Серверная доверенная виртуальная среда функционирования программных средств Numa vServer

Руководство пользователя

643.АМБН.00021-01 34 01

Листов 118

Санкт-Петербург
2023

СОДЕРЖАНИЕ

О ДОКУМЕНТЕ

Идентификация документа

Название документа	Руководство пользователя
Версия документа	Версия 1.0.4
Обозначение документа	643.АМБН.00021-01 34 01
Идентификация Изделия	Серверная доверенная виртуальная среда функционирования программных средств Numa vServer
Идентификация разработчика	ООО «НумаТех»

Аннотация документа

Настоящий документ содержит сведения, необходимые для обеспечения функционирования и настройки программного изделия Серверная доверенная виртуальная среда функционирования программных средств Numa vServer 643.АМБН.00021-01 (далее – Изделие или Numa vServer).

В документе содержатся общие сведения об Изделии, настройке, проверке, дополнительных возможностях и сообщениях, выдаваемых администратору в ходе настройки, проверки и выполнения программы, а также описание их содержания и действий, которые следует предпринять при появлении этих сообщений.

Настоящий документ соответствует требованиям к разработке эксплуатационной документации, определённым в методическом документе «Требования по безопасности информации, устанавливающие уровни доверия к средствам технической защиты информации и средствам обеспечения безопасности информационных технологий», утвержденного приказом ФСТЭК России от 02 июня 2020г. №76 по 4 уровню доверия.

Таблица 1 – Соответствие документации требованиям доверия – раздел 16 Требования к разработке Эксплуатационной документации

Требования доверия	Раздел документа, в котором представлено свидетельство
Руководство пользователя средства должно содержать описание: режимов работы средства;	пп.1.3
принципов безопасной работы средства;	Разделы 3-7 настоящего документа
функций и интерфейсов функций средства, доступных каждой роли пользователей;	пп.7.5.1., пп. 7.5.2, а также все разделы, описывающие соответствующие функции
параметров (настроек) безопасности средства, доступных каждой роли пользователей, и их безопасных значений;	Разделы 3-7 для каждой роли согласно матрице ролей, определенных в пп.7.5.1, и пп.7.5.2
типов событий безопасности, связанных с доступными пользователю функциями средства;	Разделы 3-7 для каждой роли согласно матрице ролей, определенных в пп.7.5.1, и пп.7.5.2

Требования доверия	Раздел документа, в котором представлено свидетельство
действий после сбоев и ошибок эксплуатации средства.	Соответствующие записи, находящиеся в разделах 3-7 настоящего документа

Соглашения документа

Для выделения информации в настоящем документе определены следующие типы информационных блоков:

Информация на желтом фоне – Примечание – предоставляют сведения, которые могут потребоваться для предотвращения проблем или ошибок в настройке:

Установка стороннего программного обеспечения в Numa vServer не поддерживается.

Информация на синем фоне – Заметка – информация общего характера.

Максимальные ограничения на один серверный узел Numa vServer (в скобках указаны теоретические ограничения):
до 5 Тб (16 Тб) оперативной памяти;
до 288 (4095) логических процессоров;
до 16 физических сетевых портов.

Информация на красном фоне – Предупреждение – извещают о ситуациях, которые могут нанести вред системе или оборудованию либо привести к необходимости ремонта

В процессе записи установочного образа на установочный накопитель, все содержимое накопителя будет уничтожено.

В рамке – блок кода, данные выводимые/вводимые в CLI

```
user@somepc:~# lsblk

NAME   MAJ:MIN RM   SIZE RO TYPE MOUNTPOINT
sdb      8:16   1   7,7G  0 disk
sda      8:0    0 931,5G  0 disk
├─sda2   8:2    0 930,1G  0 part /
├─sda3   8:3    0   977M  0 part [SWAP]
└─sda1   8:1    0   512M  0 part /boot/efi
```

1. ОБЩИЕ СВЕДЕНИЯ О ПРОГРАММЕ

1.1. Назначение программы

Серверная доверенная виртуальная среда функционирования программных средств Numa vServer 643.АМБН.00021-01 (далее – Numa vServer или Изделие) предназначено для построения защищенных серверных программно-аппаратных комплексов с использованием доверенной виртуальной среды функционирования механизмов разграничения и контроля доступа.

Серверы виртуализации предназначены для исполнения разнородного прикладного ПО для обработки информации, изолированного в отдельных гостевых виртуальных машинах.

1.2. Технические требования

Аппаратное обеспечение, предназначенное для исполнения Numa vServer должно соответствовать минимальным техническим требованиям (см. таблицу 2).

Таблица 2 – Технические требования к аппаратному обеспечению

№	Состав	Минимальные тех. Характеристики	Рекомендуемые тех. характеристики	Примечание
1	Процессор	1 шт., 2 ядра, базовая частота не менее 1.5 ГГц, архитектура Intel x84-64 или AMD 64, расширения виртуализации Intel-VT или AMD-V	1-2 шт., 4 ядра(или более), базовая частота не менее 2.5 ГГц, архитектура Intel x84-64 или AMD 64, расширения виртуализации Intel-VT или AMD-V	Для обеспечения прямого доступа к устройствам на шине PCI, чипсет мат. платы и процессор должны поддерживать технологии аппаратной виртуализации ввода-вывода Intel VT-d или AMD-V.
2	Оперативная память	Объем не менее 4Гб	Память с коррекцией ошибок (ECC) объемом не менее 16Гб	
3	Дисковая подсистема	Жесткий диск или твердотельный накопитель объемом не менее 128Гб. Интерфейсы подключения SATA, SAS и др.	Жесткие диски или твердотельные накопители объемом не менее 250 Гб для системы и 500 Гб для хранение данных ГВМ. Интерфейсы подключения SATA, SAS и др.	При объединение серверных узлов в пул, для хранения данных ГВМ рекомендуется использовать внешние системы хранения с блочным или файловым доступом.
4	Сетевые интерфейсы	1 шт. Ethernet-адаптер, 1 порт с базовой скоростью передачи данных 100 Мбит/с	1-2 шт. Ethernet-адаптер, 2 порта с базовой скорость передачи данных 1 Гбит/с или 10 Гбит/с	Для передачи данных ГВМ и данных управления серверным узлом не рекомендуется использовать один и тот же сетевой порт.

Помимо удовлетворения минимальным аппаратным требованиям самого Изделия, указанным в таблице 1, для функционирования каждой ГВМ серверный узел должен обладать дополнительными аппаратными ресурсами. Это относится к таким характеристикам, как объём оперативной памяти и дисковой подсистемы, тактовая частота процессоров, пропускная способность сетевых интерфейсов, шин PCI и USB и др. Минимальные требования к дополнительным ресурсам, необходимым для функционирования созданной ГВМ, рекомендуется определять исходя из минимальных системными требованиями операционной системы.

Максимальные ограничения на один серверный узел Numa vServer (в скобках указаны теоретические ограничения):

до 5 Тб (16 Тб) оперативной памяти;

до 288 (4095) логических процессоров.

до 16 физических сетевых портов.

Установка стороннего программного обеспечения в Numa vServer не поддерживается.

1.3. Режимы работы Изделия

Изделие поддерживает два основных режима работы:

Штатный режима работы. Штатный режим является основным режимом работы Изделия, Изделие выполняет свои функции согласно заявленным в ТУ. Работа с Изделием в рамках ролевой модели должна выполняться согласно настоящей инструкции.

Аварийный режим работы. В Аварийный режим работы Изделие переходит автоматически в случае нарушения контроля целостности Изделия или параметров, поставленный на контроль.

Для возвращения в штатный режим работы необходимо обратиться к администратору Изделия для расследования инцидент потенциального нарушения безопасности.

2. ОРГАНИЗАЦИОННО-РАСПОРЯДИТЕЛЬНЫЕ МЕРЫ

2.1. Процедура поставки

При поставке Изделия от среды производства до среды установки ООО «НумаТех» выполняет следующие действия:

– расчет контрольных сумм дистрибутива Изделия;

– упаковка и маркировка комплекта поставки;

– передача упакованного комплекта поставки на склад готовой продукции;

– выдача и/или отправка упакованного комплекта поставки заказчику.

2.2. Комплектность поставки Изделия

Изделие подставляется в виде установочного образа Изделия, подготовленного к установке на СВТ, и комплектуется необходимой для эксплуатации Изделия документацией (далее – Комплект Изделия).

Доступны следующие типы Комплектов Изделия:

– Комплект Изделия на материальных носителях – Изделие и документация поставляются на электронном носителе с комплектом документации в соответствии с таблицей 3;

– Комплект Изделия в электронном виде – Изделие и документация поставляются в виде файлов в соответствии с таблицей 4, которые загружаются по каналам передачи данных с сетевых ресурсов ООО «НумаТех», при условии предоставления ООО «НумаТех» соответствующего доступа.

Таблица 3 – Состав комплекта поставки сертифицированного Изделия на материальных носителях

№ п/п	Наименование составной части Изделия (документа)	Кол-во	Примечание
1	Компакт диск в составе: 1. Установочный образ Изделия 643.АМБН.00021-01; 2. Документация в составе: 643.АМБН.00021-01 32 01 Руководство администратора. Установка, настройка Numa vServer 643.АМБН.00021-01 34 01 Руководство пользователя 643.АМБН.00021-01 94 01 Инструкция по проверке контрольных сумм		На электронном носителе Идентификатор СЗИ: РОСС RU.0001.4580.xxxxxx
2	Конверт для хранения компакт-диска		
3	643.АМБН.00021-01 30 01 Формуляр		В печатном виде
4	Заверенная копия сертификата соответствия требованиям по безопасности информации Системы сертификации средств защиты информации по требованиям безопасности информации (свидетельство № РОСС RU.0001.01БИ00)		В печатном виде
5	Транспортная тара		Пластиковый пакет с застежкой типа zip-lock

Таблица 4 – Состав комплекта поставки сертифицированного Изделия в электронном виде

№ п/п	Наименование составной части Изделия (документа)	Кол-во	Примечание
1	Установочный образ Изделия 643.АМБН.00021-01		В электронном виде Идентификатор СЗИ: РОСС RU.0001.4580.xxxxxx
2	Документация в составе: 643.АМБН.00021-01 32 01 Руководство администратора. Установка, настройка Numa vServer 643.АМБН.00021-01 34 01 Руководство пользователя 643.АМБН.00021-01 94 01 Инструкция по проверке контрольных сумм 643.АМБН.00021-01 30 01 Формуляр		В электронном виде
3	Копия сертификата соответствия требованиям по безопасности информации Системы сертификации средств защиты информации по требованиям безопасности информации (свидетельство № РОСС RU.0001.01БИ00)		В электронном виде

2.3. Требования по безопасной приемке Изделия

При получении Изделия заказчик должен:

– обследовать поставку на предмет полноты комплектности. Комплект поставки должен состоять из частей, описанных в п. 2.2;

– убедиться, что в документе Формуляр заполнены все необходимые графы, стоят соответствующие печати и подписи, Формуляр Изделия промаркирован Идентификатором СЗИ, аналогичный Идентификатор СЗИ отображается в Изделии (порядок просмотра Идентификатора СЗИ в Изделии описан в разделе X);

– убедиться, что компакт-диск расположен в конверте, заклеенном наклейкой с логотипом ООО «НумаТех», отсутствуют видимые признаки вскрытия конверта (в случае поставки Изделия на материальном носителе);

– ознакомится с документацией на Изделие;

– перед установкой Изделия провести контроль целостности дистрибутива Изделия согласно документу «Инструкция по проверке контрольных сумм» 643.АМБН.00021-01 94 01, входящему в комплект поставки.

2.4. Установка Изделия

Установка Изделия осуществляется согласно разделам 3 и 4 документа «Руководство администратора. Установка, настройка Numa vServer» 643.АМБН.00021-01 32 01.

3. АДМИНИСТРИРОВАНИЕ ПУЛА РЕСУРСОВ

В данном разделе описан способ создания пула ресурсов через интерфейс командной строки xe, представлена типовая, основанная на протоколе сетевого доступа к файловым системам (Network File System, NFS) конфигурация общего хранилища, приведены примеры управления ВМ, а также описаны рекомендуемые действия при отказе физического сервера.

3.1. Определение пула ресурсов

Пул ресурсов включает в себя серверы Numa vServer, соединённые друг с другом в общую единицу, на которой возможно развертывание, запуск и исполнение виртуальных машин (ВМ). При использовании общего хранилища пул ресурсов позволяет виртуальным машинам работать с любым сервером Numa vServer, имеющим достаточный для функционирования ВМ объем оперативной памяти, и в дальнейшем динамически перемещаться между серверами Numa vServer с минимальными простоями. При сбое отдельного сервера администратор может перезапустить отказавшую ВМ на другом сервере того же пула ресурсов. Если для пула включен механизм обеспечения высокой доступности (High Availability, HA), то в случае отказа сервера ВМ будут автоматически перемещены на работающий сервер. Для одного пула ресурсов поддерживаются до 64 серверов.

Пул ресурсов имеет, по крайней мере, один физический сервер, называемый мастером. Только мастер предоставляет интерфейс для администрирования всем пулом, по мере надобности мастер передает команды отдельным участникам.

Если происходит отказ мастера пула, автоматическое переизбрание мастера происходит, только если в пуле настроены механизмы обеспечения высокой доступности.

3.2. Требования для создания пула ресурсов

Пул ресурсов - это совокупность из одного или нескольких серверов, максимум до 64. Прежде чем создавать пул или присоединять сервер к существующему пулу, необходимо убедиться, что все серверы в пуле отвечают следующим требованиям.

Описание гомогенного пула:

– поставщик ЦП (Intel, AMD) должен быть одинаковым на всех ЦП на всех серверах (разработчик, модель, характеристики);

– серверы пула должны работать под управлением одной версии программного обеспечения Numa vServer.

В дополнение к аппаратным предварительным условиям, определенным ранее, существуют некоторые другие предварительные требования для конфигурации сервера, присоединяющегося к пулу:

– сервер не является членом существующего пула ресурсов;

– на сервере не настроено общее хранилище;

– на сервере не размещены ни работающие, ни приостановленные виртуальные машины;

– на виртуальных машинах сервера не ведутся активные операции, такие как завершение работы виртуальной машины;

– системное время на сервере синхронизировано с системным временем мастера (например, с помощью NTP);

– интерфейс управления сервером не агрегирован. Вы можете настроить интерфейс управления, когда сервер успешно присоединится к пулу;

– управляющий IP-адрес является статическим, либо настроен на самом сервере, либо с помощью соответствующей конфигурации на DHCP-сервере.

Серверы Numa vServer в пулах ресурсов могут содержать различное количество физических сетевых интерфейсов и иметь локальные хранилища данных различного размера. На практике часто бывает трудно получить несколько серверов с одинаковыми процессорами, поэтому допускаются незначительные изменения. Если необходимо, чтобы в вашей среде серверы с разными ЦП находились в одном и том же пуле ресурсов, вы можете принудительно объединить пул.

Серверы, предоставляющие общее хранилище NFS, iSCSI, SAMBA для пула, должны иметь статический IP-адрес или быть адресуемыми в DNS.

Хотя это не указано в технических требованиях по созданию пула ресурсов непосредственно, но преимущества пулов (например, возможность динамически выбирать, на каком хосте Numa vServer запускать ВМ) доступны, только если имеется хотя бы одно хранилище данных с общим доступом для членов пула. Если возможно, необходимо отложить создание пула до тех пор, пока подобное хранилище не станет доступно. Когда это произойдёт, рекомендуется переместить существующие ВМ и диски, которых хранились локально, в общее хранилище пула.

3.3. Создание пула ресурсов

Когда новый сервер присоединяется к пулу ресурсов, присоединяющийся сервер синхронизирует свою локальную базу данных со всей базой данных пула и наследует некоторые параметры из пула:

– конфигурация ВМ, локального и удаленного хранилища добавляется в общую базу данных пула. Если возможно или необходимо сделать ресурсы общими, они будут собраны в объединённом хосте в пуле;

– присоединяемый сервер наследует существующие общие для пула хранилища данных, также создаются соответствующие записи физического блочного устройства (PBD), так чтобы новый сервер мог автоматически получить доступ к существующему общему хранилищу;

– сетевая информация наследуется присоединяемым сервером частично: наследуются *структурные* особенности сетевого адаптера, виртуальных сетей VLAN, агрегаций сетевых интерфейсов. Данные установленных политик *не наследуется*. К подобным свойствам, которые не могут получить значения наследованием и должны быть переконфигурированы вручную, относятся:

IP-адреса интерфейсов управления, которые сохраняются из исходной конфигурации;

местонахождение интерфейса управления, которое не меняется по сравнению с исходной конфигурацией. Например, если другие серверы пула имеют свои интерфейсы управления в агрегации интерфейсов, то добавляемый сервер должен быть в обязательном порядке включён в агрегацию после вхождения в пул;

сетевые адаптеры, выделенные для соединения с хранилищем, которые должны быть пересвязаны с новым хостом через интерфейс командной строки и физические блочные устройства, которые для правильного распределения трафика должны быть переподключены (это вызвано отсутствием в операции объединения пула операции присвоения IP-адресов; без правильной настройки, выделенные для соединения с хранилищем интерфейсы бесполезны). Более подробно о настройке с помощью команд см. пп.5.10 Настройка сетевого интерфейса, выделенного для соединения с хранилищем.

Вы можете присоединить новый сервер к пулу ресурсов только в том случае, если интерфейс управления сервером находится в той же VLAN, что и пул ресурсов.

Для объединения серверов «host1» и «host2» в пул ресурсов необходимо:

– открыть консоль на сервер «host2»;

– выполнить следующую команду:

```
xe pool-join master-address=<host1 IP-address>  master-
username=<administrator_username> master-password=<password>
```

В качестве значения параметра **master-address** должно быть указано полное его IP-адрес или его доменное имя, а параметр **master-password** должен содержать пароль администратора, присвоенный при установке Numa vServer на сервер host1.

3.4. Присвоение имени пулу ресурсов

По умолчанию, серверы Numa vServer принадлежат безымянному пулу или пулу имеющие имя мастер сервера. Для присвоения нового имени пулу ресурсов необходимо выполнить нижеприведенную команду. Для поиска нужного идентификатора **pool_uuid** можно воспользоваться функцией автодополнения (клавиша «Tab»).

```
xe pool-param-set name-label=<New_pool_name> uuid=<pool_uuid>
```

3.4.1. Создание гетерогенного пула ресурсов

Numa vServer позволяет объединять в пул отличающимся по своим характеристикам серверам. Такой пул принято называть *гетерогенным*. Создание гетерогенного пула возможно при использовании технологий в процессорах Intel (FlexMigration) и AMD (Extended migration), которые обеспечивают процессор свойствами «маскирование» (*masking*) или «выравнивание» (*leveling*). Эти свойства позволяют конфигурировать процессор так, чтобы он предоставлял нужный набор функций. Позволяя создавать пулы из серверов с отличающимися характеристиками процессоров, которые поддерживают миграцию ВМ.

Использование маскирования характеристик процессора нового сервера дает возможность согласовать ключевые характеристики имеющихся в пуле серверов, в том числе:

– процессоры сервера, вступающего в пул, должны быть от того же поставщика (AMD, Intel), что и процессоры на серверах, которые уже находятся в пуле; идентичность семейства, модели и версии не обязательна;

– процессоры сервера, вступающего в пул, должны поддерживать технологию FlexMigration фирмы Intel или технологию Extended Migration фирмы AMD;

– характеристики процессоров сервера, являющихся членами пула, должны составлять подмножество набора характеристик процессоров сервера, вступающего в пул;

– сервер, вступающий в пул, и серверы, находящиеся в пуле, должны работать с одинаковыми версиями программного обеспечения Numa vServer.

Любые изменения в наборе конфигураций пула не влияют на виртуальные машины, которые в данный момент работают в пуле. Работающая ВМ продолжает использовать набор конфигураций, который был применен при запуске. Этот набор конфигураций фиксируется при загрузке и сохраняется при переносе, приостановке и возобновлении операций. Если уровень пула падает, когда к нему присоединяется сервер с более низкими характеристиками, работающую виртуальную машину можно перенести на любой сервер в пуле, кроме недавно добавленного. При перемещении или миграции виртуальной машины на другой сервер в пределах или между пулами Numa vServer сравнивает набор функций виртуальной машины с набором функций сервера назначения. Если установлено, что наборы функций совместимы, виртуальная машина может мигрировать. Это позволяет виртуальной машине свободно перемещаться внутри и между пулами независимо от того, какие функции использует виртуальная машина.

При создании гетерогенного пула, маскирование процессора и выравнивание процессора по функциональным возможностям осуществляется автоматическом режиме, и не требует ввода дополнительных команд и конфигураций.

3.5. Добавление системы хранения с общим доступом

Полный список поддерживаемых типов хранилищ (см. раздел 6 Администрирование системы хранения). В данном пункте показано как создаётся хранилище данных с общим доступом на существующем сервере NFS.

Что бы добавить общее хранилище NFS к пулу ресурсов через интерфейс CLI, необходимо:

– открыть консоль на любом сервере в пуле;

– создать хранилище данных на <server:/path> командой:

```
xe sr-create content-type=user type=nfs name-label=<Example_SR>
shared=true device-config:server=<server IP-address> device-
config:serverpath=<path>
```

Параметр **device-config:server** содержит адрес сервера, на котором запущен сервер NFS, параметр **device-config:serverpath** – путь к папке на сервере NFS. После установки **shared**=true общее хранилище будет автоматически присоединено к каждому серверу в пуле и в дальнейшем любые подключаемые к пулу серверы будут присоединены к этому хранилищу. Идентификатор UUID созданного хранилища данных будет выведен на экран.

– найти идентификатор UUID пула можно командой:

```
xe pool-list
```

– Задать хранилище по умолчанию в качестве места хранения для участников пула можно следующей командой:

```
xe pool-param-set uuid=<pool_uuid> default-SR=<sr_uuid>
```

После выполнения данной команды все будущие ВМ будут создавать свои диски в общем хранилище см. раздел 6 Администрирование системы хранения для получения информации по созданию других типов общих хранилищ.

3.6. Вывод сервера из пула ресурсов

Перед выводом сервер Numa vServer из пула убедитесь, что все ВМ запущенные на этом сервере выключены или перенесены на другие серверы, иначе сервер нельзя будет вывести из пула.

После вывода сервера из пула, сервер будет перезагружен, переинициализирован и возвращен в исходное состояние (состояние после инсталляции).

Для извлечения сервера из пула ресурсов через интерфейс командной строки необходимо выполнить следующие действия:

– открыть консоль на любом сервере в пуле;

– найти UUID сервера командой:

```
xe host-list
```

– извлечь нужный сервер из пула:

```
xe pool-eject host-uuid=<host_uuid>
```

Сервер Numa vServer будет извлечен и будет находиться в состоянии, соответствующем конфигурации сразу после инсталляции.

Нельзя извлекать сервер, если он содержит важные данные на его локальном диске, так как все данные будут удалены после извлечения. Если нужно сохранить эти данные, следует скопировать ВМ в общее хранилище пула командой: **xe vm-copy**

После извлечения сервера из пула, сохранённые на нём ВМ будут отображены в базе данных пула и видны другими участникам пула. Они не могут быть запущены, пока местоположения ассоциированных с ними виртуальных дисков в общей хранилище не станут видны серверам пула или пока диски не будут удалены. По этой причине рекомендуется перемещать любое локальное хранилище в общее хранилище при вводе нового сервера в пул, это позволит вывести из пула любой отдельный сервер без потери данных.

3.7. Подготовка серверов пула к обслуживанию

Перед проведением на сервере, входящем в пул, операций по обслуживанию необходимо заблокировать данный сервер (это предотвратит несанкционированный старт ВМ на нем), затем переместить его ВМ на другой сервер пула.

