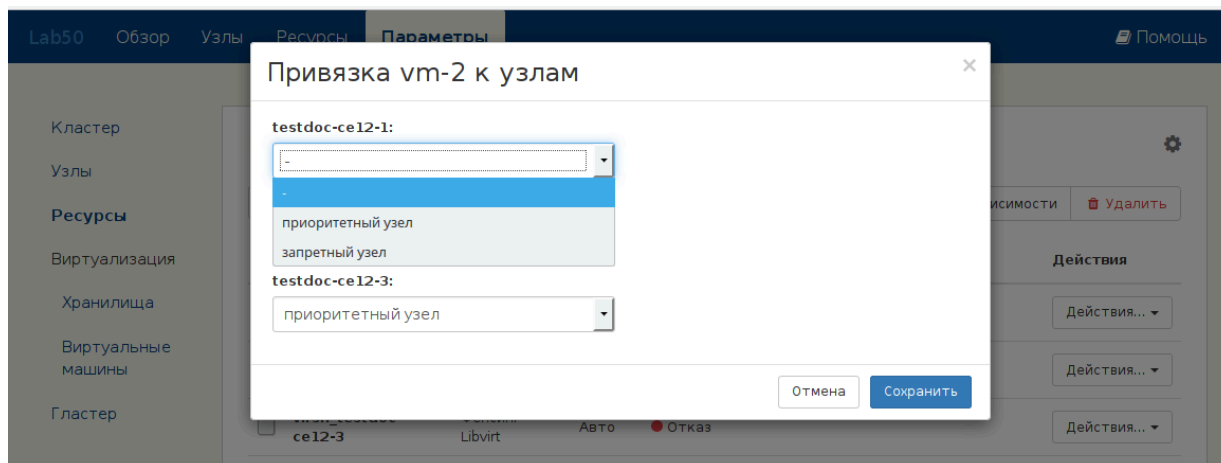


Рисунок 6.3. Сброс приоритетного/запретного узла для ресурса

6.8 Группы

Группа — это тип ресурса, который объединяет несколько простых ресурсов в определённом порядке.

Группу можно запускать и останавливать, изменять режим, устанавливать взаимозависимости с другими ресурсами. При этом ресурсы, входящие в группу, сохраняют индивидуальные свойства, такие как операции и параметры.

Создание группы

1. Выберите пункт меню «Ресурсы» на вкладке «Параметры».
2. Выберите ресурсы, которые будут входить в группу.
3. Нажмите кнопку «Создать» и выберите «Группу ресурсов».
4. В появившемся окне (рис. 6.4) введите название группы и нажмите кнопку «Создать».

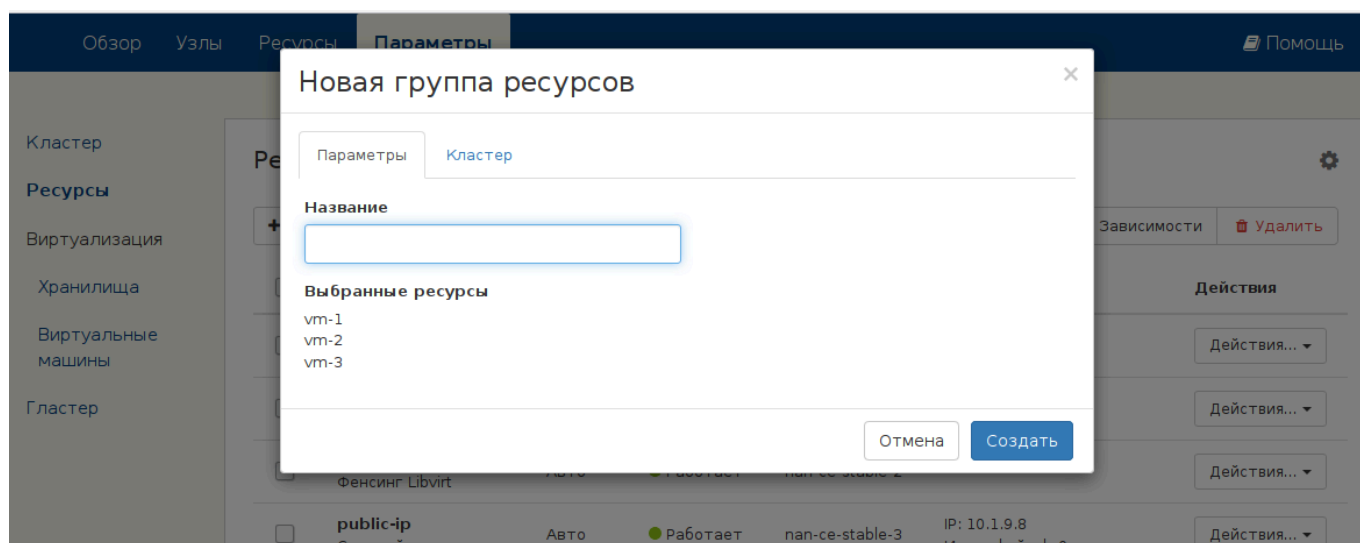


Рисунок 6.4. Создание группы ресурсов

Удаление группы

См. § 6.6.

