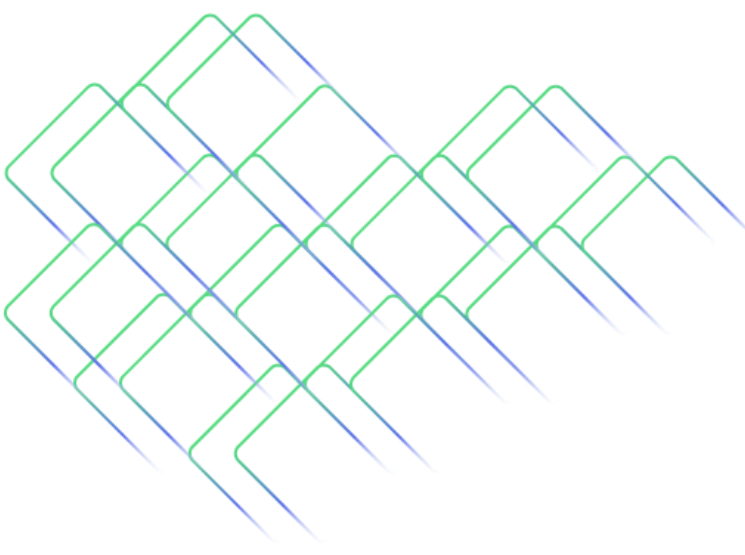
ORION SOFT

ZVIRT

Руководство по предварительному планированию инфраструктуры перед развертыванием среды виртуализации zVirt



Пояснение

Среда виртуализации zVirt состоит из взаимосвязанных компонентов, каждый из которых играет свою роль. Заблаговременное планирование и выполнению рекомендаций, позволит этим компонентам взаимодействовать и работать эффективно.

В данном руководстве рассматриваются:

- Требования к аппаратному обеспечению и безопасности;
- Опции, доступные для различных компонентов;
- Рекомендации по оптимизации вашей среды.

Внимательно прочтите данное руководство перед установкой системы безопасного управления виртуализацией zVirt, это позволит сократить время развертывания и избежать ошибок.

Глава 1. Архитектура среды виртуализации zVirt

Среда виртуализации zVirt может быть развернута как в режиме Hosted Engine, так и в режиме Standalone. Рекомендуемый вариант развертывания - Hosted Engine.

1.1. Архитектура режима установки Hosted Engine

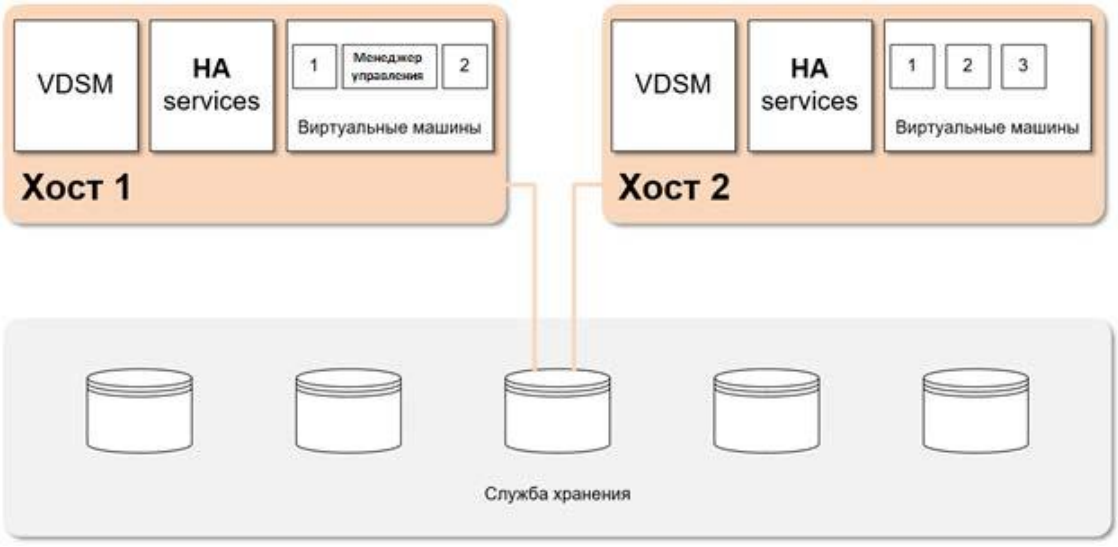
Система безопасного управления виртуализацией zVirt разворачивается в режиме Hosted Engine, при котором менеджер управления работает внутри ВМ, запущенной на хостах, управляемых этой службой управления. ВМ и менеджер управления создаются и настраиваются при первоначальной установке кластера.

Основное преимущество режима Hosted Engine состоит в том, что отсутствует необходимость в отдельном хосте с ролью менеджера управления. Кроме того, Hosted Engine может работать в режиме высокой доступности.

Минимальная конфигурация среды включает:

- Одна виртуальная машина с менеджером управления, которая размещается на хостах;
- Один (или два для режима высокой доступности ВМ HostedEngine) хоста;
- Внешняя система хранения данных (СХД) или локальное хранилище для размещения домена хранения данных (хранилища). Хранилище должно быть доступно всем хостам виртуализации.

Рисунок 1.1. Архитектура виртуализации zVirt варианта установки Hosted Engine:



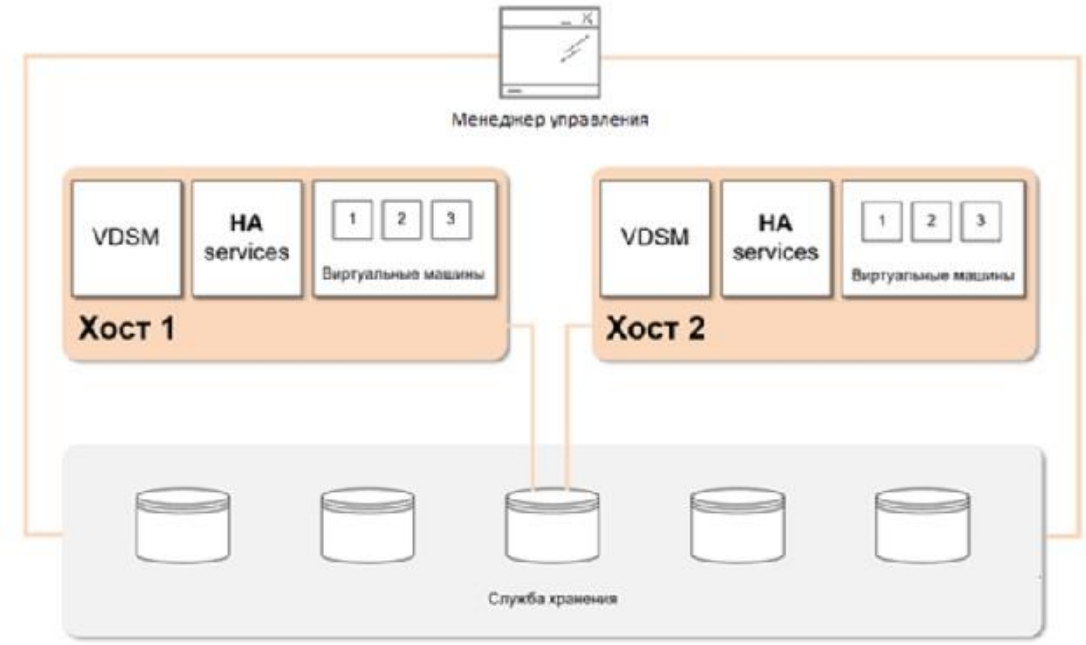
1.2. Архитектура режима установки Standalone

Система безопасного управления виртуализацией zVirt может разворачиваться в режиме Standalone. Режим Standalone проще в развертывании и управлении, но требует дополнительного физического хоста. В исключительных случаях возможно использование хоста с менеджером управления в качестве гипервизора.