Перевод в режим обслуживания мастера пула приведет к потере данных циклической базы данных за последние 24 часа для недействующих ВМ. Причина этого - синхронизация резервных копий, которая происходит каждые 24 часа.

Для подготовки участника пула к обслуживанию и возобновления его функционирования после проведения обслуживания можно использовать следующую последовательность действий:

– выполнить команды:

```
xe host-disable uuid=<host_uuid>
xe host-evacuate uuid=<host_uuid>
```

Сначала сервер блокируется, а затем перемещаются любые работающие ВМ к другим серверам в пуле.

– выполнить требуемые операции по обслуживанию сервера;

– после выполнения операций по обслуживанию разблокировать сервер:

```
xe host-enable uuid=<host_uuid>
```

После разблокировки обслуженного сервера на нем можно запустить ВМ.

4. ОБЕСПЕЧЕНИЯ ВЫСОКОЙ ДОСТУПНОСТИ (HIGH AVAILABILITY)

4.1. Описание механизма высокой доступности

Механизм обеспечения высокой доступности (далее – HA) является набором автоматических характеристик, спроектированных для учёта ситуаций (проблем), которые делают серверы недоступными, а также для безопасного восстановления функционирования после таких ситуаций. Например, механизм HA эффективен при физических разрывах передачи данных по сети или неполадок аппаратного обеспечения хоста.

Механизм HA должен ВСЕГДА использоваться в системах с многоканальным подключением к хранилищу и агрегацией сетевых интерфейсов.

В первую очередь, механизм HA позволяет запущенным ВМ при нестабильной работе хоста или его недоступности прекратить работу на нём и возобновить ее в другом месте. При этом можно избежать сценария, когда ВМ начинают работу (автоматически или вручную) на новом сервере, а предыдущий их сервер восстанавливает работу, что может привести к запуску экземпляров ВМ на разных серверах с высокой вероятностью порчи и потери данных ВМ.

Механизм HA автоматически восстанавливает административный контроль над пулом в случае, если мастер пула становится недоступным или его работа становится нестабильной.

В некоторых случаях механизм HA может также автоматизировать процесс перезапуска ВМ на исправных серверах. Для удобства последовательного запуска сервисов эти ВМ могут распределяться в группы (это даёт возможность запускать административные ВМ прежде зависящих от них ВМ – например, сервер DHCP раньше, чем зависящий от него сервер SQL).

Механизм HA спроектирован для работы с многоканальным подключением хранилищ и агрегированными сетевыми интерфейсами, и они должны быть сконфигурированы ПЕРЕД активацией HA. Если этого не было сделано, то вследствие нестабильности сетевого оборудования может возникнуть непредвиденное поведение хоста при перезагрузке (так называемое *самоизолирование*, *selffensing*)

4.1.1. Переполнение пула

Пул считается *переполненным* (*overcommitted*), если работающие ВМ не могут перезапуститься на одном из прочих серверов пула вследствие достижения в пуле критического количества отказов хостов. Данный показатель задаётся администратором.

Переполнение происходит, если не хватает свободной оперативной памяти в пуле для запуска ВМ, испытывающих отказ. Могут существовать другие малозаметные изменения, ухудшающие работу HA: изменения в виртуальных блочных устройствах (VBD) и сетях могут повлиять на выбор хоста, на котором будет перезапущена ВМ. В настоящий момент Numa vServer не может контролировать все действия, которые ведут к нарушению требований работы HA. При нарушении работы HA высылается асинхронное уведомление.

Numa vServer в реальном времени разрабатывает и осуществляет план обеспечения отказоустойчивости серверов в пуле (*failover plan*), определяющий действия в случае отказа некоторого количества серверов пула за некоторое заданное время. Важным для понимания является параметр *максимального некритического количества отказов хостов* (*host failures to tolerate*). Например, если пул ресурсов состоит из 6 серверов и некритическое количество отказов равняется 3, то пул рассчитывает план обеспечения отказоустойчивости, который позволяет при отказе любых трёх серверов продолжить работу ВМ на других серверах. Если отказывает большее количество, пул считается переполненным. План динамически пересчитывается с учётом анализа операций рабочего цикла и миграций ВМ. Если в процессе

изменений (например, добавления новых ВМ в пул) появляется опасность переполнения пула, то система может выслать предупреждение (например, по электронной почте).

4.1.1.1. Предупреждение о переполнении

Если при старте или продолжении работы ВМ пул переполняется (*overcommitted*), система предупреждает об этом. Это предупреждение отображается при отсутствии доступного графического интерфейса в терминал API. Если заданы соответствующие настройки, сообщение может быть отправлено администратору по электронной почте. Затем будет предложено закончить операцию или продолжить ее. Продолжение операции приведет к переполнению пула. Количество информации, используемой ВМ с различными приоритетами, будет отображено в пуле и на хостах.

4.1.1.2. Изоляция сервера

В случае если происходит отказ сервера, теряется структура коммутации сети или возникает проблема с управляющим стеком, сервер самоизолируется (самоограживается) для того, чтобы исключить запуск одних ВМ на двух серверах одновременно. После запуска изоляции сервер немедленно перезагружается, а работа всех ВМ прекращается. Остальные серверы регистрируют остановку ВМ и далее ВМ перезагружаются в соответствии с определенными для них приоритетами. Изолированный сервер начинает процесс перезагрузки и после перезапуска пытается войти заново в пул ресурсов.

4.1.1.3. Требования к конфигурации механизма HA

Чтобы использовать функцию высокой доступности, необходимо:

– пул из серверов (обеспечивает высокую доступность на уровне сервера в рамках одного пула ресурсов);

> Рекомендуется использовать механизм HA только в пулах, состоящих из не менее чем трех серверов.

– статические IP-адреса для всех серверов;

> Если IP-адрес сервера изменяется, когда включена высокая доступность, механизм высокой доступности предположит, что сеть хоста вышла из строя. Изменение IP-адреса может заблокировать хост и оставить его в не загружаемом состоянии. Чтобы исправить эту ситуацию, отключите механизм высокой доступности с помощью команды: xe host-emergency-ha-disable, сбросьте мастера пула с помощью команды: xe pool-emergency-reset-master, а затем снова включите высокую доступность.

– общее хранилище, доступное по протоколам iSCSI, NFS, SMB или Fibre Channel, для создания служебного томов (*heartbeat SR)* объемом 365 Мб или более;

Механизм высокой доступности создаёт два тома на *heartbeat SR*:

– том объёмом 4 Мбайта на *heartbeat;*

– том объёмом 256 Мбайт для хранения метаданных мастера пула, которые будут использованы в случае отказа мастера.

> Для большей надежности настоятельно рекомендуется для *heartbeat* SR использовать общие хранилища данных на основе протоколов NFS или iSCSI, которые не используются в других целях.

Хранилище, подключенное с использованием протоколов SMB или iSCSI, при проверке подлинности с использованием CHAP не может использоваться в качестве *heartbeat* SR.

В случае использования хранилищ на основе NetApp или EqualLogic необходимо вручную задать адрес дискового устройства (LUN) NFS или iSCSI в массиве данных для использования в качестве хранилища под *heartbeat SR.*

– для максимальной надежности рекомендуется использовать выделенный сетевой интерфейс для сети управления высокой доступностью.

Чтобы виртуальная машина была защищена механизмом высокой доступности, она должна быть мобильной. Это означает, что:

– виртуальные диски ВМ должны быть в общем хранилище. Можно использовать любой тип общего хранилища. Только для диска на основе iSCSI, NFS или Fibre Chanel, который предполагается использовать для heartbeat SR, требуется номер логического устройства (Logical Unit Number, LUN). Опционально эти диски также могут быть использованы в качестве обычных виртуальных дисков;

– ВМ могут использовать живую миграцию;

– у ВМ отсутствует соединение с физическими DVD-приводом и USB устройствами;

– у ВМ есть собственные виртуальные сетевые интерфейсы в сети пула.

При включенном механизме HA настоятельно рекомендуется агрегировать интерфейс управления на серверах пула, а также использовать многопоточные (multipath) хранилища в основе heartbeat SR.

При создании VLAN и агрегированных интерфейсов они могут оказаться не подключенными. В этой ситуации ВМ не будут под защитой механизма HA. В этом случае нужно использовать команду xe pif-plug для ввода VLAN и PIF-объектов агрегаций сервера в действие, благодаря чему ВМ смогут стать мобильными. Точно определить, почему ВМ не являются мобильными, можно, используя команду xe diagnostic-vm-status для анализа существующих ограничений.

4.2. Приоритет запуска и перезапуска ВМ

Виртуальные машины могут иметь приоритеты запуска protected (защищенная), best-effort (лучшая попытка) или unprotected (незащищенная). Приоритет задается через параметр ha-restart-priority в настройках виртуальной машины. Поведение перезапуска для виртуальных машин в каждом приоритете отличается.

4.2.1. Приоритет protected

Механизм высокой доступности гарантирует запуск защищённой виртуальной машины при сбое сервера или при ее отключении, при условии, что пул не переполнен и ВМ является мобильной.

Если защищенная виртуальная машина не может быть запущена (например, при переполнении пула), сервисы высокой доступности будут пытаться запустить виртуальную машину до тех пор, пока ВМ не запустится на освободившихся или дополнительных ресурсах пула.

Значение: ha-restart-priority: restart

4.2.2. Приоритет best-effort

При сбое виртуальных машин с приоритетом best-effort, сервисы высокой доступности буду пытаться запустить ВМ на другом сервере, но попытки запуска начнутся только после того как будут запущены защищенные ВМ.

Попытка запуска выполняется один раз, если она не удалась, то ВМ остается в выключенном состоянии и попытки запуска больше не выполняются.

Значение: ha-restart-priority: best-effort

4.2.3. Приоритет unprotected

Если незащищенная виртуальная машина или сервер, на котором она работает, остановлена, механизм высокой доступности не пытается перезапустить виртуальную машину.

Значение: ha-restart-priority: пустая строка

Механизм высокой доступности никогда не останавливает и не мигрирует работающие виртуальные машины для освобождения ресурсов в пуле для запуска виртуальных машин с приоритетами protected и best-effort.

Если в пуле возникают сбои сервера и число допустимых сбоев падает до нуля, защищенным виртуальным машинам не гарантируется запуск. В таких случаях генерируется системное предупреждение. Если происходит другой сбой, все виртуальные машины, для которых установлен приоритет protected, ведут себя в соответствии с приоритетом best-effort.

4.3. Порядок запуска

Порядок запуска - это порядок, в котором серверы с включенной высокой доступностью будут пытаться перезапустить защищенные виртуальные машины при возникновении сбоя. В значения свойства **order** для каждой защищенной виртуальной машины определяют порядок запуска.

Свойство order ВМ используется сервисами высокой доступностью, а также другими функциями, которые запускают и завершают работу ВМ. Любая ВМ может иметь набор свойств order, а не только защищенные ВМ. Однако сервисы высокой доступности используют свойство order только для защищенных виртуальных машин.

Значение свойства order является целым числом. Значение по умолчанию равно 0, что является наивысшим приоритетом. Защищенные ВМ со значением order 0 перезапускаются первыми. Чем выше значение свойства order, тем позже в последовательности перезапускается виртуальная машина.

Значение свойства order виртуальной машины можно задать с помощью интерфейса командной строки:

```
xe vm-param-set uuid=<vm_uuid> order=<число>
```

4.4. Активация механизма высокой доступности в пуле

Механизм HA может быть активирован в пуле при помощи интерфейса командной строки. Там же необходимо установить для различных ВМ уровни приоритета перезапуска при переполнении (*overcommitting*) пула.

При активации HA некоторые операции, которые могут негативно сказаться на плане перезапуска ВМ, например, извлечение сервера из пула, могут бездействовать.

Для включения механизма HA для выбранного пула следует выполнить следующую последовательность действий:

– проверить, что к пулу присоединено совместимое хранилище данных. Совместимыми являются хранилища на основе протоколов iSCSI, NFS или Fibre Channel (подробнее см.6.1.6 Создание и настройка хранилищ данных);

– для каждой виртуальной машины, которую необходимо защитить, установить приоритет перезапуска и порядок запуска, выполнив команду:

```
xe vm-param-set uuid=<vm_uuid> ha-restart-priority=restart order=1
```

– включить механизм HA в пуле и при необходимости указать время ожидания:

```
xe pool-ha-enable heartbeat-sr-uuids=<sr_uuid> ha-
config:timeout=<время в секундах>
```

Timeout - это период, в течение которого сеть или хранилище недоступны узлам в пуле. Если не указать timeout при включении высокой доступности, то timeout по умолчанию будет 30 секунд. Если какой-либо сервер не может получить доступ к сети или хранилищу в течение времени ожидания, он самостоятельно перезапустится.

– запустить команду **pool-ha-compute-max-host-failures-to-tolerate**. Эта команда вернет заданное максимальное количество хостов, отказ которого считается допустимым с точки зрения механизма HA:

```
xe pool-ha-compute-max-host-failures-to-tolerate
```

Допустимое (некритическое) количество отказов определяется моментом отправки тревоги: система пересчитывает план отказоустойчивости в зависимости от изменения состояния пула и таким образом система определяет объём памяти пула для определения количества некритических отказов для надежной работы защищенных ВМ. Система сигнализирует о тревоге, если расчетный уровень падает ниже заданного для **ha-host-failures-to-tolerate**.

– задать значение параметра допустимого количества отказов в пуле. Оно должен быть не более рассчитанного на предыдущем шаге значения:

```
xe pool-param-set ha-host-failures-to-tolerate=<число серверов>
uuid=<pool-uuid>
```

4.5. Отключение высокой доступности ВМ

Чтобы отключить функции высокой доступности для виртуальной машины, необходимо использовать команду **xe vm-param-set**, чтобы задать для параметра **ha-restart-priority** пустую строку. Установка параметра **ha-restart-priority** не сбрасывает настройки порядка запуска. При необходимости можно снова включить высокую доступность для виртуальной машины, установив для параметра **ha-restart-priority** значение **restart** или **best-effort**.

```
xe vm-param-set uuid=<vm_uuid> ha-restart-priority=<restart|best-
effort>
```

4.6. Восстановление недоступного сервера

Если по какой-то причине сервер не может получить доступ к файлу состояния высокой доступности, то возможно он стал недоступен. Для того чтобы восстановить установки необходимо деактивировать высокую доступность следующей командой:

```
xe host-emergency-ha-disable --force
```

Если сервер был мастером пула, он должен возобновить обычную работу с деактивированной высокой доступностью. Подчиненные серверы соединяются заново и автоматически деактивируют высокую доступность. Если сервер был подчиненным участником пула и не может соединиться с мастером, необходимо принудительно перезагрузить сервер как мастер пула и/или направить к новому мастеру:

```
xe pool-emergency-transition-to-master uuid=<host_uuid>
xe pool-emergency-reset-master master-
address=<new_master_hostname_or_ip-address>
```

После успешного перезапуска всех серверов следует активировать высокую доступность снова:

```
xe pool-ha-enable heartbeat-sr-uuid=<sr_uuid>
```

4.7. Выключение сервера при активированном механизме высокой доступности

Если происходит завершение работы или перезагрузка хоста, то работающий сервис высокой доступности может решить, что произошел сбой сервера. Для корректного завершения работы сервера, с включенной высокой доступностью, сначала деактивируйте сервер, затем извлеките его и завершите его работу. Для завершения работы сервера следует выполнить следующую последовательность команд:

```
xe host-disable host=<host_name>
xe host-evacuate uuid=<host_uuid>
xe host-shutdown host=<host_name>
```

4.8. Выключение ВМ с приоритетом protected

ВМ не может завершить работу при активированной высокой доступности. Для завершения работы ВМ сначала следует деактивировать высокую доступность, а затем выключить ВМ.

Если завершить работу ВМ из гостевой системы, то ВМ автоматически перезапустится. Для завершения работы следует сначала деактивировать высокую доступность в параметрах ВМ.

5. АДМИНИСТРИРОВАНИЕ СЕТИ

В данном разделе описано администрирование сетей Numa vServer, включая сети VLAN, и методы агрегирования сетевых адаптеров (*NIC bonds*), а также методы управления конфигурацией сети и устранения неполадок.

Следует заметить, что используемым по умолчанию в Numa vServer сетевым стеком является виртуальный коммутатор, однако при желании администратор может использовать стек сети на основе Linux Bridge (см. пп.5.2 Сети на основе стека виртуального коммутатора).

Для получения непосредственных практических инструкций можно использовать следующие разделы:

– о создании сетей для одиночных хостов Numa vServer (см. пп.5.9.1 Создание сетей на автономном сервере);

– об организации сети между хостами Numa vServer, объединёнными в пул (*resource pool*) (см. п.п 5.9.2 Создание сетей в пуле ресурсов);

– о создании VLAN для хостов Numa vServer (см. пп. 5.9.3 Создание виртуальных локальных сетей (VLAN));

– о настройке агрегаций сетевых интерфейсов (адаптеров) одиночных хостов Numa vServer (см. пп. 5.9.4 Агрегирование сетевых адаптеров автономного сервера);

– о настройке агрегаций сетевых интерфейсов (адаптеров) для серверов Numa vServer, объединённых в общий пул (см. пп. 5.9.5 Агрегирование сетевых адаптеров в пуле).

Термин «интерфейс управления» («управляющий интерфейс», *management interface*) используется далее для обозначения сетевого интерфейса, который служит для передачи управляющего трафика.

5.1. Поддержка сетевых интерфейсов

Numa vServer поддерживает до 16-ти физических сетевых интерфейсов (или до 8-и агрегированных сетевых интерфейсов) на одном сервере и до 7-ми виртуальных сетевых интерфейсов на одну ВМ.

Numa vServer предоставляет автоматическую настройку и управление сетевыми адаптерами посредством интерфейса командной строки. Не рекомендуется редактировать конфигурационные файлы сети напрямую.

5.2. Сети на основе стека виртуального коммутатора

Виртуальный коммутатор значительно упрощает администрирование виртуальных сетей – все настройки и данные статистики ВМ остаются связанными с ВМ, даже если она мигрирует с одного сервера общего пула на другой.

При использовании контроллера программно-определяемых сетей (SDN-controller) поддерживающего протокол Openflow, виртуальные коммутаторы обеспечивают дополнительную функциональность, например, списки управления доступом (Access Control List, ACL).

Контроллер программно-определяемых сетей не входит в стандартный состав дистрибутива Numa vServer

Чтобы определить, какой сетевой стек в настоящее время настроен, необходимо выполнить следующую команду:

```
xe host-list params=software-version
```

В полученных результатах следует искать «network_backend». Если в качестве сетевого стека настроен vSwitch, в результатах выполнения команды будет выведена следующая строка:

```
network_backend: openvswitch
```

Если в качестве сетевого стека настроен Linux-мост (Linux bridge), в результатах выполнения команды будет выведена следующая строка:

```
network_backend: bridge
```

Чтобы вернуться к стеку Linux-моста, необходимо выполнить следующую команду:

```
xe-switch-network-backend bridge
```

После её выполнения следует перезагрузить сервер.

Стек Linux-моста не поддерживает протокол openflow, межсерверные частные сети (Cross Server Private Networks) и не может управляться посредством контроллера программно-определяемых сетей.

5.3. Описание сетевых возможностей Numa vServer

В этом разделе описываются общие концепции, используемые при построении сетей в среде Numa vServer.

Во время установки Numa vServer автоматически создаётся по одной сети для каждого физического интерфейса сетевой карты. Когда администратор добавляет сервер в общий пул ресурсов, сети по умолчанию объединяются так, чтобы все физические сетевые адаптеры с одним именем устройства оказались прикреплены к одной и той же сети.

Как правило, администратор добавляет новую сеть только при необходимости создать внутреннюю сеть, создать новый VLAN или для создания агрегации сетевых адаптеров (*NIC bond*).

В Numa vServer может быть настроено четыре различных типа сетей:

– внешние сети (*external networks*), ассоциированные с физическим интерфейсом и обеспечивающие мост между виртуальными машинами и физическим сетевым интерфейсом, подключенного к физическим сетям передачи данных;

– агрегированные сети (*bonded networks*) создают связь между двумя и более физическими сетевыми интерфейсами для получения единого высокопроизводительного и высокодоступного канала между виртуальной машиной и физической сети передачи данных;

– частные сети одиночного сервера (*single-server private networks*) не имеют никакой связи с физическим сетевым интерфейсом и могут использоваться для обеспечения соединения между виртуальными машинами на данном сервере, без возможности связи со внешними сетями;

– межсерверные частные сети (*cross-server private networks*) позволяет виртуальным машинам, развёрнутым на разных серверах, общаться друг с другом при помощи программно-

определяемых частных сетей (для создания такой сети обязательно наличии контроллера программно-определяемых сетей).

Поведение некоторых сетевых функций различается в зависимости от того, используется один сервер Numa vServer или же пул серверов. Данный раздел содержит информацию о функциях, выполняющихся в обоих случаях, сопровождая описание каждой из них дополнительной информацией.

5.4. Сетевые объекты

В этом разделе описывается три типа сетевых объектов, существующих на стороне сервера и являющихся отражением трёх независимых сетевых сущностей:

– PIF (Physical Interface File), является отображением физического сетевого адаптера на сервере Numa vServer. Объекты PIF имеют имя, описание, UUID, параметры сетевой карты, которую они представляют, а также сеть и сервер, к которым они подключены;

– VIF (Virtual Interface File), является отображением виртуального сетевого адаптера на виртуальной машине. Объекты VIF имеют имя, описание, UUID, а также сеть и виртуальную машину, к которой они подключены;

– Network (сеть), представляет собой виртуальный Ethernet-коммутатор на сервере. Объекты типа Network имеют имя, описание, UUID и набор подключенных к ним VIF и PIF.

Интерфейс командной строки позволяет настраивать параметры работы в сети, управлять тем, какой из сетевых интерфейсов используется для передачи управляющего трафика, а также настраивать дополнительные возможности сети: виртуальные локальные сети (VLAN) и агрегации сетевых адаптеров (NIC bonds).

5.4.1. Сетевой объект Network (сеть)

Каждый сервер Numa vServer имеет одну или более сетей, являющихся виртуальными Ethernet-коммутаторами. Сети, не связанные с PIF, считаются внутренними и могут быть использованы для обеспечения связи только между виртуальными машинами, без возможности связи со внешними сетями. Сети, ассоциированные с PIF, считаются внешними и являются связующим звеном между VIF и PIF, обеспечивают подключение ВМ к сетям передачи данных доступным через физический интерфейс сетевого адаптера.

5.5. Виртуальные локальные сети (VLAN)

Виртуальные локальные сети (VLAN), согласно стандарту IEEE 802.1Q, позволяют создавать на одной физической сети несколько логических сетей. Сети Numa vServer могут работать с виртуальными локальными сетями различными способами.

Все поддерживаемые конфигурации виртуальной локальной сети в равной степени применимы как к пулам, так и к автономным серверам, как в случаях использования агрегации сетевых интерфейсов, так и без таковой.

5.5.1. Использование VLAN с управляющими интерфейсами

Интерфейс управления может быть настроен на VLAN на порту коммутатора, настроенного как магистральный порт (trunk port) или порт режима доступа (access port).

5.5.2. Использование VLAN с виртуальными машинами

Порты коммутатора, настроенные согласно стандарту IEEE 802.1Q в качестве магистральных для виртуальных сетей, могут быть использованы в сочетании с функциями Numa vServer VLAN для подключения гостевых виртуальных сетевых интерфейсов (VIF) к конкретным

VLAN. В этом случае сервер выполняет функции VLAN Tagging / Untagging для гостевой системы, которой недоступны никакие сведения о конфигурации VLAN.