7 Виртуализация

Защищённая виртуализация реализована на основе гипервизора KVM, входящего в состав операционной системы Astra Linux Common Edition. В состав программного комплекса виртуализации входит эмулятор оборудования (QEMU), обеспечивающий возможность запуска различных гостевых операционных систем.

В виртуальных машинах могут работать следующие 32/64-битные операционные системы:

- Альт;
- Astra Linux;
- MCBC 3 и 5;
- Linux с версиями ядра более 2.6.27;
- Windows (XP, Vista и новее);
- Windows Server (2003 и новее);
- ЗОСРВ «Нейтрино» и «QNX».

7.1 Виртуальные машины

7.1.1 Свойства виртуальных машин

Каждая виртуальная машина характеризуется следующим набором свойств:

- название;
- выделяемые ресурсы;
- источник;
- оборудование;

Название идентифицирует виртуальную машину в системе управления кластером, поэтому должно быть уникальным. Задаётся при создании машины и не может быть изменено в дальнейшем.

Ресурсы

При создании виртуальной машины пользователь задаёт необходимое количество виртуальных процессоров и некоторый объём оперативной памяти. В дальнейшем эти значения можно будет изменить, но для этого потребуются выключение виртуальной машины.

Виртуальные процессоры не привязываются к реальным ядрам или процессорам и ОС гипервизора может их перемещать по своему усмотрению. В некоторых случаях возможен *перебор ресурсов* – выделение памяти и процессоров для виртуальных машин более чем физически доступно. Защищённая виртуализация поддерживает механизм совместного использования идентичных страниц памяти.

Система управления размещает виртуальные машины на узлах в соответствии с занимаемыми ресурсами. Настройки балансировки могут задаваться системным администратором.

Источник

Для работы виртуальной необходим виртуальный диск с установленной гостевой операционной системой и прикладным программным обеспечением. В терминах ПК «Кластер высокой готовности» такой диск называется томом. Физически это может быть файл, раздел LVM или внешний том на СХД, подключаемый по протоколу iSCSI.

Управления томами в защищённой виртуализации организуется с помощью хранилищ. Хранилище — пул хранения с возможностью динамического выделения томов необходимого размера. С помощью хранилищ защищённая виртуализация абстрагирует виртуальное и физическое представление тома.

Где бы физически не располагался корневой том при создании виртуальной машины, его требуется задать. Корневой том (источник) можно сделать следующим образом:

- задать образ готового корневого тома, предварительно закачанного образа;
- установка гостевой операционной системы в новый том с существующего образа компакт-диска дистрибутива ОС.

Оборудование

Виртуальная машина эмулирует некоторый набор аппаратных средств для работы гостевой операционной системы. Различные операционные системы требуют разного совместимого оборудования. ПК «Защищённая виртуализация» предоставляет набор *артикулов* для различных гостевых операционных систем.

Артикул описывает конфигурацию эмулируемого оборудования виртуальной машины. В стандартной поставке включены артикулы для наиболее распространённых операционных систем. Системный администратор может самостоятельно создать новый артикул в целях учёта специфических особенностей создаваемых виртуальных машин.

В артикуле можно задать, к примеру, возможность проброса физического оборудования или особенности эмулируемого процессора.

Список стандартных артикулов приведён в § 7.1.3.

7.1.2 Загрузка образа

Для создания виртуальной машины необходим *источник* — корневой раздел (том) с установленной операционной системой. Можно использовать заранее созданный образ. Для этого его необходимо загрузить в кластер с помощью интерфейса командой строки.

Загрузка образа выполняется командой

```
$ hac --hosts=<сетевой адрес> image-upload <хранилище> <файл>
```

Пример:

```
$ hac --hosts=cluster \  
    image-upload default /home/user/debian-linux-http-server.qcow2
```

7.1.3 Создание

Функция создания виртуальной машины расположена в меню «Ресурсы» на вкладке «Параметры» (рис. 7.1).

В форме создания машины вам необходимо задать:

- название виртуальной машины;
- количество виртуальных процессоров и объем оперативной памяти;
- корневой том;
- тип установки: с готового образа или требуется установка гостевой операционной системы;
- артикул.

Протокол подключения к рабочему столу (VNC или SPICE) задаётся в настройках защищённой виртуализации и не может быть задан для конкретной машины.