Минимальная установка менеджера в режиме Standalone включает:

- Один хост для менеджера управления. Обычно менеджер управления развертывается на физическом хосте. Менеджер управления возможно развернуть и на виртуальной машине, если эта виртуальная машина размещена в отдельной среде виртуализации. Менеджер управления должен работать под управлением zVirt Node.
- Минимум два хоста для обеспечения высокой доступности виртуальных машин.
- Внешняя система хранения данных (СХД) или локальное хранилище для размещения домена хранения данных (хранилища). Хранилище должно быть доступно всем хостам виртуализации.

Рисунок 1.2. Архитектура виртуализации zVirt варианта установки Standalone:



Глава 2. Требования

2.1. Требования к менеджеру управления zVirt

2.1.1. Требования к аппаратному обеспечению

Минимальные и рекомендуемые требования к аппаратному обеспечению, описанные здесь, основаны на типичной установке малого и среднего размера. Точные требования варьируются в зависимости от размера и нагрузки.

Таблица 2.1. Требования к аппаратному обеспечению менеджера управления zVirt:

Ресурс	Минимальные требования	Рекомендуемые требования
Процессор	Двухъядерный процессор x86-64 с поддержкой VT-x/AMD-V	Четырехъядерный процессор или несколько двухъядерных процессоров
Память	4 ГБ	32 ГБ
Жёсткий диск	94 ГБ	1 ТБ
Сетевой интерфейс	1 сетевой интерфейс (NIC) 1 Гбит / с	2 сетевых интерфейса (NIC) 10 Гбит/с

2.1.2. Требования к браузерам

Для доступа к Порталу администрирования и Пользовательскому порталу можно использовать браузеры на основе Google Chrome или Mozilla Firefox.

2.1.3. Требования для клиента

Доступ к консолям виртуальных машин возможен с помощью клиента Remote Viewer (virt-viewer) на Linux и Windows. Для установки virt-viewer требуются права администратора.

Можно получить доступ к консолям виртуальных машин с помощью протоколов SPICE, VNC или RDP (только для Windows). Можно установить графический драйвер QXLDOD в гостевой операционной системе для улучшения функционала протокола SPICE.

В настоящее время SPICE поддерживает максимальное разрешение 2560x1600 пикселей.

Поддержка SPICE в клиентской операционной системе

Поддерживаемые драйверы QXLDOD доступны в ОС на базе CentOS Linux 7.2 и более поздних версиях, а также в Windows 10.

2.1.4. Требования к операционной системе

Менеджер управления должен быть развернут на хосте с установленной ОС из дистрибутива zVirt Node.

Не устанавливайте дополнительные пакеты после базовой установки, так как они могут вызвать проблемы с зависимостями при попытке менеджером управления установить пакеты из репозитория zVirt.

Не включайте дополнительные репозитории, кроме тех, которые предоставляются с дистрибутивом zVirt Node.

2.2. Требования к хосту

2.2.1. Требования к центральному процессору

Все процессоры должны поддерживать расширения Intel® 64 или AMD64 CPU, а также расширения аппаратной виртуализации AMD-V™ или Intel VT®. Также требуется поддержка флага No eXecute (NX).

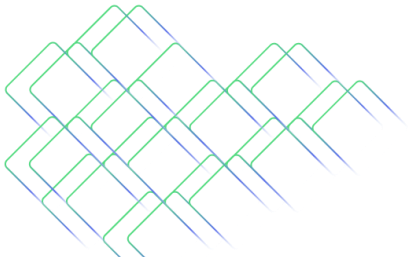


Таблица 2.2. Поддерживаемые семейства процессоров:

Intel	AMD
Nehalem	G1
Westmere G2	G2
SandyBridge	G3
IvyBridge	G4
Haswell	G5
Broadwell	G6
Skylake Client	
Skylake Server	
Cascadelake Server	
Icelake Server	

2.2.1.1. Проверка поддержки процессором необходимых флагов

Виртуализация должна быть включена в BIOS. Выключите и перезагрузите хост после внесенных изменений, чтобы убедиться, что изменение применено.

Проверить, какие расширения процессора доступны на системе, можно следующей командой:

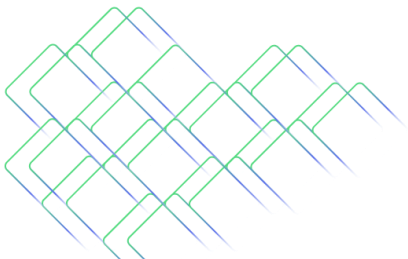
grep flags /proc/cpuinfo|head -n1|grep -Eo '(vmx|svm|nx)'

Если в выводе есть расширения - процессор поддерживает аппаратную виртуализацию. Если в выводе ничего нет, возможно, процессор поддерживает аппаратную виртуализацию, но в некоторых случаях производители отключают расширения виртуализации в BIOS. Если вы считаете, что аппаратная виртуализация выключена, обратитесь к руководству производителя материнской платы и BIOS.

2.2.2. Требования к оперативной памяти

Минимальный необходимый объем оперативной памяти составляет 4 ГБ. Максимальный поддерживаемый объем ОЗУ на хосте zVirt составляет 12 ТБ.

Объем оперативной памяти зависит от требований гостевой операционной системы (далее - виртуальная машина, ВМ), требований гостевого приложения, а также активности и использования гостевой памяти. KVM также может перезаписывать физическую оперативную память для ВМ, что позволяет вам предоставлять гостям требования к оперативной памяти, превышающие физические, при условии, что не все ВМ работают одновременно при пиковой нагрузке. KVM делает это, выделяя оперативную память только для ВМ по мере необходимости и переводя недостаточно загруженные ВМ в раздел подкачки.



2.2.3. Требования к хранилищу

Хостам требуется хранилище для хранения конфигурации, журналов, дампов ядра и для использования в качестве пространства подкачки. Хранилище может быть локальным или сетевым. zVirt Node (Хост) может загружаться с одним, несколькими или со всеми выделенными сетевыми хранилищами. Загрузка из сетевого хранилища может привести к зависанию в случае отключения сети. Если хост загружается из хранилища SAN и теряет соединение, файлы становятся доступными только для чтения, пока не восстановится сетевое соединение. Использование сетевого хранилища может привести к снижению производительности.

Минимальные требования и рекомендуемая схема разбиения хранилища:

- /(root) - 55 ГБ
- /home - 1 ГБ
- /tmp - 1 ГБ
- /boot - 1 ГБ
- /var - 15 ГБ
- /var/crash - 10 ГБ
- /var/log - 8 ГБ
- /var/log/audit - 2 ГБ
- swap - 1 ГБ

Минимальный общий объём - 94 ГБ.