Виртуальные локальные сети сервера представляются дополнительными PIF-объектами в соответствии с заданными тегами виртуальных сетей. Сети Numa vServer могут быть подключены как к PIF-объекту, представляющему физический сетевой адаптер или к PIF-объекту, связанному с виртуальной сетью, помеченный её тегом. Подробные инструкции, касающиеся создания виртуальных локальных сетей для серверов Numa vServer (автономных или входящих в пул), приведены в пп. 5.9.3 Создание виртуальных локальных сетей (VLAN).

5.5.3. Использование VLAN с сетевыми адаптерами, выделенными для соединения с хранилищем

Сетевые адаптеры, выделенные для соединения с хранилищем (также известные как *IP-enabling NIC* или просто «интерфейсы управления») могут быть сконфигурированы таким образом, чтобы использовать порты с нативной поддержкой Native VLAN (порты *в режиме доступа*) или магистральные порты и виртуальные сети Numa vServer. Подробнее о конфигурации сетевых адаптеров, выделенных для соединения с хранилищем, см. пп. 5.10 Настройка сетевого интерфейса, выделенного для соединения с хранилищем.

5.5.4. Объединение интерфейсов управления и гостевых VLAN на сетевом адаптере автономного хоста

Порт виртуального коммутатора может быть сконфигурирован для одновременной работы с Native VLAN и Tagget VLAN, позволяя одному сетевому адаптеру сервера использоваться, например, для интерфейса управления и для того, чтобы соединить VIF-объекты гостевой системы с определенными идентификаторами виртуальных сетей.

5.6. Jumbo-кадры

Jumbo-кадры (или сверхдлинные Ethernet-кадры) могут использоваться для оптимизации производительности сетевого трафика. Jumbo-кадры называются Ethernet-кадры, содержащие больше, чем 1500 байтов полезной нагрузки. Обычно они используются для того чтобы достигнуть лучшей пропускной способности, уменьшая загрузку на память системной шины и процессора.

> Numa vServer поддерживает Jumbo-кадры только при использовании vSwitch в качестве сетевого стека сервера.

Администратор должен учитывать следующие особенности использования Jumbo-кадры:

– поддержка Jumbo-кадров настраивается на уровне пула;

– vSwitch должен быть сконфигурирован в качестве сетевого стека по умолчанию на всех сетевых устройствах в пуле;

– каждое устройство в подсети должно быть соответствующим образом настроено на использование Jumbo-кадров;

– рекомендуется, чтобы администратор включал поддержку Jumbo-кадров только для выделенной сети хранения;

– поддержка Jumbo-кадров в сети управления отсутствует;

– Jumbo-кадры так же не поддерживаются для использования на BM.

Для использования Jumbo-кадров следует установить значение параметра Maximum Transmission Unit (MTU) в диапазоне между 1500 и 9216. Это может быть сделано при помощи интерфейса командной строки *xe*.

5.7. Агрегация сетевых интерфейсов

Агрегирование сетевых адаптеров (англ. *NIC bond* или *NIC teaming*), повышают доступность и/или пропускную способность, позволяя администраторам сконфигурировать два или более сетевых адаптера вместе таким образом, чтобы они логически функционировали как единый адаптер. Все связанные таким образом сетевые адаптеры будут совместно использовать один MAC-адрес.

В случае сбоя на одном из сетевых адаптеров агрегированного порта, сетевой трафик будет автоматически перенаправлен через другой сетевой адаптер. Numa vServer поддерживает до восьми сетевых интерфейсов в одной группе агрегации.

Numa vServer поддерживает следующие режимы агрегации: активно-активный (active-active), активно-пассивный (active-passive) и агрегацию с использованием протокола LACP (*Link Aggregation Control Protocol*). Число поддерживаемых адаптеров и поддерживаемый режим связывания изменяются согласно выбранному сетевому стеку:

LACP-агрегация доступна только при использовании сетевого стека vSwitch, тогда как активно-активная и активно-пассивная агрегация – как для vSwitch, так и для сетевого стека Linux Bridge.

Когда в качестве сетевого стека выступает vSwitch, имеется возможность связать вместе от двух до четырёх сетевых адаптеров, в случае использования Linux Bridge – только два сетевых адаптера.

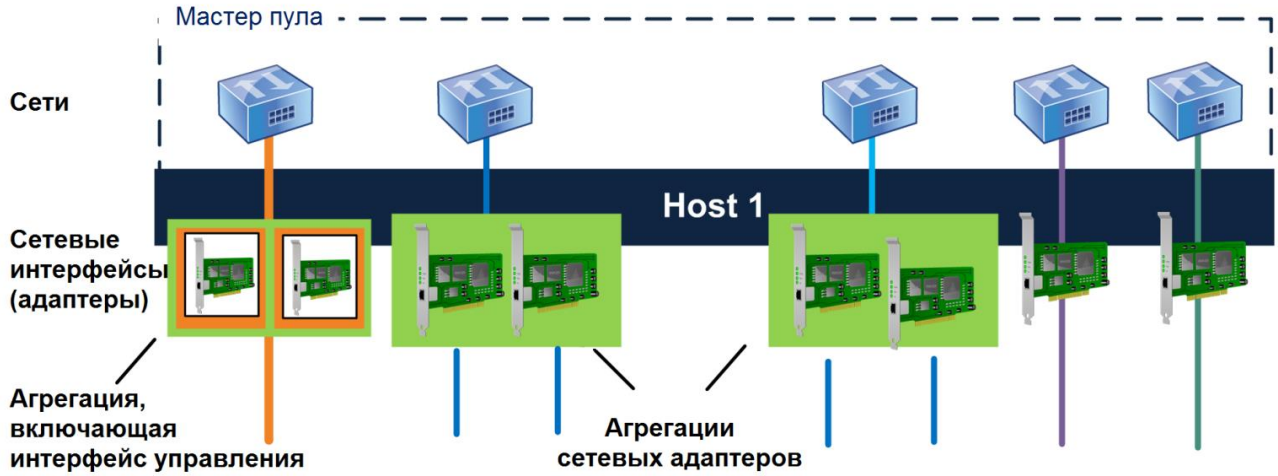

Рисунок 1 – Схема сетевого взаимодействия сервера, когда некоторые пары сетевых адаптеров агрегированы

На рисунке 1 интерфейс управления находится в группе агрегации двух сетевых адаптеров. Numa vServer будет использовать эту связь для управляющего трафика. Помимо агрегации сетевых адаптеров, предназначенной для управляющего трафика, сервер также использует другие две пары групп агрегаций и два независимых друг от друга сетевых адаптера для трафика ВМ.

Все режимы агрегации поддерживают автоматическое восстановление после сбоя (*failover*), однако не все режимы позволяют всем соединениям быть активными для всех типов трафика. Numa vServer поддерживает агрегирование следующих типов сетевых адаптеров:

– сетевые адаптеры (интерфейсы) общего назначения (не управляющие). Можно объединять сетевые адаптеры, которые Numa vServer использует исключительно для трафика ВМ. Агрегация этих сетевых адаптеров не только обеспечивает отказоустойчивость, но и также равномерно распределяет трафик многочисленных ВМ между сетевыми адаптерами;

– интерфейсы управления. Можно связать сетевой адаптер, обслуживающий трафик управления с другим сетевым адаптером так, чтобы последний обеспечил его резервирование, рекомендуемый тип такой агрегации - активно-активно;

– вторичные интерфейсы. Можно объединить сетевой адаптер, сконфигурированный в качестве вторичного интерфейса (например, для обмена данными с хранилищем). Однако для большинства серверов, выступающих в роли инициаторов обмена с системами хранения посредством протокола iSCSI, рекомендуется настройка многоканального (multipath) соединения вместо агрегации сетевых адаптеров.

Агрегацию можно создать, если VIF уже использует один из адаптеров, которые предполагается агрегировать: трафик ВМ будет автоматически перемещён на новый связанный сетевой интерфейс.

В Numa vServer агрегация сетевых адаптеров представляется дополнительным PIF-объектом. Агрегации сетевых адаптеров в Numa vServer полностью включаются в категорию базовых физических устройств (PIF).

Создание агрегации, содержащей только один сетевой адаптер, не поддерживается.

Создание агрегации не поддерживается на сетевых картах, которые передают трафик FCoE

5.7.1. Основные положения об IP-адресации агрегированных интерфейсов

Агрегированный сетевой интерфейс либо имеют один IP-адрес, либо не имеют IP-адресов, как показано ниже:

– Сети управления и сети обмена данными с хранилищем:

если агрегировать управляющий или вторичный интерфейс, единственный IP-адрес будет присвоен всей агрегации в целом. То есть отдельный интерфейс, входящий в агрегацию не имеет своего собственного IP-адреса. Агрегированное соединение рассматривается сервером как единое логическое соединение;

если агрегированные интерфейсы используются для трафика, не связанного с виртуальной машиной, например, для подключения к общему сетевому хранилищу, то ему можно настроить IP-адрес. Однако если назначен IP-адрес одному из сетевых интерфейсов (то есть создали интерфейс управления или дополнительный интерфейс), этот IP-адрес автоматически назначается всей агрегации;

если администратор агрегирует интерфейс управления или вторичный интерфейс с сетевым адаптером, не имеющим IP-адреса, агрегация принимает имеющийся IP-адрес соответствующего интерфейса автоматически;

– Сети виртуальных машин. Когда агрегированные сетевые интерфейсы используются для трафика виртуальных машин, не обязательно настраивать IP - адрес для агрегированного интерфейса. Это происходит потому, что агрегация работает на уровне 2 модели OSI, и на этом уровне не используется IP-адресация.

5.7.2. Типы агрегаций

Numa vServer предоставляет три различных типа агрегаций:

– активно-активный режим (Active-Active mode), с балансировкой трафика виртуальной машины между сетевыми интерфейсами агрегации (см. пп. 5.7.4 Активно-активный тип агрегации (Active-Active mode));

– активно-пассивный режим (Active-Passive mode), в котором только один сетевой интерфейс активно осуществляет передачу трафика (см. пп. 5.7.6 Активно-пассивный тип агрегации (Active-Passive mode));

– LACP-агрегация (LACP Link Aggregation), при которой активные и резервные сетевые интерфейсы согласованы между коммутатором и сервером (см. пп. 5.7.7 Агрегация на основе протокола LACP (Link Aggregation Control Protocol)).

При агрегировании параметры задержки Up Delay и Down Delay выставляются, соответственно, в 31 000 и 200 мс. Большие значения Up Delay являются оправданными из-за того, что некоторые коммутаторы тратят довольно существенное время на фактическое включение порта. Для получения информации об изменении задержки см. пп.5.8.2 Изменение времени задержки Up Delay трафика агрегации.

5.7.3. Состояние агрегации

Numa vServer обеспечивает регистрацию состояний для агрегированных подключений в журнале событий каждого сервера. Если один или более из агрегированных интерфейсов перестали работать или были восстановлены, в журнал заносится соответствующая запись. Аналогично, можно запросить статус интерфейсов в составе агрегации, используя параметр links-up:

```
xe bond-param-get uuid=<bond_uuid> param-name=links-up
```

Numa vServer проверяет состояние интерфейсов в агрегациях приблизительно каждые 5 секунд. Следовательно, если в пятисекундном окне произойдёт отказ ещё каких-то интерфейсов, эти отказы не будут зарегистрированы до следующей проверки состояния.

Журналы событий доступны в виде файла /var/log/journal/38855a3423a049e2be9042a56f9aa8a3/system.journal на каждом хосте.

5.7.4. Активно-активный тип агрегации (Active-Active mode)

Активно-активный (Active-Active mode) – конфигурация по принципу «активный – активный» для передачи гостевого трафика: оба сетевых адаптера могут передавать трафик ВМ одновременно. Когда агрегация используются для трафика управления, только один сетевой адаптер в связи может направить трафик: другой сетевой адаптер остается неиспользованным и обеспечивает поддержку обработки отказа(failover). Активно-активный режим является режимом по умолчанию, когда в качестве сетевого стека используется Linux Bridge или vSwitch.

Когда активно-активная агрегация используется с Linux Bridge, возможна агрегация только из двух сетевых адаптеров. При использовании vSwitch в качестве сетевого стека становится возможно агрегировать два, три, или четыре адаптера в активно-активном режиме. Однако в активно-активном режиме агрегирование трёх или четырёх сетевых адаптеров обычно эффективно только для трафика виртуальных машин.

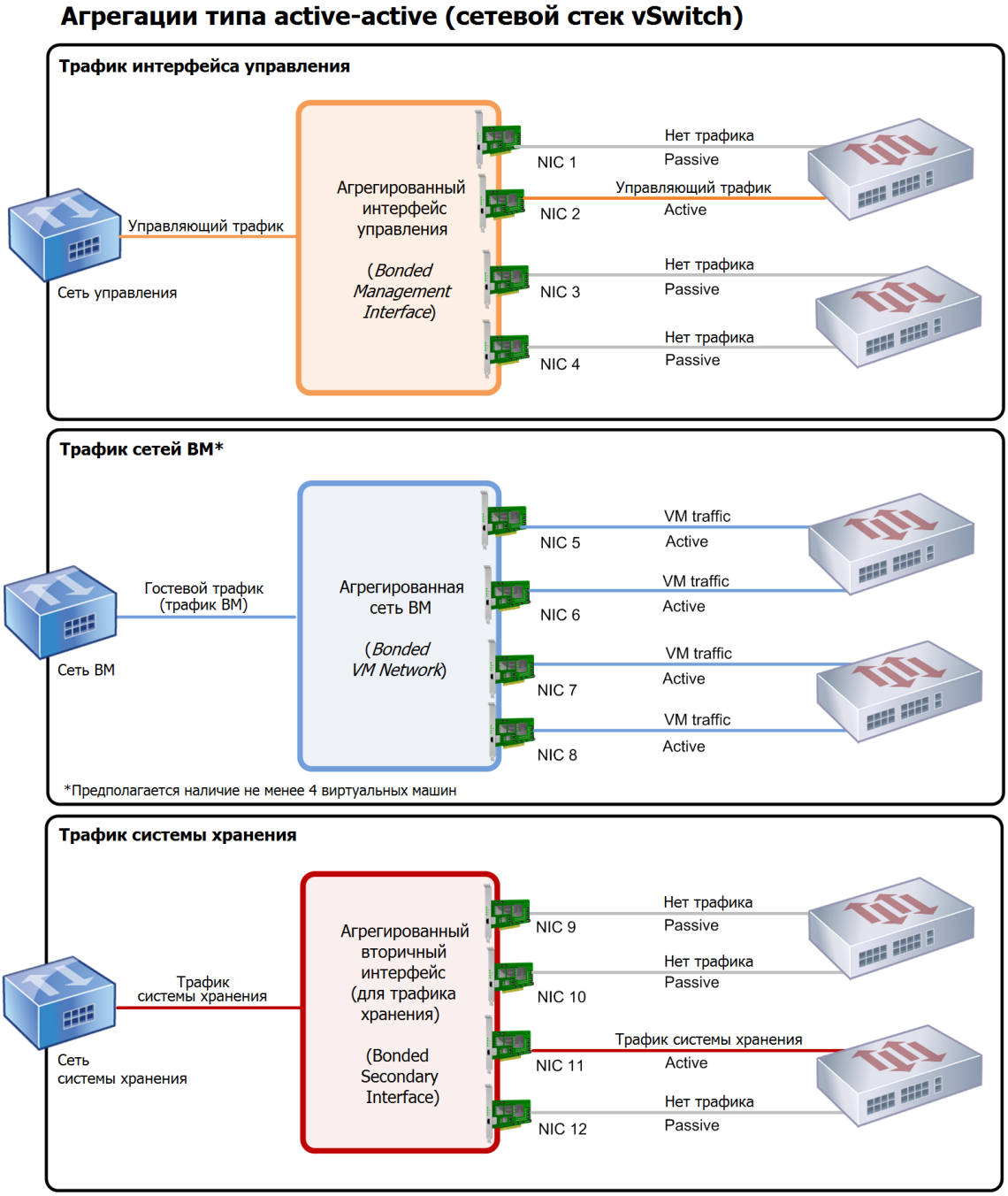

Рисунок 2 – Активно-активный тип агрегации

Numa vServer может отправлять трафик через два или большее количество сетевых адаптеров, только когда имеется более одного MAC-адреса, ассоциированного с агрегацией адаптеров. Numa vServer может использовать виртуальные MAC-адреса в VIF-объекте, чтобы распараллелить трафик через несколько соединений. В соответствии с типом трафика:

– трафик ВМ. При агрегировании сетевых адаптеров, переносящих только трафик ВМ, все соединения будут являться активными, и распределение трафика ВМ по адаптерам внутри агрегации будет балансировать между интерфейсами. Трафик отдельного VIF никогда не разделяется между сетевыми картами;

– трафик управления или обмена с системой хранения. Только одно из соединений (то есть сетевых адаптеров) в агрегации является активным, другие же остаются неиспользованными, если только активная роль не передаётся им в результате сбоя основного интерфейса. Таким образом, конфигурация, при которой интерфейс управления или вторичный интерфейс агрегируются с другими, обеспечивает отказоустойчивость.

– смешанный трафик. Если агрегация сетевых адаптеров передаёт смесь из IP-трафика обмена с хранилищем и гостевого трафика ВМ, только гостевой трафик и управляющий трафик домена пользуются возможностью балансировки трафика.

5.7.5. Балансировка трафика в режиме агрегации активно-активно

Numa vServer балансирует распределение трафика между сетевыми адаптерами на основе адреса MAC, с которого отправляется пакет. Для трафика управления существует только один MAC-адрес источника, поэтому в активном-активном режиме используется только одна сетевая карта, и трафик не распределяется. Трафик распределяется на основе двух следующих факторов:

– ВМ и связанный с ней VIF-объект, отправляющие или получающие сетевые пакеты;

– количество пересылаемых данных (в килобайтах).

Numa vServer оценивает объем данных, отправленных и полученных каждой сетевой картой, в килобайтах. Когда объем данных, отправленных на одну сетевую карту, превышает объем данных на другой сетевой карте, Numa vServer пытается перебалансировать трафик VIF-объектов между сетевыми адаптерами. Нагрузка одного VIF не делится между двумя сетевыми адаптерами.

Активно-активный режим агрегации сетевых адаптеров может обеспечить выравнивание нагрузки от трафика нескольких ВМ, этот режим не может предоставить одной ВМ пропускную способность двух сетевых адаптеров. Любой VIF-объект использует строго один из агрегированных интерфейсов за один раз. Поскольку Numa vServer периодически балансирует трафик, VIF-объекты не оказываются закрепленными за определенными сетевыми адаптерами агрегации постоянно.

Активно-активный режим иногда упоминается как Source Load Balancing-агрегация (SLB-агрегация), поскольку Numa vServer использует SLB для совместного использования потоки пакетов через агрегированные сетевые адаптеры. SLB является ответвлением Adaptive Load Balancing (ALB) с открытым исходным кодом и использует возможность ALB динамически балансировать нагрузку на сетевые адаптеры.

При повторной балансировке количество байтов, пересланных через каждое ведомое устройство (сетевой интерфейс), отслеживается в течение установленного промежутка времени. Когда пакет, который необходимо отправить, содержит MAC-адрес нового источника, этот адрес присваивается сетевому адаптеру с самым низким уровнем загруженности. Распределение трафика производится через равномерные промежутки времени.

Каждый MAC-адрес имеет соответствующую нагрузку, и Numa vServer может перемещать нагрузку между сетевыми адаптерами в зависимости от количества данных, которые ВМ отправляет и получает. Для активно-активного режима весь трафик одной ВМ может быть перераспределён только целиком на другой адаптер.

> Активно-активный режим агрегации не требует от коммутатора поддержки EtherChannel или 802.3ad (LACP).

5.7.6. Активно-пассивный тип агрегации (Active-Passive mode)

Активно-пассивная агрегация направляет трафик только по одному из сетевых адаптеров, таким образом, трафик передаётся другому сетевому адаптеру в агрегации, только если происходит сбой соединения текущего активного сетевого адаптера (failover).

Активно-пассивный тип агрегации доступен при использовании сетевого стека Linux Bridge или vSwitch. При использовании Linux Bridge можно агрегировать только два сетевых адаптера, в случае vSwitch – два, три или четыре. Однако, независимо от типа трафика, когда сетевые адаптеры агрегируются в активно-пассивном режиме, только один интерфейс является активным и отсутствует балансировка нагрузки.

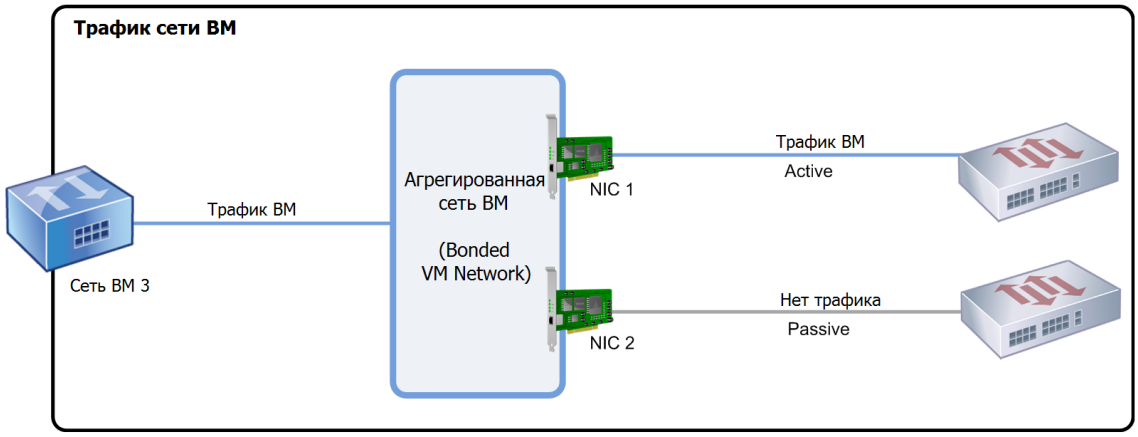

Рисунок 3 – Два сетевых адаптера, агрегированные в активно-пассивном режиме

Активно-активный тип агрегации является типом по умолчанию в среде Numa vServer, для конфигурирования агрегации активно-пассивный режиме необходимо определять этот параметр при настройке интерфейса, в противном случае интерфейс будет работать в режиме активно-активно.

Активно-пассивный режим может быть хорошим решением для обеспечения отказоустойчивости, так как предлагает несколько преимуществ. При использовании активно-пассивных агрегаций трафик передается только по одному интерфейсу, что позволяет для обеспечения избыточности использовать два коммутатора, в которых отсутствует возможность стекирования.

Активно-пассивный режим не требует от коммутатора поддержки EtherChannel или стандарта 802.3ad (LACP).

Активно-пассивный режим актуален в ситуациях, когда система не нуждается в балансировки нагрузки или когда предполагается отправка трафика через единственный сетевой адаптер.

После создания VIF-объектов или создания пула администратору следует быть чрезвычайно осторожным при внесении изменений в настройки существующих агрегаций или создания новых.

5.7.7. Агрегация на основе протокола LACP (Link Aggregation Control Protocol)

Протокол управления агрегацией каналов (Link Aggregation Control Protocol) LACP является типом агрегирования, при котором группа портов связывается вместе и обрабатывается как единый логический канал. Агрегированные каналы LACP используются как для повышения пропускной способности, так и повышения отказоустойчивости.

В отличие от других типов агрегация на основе протокола LACP требует поддержки протокола и настройки на отправителе и получателе, т.е. на сервере и коммутаторе (см. пп. 5.7.7 Агрегация на основе протокола LACP (Link Aggregation Control Protocol). Чтобы использовать LACP на сервере, в качестве сетевого стека должен использоваться vSwitch.