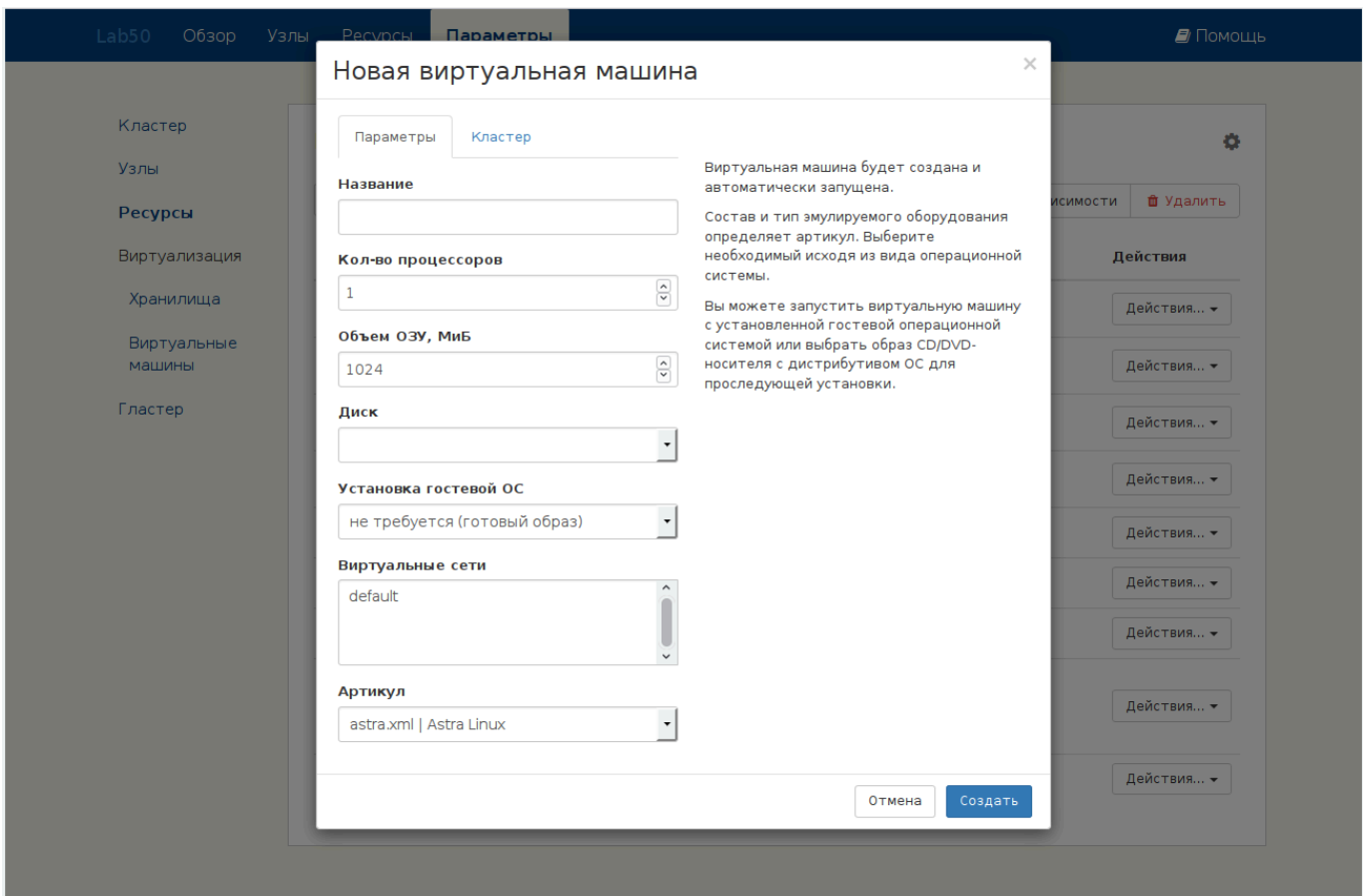


Рисунок 7.1. Создание виртуальной машины

Стандартные артикулы:

Linux

Универсальный артикул для современных операционных систем на базе Линукс: Альт, Astra Linux, Debian и др.

MCBC 3

Артикул для устаревших операционных систем на базе Линукс, главным образом MCBC 3.0.

Windows XP/Vista

Артикул совместимой с оригинальной Windows XP, Vista и Windows 2000.

Windows

Профиль оборудования для современных версий Windows и Windows Server, начиная с Windows 7 и Windows Server 2008.

Артикул	Чипсет	Диск	Сетевая карта	Паравиртуальные драйвера
Linux	i440FX	Virtio	Virtio	Да
MCBC 3	i440FX	IDE	Realtek 8139	Нет
Windows XP/Vista	i440FX	IDE	Realtek 8139	+Hyper-V, -Virtio
Windows	Q35	SATA	Intel E1000	+Hyper-V, -Virtio

После установки система управления запустит виртуальную машину на узле с учётом потребляемых и доступных ресурсов.

Если вы выбрали вариант установки гостевой операционной системы с компакт-диска, вам не необходимо подключиться к рабочему столу машины и выполнить установку.

7.1.4 Подключение к рабочему столу

Удаленное подключение к рабочему столу виртуальной машины осуществляется по протоколам VNC или SPICE. Протокол задается для всего кластера при установке ПК «Защищённая виртуализация».

Для подключения вам необходима программа-клиент, такие как:

- универсальный клиент virt-viewer (пакет virt-viewer);
- клиент vncviewer для протокола VNC (пакет tigervnc-viewer);
- клиент spice для протокола SPICE (пакет spice-client-gtk);
- универсальный клиент KRDC из состава KDE (пакет krdc).

Получить адрес подключения к выбранной виртуальной машине можно следующим образом:

- 1) Выберите пункт меню «Ресурсы» на вкладке «Параметры».
- 2) Нажмите кнопку «Действия» и выберите «Рабочий стол».