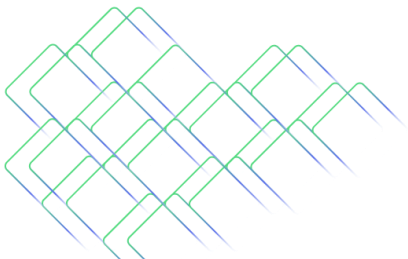
2.2.4. Требования к устройствам PCI

Хосты должны иметь как минимум один сетевой интерфейс с минимальной пропускной способностью - 1 Гбит/с. Рекомендуется, чтобы у каждого хоста было минимум два сетевых интерфейса, один из которых предназначен для поддержки интенсивных сетевых действий, таких как миграция виртуальных машин. Производительность таких операций ограничена доступной пропускной способностью.

2.2.5. Требования к проброску устройств

Если вы планируете реализовать проброс устройств, чтобы виртуальная машина могла использовать определенное PCI-устройство хоста, убедитесь, что следующие требования выполнены:

- Процессор должен поддерживать IOMMU (например, VT-d или AMD-Vi);
- Встроенное ПО должно поддерживать IOMMU;
- Используемые корневые порты CPU должны поддерживать ACS или эквивалентную ACS возможность;



- Устройства PCI должны поддерживать ACS или эквивалентную ACS возможность;
- Все коммутаторы и мосты PCIe между устройством PCI и корневым портом должны поддерживать ACS. Например, если коммутатор не поддерживает ACS, все устройства за этим коммутатором будут иметь общую группу IOMMU и могут быть назначены только одной виртуальной машине;
- Для поддержки GPU zVirt поддерживает назначение PCI-устройств на базе PCI NVIDIA K-Series Quadro (модель 2000 серии или выше), GRID и Tesla в качестве не-VGA графических устройств. В настоящее время к виртуальной машине может быть подключено до двух GPU в дополнение к одному из стандартных, эмулированных VGA интерфейсов. Эмулированный VGA используется для предварительной загрузки и установки, а графический процессор NVIDIA берет на себя управление, когда загружаются графические драйверы NVIDIA. Обратите внимание, что NVIDIA Quadro 2000 не поддерживается, как и карта Quadro K420.

Проверьте спецификацию и технические характеристики производителя, чтобы убедиться, что ваше оборудование соответствует этим требованиям. Команда lspci -v может быть использована для получения информации о PCI-устройствах, уже установленных в системе.

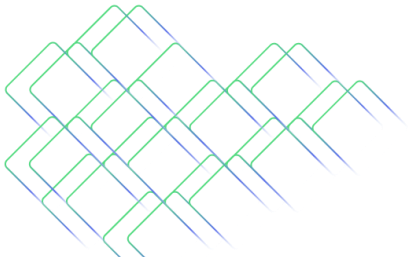
2.2.6. Требования к vGPU

Чтобы виртуальные машины на этом хосте могли использовать vGPU, хост должен отвечать следующим требованиям:

- vGPU-совместимый GPU;
- Ядро хоста с поддержкой GPU;
- Установленный GPU с правильными драйверами;
- Предопределенный тип mdev_type, соответствующий одному из типов mdev, поддерживаемых устройством;
- Драйверы с поддержкой vGPU установлены на каждом хосте в кластере;
- Операционная система виртуальной машины с поддержкой vGPU и установленными драйверами vGPU.

2.3. Сетевые требования

2.3.1. Общие требования



Для менеджера управления необходимо, чтобы IPv6 оставался включенным на физическом хосте или виртуальной машине, в зависимости от того, где запущен менеджер управления.

Не отключайте IPv6 на ВМ HostedEngine, даже если в вашей сети его не используют.

2.3.2. Диапазон сети для варианта установки Hosted Engine

В процессе развертывания zVirt в варианте Hosted Engine временно используется локальная сеть из диапазона **192.168.0.0/16**. По умолчанию используется адрес из сети **192.168.222.0/24**, если эта подсеть используется, система проверяет другие сети **192.168**, до тех пор пока не найдет свободную подсеть. Если система не найдет свободную подсеть в указанном диапазоне - установка завершится ошибкой.

С помощью командной строки можно настроить сценарий развертывания на использование альтернативного диапазона сети **/24** с помощью опции **–ansible-extra-vars=he_ipv4_subnet_prefix=PREFIX**, где **PREFIX** - префикс для диапазона по умолчанию. Например:

hosted-engine --deploy --ansible-extra-vars=he_ipv4_subnet_prefix=192.168.222

Примечание:

Задать другой диапазон можно только перед развертыванием.

2.3.3. Требования к межсетевому экрану для DNS, NTP и IPMI Fencing

DNS и NTP

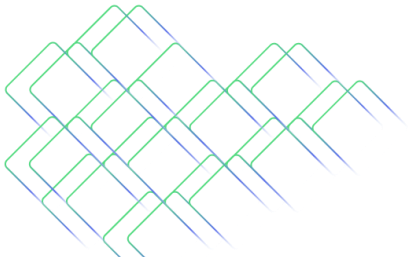
zVirt не создает сервер DNS или NTP, поэтому межсетевому экрану не нужно иметь открытые порты для входящего трафика.

По умолчанию zVirt разрешает исходящий трафик к DNS и NTP на любой адрес назначения. Если запрещается исходящий трафик, необходимо определить исключения для запросов на серверы DNS и NTP.

Важно:

- Менеджер управления zVirt и все хосты должны иметь подготовленные полные доменные имена, а также прямое и обратное разрешение имен.
- Служба DNS не должна находиться внутри среды виртуализации. Все службы DNS, используемые средой виртуализации, должны быть размещены за пределами среды виртуализации.
- Рекомендуется использовать сервер DNS вместо файла /etc/hosts для разрешения имен.

IPMI и другие механизмы Fencing (опционально)



Для IPMI (Intelligent Platform Management Interface) и других механизмов Fencing (ограждения) межсетевому экрану не нужно иметь открытые порты для входящего трафика.

По умолчанию zVirt разрешает исходящий трафик IPMI на порты с любым адресом назначения. Если вы запрещаете исходящий трафик, сделайте исключения для запросов IPMI или Fencing.

Каждый хост в кластере должен иметь возможность подключаться к устройствам ограждения всех остальных хостов в кластере. Если хосты в кластере получают ошибку (например: сетевая ошибка, ошибка хранилища...) и не могут функционировать как гипервизор, они должны иметь возможность подключения к другим хостам в центре данных.

2.3.4. Требования к межсетевому экрану менеджера управления

Менеджеру управления требуется, чтобы следующие номера портов были открыты, чтобы пропускать сетевой трафик через межсетевой экран системы.

Сценарий установки автоматически настроит firewalld во время развертывания, но перезапишет старую конфигурацию, если вы используете iptables. Если вы хотите использовать iptables и оставить существующую конфигурацию, вы должны настроить его самостоятельно. Описанная здесь конфигурация межсетевого экрана предполагает конфигурацию по умолчанию.