В таблице 2 приведено сравнение активно-активного типа агрегации SLB (балансировка на основе источника) и LACP-агрегации.

Таблица 5 – Сравнение активно-активного типа агрегации SLB и LACP-агрегации

Режим	**Преимущества**	**Ограничения**
Активно-активная агрегация SLB	не требует от коммутаторов поддержки стекирования поддерживает агрегацию до четырёх	оптимальное выравнивание нагрузки требует наличия, по меньшей мере, одного сетевого адаптера на один VIF-

Режим	Преимущества	Ограничения
	сетевых адаптеров	объект управляющий трафик, равно как и трафик обмена с хранилищем, не может быть распараллелен между несколькими сетевыми адаптерами выравнивание нагрузки происходит только при наличии нескольких MAC-адресов
LASP-агрегация	все соединения могут быть активными, независимо от типа трафика балансирование трафика не подразумевает зависимости от исходных MAC-адресов – таким образом, все типы трафика могут быть сбалансированы	коммутаторы должны поддерживать стандарт IEEE 802.3ad требует наличие управляемого коммутатора поддерживается только в сетевом стеке vSwitch требует один коммутатора или коммутаторов в стеке

5.7.8. Балансировка трафика в режиме агрегации LACP

Numa vServer поддерживает при использовании LACP-агрегации два типа хеширования (термин *хеширование* здесь относится к способу, согласно которому сетевые адаптеры и коммутатор распределяют трафик):

– балансировка нагрузки, основанная на IP-адресе и используемых портах источника и получателя;

– балансировка нагрузки, основанная на MAC-адресе источника.

В зависимости от типа хеширования и типа трафика, LACP-агрегация потенциально позволяет распределять трафик более равномерно, чем активно-активная агрегация.

Необходимо задать настройки для входящего и исходящего трафика – и на сервере, и на коммутаторе: конфигурация не обязательно должна совпадать с обеих сторон.

5.7.8.1. Балансировка нагрузки, основанная на IP-адресе и используемых портах источника и получателя

Этот тип хеширования используется по умолчанию при использовании LACP-агрегации.

Трафик, поступающий от одной виртуальной машины, может быть распределен по двум каналам при условии, что есть различия в исходном или целевом IP-адресе или номерах портов.

Если в ВМ исполняется несколько приложений, использующих различные IP-адреса или номера портов, данный тип хеширования распределяет трафик по нескольким каналам, давая гостевой системе возможность использования совокупной пропускной способности агрегированного интерфейса.

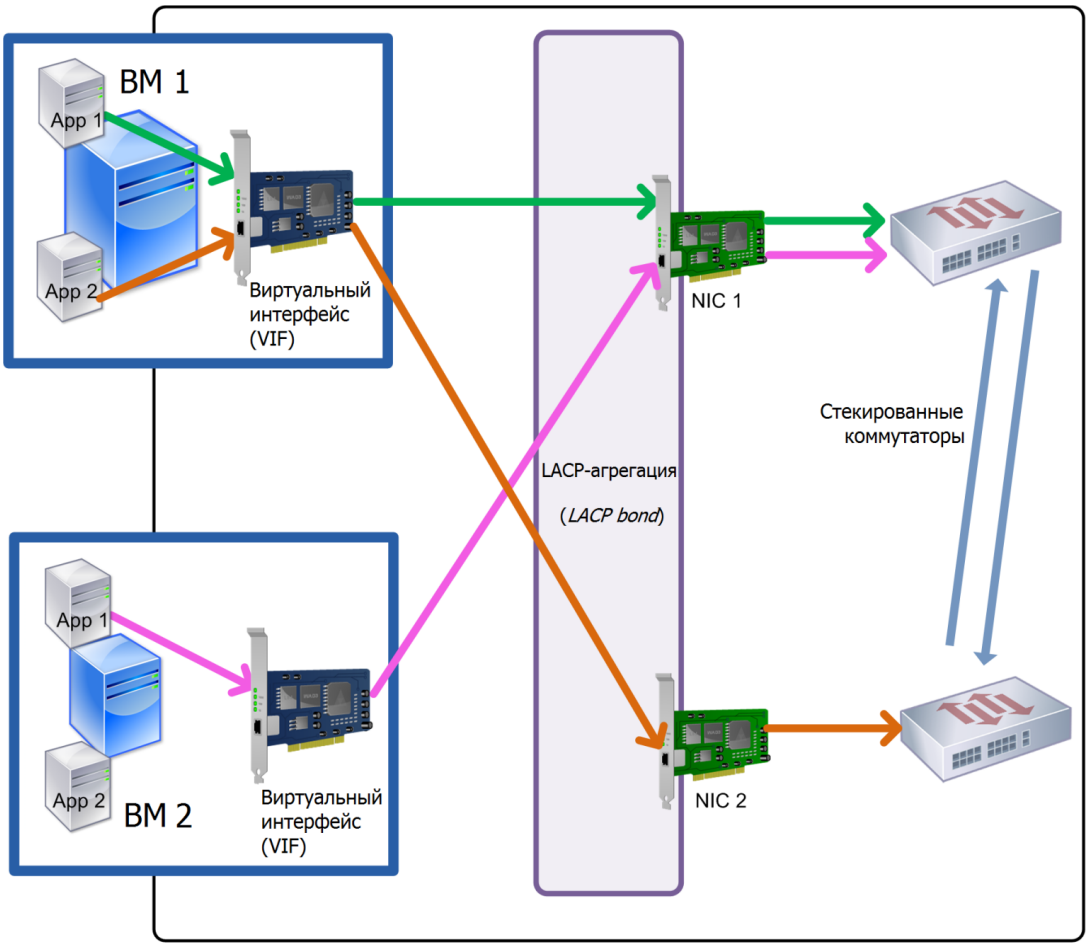

Рисунок 4 – Балансировка нагрузки, основанная на IP-адресе и используемых портах источника и получателя (стек коммутаторов)

Такой тип конфигурации агрегаций выгодно использовать при необходимости балансировать трафик двух различных приложений, исполняемых в одной ВМ.

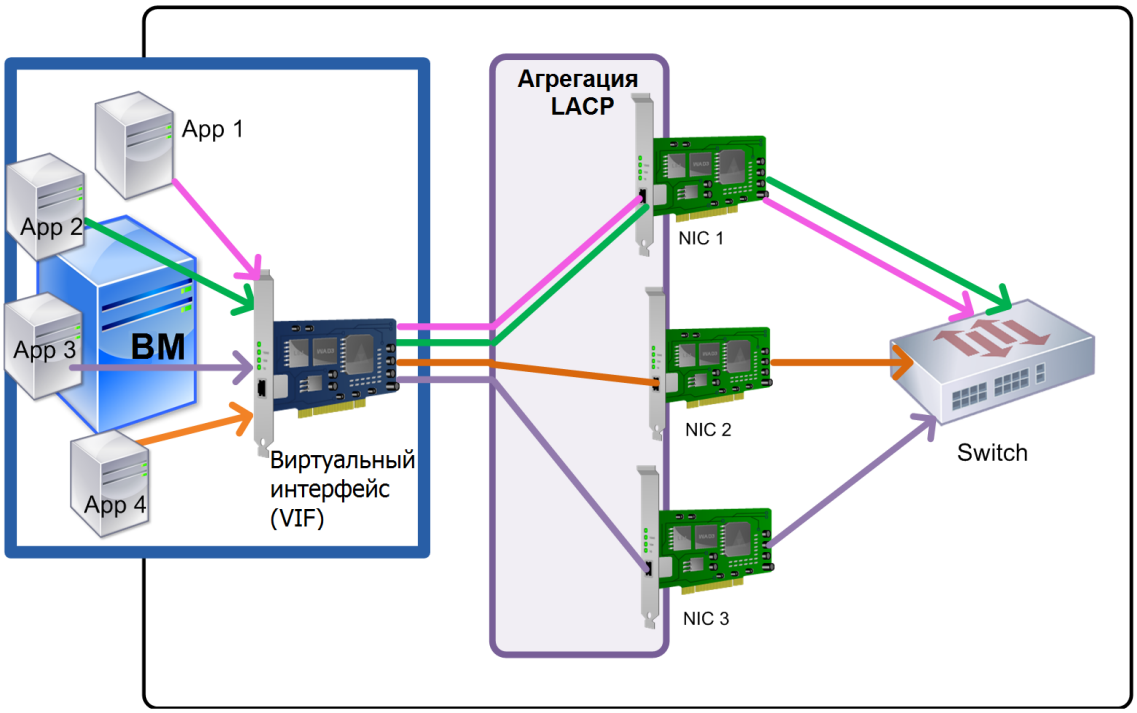

Рисунок 5 – Балансировка нагрузки, основанная на IP-адресе и используемых портах источника и получателя (один коммутаторов)

Алгоритм балансировки для этого типа хеширования использует пять факторов для распределения трафика между сетевыми интерфейсами: IP-адрес источника, номер порта источника, IP-адрес назначения, номер порта назначения и MAC-адрес источника.

5.7.8.2. Балансировка нагрузки, основанная на MAC-адресе источника

Этот тип балансировки нагрузки работает хорошо, когда есть несколько виртуальных машин на одном сервере. Трафик балансируется на основе виртуального MAC-адреса виртуальной машины, с которой исходит трафик. Numa vServer отправляет исходящий трафик, используя тот же самый алгоритм, как в случае активно-активного агрегирования. Трафик от одной виртуальной не распределяется между несколькими сетевыми адаптерами. В результате этот тип хеширования не подходит для ситуаций, когда виртуальных интерфейсов меньше, чем физических сетевых адаптеров.

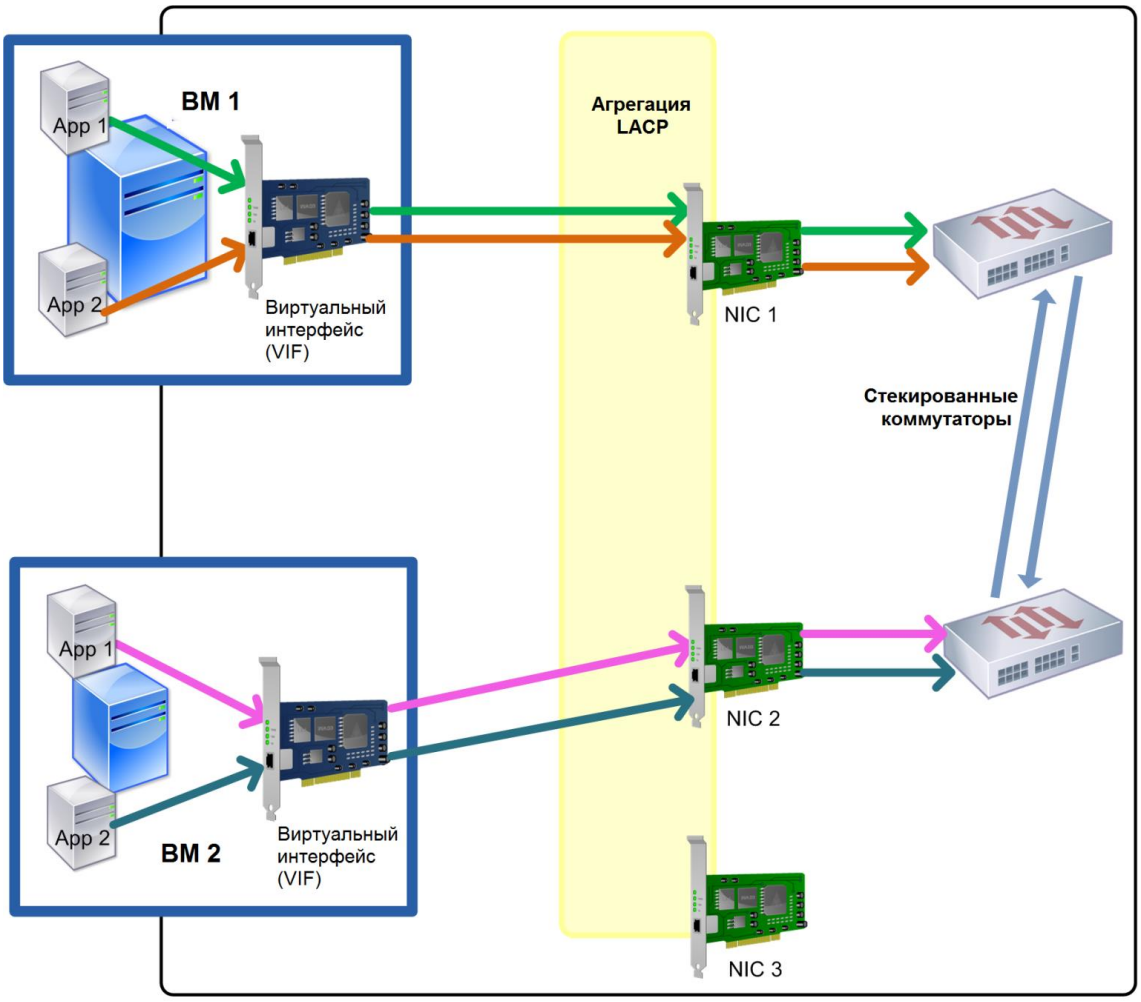

Рисунок 6 – Балансировка нагрузки, основанная на MAC-адресе источника

5.7.9. Настройка коммутатора

В зависимости от требований избыточности можно подключить агрегацию сетевых адаптеров к одному или к нескольким стекированным коммутаторам. Если вы подключаете один из сетевых адаптеров ко второму резервному коммутатору, и сетевая карта или коммутатор выходит из строя, трафик переключается на другую сетевую карту. Добавление второго коммутатора предотвращает возникновение единой точки отказа в вашей конфигурации.

Используйте стековые коммутаторы, если вы хотите подключить агрегированные сетевые адаптеры к нескольким коммутаторам и настроили тип агрегации LACP. Термин «стековые коммутаторы» относится к ситуации, когда несколько физических коммутаторов сконфигурированы так, чтобы функционировать как один логический коммутатор. Коммутаторы в стеке должны быть соединены физически, и их ПО должно обеспечивать работу коммутаторов как единого логического коммутатора. Осуществляйте настройку стека в соответствии с документацией производителя оборудования.

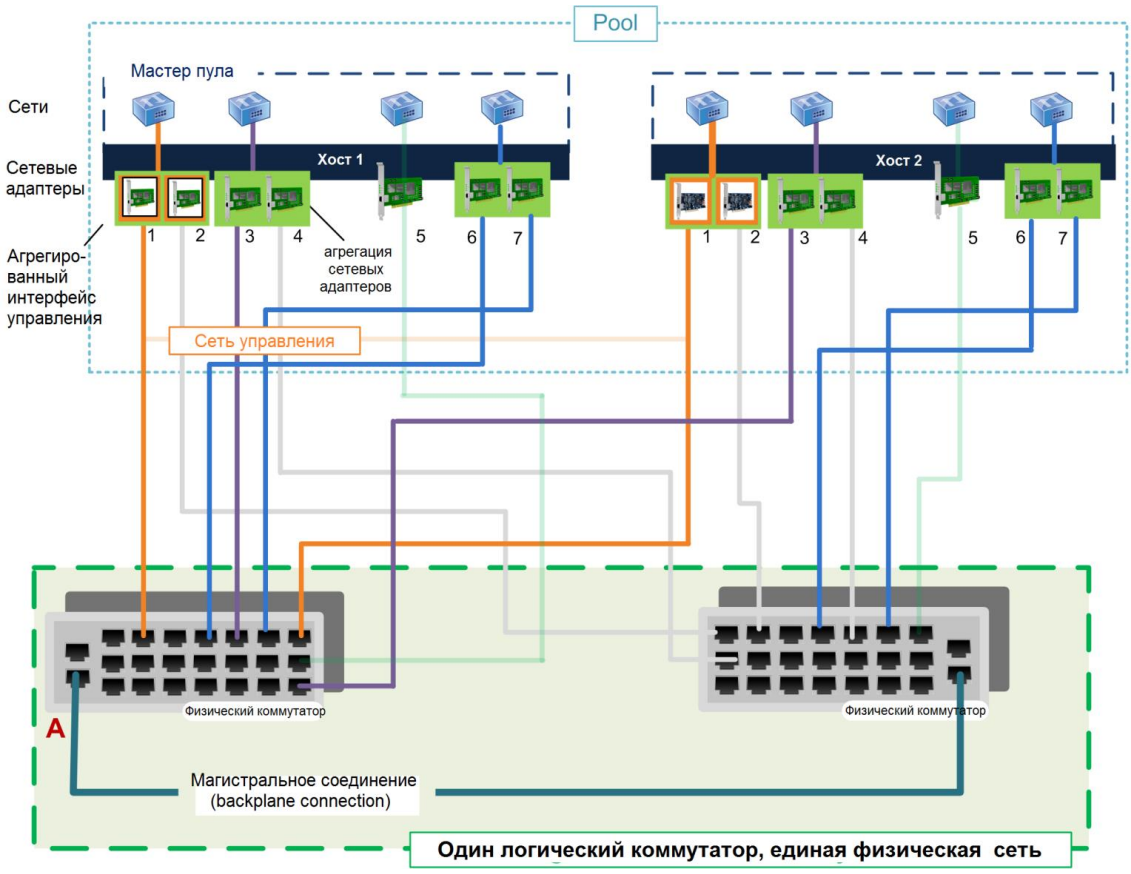

Рисунок 7 – Вариант схемы соединения серверов и стека коммутаторов

5.7.9.1. Конфигурация коммутатора для агрегаций типа LACP

Несмотря на то, что некоторые детали конфигурации коммутаторов зависят от производителя, есть несколько ключевых пунктов, которые следует помнить, настраивая коммутаторы для использования с агрегациями типа LACP:

– сам коммутатор должен поддерживать LACP и стандарт IEEE 802.3ad;

– при создании LAG-группы на коммутаторе необходимо создать одну LAG-группу для каждой агрегации LACP на сервере.

– возможно, также следует добавить к LAG-группе идентификатор VLAN;

– каналы LACP на Numa vServer требуют, чтобы параметр режима Static Mode в LAG-группе был установлен в «Disabled» (отключено).

Как упоминалось ранее, наличие стекирования в коммутаторах является необходимым условием при подключении агрегации LACP к нескольким коммутаторам.

5.8. Первоначальная конфигурация сети после установки

Некоторые сетевые параметры сервера указываются ещё во время установки: IP-адрес (DHCP), сетевой адаптер, используемый в качестве интерфейса управления.

Если сервер содержит несколько сетевых адаптеров/интерфейсов, конфигурация сети, после установки, будет зависеть от того, какой из них был выбран для администрирования во время установки:

– для каждого адаптера сервера создается отдельный ссылочный объект на физический сетевой интерфейс (PIF или PIF-объект);

– физическому сетевому интерфейсу сетевого адаптера, используемому в качестве интерфейса управления, присваиваются настройки IP-адреса, указанные во время установки;

– для каждого физического сетевого интерфейса создается сеть («Network 0», «Network 1» и т. д.);

– каждая сеть связана с одним PIF-объектом;

– настройки IP-адреса всех остальных физических интерфейсов остаются неустановленными.

В случае если сервер содержит единственный сетевой адаптер, после установки настройки будут следующими:

– будет создан один PIF-объект, соответствующий единственному сетевому адаптеру;

– PIF-объекту будут присвоены настройки IP-адресации, указанные во время установки, чтобы обеспечить возможность управления;

– физический сетевой интерфейс будет настроен в качестве интерфейса управления;

– будет создана единственная сеть «Network 0»;

– сеть «Network 0» будет подключена к PIF с целью обеспечения внешнего подключения к ВМ.

В обоих случаях полученные настройки сети обеспечат подключение xe CLI (интерфейс командной строки) и любого иного управляющего ПО. Эта конфигурация также позволит создавать внешние сети для виртуальных машин, размещенных на сервере.

Физический сетевой интерфейс PIF, используемый для управления, является единственным физическим сетевым интерфейсом, которому присваиваются настройки IP-адреса во время установки Numa vServer. Создание внешних сетей для виртуальных машин достигается путем соединения физического сетевого интерфейса с виртуальным с помощью сетевого объекта, выполняющего роль виртуального Ethernet-коммутатора.

5.8.1. Изменение конфигурации сети

Изменить конфигурацию сети можно посредством изменения сетевого объекта. Чтобы сделать это, необходимо выполнить команду, затрагивающую либо сетевой объект, либо виртуальный сетевой интерфейс (VIF).

5.8.1.1. Изменение сетевого объекта

Изменять параметры сети, такие как размер фрейма (максимальный блок передачи), метка имени, описание имени и прочие значения, можно посредством команды **xe network-param-set** и соответствующих параметров.

Единственным обязательным параметром является **uuid**.

Необязательные параметры включают в себя:

– default_locking_mode (см. пп. 5.14.2.7 Упрощённая настройка режима блокировки виртуального сетевого интерфейса);

– name-label (название);

– name-description (описание);

– MTU (размер фреймов);

– purpose (добавление цели, см. пп. 5.14.2.9 Назначение цели для сети);

– other-config (прочие настройки).

Если значение параметра не задано, то его значение приравнивается к нулю. Для установки пары «ключ-значение» в параметре адаптера, необходимо использовать синтаксическую конструкцию 'map-param:key=value'.

5.8.2. Изменение времени задержки Up Delay трафика агрегации

Как было описано в разделе «Агрегация сетевых интерфейсов», по умолчанию параметр задержки Up Delay равен 31 000 мс, такое значение выбрано, чтобы избежать перебалансировки трафика в случае кратковременных отказов. Несмотря на то, что время ожидания может показаться очень большим, такая задержка необходима для всех режимов агрегации, а не только для активно-активного.

Администратор может изменить время ожидания перед группировкой с помощью выполнения следующих команд:

– установка времени ожидания в миллисекундах:

```
xe pif-param-set uuid=<uuid главного интерфейса агрегации (PIF)>
other-config:bond-updelay=<Время задержки в мс>
```

– чтобы изменения вступили в силу, необходимо отключить, а затем снова включить физический сетевой интерфейс:

```
xe pif-unplug uuid=<uuid главного интерфейса агрегации (PIF)>
xe pif-plug uuid=<uuid главного интерфейса агрегации (PIF)
```

5.9. Управление конфигурацией сети

Процедуры настройки сети в этом разделе различаются в зависимости от того, настраивается сеть на отдельном сервере или на сервере, который является частью пула ресурсов.

5.9.1. Создание сетей на автономном сервере

Так как внешние сети создаются для каждого физического сетевого интерфейса во время установки, создание дополнительной сети обычно может потребоваться только в случае:

– создания частной сети;

– поддержки дополнительных операций, таких как VLAN и объединение сетевых адаптеров.

Добавление новой сети с помощью командной строки:

– открыть консоль управления сервера;

– с помощью команды создания сети создать новую сеть:

```
xe network-create name-label=<mynetwork>
```

После ввода команды на экране будет отображён UUID новой сети.

На данном этапе сеть не подключена к физическому интерфейсу, т.е. является внутренней.

5.9.2. Создание сетей в пуле ресурсов

Все серверы в рамках одного пула должны иметь одинаковое количество сетевых адаптеров (NIC), хотя это требование не является строгим.

Идентичность физической конфигурации сервера в рамках одного пула важна, потому что все серверы одного пула имеют общий набор сетей. Физические сетевые интерфейсы отдельных серверов подключаются к общим сетям пула по имени устройства. Например, все серверы в рамках пула с сетевым интерфейсом eth0 будут иметь соответствующий физический сетевой интерфейс, подключенный к общей сети пула Network 0. То же самое будет справедливо для серверов с интерфейсами eth1 и Network 1, так же, как и для иных адаптеров, присутствующих как минимум на одном сервере в пуле.