Таблица 2.3. Требования к межсетевому экрану менеджера управления

Номер правила	Порт	Протокол	Источник	Назначение	Применение	Шифрование по умолчанию
M1	-	ICMP	Хосты виртуализации	Менеджер управления	Необязательно. Может помочь при диагностике.	Нет
M2	22	TCP	Система (ы), используемая для обслуживания менеджера управления	Менеджер управления	Безопасный доступ Secure Shell (SSH). Необязательно.	Да
M3	2222	TCP	Клиенты, получающие доступ к консолям ВМ	Менеджер управления	Доступ через Secure Shell (SSH) для подключения к консолям виртуальной машины.	Да

Номер правила	Порт	Протокол	Источник	Назначение	Применение	Шифрование по умолчанию
M4	80, 443	TCP	Клиенте портала администрирования. Клиенты пользовательского портала, хосты виртуализации, клиенты REST API	Менеджер управления	Предоставляет HTTP (порт 80, не зашифрованный) и HTTPS (порт 443, зашифрованный) доступ к менеджеру zVirt. HTTP перенаправляет соединения на HTTPS.	Да
M5	6100	TCP	Клиенты портала администрирования. Пользовательский портал	Менеджер управления	Предоставляет доступ через прокси-сервер websocket для веб-консольного клиента noVNC, когда прокси-сервер websocket работает на менеджере zVirt. Однако, если прокси-сервер websocket работает на другом хосте, этот порт не используется.	Нет
M6	7410	UDP	Хосты виртуализации	Менеджер управления	Если Kdump включен на хостах, откройте порт для fence_kdump на менеджере управления zVirt. Fence_kdump не поддерживает шифрованное соединение. Вы можете вручную настроить этот порт, чтобы заблокировать доступ от	Нет

Номер правила	Порт	Протокол	Источник	Назначение	Применение	Шифрование по умолчанию
					хостов, которые не соответствуют требованиям.	
M7	54323	TCP	Клиенты портала администрирования	Менеджер управления (прокси-сервер ImageIO)	Требуется для связи с ImageIO Proxy (ovirtimageioproxy).	Да
M8	6442	TCP	Хосты виртуализации	Open Virtual Network (OVN)	Требуется для подключения к базе данных Open Virtual Network (OVN).	Да
M9	9696	TCP	Клиенты провайдера внешней сети для OVN	Внешний сетевой провайдер для OVN	OpenStack Networking API	Да, с конфигурацией, сгенерированной при установке менеджера.
M10	35357	TCP	Клиенты провайдера внешней сети для OVN	Внешний сетевой провайдер для OVN	OpenStack Identity API	Да, с конфигурацией, сгенерированной при установке менеджера.
M11	53	TCP, UDP	Менеджер управления zVirt	DNS - сервер	DNS-запросы поиска от портов с номерами более, чем 1023 к порту 53 и ответы на них. Открыты по умолчанию.	Нет
M12	123	UDP	Менеджер управления zVirt	NTP - сервер	NTP-запросы от портов с номерами более, чем 1023 к порту 123 и ответы на них. Открыты по умолчанию.	Нет

Примечание:

- Порт для базы данных OVN northbound (6641) не указан, поскольку в конфигурации по умолчанию, единственным клиентом для базы данных OVN northbound (6641) является ovirt-provider-ovn. Поскольку они работают на одном хосте, их взаимодействие не видно в сети.

2.3.5. Требования к межсетевому экрану хоста виртуализации

Хостам виртуализации необходимо, чтобы номера портов, перечисленные в таблице № 2.4 были открыты, чтобы пропускать сетевой трафик через межсетевой экран системы.

Таблица 2.4. Требования к межсетевому экрану хоста виртуализации

Номер правила	Порт	Протокол	Источник	Назначение	Применение	Шифрование по умолчанию
X1	22	TCP	Менеджер управления zVirt	Хосты виртуализации	Secure Shell (SSH). Необязательно.	Да
X2	2223	TCP	Менеджер управления zVirt	Хосты виртуализации	Доступ через Secure Shell (SSH) для подключения к консолям виртуальной машины.	Да
X3	161	UDP	Хосты виртуализации	Менеджер zVirt	Простой протокол управления сетью (SNMP). Требуется только в том случае, если вы хотите, чтобы прерывания Simple Network Management Protocol отправлялись с хоста одному или нескольким внешним SNMP-менеджерам.	Нет

Номер правила	Порт	Протокол	Источник	Назначение	Применение	Шифрование по умолчанию
					Необязательно.	
X4	111	TCP	NFS сервер	Хосты виртуализации	NFS соединения. Необязательно.	Нет
X5	5900 - 6923	TCP	Портал администрирования. Пользовательский портал.	Хосты виртуализации	Удаленный доступ к гостевой консоли через VNC и SPICE. Эти порты должны быть открыты для обеспечения доступа клиентов к виртуальным машинам.	Да (опционально)
X6	5989	TCP, UDP	Менеджер объектов общей информационной модели (CIMOM)	Хосты виртуализации	Используется менеджерами объектов общей информационной модели (CIMOM) для мониторинга виртуальных машин, работающих на хосте. Требуется только в том случае, если вы хотите использовать CIMOM для мониторинга виртуальных машин в вашей среде виртуализации.	Нет

Номер правила	Порт	Протокол	Источник	Назначение	Применение	Шифрование по умолчанию
X7	9090	TCP	Менеджер управления zVirt. Клиентские машины.	Хосты виртуализации	Требуется для доступа к веб-интерфейсу Cockpit, если он установлен.	Да
X8	16514	TCP	Хосты виртуализации	Хосты виртуализации	Миграция виртуальных машин с использованием libvirt.	Да
X9	49152 - 49215	TCP	Хосты виртуализации	Хосты виртуализации	Миграция и ограждение (fencing) виртуальных машин с использованием VDSM. Эти порты должны быть открыты для облегчения как автоматической, так и ручной миграции виртуальных машин.	Да. В зависимости от агента для ограждения, миграция осуществляется через libvirt.
X10	54321	TCP	Менеджер управления zVirt	Хосты виртуализации	Хосты виртуализации.	Связь VDSM с менеджером zVirt и другими хостами виртуализации.
X11	54322	TCP	Менеджер управления zVirt (прокси-сервер ImageIO)	Хосты виртуализации	Требуется для связи с демоном ImageIO.	Да
X12	6081	UDP	Хосты виртуализации	Хосты виртуализации	Требуется, когда в качестве сетевого поставщика используется открытая	Нет

Номер правила	Порт	Протокол	Источник	Назначение	Применение	Шифрование по умолчанию
					виртуальная сеть (OVN), чтобы OVN мог создавать туннели между хостами.	
X13	53	TCP, UDP	Хосты виртуализации	DNS	DNS-запросы поиска от портов с номерами более, чем 1023 к порту 53 и ответы на них. Открыты по умолчанию.	Нет
X14	123	UDP	Хосты виртуализации	NTP	NTP-запросы от портов с номерами более, чем 1023 к порту 123 и ответы на них. Открыт по умолчанию.	Нет
X15	4500	TCP, UDP	Хосты виртуализации	Хосты виртуализации	Internet Security Protocol (IPSec)	Да
X16	500	UDP	Хосты виртуализации	Хосты виртуализации	Internet Security Protocol (IPSec)	Да
X17	-	FY, ESP	Хосты виртуализации	Хосты виртуализации	Internet Security Protocol (IPSec)	Да

Примечание:

Правила межсетевого экрана автоматически настраиваются по умолчанию при добавлении нового хоста в среду виртуализации, перезаписывая любую существующую конфигурацию межсетевого экрана. Чтобы отключить автоматическую настройку межсетевого экрана при добавлении нового хоста, снимите флажок «Автоматически настраивать межсетевой экран хоста» в разделе Дополнительные параметры.

2.3.6. Требования к межсетевому экрану сервера баз данных

СУВ zVirt поддерживает использование удаленного сервера для базы данных менеджера zVirt и базы данных Data Warehouse (ovirt-engine-history). Если вы планируете использовать удаленный сервер, он должен разрешать соединения менеджера управления с службой Data Warehouse (которая может быть расположена отдельно от менеджера управления).