Если у одного из серверов количество сетевых адаптеров отличается от всех остальных в пуле, могут возникнуть сложности, так как не все сети пула будут действительны для всех его участников. Например, если серверы host1 и host2 находятся в одном пуле, и при этом у host1 четыре сетевых адаптера, а у host2 всего два, то для host2 будут действительны только сети, подключенные к физическим интерфейсам, соответствующим eth0 и eth1. Виртуальные машины на host1 с виртуальными сетевыми интерфейсами, подключенными к сетям, соответствующим eth0 и eth1 не смогут мигрировать на сервер host2.

5.9.3. Создание виртуальных локальных сетей (VLAN)

Для серверов в рамках пула ресурсов можно воспользоваться командой **xe pool-vlan-create**. Эта команда создает VLAN и автоматически создает и подключает необходимые физические интерфейсы.

Чтобы создать сеть для использования с VLAN, необходимо выполнить следующие действия:

– открыть консоль управления сервера;

– создать новую сеть для использования с VLAN, выполнив команду:

```
xe network-create name-label=network5
```

На экране будет отображен UUID новой сети.

– чтобы найти UUID физического сетевого интерфейса, соответствующего физическому сетевому адаптеру, поддерживающему желаемый VLAN-тэг, необходимо использовать команду pif-**list**. Будут показаны UUID и имена устройств всех физических интерфейсов, включая все существующие VLAN:

```
xe pif-list
```

– создать объект VLAN и указать желаемый физический сетевой интерфейс и VLAN-тэг:

```
xe vlan-create network-uuid=<network_uuid> pif-uuid=<pif_uuid>
vlan=5
```

На экране будет отображен UUID новой сети VLAN;

– после создания новой сети можно подключить виртуальные сетевые интерфейсы ВМ к этой сети.

5.9.4. Агрегирование сетевых адаптеров автономного сервера

В данном разделе описывается порядок объединения интерфейсов сетевых адаптеров на сервере Numa vServer, не объединённых в пул с другими серверами, при помощи командной строки.

5.9.4.1. Агрегирование сетевых адаптеров

Для агрегации двух или четырёх сетевых адаптеров следует:

– создать новую сеть для использования с объединенным сетевым адаптером командой **network-create:**

```
xe network-create name-label=<bond0>
```

На экране будет отображен UUID новой сети.

– узнать идентификаторы UUID физических интерфейсов, которые предполагается агрегировать, можно командой **pif-list**:

```
xe pif-list
```

– настроить активно-активный тип агрегации (это тип по умолчанию) командой **bond-create,** чтобы создать агрегацию. Отделите параметры запятыми, укажите UUID только что созданной сети и UUID физических интерфейсов, которые необходимо объединить:

```
xe bond-create network-uuid=<network_uuid> pif-
uuids=<pif_uuid_1>,<pif_uuid_2>,<pif_uuid_3>
```

Ввести два UUID, если вы объединяете два сетевых адаптера, и четыре UUID – если четыре. После выполнения команды будет выведен UUID агрегации.

– или настроить агрегацию типа «активно-пассивная» или типа LACP с использованием того же синтаксиса, указав необязательный параметр **mode** и задав для него значение lacp или active-backup:

```
xe bond-create network-uuid=<network_uuid> pif-
uuids=<pif_uuid_1>,<pif_uuid_2>,<pif_uuid_3> mode=<balance-slb |
active-backup | lacp>
```

5.9.4.2. Управление MAC-адресом агрегации

При агрегировании физического сетевого интерфейса используемого в данный момент в качестве интерфейса управления, роль интерфейса управления переносится на всю агрегацию. Если на сервере используется DHCP, то в большинстве случаев MAC-адрес агрегации будет совпадать с таковым у физического сетевого интерфейса, используемого в настоящий момент, а IP-адрес интерфейса управления останется неизменным.

MAC-адрес агрегации можно изменить, чтобы он отличался от MAC-адреса (текущего) адаптера интерфейса управления. Однако после объединения и изменения MAC/IP-адреса текущие сессии сети на сервере будут сброшены.

Управлять MAC-адресом агрегации можно двумя способами:

– в команде **xe bond-create** можно указать необязательный параметр **mac**. Этот параметр можно использовать, чтобы сделать MAC-адрес произвольным;

– если параметр **mac** не указан, будет использован MAC-адрес управляющего интерфейса, если он является одним из интерфейсов агрегации. Если в агрегацию не входит управляющий интерфейс, а входит другой интерфейс, то агрегация использует MAC-адрес (а также IP-адрес) интерфейса управления. Если ни один из сетевых интерфейсов агрегации не является интерфейсом управления, агрегация использует MAC-адрес первого из указанных сетевых адаптеров.

5.9.4.3. Отключение агрегации сетевых адаптеров

При возврате сервера к неагрегированной конфигурации следует обратить внимание, что команда **xe bond-destroy** автоматически настроит интерфейс, указанный в параметре **primary-slave**, в качестве интерфейса управления. Впоследствии, все виртуальные интерфейсы будут перемещены на него.

Термин «primary-slave» здесь соответствует физическому сетевому интерфейсу, настройки MAC- и IP-адресов которого были использованы при создании агрегации. При объединении двух сетевых адаптеров интерфейсом primary-slave будет являться:

– интерфейс управления (если он является одним из агрегируемых);

– любой другой сетевой адаптер с IP-адресом (если интерфейс управления не является одним из агрегируемых);

– первый указанный сетевой адаптер. Узнать, какой это интерфейс, можно выполнив команду:

```
xe bond-list params=all
```

5.9.5. Агрегирование сетевых адаптеров в пуле

Если это возможно, рекомендуется создавать агрегации сетевых адаптеров в процессе создания пула ресурсов, до добавления к пулу дополнительных серверов к пулу и создания ВМ. Это позволит автоматически дублировать настройки агрегации на серверы по мере их подключения к пулу и уменьшит объёмы работы.

Для добавления агрегации сетевых адаптеров к существующему пулу можно выбрать один из двух путей:

– с помощью интерфейса CLI настроить агрегации на мастере пула, а затем на каждом рядовом члене пула;

– с помощью интерфейса CLI настроить группы на мастере пула, чтобы он получил настройки от мастера.

Данный раздел описывает процесс создания агрегаций сетевых адаптеров на серверах Numa vServer входящих в пул с помощью командной строки.

Не пытайтесь агрегировать сети при включенной функции высокой доступности HA (High Availability). Процесс агрегирования нарушит текущие процессы функции HA и вызовет отключение серверов; после этого, скорее всего, не удастся выполнить их перезагрузку правильно, а для восстановления потребуется выполнить команду **xe host-emergency-ha-disable**.

5.9.5.1. Добавление агрегаций сетевых адаптеров к новому пулу ресурсов

Для агрегации двух или четырёх сетевых адаптеров следует:

– выбрать сервер, который должен стать мастером пула. По умолчанию любой сервер является мастером собственного безымянного пула. Чтобы создать пул ресурсов с помощью командной строки, необходимо переименовать существующий безымянный пул:

```
xe pool-param-set name-label=<New Pool> uuid=<pool_uuid>
```

– процесс агрегирования сетевых адаптеров описан в пп. 5.9.4 Агрегирование сетевых адаптеров автономного сервера;

– открыть консоль сервера, который необходимо добавить к пулу, и выполнить команду:

```
xe pool-join master-address=<host1> master-username=root master-
password=<password>
```

Данные сети и агрегации будут автоматически перенесены на новый сервер. Интерфейс управления автоматически переносится с сетевого адаптера сервера, где он был изначально настроен, на физический сетевой интерфейс, включенный в агрегацию (то есть интерфейс управления теперь поглощён агрегацией, вся она выполняет роль интерфейса управления).

Чтобы найти идентификатор UUID настраиваемого сервера, необходимо выполнить следующую команду:

```
xe host-list
```

5.9.5.2. Добавление агрегаций сетевых адаптеров к существующему пулу ресурсов

Не пытайтесь агрегировать сети при включенном HA. Процесс создания агрегации нарушит текущие процесс агента сервера и вызовет отключение сервера; после этого, скорее всего, не удастся выполнить их перезагрузку правильно, а для восстановления потребуется выполнить команду **xe host-emergency-ha-disable**.

В отсутствии графического интерфейса для объединения сетевых адаптеров самым быстрым способом создать агрегацию адаптеров, общую для всего пула, будет следующий: нужно создать эту агрегацию на сервере, являющемся мастером пула, а затем перезапустить прочие серверы пула.

В качестве альтернативы перезагрузки сервера можно использовать команду **sytemctl restart xapi**.

В результате настройки агрегации и настройки VLAN на мастере пула будут переданы каждому серверу. Интерфейс управления каждого сервера придётся перенастраивать вручную.

Процесс агрегирования сетевых адаптеров описан в пп. 5.9.4 Агрегирование сетевых адаптеров автономного сервера.

5.10. Настройка сетевого интерфейса, выделенного для соединения с хранилищем

С помощью интерфейса CLI можно присвоить сетевому интерфейсу IP-адрес и выделить его под определенные функции, например, под трафик сети хранения данных. Чтобы присвоить сетевому интерфейсу IP-адрес необходимо создать *вторичный* интерфейс (основной интерфейс с доступным IP-адресом, используемый Numa vServer для управляющего трафика, называется *интерфейсом управления*).

При необходимости выделить вторичный интерфейс под определенные функции, необходимо проверить конфигурацию сети и убедиться, что сетевой адаптер используется исключительно для необходимого трафика. Например, чтобы выделить сетевой интерфейс под трафик системы хранения, то сетевой интерфейс, запоминающее устройство, коммутатор и/или VLAN должны иметь такую конфигурацию, чтобы к запоминающему устройству, получающему трафик, доступ можно было получить только через выделенный сетевой интерфейс. Если физическая или IP-конфигурация не ограничивают трафик, который может быть пропущен через сетевой интерфейс системы хранения, то прочий трафик, как например управляющий трафик, можно пропустить через вторичный интерфейс.

При создании нового интерфейса для трафика системы хранения данных, необходимо присвоить ему IP-адрес, который будет расположен:

– в той же подсети, что и интерфейс системы хранения, если этот применимо;

– будет расположен в отдельной подсети от прочих вторичных интерфейсов или интерфейса управления.

При конфигурации вторичных интерфейсов следует обратить внимание, что каждый интерфейс должен быть расположен в отдельной подсети. Например, при необходимости выделить два дополнительных вторичных интерфейса под хранение, понадобятся IP-адреса в трёх разных подсетях: одна подсеть для интерфейса управления, вторая – для вторичного интерфейса №1 и третья – для вторичного интерфейса №2.

При использовании агрегирования интерфейсов с целью повысить отказоустойчивость передачи трафика системы хранения, рекомендуется рассмотреть вариант использования LACP на основе vSwitch.

При выборе сетевого адаптера для использования с хранилищами на основе iSCSI или NFS, необходимо убедиться, что выделенный адаптер использует отдельную IP-подсеть, которая не маршрутизируется с интерфейса управления. Если это не обеспечено, то трафик системы хранения может быть направлен через интерфейс управления после перезапуска сервера, согласно порядку инициализации сетевых интерфейсов.

Чтобы присвоить функции сетевому адаптеру необходимо:

– убедиться, что физический сетевой интерфейс находится в отдельной подсети или что маршрутизация настроена согласно топологии имеющейся сети, чтобы обеспечить передачу необходимого трафика через выбранный физический сетевой интерфейс;

– установить IP-конфигурацию для физического сетевого интерфейса, внеся соответствующие величины в параметры режима, а при использовании статического IP-адреса – такие параметры как IP, маску сети, шлюз и DNS:

```
xe pif-reconfigure-ip mode=<DHCP | Static> uuid=<pif-uuid>
```

– присвоить значение **true** параметру **disallow-unplug**:

```
xe pif-param-set disallow-unplug=true uuid=<pif-uuid>
xe pif-param-set other-config:management_purpose="Storage"
uuid=<pif-uuid>
```

При необходимости использования вторичного интерфейса для трафика системы хранения, который также может быть маршрутизирован с интерфейса управления (принимая во внимание, что такая конфигурация не является наилучшим решением), можно воспользоваться двумя способами:

– после перезапуска хоста убедиться в правильности конфигурации вторичного интерфейса, а затем с помощью команд **xe pbd-unplug** и **xe pbd-plug** повторно инициализировать подключение системы хранения к хосту. В результате будет переустановлено соединение с системой хранения, а трафик будет направлен через нужный интерфейс;

– в качестве альтернативы можно использовать команду **xe pif-forget**, чтобы удалить интерфейс из базы данных и вручную настроить его в управляющем домене. Эта рекомендация является дополнительной и потребует навыка ручной конфигурации сетей Linux.

5.11. Использование сетевых адаптеров с поддержкой SR-IOV

Технология *Single Root I/O Virtualization* (*SR-IOV*) является технологией виртуализации PCI-устройств, которая позволяет одному PCI-устройству выполнять функцию нескольких аналогичных устройств на шине PCI. Само по себе физическое устройство называется *физической функцией* (*Physical Function, PF*), в то время как все остальные называются *виртуальными функциями* (*Virtual Functions*, *VF*). Целью является предоставление гипервизору возможности напрямую присваивать один или несколько виртуальных сетевых адаптеров виртуальной машине с помощью технологии SR-IOV: пользователь может использовать виртуальные адаптеры наравне со всеми прочими напрямую подключенными PCI-устройствами.

Присваивание одного или нескольких виртуальных адаптеров виртуальной машине позволяет ей напрямую использовать аппаратное обеспечение. После конфигурации, каждая ВМ ведет себя так, как если бы она напрямую использовала сетевой адаптер, что сокращает издержки на обработку и увеличивает производительность.

Если виртуальной машине требуется мобильность, и она имеет присвоенные виртуальные адаптеры SR-IOV, то выполнение таких функций, как Live Migration, High Availability и Disaster Recovery невозможно. Это происходит из-за того, что ВМ напрямую связана с виртуальной функцией физического сетевого адаптера с поддержкой SR-IOV. Кроме того, сетевой трафик ВМ, направляемый через виртуальные сетевые адаптеры функции SR-IOV, минует vSwitch, что делает невозможным создание ACL или просмотр QoS.

SR-IOV имеет лучшую производительность, чем VIF. Он может обеспечить аппаратное разделение трафика между разными виртуальными машинами через один и тот же сетевой адаптер в обход сетевого стека Numa vServer.

5.11.1. Преимущества SR-IOV

Используя эту функцию, можно:

– включить SR-IOV на сетевых картах, которые поддерживают SR-IOV;

– отключить SR-IOV на сетевых картах, поддерживающих SR-IOV;

– управлять SR-IOV VFs как пулом ресурсов VF;

– назначить VF SR-IOV виртуальной машине;

– настроить VFs SR-IOV (например, MAC-адрес, VLAN, скорость);

– запустить тест для проверки поддержки SR-IOV.

5.11.2. Конфигурация системы для работы с SR-IOV

Для корректной работы SR-IOV необходима поддержка оборудованием следующих технологии:

– виртуализация ввода-вывода MMU (AMD-Vi и Intel VT-d);

– альтернативная интерпретация идентификатора маршрутизации (ARI);

– услуги по переводу адресов (ATS);

– службы контроля доступа (ACS).

5.11.2.1. Ограничения при работе с SR-IOV

Технология SR-IOV следующие ограничения:

– для некоторых сетевых адаптеров, использующих устаревшие драйверы (например, семейство Intel I350), необходимо перезагрузить сервер, чтобы включить или отключить SR-IOV на этих устройствах;

– только гостевые ВМ с поддержкой аппаратной виртуализации (HVM) поддерживают работу с технологией SR-IOV;

– сеть уровня пула, имеющая разные типы сетевых карт, не поддерживается;

– VF SR-IOV и обычный VIF одного и того же сетевого адаптера могут не иметь возможности связываться друг с другом из-за аппаратных ограничений сетевого адаптера. Чтобы эти хосты могли обмениваться данными, убедитесь, что для связи используется шаблон VF для VF или VIF для VIF, а не VF для VIF;

– настройки качества обслуживания для некоторых VF SR-IOV не вступают в силу, поскольку они не поддерживают ограничение скорости сети;

– выполнение динамической миграции, приостановки и создание снимка состояния не поддерживается на виртуальных машинах, использующих SR-IOV VF;

– VF SR-IOV не поддерживают горячее подключение;

– для некоторых сетевых карт с устаревшими драйверами может потребоваться перезагрузка даже после перезапуска хоста, это указывает на то, что сетевая карта не может включить SR-IOV;

– аппаратное ограничение: функция SR-IOV полагается на контроллер для сброса функций устройства в исходное состояние в течение 100 мс по запросу гипервизора с использованием сброса функционального уровня (FLR);

– SR-IOV может использоваться в пуле, который использует механизм HA (высокая доступность). Виртуальные машины, которым назначены SR-IOV VF, перезапускаются, когда в пуле есть сервер, имеющий соответствующие ресурсы.

5.11.2.2. Настройка SR-IOV для устаревших драйверов

Обычно максимальное количество VF, которые может поддерживать сетевая карта, может быть определено автоматически. Для сетевых карт, использующих устаревшие драйверы (например, семейство Intel I350), ограничение определяется в файле конфигурации модуля драйвера. Может потребоваться корректировка лимита вручную. Чтобы установить его на максимум, необходимо:

– открыть файл редактором:

```
/etc/modprobe.d/igb.conf
```

– установить максимальное количество VF на 7 в поле «VFs-maxvfs-by-user»:

```
## VFs-param: max_vfs
## VFs-maxvfs-by-default: 7
## VFs-maxvfs-by-user: 7
options igb max_vfs=0
```

– сохранить изменения;

– выгрузить и загрузить драйвер для применения конфигурации:

```
rmmod igb && modprobe igb
```

Вносимое в файл конфигурации драйвера значение должно быть меньше или равно значению VFs-maxvfs-by-default.

Не рекомендуется изменять никакие другие строки в этом файле.

5.11.2.3. Присвоение ВМ виртуального адаптера SR-IOV

– открыть CLI севера;

– выполнить команду lspci, чтобы отобразить список виртуальных функций (VF). Например:

```
07:10.0 Ethernet controller: Intel Corporation 82559 Ethernet
Controller Virtual Function (rev 01)
```

В примере выше 07:10.0 является (bus:device.function) - адресом виртуального адаптера.

– виртуальный адаптер присваивается ВМ с помощью следующей команды:

```
xe vm-param-set other-config:pci=0/0000:<bus:device.function>
uuid=<vm-uuid>
```

– запустить ВМ и установить в ОС драйвер соответствующей виртуальному адаптеру для конкретного устройства.

Одной ВМ может быть присвоено несколько виртуальных адаптеров, однако один виртуальный адаптер **не может** использоваться несколькими ВМ.

5.12. Ограничение базовой скорости передачи данных (QoS limit)

Чтобы ограничить объём трафика, который ВМ может отправлять за секунду времени, можно присвоить значение базовой скорости передачи данных (*Quality of Service limit*) для виртуального сетевого интерфейса. Это позволяет установить максимальный уровень передачи исходящих пакетов данных в Кбайт/сек.

Значение QoS ограничивает уровень исходящих данных из виртуальной машины, и не ограничивает объём входящих данных.

В зависимости от сетевого стека, настроенного для пула, можно установить значение параметра QoS для виртуального сетевого интерфейса виртуальной машины в одном из двух мест: либо на внешнем контроллере, совместимым с vSwitch, либо на сервер.

Таблица 6 - Настройка QoS в зависимости от используемого сетевого стека

Сетевой стек	Доступные методы конфигурации
vSwitch	Внешний контроллер. Этот метод является предпочтительным методом настройки QoS на виртуальном интерфейсе, когда vSwitch выполняет роль сетевого стека; Командная строка. Возможно установить уровень передачи QoS с помощью команд, как показано в примере ниже.
Linux bridge	Командная строка. Установить уровень передачи QoS также можно с помощью командной строки, выполнив команды, описанные ниже.

Когда vSwitch используется в качестве сетевого стека, можно непреднамеренно установить значение QoS как на внешнем контроллере, так и внутри сервера Numa vServer. В таком случае, Numa vServer ограничит исходящий трафик согласно более низкому из установленных значений.

Пример команд интерфейса командной строки для установки QoS. Чтобы ограничить виртуальный сетевой интерфейс до максимального уровня передачи в 100 Кбайт/с с помощью командной строки, необходимо воспользоваться следующей командой:

```
xe vif-param-set uuid=<vif_uuid> qos_algorithm_type=ratelimit
xe vif-param-set uuid=<vif_uuid> qos_algorithm_params:kbps=100
```

Если используется внешний контроллер, рекомендуется установить ограничение скорости передачи в контроллере вместо выполнения этой команды CLI.

5.13. Изменение параметров конфигурации сети

В данном разделе рассматриваются возможности изменения сетевой конфигурации сервера Numa vServer. Это включает в себя:

– изменение имени хоста (hostname);
– добавление или удаление DNS-серверов;
– изменение IP-адреса;
– изменение сетевого адаптера, используемого в качестве интерфейса управления;
– добавление нового физического сетевого адаптера к серверу;
– включение ARP-фильтрации (блокировка порта коммутатора).

5.13.1. Изменение имени хоста

Системное имя хоста, также известное как DNS-имя, определяется в рамках общей базы данных пула и управляется с помощью следующей команды:

```
xe host-set-hostname-live host-uuid=<host_uuid> host-name=<host-
name>
```

Базовое имя хоста управляющего домена изменяет в динамическом режиме и отображает действующее имя хоста.

5.13.2. DNS-серверы

Чтобы добавить или удалить DNS-сервер в конфигурации IP-адресации хоста Numa vServer, следует пользоваться командой **xe pif-reconfigure-ip**. Например, для физического сетевого интерфейса со статическим IP-адресом:

– Настройте порядок поиска суффиксов DNS вашего домена для разрешения неполных доменных имен:

```
xe pif-param-set uuid=<pif-uuid_in_the_dns_subnetwork> other-
config:domain=suffix.com
```

– Настройте DNS-сервер для использования на хостах vServer:

```
xe pif-reconfigure-ip mode=static dns=<dnshost> ip=<ip>
gateway=<gateway> netmask=<netmask> uuid=<uuid>
```

– Вручную настройте интерфейс управления на использование PIF, который находится в той же сети, что и ваш DNS-сервер:

```
xe host-management-reconfigure pif-uuid=<pif_in_the_dns_subnetwork>
```

Альтернативный способ добавления DNS-сервера:

– xe pif-reconfigure-ip mode=static dns=<dnshost> ip=<ip> gateway=<gateway> netmask=<netmask> uuid=<uuid>

– resolvectl domain xenbr0 suffix.com

Изменение конфигурации IP-адреса сервера

Конфигурацию сетевого интерфейса можно изменить с помощью командной строки. Чтобы изменить конфигурацию IP-адреса физического сетевого интерфейса следует воспользоваться командой **xe pif-reconfigure-ip**.

5.13.3. Изменение конфигурации IP-адреса в пуле ресурсов

Серверы Numa vServer в пуле ресурсов имеют один административный IP-адрес, используемый для управления, связи и взаимодействия с другими серверами в пуле. Процедура изменения IP-адреса интерфейса управления для мастера пула отличается от аналогичной процедуры для остальных хостов.

> Изменять IP-адрес и прочие параметры сервера необходимо с осторожностью. В зависимости от топологии сети и вносимых изменений, соединение с сетевой системой хранения данных может быть потеряно. Если это произойдет, систему хранения необходимо подключить заново с помощью команды **xe pbd-plug** командной строки. Рекомендуется мигрировать ВМ с сервера до внесения изменений в IP-конфигурацию.

Для изменения IP-адреса рядового участника пула следует:

– установить желаемый IP-адрес посредством интерфейса командной строки выполнив команду **xe pif-reconfigure-ip**. Например, для получения IP-адреса по DHCP:

```
xe pif-reconfigure-ip uuid=<pif_uuid> mode=DHCP
```

– выполнить команду **xe host-list**, чтобы убедиться, что участник пула был успешно подключен к мастеру, проверив, что все остальные серверы пула отображаются корректно.

```
xe host-list
```

Изменение IP-адреса мастера пула потребует дополнительных действий, поскольку каждый рядовой участник пула использует IP-адрес мастера для связи и взаимодействия с ним, и не будет знать, как связаться с мастером после изменения этого адреса.

По возможности для мастера пула необходимо использовать выделенный IP-адрес, который с малой вероятностью будет подвержен изменениям на протяжении всего срока службы пула.

Для изменения IP-адреса мастера пула следует выполнить следующую последовательность действий:

– установить желаемый IP-адрес посредством интерфейса командной строки выполнив команду **xe pif-reconfigure-ip**. Например, для получения IP-адреса по DHCP:

```
xe pif-reconfigure-ip uuid=<pif_uuid> mode=DHCP
```

– после изменения IP-адреса мастера пула, все серверы рядовых участников перейдут в аварийный режим, по причине невозможности установить с ним связь. Чтобы принудительно подключить мастер к остальным членам пула и сообщить им его новый IP-адрес, необходимо с сервера-мастера выполнить следующую команду:

```
xe pool-recover-slaves
```

5.13.4. Смена интерфейса управления

В случае если Numa vServer установлен на сервере с несколькими сетевыми адаптерами, то один из сетевых интерфейсов адаптеров должен быть выбран в качестве интерфейса управления. Интерфейс управления используется для подключения к серверу клиента управления и для межсерверного взаимодействия.

Изменение сетевого адаптера, используемого в качестве интерфейса управления:

– чтобы определить, какой физический сетевой интерфейс соотнесен с сетевым адаптером, и должен использоваться в качестве интерфейса управления, выполнить следующую команду:

```
xe pif-list
```

Будет показан UUID каждого физического сетевого интерфейса.

– чтобы просмотреть IP-адрес для физического сетевого интерфейса, который будет использоваться в качестве интерфейса управления, выполнить команду **xe pif-param-list**. При необходимости изменить IP-адрес, выполнив команду **xe pif-reconfigure-ip**:

```
xe pif-param-list uuid=<pif_uuid>
```

– выполнить команду **xe host-management-reconfigure** для изменения сетевого интерфейса используемого в качестве интерфейса управления. Если этот сервер входит в состав пула, эту команду необходимо выполнить рядового участника пула:

```
xe host-management-reconfigure pif-uuid=<pif_uuid>
```

– для изменения интерфейса управления в пуле необходимо выполнить команду:

```
xe pool-management-reconfigure network-uuid=<network_uuid>
```

5.13.4.1. Отключение интерфейса управления

Чтобы полностью отключить удаленный доступ к административной консоли, необходимо выполнить команду:

```
xe host-management-disable
```

После того как интерфейс управления был отключен, для выполнения административных задач, необходимо будет войти в физическую консоль управления на сервере.

5.13.5. Добавление нового сетевого адаптера

После установки нового физического сетевого адаптера в сервер он не будет автоматически определён в системе. Для того что бы новый сетевой адаптер определился в системе необходимо выполнить команду **xe pif-scan.**

5.14. Использование блокировки порта коммутатора

Функция блокировки порта коммутатора Numa vServer позволяет контролировать трафик, отправляемый неизвестными, ненадежным или потенциально опасными ВМ посредством ограничения их способности имитировать наличие MAC-адресов и IP-адресов, которые не были им присвоены. Функции позволяет воспользоваться командами блокировки порта, чтобы заблокировать весь трафик в сети по умолчанию или определить конкретные IP-адреса, с которых отдельным ВМ будет разрешено отправлять трафик.

Блокировка порта коммутатора является функцией, разработанной для провайдеров общедоступных облачных услуг в средах подверженным внутренним угрозам. Эта функция может помочь провайдерам общедоступных облачных услуг, обладающих сетевой архитектурой, где каждая ВМ имеет открытый и связанный с Internet IP-адрес. Поскольку арендаторы облака всегда считаются ненадежными, применение таких мер безопасности, как защита от несанкционированного получения доступа к ресурсам сети за счёт использования чужого IP-адреса, может быть необходимым для защиты виртуальных машин от атак других арендаторов облака.

Использование блокировки порта коммутатора позволяет упростить конфигурацию сети, ограничив действия всех арендаторов и незарегистрированных пользователей одним уровнем сети (L2).

Одной из наиболее важных функций команд блокировки порта коммутатора является ограничение трафика, который может быть получен от ненадежного незарегистрированного пользователя, что, в свою очередь, ограничивает способность такого пользователя имитировать наличие MAC и IP-адреса, которым он на самом деле не обладает. В частности, можно использовать эти команды, чтобы предотвратить такие действия незарегистрированного пользователя, как:

– сообщение IP-адреса или MAC-адреса, который не входит в список разрешенных адресов, указанных администратором Numa vServer;

– перехват, несанкционированное получение или перебой трафика других ВМ.

5.14.1. Требования функции блокировки

Функция блокировки порта коммутатора Numa vServer поддерживается в сетевых стеках Linux bridge и vSwitch.

Если в среде активировано ролевое управление доступом (RBAC), пользователь, настраивающий блокировку порта коммутатора, должен быть авторизован в системе посредством учётной записи с функциями не ниже Оператора пула или Администратора пула. Если RBAC не активирован, пользователь должен быть авторизован в системе мастера пула посредством учётной записи с правами администратора.

При выполнении команд блокировки порта коммутатора, сеть может быть как онлайн, так и оффлайн.

В виртуальных машинах под управлением ОС Windows значок отключенной сети появляется только в том случае, если в гостевой системе установлены vServer VM Tools.

5.14.2. Примечания

Когда значения функции блокировки порта коммутатора не установлено, виртуальные интерфейсы имеют параметр «network_default», а Сети – «unlocked».

Конфигурация блокировки порта коммутатора не поддерживается, когда в контроллер программно-определяемых сетей или иные адаптеры.

Блокировка порта коммутатора не защитит от таких действий арендаторов облака, как:

– осуществление атаки на другого арендатора/пользователя на том же IP-уровне. Однако блокировка порта коммутатора может предотвратить атаки в рамках одного IP-уровня, если они осуществляются следующим образом:

действия под видом законного пользователя/арендатора облака;

попытка перехвата трафика, предназначенного другому пользователю;

– истощение ресурсов сети;

– получение трафика, предназначенного другим ВМ посредством чрезмерной лавинной маршрутизации (для широковещательных MAC-адресов или MAC-адресов с неизвестным назначением).

И соответственно блокировка порта коммутатора не контролирует то, куда ВМ может отправлять трафик.

5.14.2.1. Примечания по реализации функции

Применить функцию блокировки порта коммутатора можно либо с помощью командной строки, либо с помощью ПО совместимого с API. Однако в крупных средах, где автоматизация играет первостепенную роль, наиболее типичным методом применения может являться использование API.

5.14.2.2. Принцип блокировки порта коммутатора

Функция блокировки порта коммутатора позволяет контролировать фильтрацию пакетов на одном или двух уровнях:

– на уровне виртуального сетевого интерфейса. То, каким образом должны фильтроваться пакеты данных, определяется установками виртуального сетевого интерфейса. Можно настроить виртуальный сетевой интерфейс так, чтобы в целом запретить виртуальной машине отправлять трафик, ограничить возможность виртуальной машины отправлением трафика, использующего только заданный IP-адрес, или разрешить виртуальной машине отправлять трафик на любой IP-адрес сети, подключенной к данному виртуальному сетевому интерфейсу;

– на уровне сети. То, каким образом должны фильтроваться пакеты данных, определяется сетью Numa vServer. Когда режим блокировки виртуального сетевого интерфейса установлен на network_default, это говорит о том, что, то какой трафик можно пропускать, определяется настройками блокировки на уровне сети.

Работа функции не зависит от того, какой сетевой стек используется. Однако Linux bridge не полностью поддерживает блокировку порта коммутатора в IPv6 сетях.

5.14.2.3. Варианты режима блокировки виртуального сетевого интерфейса

Функция блокировки порта коммутатора Numa vServer представляет собой режим блокировки с четырьмя вариантами настройки виртуального сетевого интерфейса. Эти варианты могут быть применимы только в том случае, когда виртуальный сетевой интерфейс подключен к работающей ВМ.

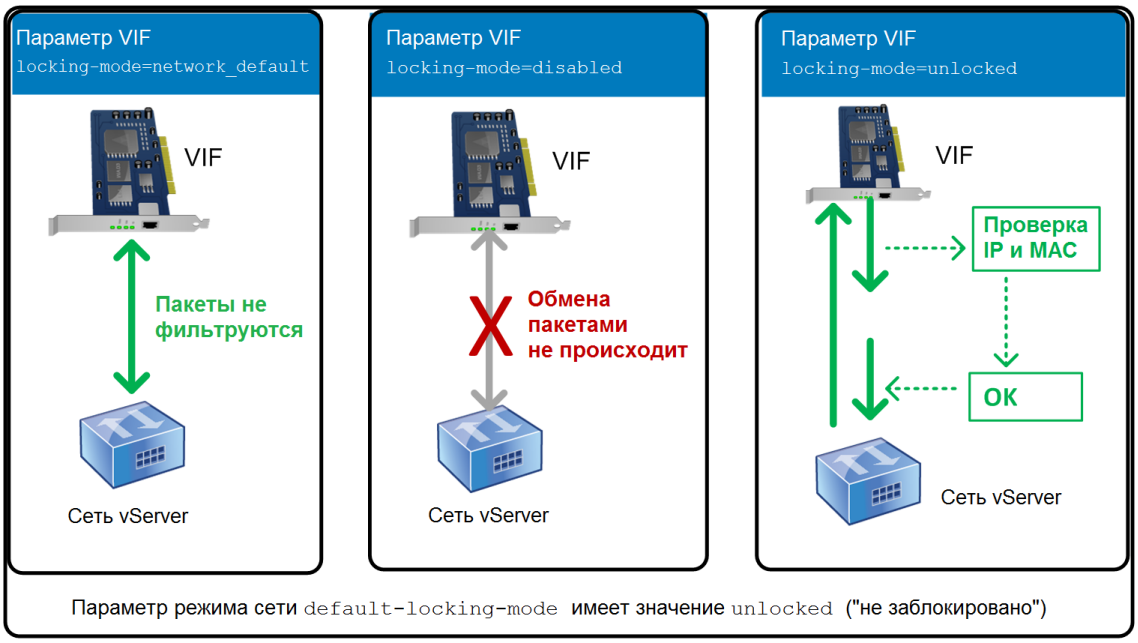

Рисунок 8 – Блокировка порта коммутатора

Рисунок 8 показывает, как ведет себя виртуальный сетевой интерфейс (VIF) в трёх различных вариантах режима блокировки, когда параметр, отвечающий за режим блокировки самой сети (параметр **default-locking-mode**) находится в состоянии *unlocked*.

На первом изображении рисунка, виртуальный сетевой интерфейс настроен по умолчанию (**locking-mode**=network_default), поэтому трафик, исходящий от ВМ, не фильтруется.

На втором изображении, виртуальный сетевой интерфейс не отправляет и не получает пакеты трафика, поскольку режим блокировки ***disabled*** запрещает обмен трафиком.

На третьей изображении рисунка, виртуальный сетевой интерфейс находится в состоянии ***unlocked***, поэтому он может отправлять только пакеты трафика с заданными MAC- и IP-адресами.

Подробнее значения параметра **locking-mode** описаны ниже:

– **network_default** («для сети по умолчанию»). Когда режим виртуальной сети установлен в значение ***network_default***, сервер использует параметр сети **default-locking-mode** чтобы определить, следует ли фильтровать пакеты, проходящие через виртуальную сеть и как это делать. Соответственно, поведение отличается в зависимости то того, какое из значений параметра установлено для сети:

default-locking-mode=disabled («запрещено»), Numa vServer применяет правило фильтрации, по которому виртуальный сетевой интерфейс удаляет весь трафик;

default-locking-mode=unlocked («не заблокировано»), Numa vServer снимает все правила фильтрации по отношению к виртуальному сетевому интерфейсу. По умолчанию этот параметр режима блокировки установлен на значение unlocked («разблокирован»).

Режим блокировки сети по умолчанию не влияет на подключенные VIF, состояние блокировки которых отличается от network_default.

Параметр **default-locking-mode** сети с подключенным активным виртуальным сетевым интерфейсом, не подлежит изменению.

– **locked («заблокирован)**. Numa vServer применяет правила фильтрации, согласно которым через виртуальный сетевой интерфейс может проходить трафик, отправленный только с определённых MAC- и IP-адресов или только на определённые MAC- и IP-адреса. В данном режиме, если не указано ни одного IP-адреса, виртуальная машина не может передавать трафик через данный виртуальный сетевой интерфейс (данной сети).

Чтобы указать IP-адреса, с которых VIF принимает трафик, используйте IP-адреса IPv4 или IPv6 с помощью параметров ipv4_allowed или ipv6_allowed. Однако если у вас настроен мост Linux, не вводите адреса IPv6.

Numa vServer позволяет вводить адреса IPv6, когда Linux bridge активен. Однако не может фильтровать на основе введенных адресов IPv6. Причина в том, что у Linux Bridge нет модулей для фильтрации пакетов протокола обнаружения соседей (NDP). Следовательно, полная защита не может быть реализована, и ВМ смогут выдавать себя за другие ВМ путем подделки пакетов NDP. В результате, если вы укажете хотя бы один адрес IPv6, то Numa vServer пропускает весь трафик IPv6 через VIF. Если вы не укажете адреса IPv6, то не пропускает трафик IPv6 в VIF.

– **unlocked («не заблокирован»).** Через виртуальный сетевой интерфейс может проходить весь трафик. То есть к трафику, проходящему через виртуальный сетевой интерфейс, фильтрация не применяется.

– **disabled («запрещён»).** Трафик вообще не может проходить через виртуальный сетевой интерфейс, то есть происходит его удаление.

5.14.2.4. Настройка блокировки порта коммутатора

В данном пункте описываются три различные процедуры:

– ограничение работы виртуальных сетевых интерфейсов использованием определенного IP-адреса;

– добавление IP-адреса к существующему списку ограничений (например, когда необходимо добавить IP-адрес к виртуальному сетевому интерфейсу);

– удаление IP-адреса из существующего списка ограничений.

Если режим виртуального сетевого интерфейса находится в состоянии ***locked***, то он сможет использовать только адреса, указанные в параметрах **ipv4_allowed** или **ipv6_allowed**.

Потому как в некоторых относительно редких случаях, виртуальные сетевые интерфейсы могут иметь более одного IP-адреса, то возможно указать несколько IP-адресов для одного виртуального сетевого интерфейса.

Эти процедуры можно выполнить как до, так и после подключения виртуального сетевого интерфейса (или запуска виртуальной машины).

Чтобы ограничить работу виртуальных сетевых интерфейсов с определённым IP-адресом необходимо:

– перевести параметр режима default-locking в состояние locked, если этот режим в данный момент не используется, команда:

```
xe vif-param-set uuid=<vif-uuid> locking-mode=locked
```

Команда **xe vif-uuid** выведет UUID виртуального сетевого интерфейса, которому вам необходимо разрешить отправлять трафик. Чтобы получить UUID, выполните команду **xe vif-list** на сервере. Команда **xe vm-uuid** укажет виртуальную машину, в отношении которой выведена данная информация.

– чтобы указать IP-адреса, с которых виртуальной машине можно отправлять трафик, необходимо указать один или несколько желаемых IP-адресов версии IPv4. Например:

```
xe vif-param-set uuid=<vif-uuid> ipv4-allowed=<список ipv4-адресов
через запятую>
```

– или указать один или несколько желаемых IP-адресов версии IPv6. Например:

```
xe vif-param-set uuid=<vif-uuid> ipv6-allowed=<список ipv6-адресов
через запятую>
```

Через запятую можно указать несколько IP-адресов.

Выполнив предыдущую процедуру ограничения работы виртуальных сетевых интерфейсов с определённым IP-адресом, можно добавить к этому ограничению один или несколько IP-адресов, которые сможет использовать виртуальный сетевой интерфейс.

– чтобы добавить IP-адрес к существующему списку, следует указать IP-адрес версии IPv4. Например:

```
xe vif-param-add uuid=<vif-uuid> param-name=ipv4-allowed param-
key=<ipv4-адрес>
```

– или указать IP-адрес версии IPv6. Например:

```
xe vif-param-add uuid=<vif-uuid> name=ipv6-allowed param-key=<ipv6-
адрес>
```

Если администратор ограничивает работу виртуального сетевого интерфейса использованием двух или более IP-адресов, то можно удалить один из таких IP-адресов из списка.

Чтобы удалить IP-адрес из существующего списка, необходимо указать IP-адрес версии IPv4, которые необходимо удалить. Например:

```
xe vif-param-remove uuid=<vif-uuid> param-name=ipv4-allowed param-
key=<ipv4-адрес>
```

или указать IP-адрес версии IPv6, который необходимо удалить. Например:

```
xe vif-param-remove uuid=<vif-uuid> param-name=ipv6-allowed param-
key=<ipv6-адрес>
```

5.14.2.5. Запрет ВМ отправлять или получать трафик из определенной сети

Описанная ниже процедура запрещает виртуальной машине пропускать трафик через виртуальный сетевой интерфейс. Эту процедуру можно использовать, чтобы запретить взаимообмен трафиком между виртуальной машиной и определенной сетью. Это предоставляет возможность более тонкого контроля, чем полный запрет на обмен трафиком с сетью.

Нет необходимости отключать виртуальный сетевой интерфейс, чтобы установить режим его блокировки; команда изменит правила фильтрования, не отключая его. В этом случае сетевое соединение по-прежнему присутствует, однако VIF отбрасывает все пакеты, которые виртуальная машина пытается отправить.

Чтобы получить UUID виртуального сетевого интерфейса, необходимо выполнить команду **xe vif-list**. Поле **device** показывает номер устройства виртуального сетевого интерфейса.

Чтобы запретить виртуальному сетевому интерфейсу принимать трафик из сети, необходимо ввести команду:

```
xe vif-param-set uuid=<vif-uuid> locking-mode=disabled
```

5.14.2.6. Снятие ограничения работы виртуального сетевого интерфейса

Чтобы вернуться к разблокированному состоянию виртуального сетевого интерфейса, необходимо перевести режим **default-locking** виртуального сетевого интерфейса в состояние unlocked (если это состояние уже в данный момент не используется), выполнив следующую команду:

```
xe vif-param-set uuid=<vif_uuid> locking-mode=unlocked
```

5.14.2.7. Упрощённая настройка режима блокировки виртуального сетевого интерфейса

Вместо того чтобы выполнять команды режима блокировки виртуального сетевого интерфейса для каждого в отдельности, можно заблокировать все виртуальные сетевые интерфейсы по умолчанию. Чтобы добиться этого, необходимо внести изменения в фильтрацию пакетов трафика на уровне сети, благодаря которой сеть Numa vServer определяет и какие пакеты необходимо фильтровать, согласно процедуре, описанной в пп. 5.14 Использование блокировки порта коммутатора.

В частности, настройка сетевого параметра **default-locking-mode** определяет поведение новых виртуальных сетевых интерфейсов с настройками по умолчанию. Если **locking-mode** виртуального сетевого интерфейса установлен по умолчанию (default), то виртуальный сетевой интерфейс обращается к сетевому режиму блокировки (default-locking-mode), чтобы определить, фильтруются ли пакеты трафика, проходящие через виртуальный сетевой интерфейс, а также то, каким образом происходит фильтрация:

– unlocked («разблокирован»). Numa vServer разрешает ВМ отправлять трафик на любой IP-адрес сети, подключенной к данному виртуальному сетевому интерфейсу;

– disabled («запрещён»). Numa vServer применяет правило фильтрации, по которому виртуальный сетевой интерфейс удаляет весь трафик.

По умолчанию параметр **default-locking-mode** установлен на unlocked для всех сетей.

Устанавливая режим блокировки виртуального сетевого интерфейса в состояние по умолчанию (network_default), можно использовать эту настройку в качестве базовой конфигурации (на уровне сети) для всех вновь созданных виртуальных сетевых интерфейсов, подключенных к определенной сети.

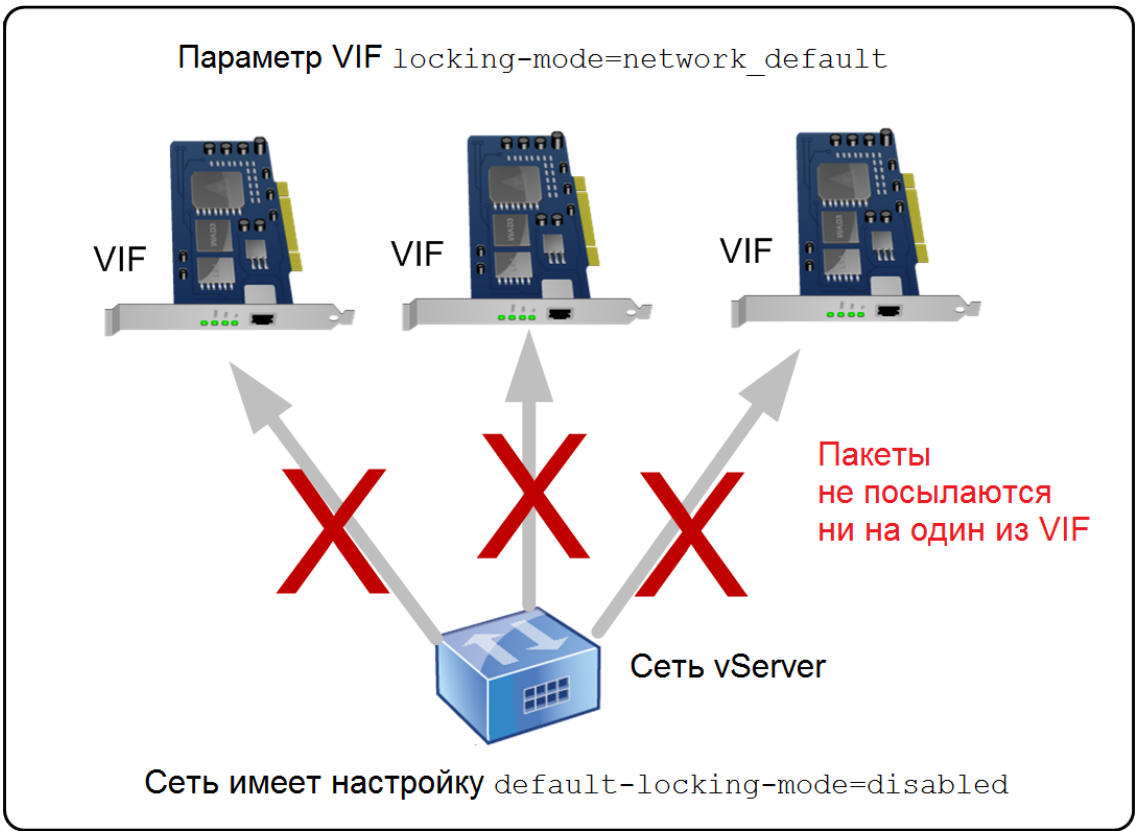

Рисунок 9 – Блокировка виртуального сетевого интерфейса в пуле ресурсов

Рисунок 9 показывает, как виртуальный сетевой интерфейс, в случае если он настроен по умолчанию (**locking-mode**=network_default), руководствуется значением параметра **default-locking-mode** сети. Ко всей сети применен параметр **default-locking-mode**=disabled, поэтому виртуальные сетевые интерфейсы не могут пропускать трафик.