Аналогично, если вы планируете получить доступ к локальной или удаленной базе данных Data Warehouse из внешней системы, база данных должна разрешать эти соединения.

Важно:

Доступ к базе данных менеджера zVirt из внешних систем не поддерживается.

Таблица 2.5. Требования к межсетевому экрану сервера базы данных

Номер правила	Порт	Протокол	Источник	Назначение	Применение	Шифрование по умолчанию
БД1	5432	TCP, UDP	Менеджер управления, сервис Data Warehouse	Хосты виртуализации	Secure Shell (SSH). Необязательно.	Да
БД2	5432	TCP, UDP	Внешние системы	Хосты виртуализации	Доступ через Secure Shell (SSH) для подключения к консолям виртуальной машины.	Да

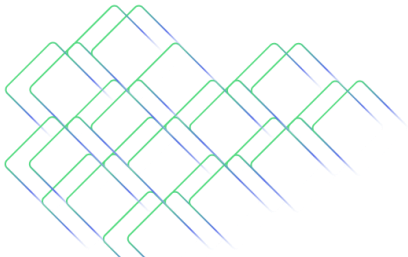
Глава 3. Пояснения

В этой главе описаны преимущества, ограничения и доступные опции для различных компонентов среды виртуализации zVirt.

3.1. Хосты

Все хосты в кластере должны иметь одинаковый тип процессора. Процессоры Intel и AMD не могут находится в одном кластере.

3.2. Центр данных



Каждый центр данных должен иметь как минимум один домен хранения данных. Также центр данных может иметь не более одного домена хранения Экспорт. Домены типа Экспорт и ISO устарели, но при необходимости их можно создать.

Домен хранения может состоять либо из блочных устройств (iSCSI или Fibre Channel), либо из файловой системы (POSIX).

По умолчанию домены GlusterFS и локальные домены хранения поддерживают размер блока 4K. Размер блока 4K может обеспечить более высокую производительность, особенно при использовании больших файлов, а также необходим при использовании инструментов, требующих совместимости с 4K, таких как VDO.

Важно:

В настоящее время zVirt не поддерживает блочное хранилище с размером блока 4K. Вы должны настроить блочное хранилище в режиме (512b block).

Типы хранилищ, описанные в следующих разделах, поддерживаются для использования в качестве доменов хранения данных. Домены хранения ISO и Экспорт поддерживают только файловые типы хранения. Домен ISO поддерживает локальное хранение при использовании в локальном центре данных.

3.2.1. NFS

zVirt поддерживает NFS версий 3 и 4. Для производственных рабочих нагрузок требуется сервер NFS корпоративного уровня, если NFS не используется только в качестве домена хранения ISO. Когда корпоративная NFS развернута на 10GbE, разделена с помощью VLAN, а отдельные службы настроены на использование определенных портов, она является одновременно быстрой и безопасной.

При расширении NFS хранилища СУВ zVirt сразу же распознает изменение размера хранилища данных. Никакой дополнительной настройки на хостах или менеджере управления не требуется. Это дает NFS небольшое преимущество перед блочным хранилищем с точки зрения масштабирования и эксплуатации.

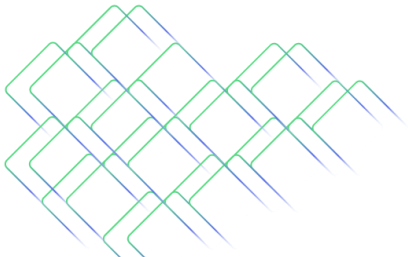
3.2.2. iSCSI

Для производственных рабочих нагрузок требуется сервер iSCSI корпоративного уровня. Если корпоративный iSCSI развернут на 10GbE, разделен на виртуальные локальные сети и использует аутентификацию CHAP, он является одновременно быстрым и безопасным. iSCSI также может использовать многоканальность (multipathing) для повышения высокой доступности.

zVirt поддерживает 1500 логических томов на блочный домен хранения. Разрешается использовать не более 300 LUN.

3.2.3. Fibre Channel

Fibre Channel является одновременно быстрым и безопасным, и его следует использовать, если он уже используется в целевом центре данных. Его преимущество



заключается в низкой нагрузке на процессор по сравнению с iSCSI и NFS. Fibre Channel также может использовать многоканальность (multipathing) для повышения высокой доступности.

zVirt поддерживает 1500 логических томов на блочный домен хранения. Разрешается использовать не более 300 LUN.

3.2.4. Fibre Channel over Ethernet

Чтобы использовать Fibre Channel over Ethernet (FCoE) необходимо включить ключ fcoe в менеджере управления и установить пакет vdsm-hook-fcoe на хосты.

zVirt поддерживает 1500 логических томов на блочный домен хранения. Разрешается использовать не более 300 LUN.

3.2.5. Gluster Storage

Gluster Storage (GS) - это POSIX-совместимая файловая система с открытым исходным кодом. Три или более серверов конфигурируются как кластер Gluster Storage, вместо сетевых устройств хранения данных (NAS) или массива сети хранения данных (SAN).

Gluster Storage следует использовать по 10GbE и разделять с помощью виртуальных локальных сетей.

3.2.6. Гиперконвергентная инфраструктура

zVirt поддерживает гиперконвергентный вариант развертывания с помощью Gluster. Вместо того чтобы подключать zVirt к внешнему хранилищу Gluster, существует возможность объединить zVirt и Gluster в одной инфраструктуре, что позволяет снизить эксплуатационные расходы и накладные расходы.

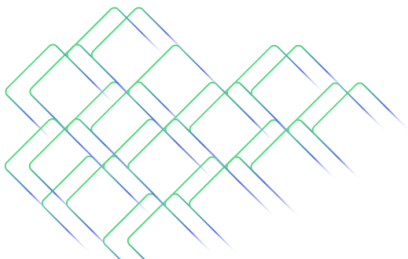
3.2.7. POSIX-совместимые файловые системы

Другие POSIX-совместимые файловые системы могут использоваться в качестве доменов хранения в zVirt, если они являются кластерными файловыми системами, такими как Global File System 2 (GFS2), и поддерживают разреженные файлы и прямой ввод-вывод. Файловая система Common Internet File System (CIFS), например, не поддерживает прямой ввод/вывод, что делает ее несовместимой с zVirt.

3.2.8. Локальное хранилище

Локальное хранилище создается на отдельном хосте, используя собственные ресурсы хоста. Когда вы настраиваете хост на использование локального хранилища, он автоматически добавляется в новый локальный центр данных и кластер, в который не могут быть добавлены другие хосты. Виртуальные машины, созданные в кластере с одним хостом, не могут быть перемещены, ограждены или запланированы.

Для хостов локальное хранилище всегда должно быть определено в файловой системе, отдельной от / (root). Используйте отдельный логический том или диск.



3.3. Пояснения по сети

При планировании и настройке сетей в среде СУВ zVirt настоятельно рекомендуется ознакомиться с концепциями сетей и их использованием. Прочитайте руководства производителя сетевого оборудования для получения дополнительной информации об управлении сетями.

Логические сети могут поддерживаться с помощью физических устройств, таких как сетевые карты, или логических устройств, таких как bond. Bonding улучшает высокую доступность и обеспечивает повышенную отказоустойчивость, поскольку все объединенные сетевые карты должны выйти из строя, чтобы сам bond вышел из строя. Режимы объединении 1, 2, 3 и 4 могут использоваться для сетей виртуальных машин. Режимы 0, 5 и 6 не предназначены для сети виртуальных машин. Система виртуализации по умолчанию использует режим 4.