Чтобы изменить настройку режима блокировки по умолчанию для сети необходимо создать сеть и изменить параметр default-locking-mode, выполнив следующую команду:

```
xe network-param-set uuid=<network-uuid> default-locking-
mode=[unlocked|disabled]
```

Чтобы получить идентификатор UUID сети, нужно выполнить команду **xe network-list**. Эта команда отобразит идентификаторы всех сетей на сервере, где была выполнена команда.

Чтобы узнать значение по умолчанию сетевого режима блокировки, следует выполнить одну из следующих команд:

```
xe network-param-get uuid=<network-uuid> param-name=default-locking-
mode
```

или

```
xe network-list uuid=<network-uuid> params=default-locking-mode
```

5.14.2.8. Применение настроек сети для фильтрации трафика к виртуальному сетевому

интерфейсу

Чтобы использовать настройки сети для фильтрации трафика к определенному VIF необходимо выполнить следующие шаги:

– перевести режим блокировки виртуального сетевого интерфейса в состояние network_default (если этот режим в данный момент уже не используется), выполнив следующую команду:

```
xe vif-param-set uuid=<vif_uuid> locking-mode=network_default
```

– перевести параметр default-locking в состояние unlocked (если этот режим в данный момент уже не используется), выполнив следующую команду:

```
xe network-param-set uuid=<network-uuid> default-locking-
mode=unlocked
```

5.14.2.9. Назначение цели для сети

Назначение цели для сети может использоваться для добавления дополнительных функций. Например, возможность использовать сеть для создания соединений NBD.

Чтобы назначить цель необходимо выполнить команду:

```
xe network-param-add param-name=purpose param-key=<purpose>
uuid=<network-uuid>
```

Для удаления цели из сети выполнить команду:

```
xe network-param-remove param-name=purpose param-key=<purpose>
uuid=<network-uuid>
```

В настоящее время доступными значениями для сетевых целей являются **nbd** и **insecure_nbd**.

5.15. Устранение неполадок сети

5.15.1. Обнаружение ошибок и неисправностей сети

Некоторые модели сетевых карт требуют обновления встроенного программного обеспечения от поставщика для надежной работы под нагрузкой или при включении определенных оптимизаций. Если вы видите поврежденный трафик на виртуальных машинах, попробуйте получить последнюю версию прошивки от вашего поставщика.

Если проблема все еще сохраняется, вы можете использовать CLI для отключения оптимизации приема или передачи разгрузки на физическом интерфейсе:

Отключение оптимизации приема или передачи разгрузки может привести к потере производительности и увеличению загрузки ЦП.

– определить UUID физического интерфейса:

```
xe pif-list
```

– отключить разгрузку для передачи данных (TX) физического интерфейса:

```
xe pif-param-set uuid=<pif_uuid> other-config:ethtool-tx=off
```

– переподключить физический интерфейс или перезапустить сервер, чтобы изменения вступили в силу.

5.15.2. Аварийный перезапуск сети

Неправильные сетевые настройки могут вызвать потерю соединения с сетью, а сервер может стать недоступным при подключении по SSH. Аварийный перезапуск сети представляет собой простой механизм восстановления и перезагрузки сети сервера.

Эта функция доступна из интерфейса командной строки с помощью команды **xe-reset-networking**.

Неправильные настройки, вызывающие потерю соединения с сетью, могут также включать в себя переименование сетевых интерфейсов, объединение адаптеров или VLAN, или ошибки при изменении интерфейса управления (например, неправильно введённый IP-адрес).

Эта функция должна использоваться только в случае экстренной ситуации, так как она удаляет конфигурацию всех физических сетевых интерфейсов, агрегаций, VLAN и туннелей, имеющих отношение к данному серверу. Гостевые сети и виртуальные сетевые интерфейсы сохраняются. При выполнении этой функции, виртуальные машины будут принудительно выключены, поэтому рекомендуется перед выполнением команды, выключить виртуальные машины. Перед перезапуском, администратор может внести изменения в интерфейс управления и указать, какую IP-конфигурацию следует использовать: DHCP или статическую.

Если мастер пула требует перезапуска сети, она должна быть выполнена до перезапуска сетей всех остальных рядовых участников пула. Затем следует выполнить перезапуск сетей всех остальных серверов сети, чтобы обеспечить однородность сетевой конфигурации пула.

Если IP-адрес мастера пула (интерфейс управления) изменяется в результате сброса сети или **xe host-management-reconfigure**, примените команду сброса сети к другим серверам в пуле. Это необходимо для того, чтобы участники пула могли повторно подключиться к мастеру пула с его новым IP-адресом. В этой ситуации необходимо указать IP-адрес мастера пула.

Перезапуск сети НЕ поддерживается при включенной функции High Availability (HA). Чтобы при таком сценарии выполнить перезапуск конфигурации сети, необходимо сначала в ручную отключить HA, а затем выполнить команду перезапуска сети.

5.15.2.1. Проверка параметров сети после их сброса

Указав режим конфигурации для перезапуска сети, в интерфейсе командной строки отобразятся настройки, которые будут применены к серверу после перезагрузки. Это будет последней возможностью внести изменения перед применением команды аварийного перезапуска сети.

Аварийный перезапуск сети также необходимо применить к остальным серверам пула, чтобы продублировать агрегации, VLAN и туннели в соответствии с новой конфигурации мастера пула.

5.15.2.2. Перезапуск сети с помощью интерфейса командной строки

В таблице 4 показаны доступные необязательные параметры, которые можно применить во время выполнения команды **xe**-**reset**-**networking**.

Ответственность за правильность параметров, указанных для команды **xe**-**reset**-**networking**, лежит на администраторе. Необходимо внимательно проверить все указанные параметры. При указании неверных параметров связь с сетью будет потеряна. В такой ситуации рекомендуется повторно выполнить команду **xe**-**reset**-**networking**, не указывая никаких параметров.

Перезапуск сетевой конфигурации всего пула необходимо начинать с мастера пула, а затем переходить к перезапуску сети всех остальных серверах пула.

Таблица 7 – Параметры xe-reset-networking

Параметр	**Обязательный/ необязательный**	**Описание**
-m, --master	необязательный	IP-адрес интерфейса управления мастера пула. Сбрасывает на последнее известное значение IP-адреса мастер пула
--device	необязательный	Имя устройства административного интерфейса. Сбрасывает на значение имени устройства, указанное при установке
-mode=static	необязательный	Позволяет использовать следующие четыре параметра для статической IP-конфигурации административного интерфейса. Если этот параметр не указан, по умолчанию используется значение параметра DHCP
--ip	обязательный, если mode=static	IP-адрес интерфейса управления сервера. Действителен только, если mode=static
--netmask	обязательный, если mode=static	Маска сети интерфейса управления. Действителен только, если mode=static
--gateway	необязательный	Шлюз интерфейса управления. Действителен только, если mode=static
--dns	необязательный	DNS-сервер интерфейса управления. Действителен только, если mode=static

5.15.2.3. Примеры аварийного сброса настроек на мастере пула

Ниже описаны примеры команд, которые могут быть применены к мастеру пула:

– сброс настроек сети для DHCP-конфигурации:

```
xe-reset-networking
```

– сброс настроек сети для статической IP-конфигурации:

```
xe-reset-networking --mode=static --ip=<ip-address> --
netmask=<netmask> --gateway=<gateway> --dns=<dns>
```

– сброс настроек сети для DHCP-конфигурации, если после первоначальной установки произошла смена интерфейса управления:

```
xe-reset-networking --device=<device-name>
```

– сброс настроек сети для статической IP-конфигурации, если после первоначальной установки произошла смена интерфейса управления:

```
xe-reset-networking --device=<device-name> --mode=static --ip=<ip-
address> --netmask=<netmask> --gateway=<gateway> --dns=<dns>
```

5.15.2.4. Примеры аварийного сброса настроек для рядового участника пула

Все примеры, приведённые в предыдущем пункте, также применимы и для рядовых серверов пула. Кроме того, можно указать IP-адрес мастера пула (что будет необходимым в случае его изменения). Ниже приведены примеры управления конфигурацией рядовых участников пула:

– сброс настроек сети для конфигурации DHCP:

```
xe-reset-networking
```

– сброс настроек сети для конфигурации DHCP при изменении IP-адреса мастера пула:

```
xe-reset-networking --master=<master-ip-address>
```

– сброс настроек сети для статической IP-конфигурации, при условии, что IP-адрес мастера пула не менялся:

```
xe-reset-networking --mode=static --ip=<ip-address> --
netmask=<netmask> --gateway=<gateway> --dns=<dns>
```

– сброс настроек сети для DHCP-конфигурации, если после первоначальной установки произошла смена интерфейса управления и IP-адреса мастера пула:

```
xe-reset-networking --device=<device-name> --master=<master-ip-
address>
```

6. АДМИНИСТРИРОВАНИЕ СИСТЕМЫ ХРАНЕНИЯ

В этом разделе описывается, как оборудование физического хранилища сопоставляется с виртуальными машинами (ВМ) и программными объектами, используемыми API управления для выполнения задач, связанных с хранилищем. Подробные разделы по каждому из поддерживаемых типов хранения содержат следующую информацию:

– процедуры создания хранилища для виртуальных машин с помощью интерфейса командной строки с параметрами конфигурации устройств, зависящими от типа хранилища;

– создание моментальных снимков для целей резервного копирования;

– рекомендации по управлению хранилищем;

– настройка параметров QoS для виртуального диска.

6.1. Хранилище данных

Хранилище данных (Storage Repsitory или SR) - это определенный целевой объект хранения, в котором хранятся образы виртуальных дисков виртуальных машин (VDI). VDI является дисковой абстракцией, содержащей контент виртуального диска.

Образы VDI поддерживаются большим количеством различных типов хранилищ.

Хранилища Numa vServer имеют встроенную поддержку дисков, как локальных:

– IDE;

– SATA;

– SCSI;

– SAS.

Так и удалённых:

– iSCSI;

– NFS;

– CIFS/SMB;

– SAS;

– Fibre Channel.

Хранилища данных и абстракции VDI предоставляют поддержку усовершенствованных функций, таких как *thin-provisioning* (возможность экономичного выделения места для нужд хранения данных, далее по тексту термин даётся в англоязычном варианте либо как «экономичное выделение»), поддержка снимков состояния VDI (*snapshots*) и быстрого клонирования, которые будут выполняться на поддерживающих их подсистемах хранения. Для подсистем хранения, не поддерживающих эти операции непосредственно, предоставляется программный стек на основе спецификации механизма *предоставления виртуального жесткого диска* (*Virtual Hard Disk*, *VHD*) реализующий эти операции программно.

Каждый сервер может использовать множество хранилищ данных и различные их типы одновременно. Хранилище также может использоваться различными серверами совместно или быть выделенным определенному серверу. Совместно используемая система хранения используется участниками определенного ресурсного пула. Совместно используемое хранилище должно быть доступным по сети для каждого сервера. Все участники пула рекомендуется иметь, по крайней мере, одно совместно используемое хранилище данных.

Хранилища данных не только являются местом хранения образов виртуальных дисков (VDI), но также поддерживают операции для создания, удаления, изменения размеров, клонирования, присоединения и обнаружения содержащихся в них отдельных образов.

Хранилища являются постоянной структурой данных на диске. Для типов хранилищ *на базе блочных устройств* процесс создания нового хранилища подразумевает затирание любых существующих данных в указанном месте хранения.

6.1.1. Образы виртуальных дисков (VDI)

Образы виртуальных дисков (VDI) являются абстрактными (логическими) объектами хранения, которые предоставляются виртуальным машинам как физические диски. Образ VDI является основной единицей виртуализированного хранения в Numa vServer. Подобно хранилищам данных, образы VDI являются персистентными дисковыми объектами и существуют независимо от серверов. Фактическое дисковое представление данных образа отличается в зависимости от типа хранилища и управляется при помощи отдельного интерфейса программного модуля, называемого SM API (для каждого хранилища отдельно).

6.1.2. Физические блочные устройства (PBD)

Физические блочные устройства (*Physical Block Devices*, *PBD*) представляют собой интерфейс между физическим сервером и присоединённым SR. Физические блочные устройства являются соединительными объектами, которые позволяют сопоставить SR с сервером. PBD хранят поля конфигурации устройства, используются для соединения и взаимодействия с выбранным местом назначения трафика хранения (англ. *storage target*). Например, конфигурация устройства NFS включает IP-адрес сервера NFS и связанного пути, который хост Numa vServer использует для монтирования диска. Объекты PBD управляют присоединением данного хранилища к данному хосту Numa vServer.

6.1.3. Виртуальные блочные устройства (VBD)

Виртуальные блочные устройства, как и PBD, являются связующими объектами, позволяющими задавать соответствие между образами VDI и ВМ.

В дополнение к обеспечению механизма для присоединения (также названного подключением, англ. *plugging*) VDI в ВМ, VBD позволяет выполнить настройку параметров QoS, статистики и возможности загрузки с данного VDI.

6.1.4. Взаимосвязь объектов хранения

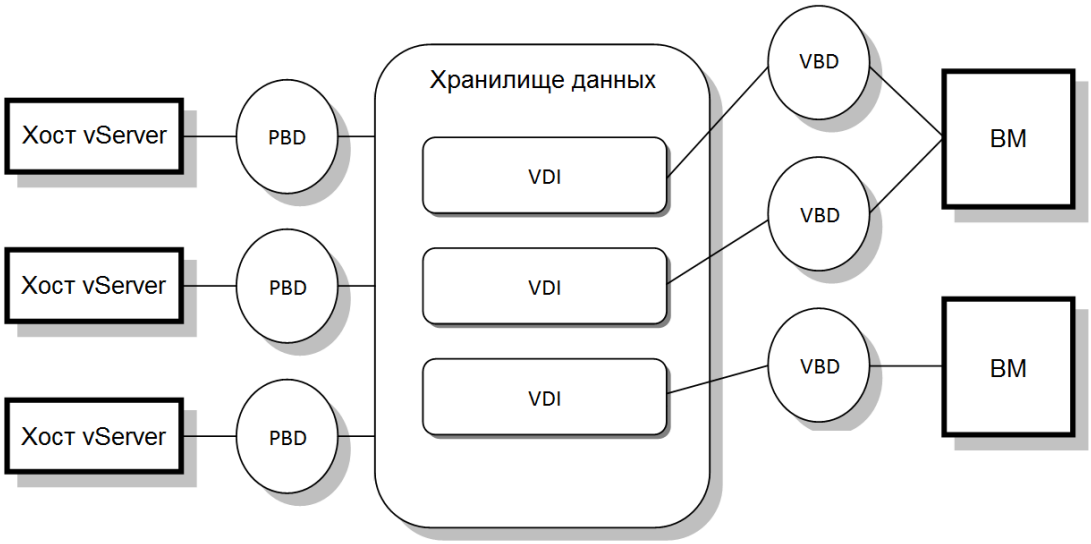

Рисунок 10 – Взаимосвязь физических, логических и виртуальных элементов системы хранения

На рисунке 10 показано, как связаны объекты хранения.

6.1.4.1.1. Форматы виртуальных дисков

Существуют следующие типы сопоставления физического хранилища с VDI:

– VHD на базе логического тома на LUN (Logical Unit – логический блок, логическая единица, на которую разбивается контейнер хранилища данных). По умолчанию в Numa vServer систему хранения на базе блочных устройств добавляет менеджер логического тома (англ. Logical Volume Manager, LVM) на диск – либо на локально подключенное устройство (хранилище типа LVM), либо на LUN, присоединённый к SAN по протоколу Fibre Channel (хранилище типа

LVMoHBA), iSCSI (хранилище типа LVMoISCSI) или SAS (хранилище типа LVMoHBA). Образы VDI представляются как отдельные тома с менеджерами LVM и хранятся в формате VHD для возможности экономичного выделения памяти (thin provisioning) для связанных узлов при создании снимков и клонировании.

– VHD на базе файла в файловой системе. Образы ВМ хранятся как файлы формата VHD с возможностью thin-provisioning в локальной (не совместно используемой) файловой системе (хранилища типа EXT), или в совместно используемой NFS (хранилище NFS).

6.1.4.1.2. Типы VDI

В общем случае создаётся VDI формата VHD. Администратор может решить использовать «сырой» («неразмеченный», англ. *raw*) тип при создании VDI (с помощью xe CLI). Чтобы проверить, был ли VDI создан с **type**=raw, нужно проверить его map-параметр **sm-config**. С этой целью могут использоваться, соответственно, команды **sr-param-list** и **vdi-param-list**.

6.1.4.1.3. Создание не размеченного виртуального диска при помощи интерфейса CLI

– выполнить следующую команду для создания VDI, указав UUID хранилища, в которое требуется поместить виртуальный диск:

```
xe vdi-create sr-uuid=<sr-uuid> type=user virtual-size=<virtual-
size> name-label=<VDI_name> sm-config:type=raw
```

– присоединить новый виртуальный диск к ВМ и использовать в ней обычные инструменты для создания и форматирования разделов (или использовать новый диск иным образом). Можно использовать команду **vbd-create** для создания нового VBD для отображения виртуального диска в ВМ.

6.1.4.1.4. Преобразование между форматами VDI

Невозможно сделать прямое преобразование между форматом VHD и raw. Вместо этого можно создать новый VDI (неразмеченный, как описано выше, или VHD), а затем скопировать данные в него из существующего тома. Рекомендуется использовать xe CLI, чтобы быть уверенным, что новый VDI имеет виртуальный размер не меньше копируемого VDI (путем проверки его поля виртуального размера, например, при помощи команды **vdi-param-list**). Затем можно присоединить этот новый VDI к ВМ и использовать в ней соответствующие инструменты (стандартные инструменты управления дисками в Windows или команда **dd** в Linux), для создания поблочной копии данных. Если новый том является томом VHD, важно использовать инструмент, который поможет избежать записи пустых секторов на диск – чтобы пространство было использовано в базовом хранилище оптимально – в этом случае основанный на файле подход копирования может подойти больше.

6.1.4.1.5. Образы виртуальных дисков на основе VHD

Файлы VHD могут быть объединены в *цепь*, позволяя двум образам виртуальных дисков совместно использовать общие данные. В случаях, когда ВМ с VHD клонируется, получающиеся виртуальные машины совместно используют общие дисковые данные во время клонирования. Каждая ВМ продолжает вносить свои собственные изменения в собственную отдельную версию VDI (механизм *copy-on-write*, *CoW*). Эта функция позволяет быстро клонировать основанные на VHD ВМ из шаблонов, упрощая и ускоряя настройку и развертывание новых ВМ.

Это приводит к созданию со временем деревьев объединённых в цепь образов VDI, так как ВМ и связанные с ними образы виртуальных дисков клонируются. Когда один из VDI в цепи удаляется, Numa vServer рационализирует другие VDI в ней для удаления ненужных образов. Этот процесс объединения (англ. *coalescing*) работает асинхронно. Сумма исправленного дискового пространства и время, необходимое для выполнения процесса, зависят от размера

VDI и суммы совместно используемых данных. Только один подобный процесс объединения может быть активен в хранилище в каждый момент времени. Поток этого процесса запускается на главном хосте (мастере) хранилища.

Если имеет место критическая работа ВМ мастера пула и медленный случайный ввод-вывод из-за этого процесса, можно предпринять следующие шаги:

– переместить ВМ на другой хост (миграция);

– установить дисковый приоритет ввода-вывода в более высокий уровень и скорректировать настройки планировщика (см. пп. 6.1.6.11 Настройки QoS для виртуальных дисков).

Формат VHD, используемый в Numa vServer хранилищами на основе LVM или файлов, использует механизм экономичного выделения места *thin-provisioning*. Файл образа автоматически расширяется (блоками, по 2 Мбайта), по мере того как ВМ записывает на диск данные. Для VHD на базе файла это имеет значительное преимущество: файлы образов ВМ занимают не больше пространства в физической системе хранения, чем реально требуется для записанных на них данных. В случае VHD, основанного на LVM, объём логического контейнера должен быть изменён до виртуального размера VDI, однако при создании снимка или клона неиспользуемое место на диске CoW-экземпляра исправляется.

Различие между описанными двумя вариантами поведения может быть характеризовано следующим образом:

– для VHD, основанных на LVM, «разностные» узлы в цепочке используют ровно столько данных, сколько было записано на диск, но узлы-«листья» такой древоподобной структуры (то есть клоны VDI) занимают виртуальный диск полностью. Узлы-«листья», представляющие собой снимки VDI, продолжают занимать минимальное требуемое пространство, когда не используются, и для сохранения своего размера могут быть присоединены в режиме «только для чтения». Когда узлы-снимки присоединяются в режиме «чтение и запись», они будут полностью «растянуты» на всё свободное пространство диска при присоединении и возвращены к реальным размерам при отсоединении;

– в случае VHD на базе файлов все узлы используют лишь объёмы памяти, соразмерные объёмам записанных данных, и размеры файлов, соответствующих узлам-«листьям», растут в соответствии с темпами активной записи. Если для новой ВМ выделяется VDI на 100 Гбайт, и устанавливается ОС, файл VDI физически будет иметь размер, равный совокупному объёму данных ОС, записанных на диск, плюс небольшие издержки метаданных.

При клонировании ВМ на основе единственного шаблона VHD каждая дочерняя ВМ формирует цепочку, где записаны новые изменения к новой ВМ, а старые блоки считываются непосредственно из родительского шаблона. Если новая ВМ была преобразована в дальнейшем в шаблон и из него появляется ещё больше клонов, получающаяся цепочка приведёт к ухудшению производительности. Numa vServer поддерживает максимальную длину цепочки 30, но обычно не рекомендуется приближаться к этому пределу без серьезных оснований. При наличии сомнений лучше *скопировать* ВМ, используя команду **vm**-**copy**, которая автоматически сбросит длину цепочки в «0».

6.1.5. Форматы хранилищ

Новые хранилища могут быть созданы через интерфейс CLI командой **sr**-**create**. Эта команда создаёт новое хранилище данных на аппаратных средствах хранения (потенциально уничтожая любые уже существующие данные на них) и создает программный объект «хранилище» и соответствующую запись PBD, позволяя виртуальным машинам использовать хранилище. К успешно созданному хранилищу автоматически подключается PBD. Если для хранилища установлен флаг shared=true, запись PBD создаётся и подключается для каждого хоста Numa vServer в пуле.

При создании хранилища для системы хранения с поддержкой протокола IP (англ. *IP-based storage*) (iSCSI или NFS), можно сконфигурировать под его нужды сетевой адаптер/интерфейс, обрабатывающий управляющий трафик, или новый сетевой адаптер. Процесс присвоения IP-адреса сетевому интерфейсу (адаптеру) описан в пп. 5.10 Настройка сетевого интерфейса, выделенного для соединения с хранилищем.

Все типы хранилищ в Numa vServer поддерживают изменение размеров VDI, быстрое клонирование и снимки состояний. Хранилища на основе LVM (локальные, iSCSI или HBA) обеспечивают *thin-provisioning* для снимка и скрытых родительских узлов. Другие типы хранилищ имеют полную поддержку *thin-provisioning*, включая таковую и для активных виртуальных дисков.

Когда VHD VDI не присоединены, например, в случае снимка VDI, они сохраняются по умолчанию с поддержкой *thin-provisioning*. Из-за этого обязательно необходимо обеспечить для VDI достаточный объём доступного дискового пространства, для превращения его в диск с *неэкономичным выделением* (*thick-provisioning*) при попытке присоединить его.

Клоны VDI будут поддерживать *thick-provisioning*.

Максимальные поддерживаемые размеры VDI описаны в таблице 8.