Нет необходимости иметь одно устройство для каждой логической сети, поскольку несколько логических сетей могут совместно использовать одно устройство с помощью тегов виртуальных локальных сетей (VLAN) для изоляции сетевого трафика. Чтобы использовать эту функцию, тегирование VLAN должно поддерживаться на уровне коммутатора.

Ограничения на количество логических сетей в СУВ zVirt:

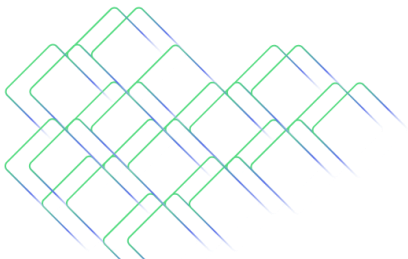
- Количество логических сетей, подключенных к хосту, ограничено количеством доступных сетевых устройств в сочетании с максимальным количеством виртуальных локальных сетей (VLAN), которое составляет 4096;
- Количество сетей, которые могут быть присоединены к хосту за одну операцию, в настоящее время ограничено 50;
- Количество логических сетей в кластере ограничено количеством логических сетей, которые могут быть присоединены к хосту, поскольку сетевое взаимодействие должно быть одинаковым для всех хостов в кластере;
- Количество логических сетей в центре данных ограничено только количеством содержащихся в нем кластеров в сочетании с количеством логических сетей, разрешенных для каждого кластера.

Важно:

Будьте особенно внимательны при изменении свойств сети управления ovirtmgmt. Неправильные изменения свойств сети ovirtmgmt могут привести к тому, что хосты станут недоступными.

3.4. Поддержка серверов каталога

Во время установки менеджер zVirt по умолчанию создает пользователя admin в домене internal. Эта учетная запись предназначена для использования при первоначальной настройке среды и для устранения неполадок. Вы можете создать



дополнительных пользователей во внутреннем домене с помощью утилиты ovirt-aaa-jdbc-tool. Учетные записи пользователей, созданные в локальных доменах, называются локальными пользователями.

Вы также можете подключить внешний сервер каталогов к системе виртуализации и использовать его в качестве внешнего домена. Учетные записи пользователей, созданные во внешних доменах, называются пользователями каталога. Поддерживается использование более одного сервера каталогов.

Для использования с СУВ zVirt поддерживаются следующие серверы каталогов:

- Active Directory https://docs.microsoft.com/en-us/windows-server/identity/identity-and-access
- Identity Management (IdM - на основе IPA) https://access.redhat.com/documentation/en-us/red_hat_enterprise_linux/8/html-single/planning_identity_management/index#intro-to-ipa-overview-of-planning-idm-and-access-control
- Red Hat Directory Server 9 (RHDS 9 - основан на 389DS) https://access.redhat.com/documentation/en-us/red-hat-directory-server/
- OpenLDAP http://www.openldap.org/doc/
- IBM Security (Tivoli) Directory Server https://www.ibm.com/support/knowledgecenter/SSVJJU_6.4.0/com.ibm.IBMDS.doc_6.4/welcome.ht

Важно:

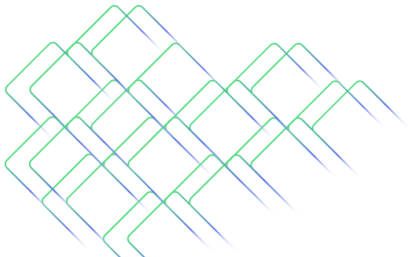
Пользователь с правами на чтение всех пользователей и групп должен быть создан в сервере каталогов специально для использования в качестве сервисной учётной записи для менеджера управления zVirt. Не используйте пользователя с правами администратора на сервере каталогов в качестве сервисной учётной записи для менеджера управления zVirt.

3.5. Инфраструктурные пояснения

3.5.1. Локальный или удаленный хостинг

Следующие компоненты могут быть размещены как на менеджере управления, так и на удаленном хосте. Размещение всех компонентов на менеджере управления проще и требует меньше ресурсов для обслуживания. Перемещение компонентов на удаленный хост требует больших ресурсов, но может повысить производительность как менеджера управления, так и Data Warehouse.

База данных и служба Data Warehouse:



- Чтобы разместить Data Warehouse на менеджере управления, выберите Yes, когда появится соответствующий запрос от утилиты engine-setup.
- Чтобы разместить Data Warehouse на удаленной машине, выберите No, когда появится соответствующий запрос от утилиты engine-setup.

Вы также можете разместить службу Data Warehouse и базу данных отдельно друг от друга.

Websocket proxy:

- Чтобы разместить прокси-сервер на менеджере управления, выберите Yes, когда появится соответствующий запрос от утилиты engine-setup.

3.5.2. Только удаленный хостинг

Следующие компоненты должны быть размещены на удаленной машине:

DNS:

Использование службы DNS внутри ВМ в той же среде виртуализации не рекомендуется.

Хранилище:

За исключением локального хранилища, служба хранения не должна находиться на одной машине с менеджером управления или любым хостом.

Управление идентификацией:

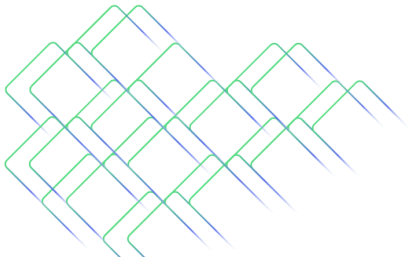
IdM (ipa-server) несовместим с пакетом mod_ssl, который требуется для менеджеру управления.

3.6. Моментальные снимки

Моментальные снимки — это функция, которая позволяет администратору создавать точки восстановления операционной системы, приложений и данных виртуальной машины на определенный момент времени.

Моментальные снимки сохраняют данные, присутствующие в образе жесткого диска виртуальной машины, в виде тома COW и позволяют восстановить данные, существовавшие на момент создания моментального снимка.

При создании моментального снимка, происходит создание новой дельта-копии диска виртуальной машины (слоя COW поверх текущего), после чего все данные будут записывать в эту дельту (новый слой COW). Таким образом, чем больше данных записывается после создания моментального снимка, тем больше времени потребуется для фиксации и консолидации их обратно в родительский образ, поэтому не рекомендуется использовать в продуктивной среде моментальные снимки, особенно в высоконагруженных виртуальных машинах.



Важно понимать, что образ жесткого диска виртуальной машины представляет собой цепочку из одного или нескольких слоёв. С точки зрения виртуальной машины эти слои выглядят как один образ диска. Виртуальная машина работает только с диском и не имеет доступа к слоям, из которых состоит диск.

Важно!

Каждый моментальный снимок создается для того, чтобы администратор мог отменить изменения в виртуальной машине, внесенные после создания моментального снимка. Снимки обеспечивают функциональность, аналогичную функции точки восстановления.

Моментальные снимки - это не резервные копии, а фиксация состояния образа виртуальной машины в определенный момент времени, к которому можно вернуться при необходимости. Не полагайтесь на моментальные снимки как на полноценный функционал резервного копирования.
Не храните продолжительное время моментальные снимки виртуальных машины. Как только убедитесь, что возврат к состоянию на момент создания моментального снимка больше не требуется - удалите моментальные снимки.

Ограничьте количество моментальных снимков. Создание нескольких снимков подряд может снизить производительность виртуальной машины и хоста, так как qemu придется просматривать каждый образ в цепочке моментальных снимков, чтобы считать новый файл из базового образа (base_image).

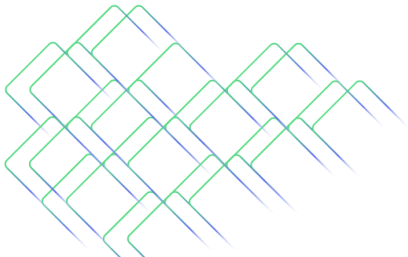
Перед созданием моментальных снимков на виртуальной машине должны быть установлены гостевые дополнения.

Три основные операции со снимками:

- Создание, которое включает в себя первый снимок, созданный для виртуальной машины.
- Предварительный просмотр, который включает предварительный просмотр моментального снимка, чтобы определить, следует ли восстанавливать данные на момент времени, когда был сделан снимок.
- Удаление, которое включает удаление точки восстановления, которая больше не требуется.

Моментальные снимки дисков виртуальных машин, помеченных как общие, и те, которые основаны на прямом подключении LUN, не поддерживаются. Моментальный снимок любой другой виртуальной машины, которая не клонируется или не мигрирует, может быть сделан во время работы, приостановки или после остановки.

Глава 4. Рекомендации

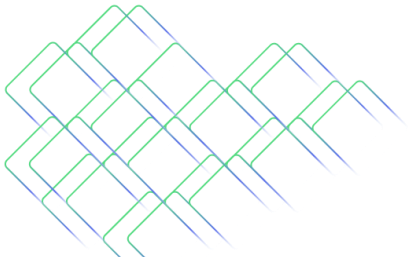


4.1. Общие рекомендации

- Сразу после завершения развертывания создайте полную резервную копию и сохраните ее в отдельном месте. После этого регулярно создавайте резервные копии.
- Избегайте запуска любой службы, от которой зависит менеджер управления, в качестве виртуальной машины в той же среде виртуализации. Если так делается, необходимо тщательно спланировать это, чтобы минимизировать время простоя, если виртуальная машина, содержащая эту службу, выйдет из строя.
- Убедитесь, что хост или виртуальная машина, на котором будет производится развертывание менеджера управления, имеет достаточную энтропию. Значения ниже 200 могут привести к сбою при развертывании менеджера. Чтобы проверить значение энтропии, выполните команду cat /proc/sys/kernel/random/entropy_avail. Чтобы увеличить энтропию, установите пакет rng-tools.
- Вы можете автоматизировать развертывание хостов и виртуальных машин с помощью PXE, Kickstart, CloudForms, Ansible или их комбинации. Обратите внимание, что установка в режиме Hosted engine с помощью PXE не поддерживается.
- Используйте протокол сетевого времени (NTP) на всех хостах и виртуальных машинах в среде для синхронизации времени. Аутентификация и сертификаты особенно чувствительны к разнице во времени. В zVirt поддерживается только сервер chrony.
- Документируйте всё, чтобы все, кто работает с окружением, знали о его текущем состоянии и необходимых процессах.

4.2. Рекомендации по безопасности

- Не отключайте функции безопасности (такие как HTTPS, SELinux и межсетевой экран) на хостах или виртуальных машинах.
- Создайте индивидуальные учетные записи администраторов, вместо того чтобы допускать использование одной учетной записи администратора несколькими сотрудниками. Это также полезно для отслеживания действий.
- Ограничьте доступ к хостам и создайте отдельные учетные записи. Не используйте одну учётную запись с правами root на всех хостах виртуализации.



- На хостах виртуализации должны использоваться только пакеты и службы, необходимые для виртуализации, производительности, безопасности и мониторинга. Хосты не должны иметь дополнительных пакетов, таких как анализаторы, компиляторы или других компонентов, которые добавляют риск для безопасности и стабильности.

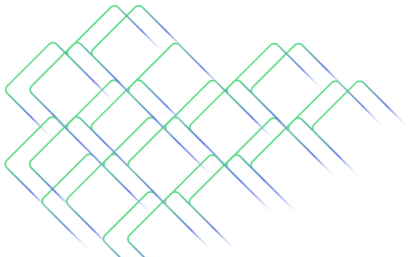
4.3. Рекомендации по хостам

- Стандартизируйте хосты в одном кластере. Это включает в себя использование одинаковых моделей оборудования и версий микропрограммного обеспечения. Смешивание различного серверного оборудования в одном кластере может привести к нестабильной производительности от хоста к хосту.
- Настройте устройства ограждения во время развертывания. Устройства ограждения необходимы для обеспечения высокой доступности.
- Используйте отдельные аппаратные коммутаторы для ограждения трафика. Если мониторинг и ограждение проходят через один коммутатор, этот коммутатор становится единой точкой отказа для обеспечения высокой доступности.

4.4. Рекомендации по работе с сетью

- Объединяйте сетевые интерфейсы, особенно на продуктивных хостах. Объединение улучшает общую доступность, а также пропускную способность сети.
- Стабильная сетевая инфраструктура использующая DNS и DHCP.
- Если объединения сетевых интерфейсов (bonding) будут использоваться совместно с другим сетевым трафиком, необходимо обеспечить надлежащее качество обслуживания (QoS) для хранилища и другого сетевого трафика.
- Для оптимальной производительности и упрощения поиска и устранения неисправностей используйте виртуальные локальные сети для разделения различных типов трафика и оптимального использования сетей 10 GbE или 40 GbE.
- Если базовые коммутаторы поддерживают jumbo frames, установите MTU на максимальный размер (например, 9000), который поддерживают базовые коммутаторы. Эта настройка обеспечивает оптимальную пропускную способность, более высокую пропускную способность и меньшее использование ЦП для большинства приложений. MTU по умолчанию определяется минимальным размером, поддерживаемым базовыми коммутаторами. Если у вас включен LLDP, вы можете увидеть MTU, поддерживаемый хостом, в подсказках сетевой карты в окне Установка сетей хоста.

  Важно:

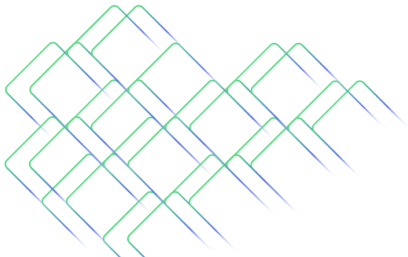


Если вы измените параметры MTU сети, вы должны распространить эти изменения на работающие виртуальные машины в сети: “Горячее” отключение и повторное подключение vNIC каждой виртуальной машины, которая должна применить настройки MTU, или перезапуск виртуальных машин. В противном случае эти интерфейсы выйдут из строя при миграции виртуальной машины на другой хост.

- Сети 1 GbE следует использовать только для трафика управления. Используйте 10 GbE или 40 GbE для виртуальных машин и хранилищ на базе Ethernet.
- Если на хост добавляются дополнительные физические интерфейсы для использования хранилища, снимите флажок Сеть ВМ, чтобы VLAN назначалась непосредственно физическому интерфейсу.

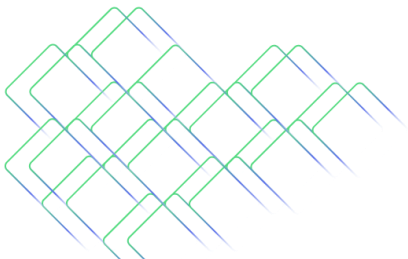
Рекомендации по настройке сетей хоста:

- Всегда используйте менеджер управления для изменения сетевой конфигурации хостов в кластерах. В противном случае вы можете создать неподдерживаемую конфигурацию.
- Если ваша сетевая инфраструктура сложная, вам может потребоваться настроить сеть хоста вручную перед добавлением хоста в среду виртуализации.
- Настроить сеть можно с помощью Cockpit. В качестве альтернативы можно использовать nmtui или nmcli.
- Если сеть не требуется для развертывания в режиме Hosted Engine или для добавления хоста в менеджер управления, настройте сеть на Портале администрирования после добавления хоста в менеджер управления.
- Используйте следующие соглашения об именовании:
  - Устройства VLAN: VLAN_NAME_TYPE_RAW_PLUS_VID_NO_PAD
  - Интерфейсы VLAN: physical_device.VLAN_ID (например, eth0.23, eth1.128, enp3s0.50).
  - Интерфейсы bond: bondnumber (например, bond0, bond1).
  - VLAN на объединенных интерфейсах: bondnumber.VLAN_ID (например, bond0.50, bond1.128).
- Используйте объединение сетей (bonding). Teaming не поддерживается в zVirt и приведет к ошибкам.
- Используйте рекомендуемые режимы объединения:
  - Режим 0 (round-robin) - передача пакетов через сетевые интерфейсы в последовательном порядке. Пакеты передаются в цикле, который начинается с первого доступного сетевого интерфейса на хосте и заканчивается последним доступным сетевым интерфейсом на хосте.



Все последующие циклы начинаются с первой доступной карты сетевого интерфейса. Режим 0 обеспечивает отказоустойчивость и распределяет нагрузку между всеми сетевыми интерфейсными картами в связке. Обратите внимание, что режим 0 не может использоваться в сочетании с bridge и поэтому не совместим с логическими сетями виртуальных машин.

- Режим 1 (active-backup) - переводит все сетевые интерфейсы в резервное состояние, в то время как один сетевой интерфейс остается активным. В случае отказа активного сетевого интерфейса один из резервных интерфейсов заменяет сбойный интерфейс в качестве единственного активного сетевого интерфейса в бонде. MAC-адрес соединения в режиме 1 виден только на одном порту, чтобы предотвратить путаницу, которая может возникнуть, если MAC-адрес соединения изменится на MAC-адрес активной сетевой интерфейсной карты. Режим 1 обеспечивает отказоустойчивость и поддерживается в zVirt.
- Режим 2 (XOR) - выбирает сетевой интерфейс, через который будут передаваться пакеты, на основе результата операции XOR над MAC-адресами источника и назначения по модулю количества сетевых интерфейсов. Этот расчет гарантирует, что для каждого используемого MAC-адреса назначения будет выбрана одна и та же карта сетевого интерфейса. Режим 2 обеспечивает отказоустойчивость и балансировку нагрузки и поддерживается в zVirt.
- Режим 3 (broadcast) - передает все пакеты всем сетевым интерфейсам. Режим 3 обеспечивает отказоустойчивость и поддерживается в zVirt.
- Режим 4 (IEEE 802.3ad) - создает группы агрегации, в которых интерфейсы имеют одинаковые настройки скорости и дуплекса. Режим 4 использует все сетевые интерфейсы в активной группе агрегации в соответствии со спецификацией IEEE 802.3ad и поддерживается в zVirt.
- Режим 5 (adaptive transmit load balancing) - обеспечивает распределение исходящего трафика с учетом нагрузки на каждый сетевой интерфейс в связке и то, что текущий сетевой интерфейс получает весь входящий трафик. Если сетевой интерфейс, назначенный для приема трафика, выходит из строя, роль приема входящего трафика возлагается на другой сетевой интерфейс. Режим 5 нельзя использовать в сочетании с bridge, поэтому он не совместим с логическими сетями виртуальных машин.
- Режим 6 (adaptive load balancing) - объединяет режим 5 (adaptive transmit load balancing) с балансировкой нагрузки при приеме для трафика IPv4 без каких-либо специальных требований к коммутатору. Для балансировки принимаемой нагрузки используется согласование ARP. Режим 6 нельзя



использовать в сочетании с bridge, поэтому он не совместим с логическими сетями виртуальных машин.

- Если сеть ovirtmgmt не используется виртуальными машинами, сеть может использовать любой поддерживаемый режим объединения.
- Если сеть ovirtmgmt используется виртуальными машинами, сеть должна использовать режимы объединения 1, 2, 3 или 4.
- По умолчанию в zVirt используется режим объединения 4 Dynamic Link Aggregation. Если ваш коммутатор не поддерживает протокол Link Aggregation Control Protocol (LACP), используйте режим 1 Active-Backup.

  Пример настройки VLAN на физической сетевой карте (в примере используется nmcli, но вы можете использовать любой инструмент):

  nmcli connection add type vlan con-name vlan50 ifname eth0.50 dev eth0 id 50
  nmcli con mod vlan50 +ipv4.dns 8.8.8.8 +ipv4.addresses 123.123.0.1/24 +ivp4.gateway 123.123.0.254

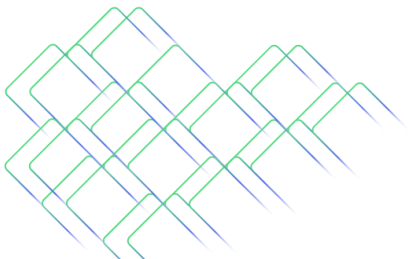
  Пример настройки VLAN поверх объединения (в примере используется nmcli, но вы можете использовать любой инструмент):

  nmcli connection add type bond con-name bond0 ifname bond0 bond.options "mode=active-backup,miimon=100" ipv4.method disabled ipv6.method ignore
  nmcli connection add type ethernet con-name eth0 ifname eth0 master bond0 slave-type bond
  nmcli connection add type ethernet con-name eth1 ifname eth1 master bond0 slave-type bond
  nmcli connection add type vlan con-name vlan50 ifname bond0.50 dev bond0 id 50
  nmcli con mod vlan50 +ipv4.dns 8.8.8.8 +ipv4.addresses 123.123.0.1/24 +ivp4.gateway 123.123.0.254
- Не отключайте сервис firewalld (межсетевой экран).

4.5. Рекомендации по развертыванию в режиме Hosted Engine

Создайте отдельный центр данных и кластер для ВМ HostedEngine и других служб уровня инфраструктуры, если ваша инфраструктура может позволить это. Хотя виртуальная машина с менеджером управления может работать на хостах в обычном кластере, отделение от остальных виртуальных машин помогает упростить план резервного копирования, производительность, доступность и безопасность.

Домен хранения, предназначенный для ВМ HostedEngine, создается во время развертывания. Не используйте этот домен хранения для других виртуальных машин.



Если ожидается большая нагрузка на хранилище, разделите сети миграции, управления и хранения, чтобы уменьшить влияние на работоспособность ВМ HostedEngine.

Все хосты способные поддерживать работу ВМ HostedEngine должны иметь одинаковое семейство процессоров, чтобы виртуальная машина могла безопасно мигрировать между ними. Если вы планируете создание кластера с хостами, имеющими различные семейства процессоров, необходимо начинать установку с самого раннего семейства.

Если ВМ HostedEngine выключается или нуждается в миграции, на хосте должно быть достаточно памяти, чтобы виртуальная машина могла перезапуститься или мигрировать на него.