Таблица 8 – Поддерживаемые максимальные размеры VDI

Формат хранилища данных	**Максимальный размер VDI**
EXT3	2 Тбайта
LVM	2 Тбайта
NFS	2 Тбайта
LVMoFCOE	2 Тбайта
LVMoiSCSI	2 Тбайта
LVMoHBA	2 Тбайта

6.1.5.1. Локальный LVM

Хранилища типа «локальный LVM» представляют собой диски в локально присоединенной *группе томов*. По умолчанию Numa vServer использует локальный диск на том физическом хосте, на котором он установлен. Менеджер логического тома (LVM) Linux используется для управления хранением ВМ. Образы VDI представляются в формате VHD на логическом томе LVM указанного размера.

6.1.5.1.6. Особенности работы с LVM

Реализованные в Numa vServer возможности создания снимков состояния и быстрого клонирования для основанных на LVM хранилищ влекут соответствующие потери производительности. В случаях, где важна оптимальная производительность, Numa vServer поддерживает создание из образов VDI в не размеченном (*raw*) формате в дополнение к формату VHD по умолчанию. Функциональность снимка Numa vServer не поддерживается на необработанных образах VDI.

Нетранспортабельные снимки, использующие по умолчанию провайдер Windows VSS, будут работать с любым типом из VDI.

Не следует делать снимок ВМ, имеющей присоединенные диски с типом **type**=raw. Это может привести к созданию частичного снимка. Можно идентифицировать такой снимок образа VDI путём проверки поля **snapshot-of** и затем удалить его.

6.1.5.1.7. Создание локального хранилища на основе LVM

Хранилище на основе LVM создаётся по умолчанию при установке хоста.

Для создания локального хранилища LVM на /dev/sdb используется следующая команда:

```
xe sr-create host-uuid=<valid_uuid> content-type=user name-
label=<Example Local LVM SR> shared=false device-
config:device=/dev/sdb type=lvm
```

6.1.5.2. Локальный EXT3

Использование EXT3 включает *экономичное выделение места* (*thin-provisioning*) на локальных хранилищах. Однако типом хранилища по умолчанию является LVM, обеспечивающий непротиворечивую производительность записи и предотвращающий переполнение хранения (англ. *overcommitting*). Клиенты, использующие EXT3, могут заметить уменьшение производительности:

– при выполнении операций жизненного цикла ВМ (создание, приостановка и возобновление ВМ;

– создание больших файлов в файловой системе ВМ.

Локальное хранилище на основе EXT должно быть сконфигурировано с помощью интерфейса CLI.

6.1.5.2.8. Создание локального хранилища на основе EXT3 (ext)

Для создания локального хранилища EXT3 на /dev/sdb используется следующая команда:

```
xe sr-create host-uuid=<valid_uuid> content-type=user name-
label=<Example Local EXT3 SR> shared=false device-
config:device=/dev/sdb type=ext
```

6.1.5.3. Udev

Тип **udev** подразумевает устройства, подключенные как образы VDI при помощи диспетчера устройств **udev**.

Numa vServer имеет два хранилища типа **udev**, которые представляют собой съемные устройства хранения: первое используется для CD- или DVD-диска в физическом приводе CD-ROM или DVD-ROM хоста Numa vServer, второе – для USB-устройства, подключенного к порту USB Numa vServer. Образы VDI, соответствующие этим носителям, подключаются и отключаются в соответствии с установкой и отсоединением CD-дисков и USB накопителей.

6.1.5.4. ISO

Тип ISO относится к образам CD-дисков, сохраненным в виде файлов в формате ISO. Этот тип хранилищ полезен для создания совместно используемых библиотек ISO.

Для хранилищ, хранящих библиотеку образов ISO, параметр content-type (тип контента) должен быть установлен в значение **ISO**.

Например:

```
xe sr-create host-uuid=<valid_uuid> content-type=iso type=iso name-
label=<Example ISO SR> device-config:location=<nfs server:path>
```

Рекомендуется использовать SMB версии 3.0 для монтирования ISO SR на файловом сервере Windows. Версия 3.0 выбрана по умолчанию, потому что она более безопасна и надежна, чем SMB версии 1.0. Однако можно смонтировать ISO SR, используя SMB версии 1.0, используя следующую команду:

```
xe sr-create content-type=iso type=iso shared=true device-
config:location=<valid location> device-config:username=<username>
device-config:cifspassword=<password> device-config:type=cifs
device-config:vers=<Choose either 1.0 or 3.0> name-label=<Example ISO
SR>
```

6.1.5.5. Программная поддержка iSCSI

Numa vServer предоставляет поддержку для совместно используемых хранилищ на iSCSI LUN. Поддержка реализуется при помощи iSCSI-инициатора «Open-iSCSI» (программное обеспечение) или при помощи поддерживаемого *адаптера шины хоста* (англ. *Host Bus Adapter*, *HBA*) iSCSI. Шаги, которые необходимо выполнить для использования HBA iSCSI, идентичны описанным для адаптеров Fibre Channel HBA (пп. 6.1.5.7.12 Удаление записей устройств SAS, FC или iSCSI основанных на HBA).

Поддержка совместного использования iSCSI при помощи программного обеспечения инициатора iSCSI реализована на основе Менеджера томов Linux (LVM) и обеспечивает тот же выигрыш в производительности, что и использование образов VDI на LVM в случае локальных дисков. Совместное использование хранилищ iSCSI, основанное на программном обеспечении инициатора хоста, допускает использование миграции ВМ: ВМ могут быть запущены на любых серверах пула и перемещаться между ними без существенной потери времени.

Хранилища iSCSI полностью используют LUN, определённый в процессе создании хранилища, и не могут охватывать более одного LUN. Поддержка CHAP предоставляется для аутентификации клиента – как во время инициализации информационного канала, так и во время фаз открытия LUN (*LUN discovery phases*).

> Размер блока iSCSI LUN должен составлять 512 байт

6.1.5.5.9. Настройка iSCSI для хостов Numa vServer

Для однозначного определения в сети у всех инициаторов iSCSI и мест назначения трафика хранения («целей») должны быть уникальные имена. Инициатор имеет адрес инициатора iSCSI, и цель имеет целевой адрес iSCSI. В совокупности они дают так называемые *iSCSI Qualified Names*, *IQN* («уточнённые имена iSCSI»).

Хосты Numa vServer поддерживают наличие единственного инициатора iSCSI, автоматически создаваемого и конфигурирующегося со случайным IQN при установке хоста. Один инициатор может использоваться для соединения с несколькими целевыми iSCSI одновременно.

Целевые iSCSI обычно обеспечивают управление доступом с помощью списков IQN инициатора iSCSI таким образом, что все целевые iSCSI/LUN для возможности получения доступа к ним со стороны хоста Numa vServer должны быть сконфигурированы для предоставления доступа на основе IQN инициатора хоста. Точно так же целевые LUN, которые планируются

использовать в качестве совместно используемых iSCSI-хранилищ, должны быть сконфигурированы для предоставления доступа для IQN всех хостов пула.

Целевые объекты iSCSI, которые не обеспечивают управление доступом, обычно по умолчанию ограничивают доступ LUN одним инициатором для обеспечения целостности данных. Если iSCSI LUN используется в качестве общего SR на нескольких серверах в пуле, убедитесь, что для указанного LUN включен доступ с несколькими инициаторами

Значение IQN хоста Numa vServer может быть скорректировано следующей командой при использовании программного инициатора iSCSI:

```
xe host-param-set uuid=<valid_host_id> other-
config:iscsi_iqn=<new_initiator_iqn>
```

Обязательно, чтобы у каждой цели iSCSI и инициатора было уникальное IQN. Если используется групповой идентификатор IQN, могут произойти повреждения данных или отказ доступа к LUN.

Не следует изменять IQN хоста Numa vServer, имеющего хранилища, присоединенные по iSCSI. Такое изменение может привести к отказам соединения с новыми местами назначения (*targets*) или существующими хранилищами.

6.1.5.6. Программное хранилище FCoE

Программное обеспечение FCoE предоставляет стандартную структуру, к которой поставщики оборудования могут подключить свои сетевые адаптеры с поддержкой FCoE и получить те же преимущества, что и аппаратный FCoE. Эта функция исключает необходимость использования дорогих адаптеров HBA.

Перед созданием программного хранилища FCoE вручную выполните настройку, необходимую для предоставления LUN хосту. Эта конфигурация включает в себя настройку структуры FCoE и выделение LUN для общедоступного имени вашей сети SAN (PWWN). После завершения этой конфигурации доступный LUN подключается к CNA хоста как устройство SCSI. Устройство SCSI может затем использоваться для доступа к LUN, как если бы оно было локально подключенным устройством SCSI. Для получения информации о настройке физического коммутатора и массива для поддержки FCoE см. *** Документацию, предоставленную поставщиком.

Программное обеспечение FCoE может использоваться с Open vSwitch и Linux bridge в качестве серверной сети

6.1.5.6.10. Настройка программного хранилища FCoE

Перед созданием программного хранилища FCoE необходимо убедиться, что к хосту подключены сетевые карты с поддержкой FCoE.

Команда для создания программного хранилища FCoE:

```
xe sr-create type=lvmofcoe name-label="FCoE SR" shared=true device-
config:SCSIid=<SCSI_id>
```

6.1.5.7. Аппаратные контроллеры шин хоста (Hardware HBA)

Этот подраздел описывает различные операции, требуемые для управления аппаратными контроллерами шин SAS, Fibre Channel и iSCSI хостов Numa vServer.

6.1.5.7.11. Пример настройки QLogic iSCSI HBA

Полное руководство по конфигурированию HBA для Fibre Channel QLogic и iSCSI

Как только HBA физически установлен на хосте Numa vServer, возможно использовать следующие шаги для конфигурирования HBA:

– настроить сетевую конфигурацию протокола IP для HBA. В этом примере предполагается использовать порт 0 (*port 0*) для DHCP и HBA. Необходимо определить надлежащие значения для использования статической IP-адресации или многопортового HBA.

– добавить постоянную iSCSI-цель для порта 0 HBA:

– использовать команду xe **sr-probe** для запуска повторного сканирования контроллера HBA и вывода на экран списка доступных LUN.

6.1.5.7.12. Удаление записей устройств SAS, FC или iSCSI основанных на HBA

Данный шаг не является обязательным и должен выполняться лишь опытными администраторами при возникновении такой необходимости

Каждому основанному на HBA LUN соответствует запись, содержащая глобальный путь к устройству *** /dev/disk/by-scsibus в формате <SCSIid>-<adapter>:<bus>:<target>:<lun> и стандартный путь к устройству в /dev. Для демонтирования записи о LUN, который больше не планируется использовать в качестве хранилища, следует выполнить следующие действия:

– использовать **sr-forget** или **sr-destroy** для хранилищ, которые требуется удалить из базы данных хоста Numa vServer (см. пп. 6.1.6.3 Удаление хранилища);

– удалите конфигурацию зонирования в пределах SAN для требуемого LUN на хостах;

– использовать команду **sr-probe** для определения значений ADAPTER, BUS, TARGET и LUN, соответствующих удаляемому LUN;

– удалить записи устройства следующей командой:

```
echo "1" >
/sys/class/scsi_device/<adapter>:<bus>:<target>:<lun>/device/delete
```

Необходимо убедиться, какой именно LUN удалять. Случайное удаление LUN, необходимого для работы хоста, такого как загрузочное или корневое устройство, делает хост непригодным для использования.

6.1.5.8. Совместно используемое хранилище типа LVM

Совместно используемый тип LVM представляет диски как логические тома в группе томов, созданной в LUN на основе iSCSI (FC или SAS).

Размер блока iSCSI LUN должен составлять 512 байт

6.1.5.8.13. Создание совместно используемого LVM поверх хранилища iSCSI с использованием программного инициатора iSCSI (lvmoiscsi)

Таблица 9 – Параметры конфигурации устройства для хранилищ lvmoiscsi

Имя параметра	Описание	Обязательность
target	IP-адрес или имя хоста iSCSI-файлера (*iSCSI-filer*), на котором расположено хранилище	Да
targetIQN	IQN-target-адрес iSCSI-файлера, на котором расположено хранилище	Да
SCSIid	Идентификатор шины SCSI целевого LUN	Да
chapuser	Имя пользователя, которое будет использоваться для аутентификации CHAP	Нет
chappassword	Пароль, который будет использоваться для аутентификации CHAP	Нет
port	Номер сетевого порта, на который посылается запроса к целевому хранилищу	Нет
usediscoverynumber	Определенный индекс записи iSCSI для использования	Нет
incoming_chapuser	Имя пользователя, которое фильтр iSCSI будет использовать для аутентификации на хосте	Нет
incoming_chappassword	Пароль, который фильтр iSCSI будет использовать для аутентификации на хосте	Нет

Для создания совместно используемого хранилища типа lvmoiscsi на определенном LUN iSCSI используют следующую команду:

```
xe sr-create host-uuid=<valid_uuid> content-type=user name-
label=<"Example shared LVM over iSCSI SR"> shared=true device-
config:target=<target_ip=> device-config:targetIQN=<target_iqn>
device-config:SCSIid=<scsi_id> type=lvmoiscsi
```

6.1.5.8.14. Создание совместно используемого LVM поверх хранилища на базе Fibre Channel/Fibre Channel over Ethernet/iSCSI HBA или SAS (lvmohba)

Хранилища типа lvmohba создаются и управляются через интерфейс CLI.

Для создания совместно используемого хранилища lvmohba необходимо выполнить следующие шаги на каждом хосте пула:

– внести в зону один или более LUN каждому хосту Numa vServer пула. Детали процесса зависят от используемого оборудования SAN;

– использовать команду **sr**-**probe** для определения глобального пути устройства LUN HBA. Команда **sr**-**probe** осуществляет пересканирование установленных HBA в системе для обнаружения любых новых LUN, призонированных к хосту, и возвращает список свойств для каждого найденного LUN. Следует определить параметр **host**-**uuid**, чтобы убедиться, что пересканирование происходит на требуемом хосте.

Глобальный путь устройства, возвращённый в свойстве <path>, будет годен для всех хостов пула и поэтому должен использоваться в качестве значения для параметра **device**-**config:device** при создании хранилища. Если присутствует несколько LUN, используют имя

поставщика, размер LUN, порядковый номер LUN или идентификатор SCSI для включения в состав <path> для однозначной идентификации требуемого LUN:

```
xe sr-probe type=lvmohba host-uuid=1212c7b3-f333-4a8d-a6fb-
80c5b79b5b31

    Error code: SR_BACKEND_FAILURE_90
    Error parameters: , The request is missing the device parameter,
\
    <?xml version="1.0" ?>
    <Devlist>
        <BlockDevice>
            <path>
                /dev/disk/by-id/scsi-
360a98000686669496734463876653 36f
            </path>
            <vendor>
                HITACHI
            </vendor>
            <serial>
                730157980002
            </serial>
            <size>
                80530636800
            </size>
            <adapter>
                4
            </adapter>
            <channel>
                0
            </channel>
            <id>
                4
            </id>
            <lun>
                2
            </lun>
            <hba>
                qla2xxx
            </hba>
        </BlockDevice>
        <Adapter>
            <host>
                Host4
            </host>
            <name>
                qla2xxx
            </name>
            <manufacturer>
                QLogic HBA Driver
            </manufacturer>
```

```
            <id>
                4
            </id>
        </Adapter>
    </Devlist>
```

– на основном хосте пула создать хранилище, задав глобальный путь устройства в соответствии с путём, возвращённым в свойстве <path> в качестве результата команды sr-probe. PBD будут созданы и включены для каждого хоста в пуле автоматически:

```
xe sr-create host-uuid=<valid_uuid> content-type=user name-
label=<"Example shared LVM over HBA SR"> shared=true device-
config:SCSIid=<device_scsi_id> type=lvmohba
```

6.1.5.9. NFS и SMB

Общие ресурсы на серверах NFS (которые поддерживают NFSv4 или NFSv3) или на серверах SMB (которые поддерживают SMB 3.0) можно сразу использовать в качестве SR для виртуальных дисков. VDI хранятся только в формате Microsoft VHD. Кроме того, поскольку эти SR могут совместно использоваться, VDI, хранящиеся в совместно используемых SR, позволяют:

– запускать виртуальные машины на любых серверах Numa vServer в пуле ресурсов;

– использовать миграцию виртуальных машин между серверами Numa vServer в пуле ресурсов с использованием динамической миграции (без заметных простоев)

Поддержка SMB 3.0 ограничивается возможностью подключения к общему ресурсу по протоколу 3.0. Дополнительные функции, такие как прозрачная отработка отказа, зависят от доступности функций в вышестоящем ядре Linux и не поддерживаются в Numa vServer

Для NFSv4 поддерживается только **AUTH_SYS** тип аутентификации

VDI, хранимые на файловых SR, имеют *thin-provisioning* (экономичное выделение места хранения). Дисковое пространство для файла образа выделяется, когда виртуальная машина записывает данные на диск. Этот подход имеет значительное преимущество в том, что файлы образов виртуальных машин занимают столько места в хранилище, сколько требуется. Например, если VDI 100 Гбайт выделен для виртуальной машины и установлена ОС, файл VDI отражает только размер данных ОС, записанных на диск, а не все 100 Гбайт.

Файлы VHD также могут быть объединены в цепочку, что позволяет двум VDI обмениваться общими данными. В случаях, когда файловая виртуальная машина клонирована, полученные виртуальные машины совместно используют общие данные на диске во время клонирования. Каждая виртуальная машина продолжает вносить свои собственные изменения в изолированную версию VDI с копированием при записи. Эта функция позволяет быстро клонировать виртуальные машины на основе файлов из шаблонов, обеспечивая очень быструю подготовку и развертывание новых виртуальных машин.

Максимальная поддерживаемая длина цепочек VHD равняется 30.

Реализации SR и VHD на основе файлов в Numa vServer предполагают, что они имеют полный контроль над каталогом SR на файловом сервере. Администраторы не должны изменять содержимое каталога SR, так как это может привести к повреждению содержимого VDI.

Поскольку VDI на файловых SR создаются с *thin-provisioning*, администраторы должны убедиться, что на файловых SR достаточно дискового пространства для всех необходимых VDI. Сервер Numa vServer не проверяет наличие пространства, необходимого для VDI на файловых SR.

6.1.5.9.15. Настройка общего хранилища NFS

Чтобы создать NFS SR, необходимо указать имя хоста или IP-адрес сервера NFS. Можно создать SR на любом допустимом пути назначения; используя команду **sr-probe** для отображения списка допустимых путей назначения, экспортируемых сервером.

В тех случаях, когда Numa vServer используется с хранилищем нижнего уровня, он осторожно ожидает подтверждения всех записей перед передачей подтверждений на виртуальные машины. Этот подход требует заметных затрат производительности и может быть решен путем установки хранилища для представления точки монтирования SR как экспорта в асинхронном режиме. Записи асинхронного экспорта подтверждают, что на самом деле их нет на диске.

Сервер NFS должен быть настроен на экспорт указанного пути на всех серверах в пуле. Если эта конфигурация не выполнена, создание SR и подключение PBD завершится неудачей.

В реализации Numa vServer NFS по умолчанию используется TCP. Если ситуация позволяет, можно настроить реализацию на использование UDP в сценариях, где может быть выигрыш в производительности. Для этого при создании SR необходимо указать параметр **device-config:useUDP=true**.

Для создания совместно используемого хранилища NFS на 192.168.1.10:/export1 можно использовать следующую команду:

```
xe sr-create content-type=user name-label="shared NFS SR"
shared=true device-config:server=192.168.1.10 device-
config:serverpath=/export1 type=nfs nfsversion=<"3", "4">
```

Для создания NFS SR без общего доступа необходимо выполнить следующую команду:

```
xe sr-create host-uuid=host_uuid content-type=user name-label="Non-
shared NFS SR" device-config:server=192.168.1.10 device-
config:serverpath=/export1 type=nfs nfsversion=<"3", "4">
```

6.1.5.9.16. Настройка общего хранилища SMB

Чтобы создать SR SMB, необходимо указать имя хоста или IP-адрес сервера SMB, полный путь экспортируемого общего ресурса и соответствующие учетные данные.

Например, чтобы создать общий SR SMB 192.168.1.10:/share1, используется следующая команда:

```
xe sr-create content-type=user name-label="Example shared SMB SR"
shared=true device-config:server=//192.168.1.10/share1 device-
config:username=<valid_username> device-config:password=<valid_password>
type=smb
```

Чтобы создать не общий SR SMB, используется следующая команда:

```
xe sr-create host-uuid=host_uuid content-type=user name-label="Non-
shared SMB SR" device-config:server=//192.168.1.10/share1 device-
config:username=<valid_username> device-config:password=<valid_password>
type=smb
```

6.1.5.10. LVM поверх аппаратных HBA

Хранилища типа LVM поверх аппаратных HBA представляют диски как устройства VHD на логических томах (*logical volumes*) в созданной *группе томов* на обеспечении LUN HBA, например, основанном на аппаратных средствах iSCSI или поддержке FC.

Хосты Numa vServer поддерживают SAN (*storage area networks*) на Fibre Channel (*FC*) через адаптеры HBA от Emulex или QLogic. Все настройки FC, необходимые для предоставления LUN FC хосту, должны быть выполнены вручную, включая устройства хранения, сетевые устройства и адаптеры HBA на хосте. Как только вся конфигурация FC будет настроена, HBA представит хосту устройство SCSI, работающее «поверх» FC LUN. Устройство SCSI можно будет использовать для получения доступа к FC LUN так, будто это локально присоединённое устройство SCSI.

Для вывода списка поддерживаемых в настоящий момент на хосте устройств SCSI на основе LUN следует использовать команду sr-probe. Эта команда вызывает сканирование для новых подобных устройств SCSI. Значение пути, возвращаемое sr-probe для устройств SCSI, пригодно для всех хостов с доступом к LUN и потому должно указываться при создании совместно используемого хранилища в пуле.

Те же функции применимы к iSCSI HBA от QLogic.

См. пп. 6.1.6.1 Создание хранилищ для получения дополнительной информации о создании совместно используемых хранилищ на базе HBA FC и iSCSI.

> Numa vServer для Fibre Channel не поддерживает прямое сопоставление LUN с виртуальной машиной. LUN на основе HBA должны быть сопоставлены с хостом и указаны для использования в SR. VDI в SR выставляются виртуальным машинам как стандартные блочные устройства.

Numa vServer для Fibre Channel не поддерживает прямое сопоставление LUN с виртуальной машиной. LUN на основе HBA должны быть сопоставлены с хостом и указаны для использования в SR. VDI в SR выставляются виртуальным машинам как стандартные блочные устройства.

6.1.6. Создание и настройка хранилищ данных

В этом разделе описываются процессы создания хранилищ различных типов и предоставления доступа к ним хостам Numa vServer. Примеры показывают, как сделать это с использованием интерфейса командной строки.

6.1.6.1. Создание хранилищ

В этом разделе объясняется, как создать хранилища (SR) разных типов и сделать их доступными для сервера Numa vServer. Приведенные примеры охватывают создание SR с помощью xe CLI.

Создание нового хранилища данных для использования на хосте Numa vServer включает два основных шага, выполняемых с помощью CLI:

– сканирование хранилищ для определения значений любых требуемых параметров;

– создание хранилища. Для инициализации объектов хранилища и связанных объектов PBD следует включить PBD и активировать хранилище.

Эти шаги отличаются в зависимости от типа создаваемого хранилища, в случае успеха команда **sr**-**create** возвращает UUID созданного хранилища.

Когда хранилища долго не используются, они могут также быть удалены, чтобы освободить физическое устройство, или «деактивированы» (*forgotten*) для отсоединения хранилища от одного хоста Numa vServer и присоединения к другому (см. пп.6.1.6.3 Удаление хранилища).

6.1.6.2. Сканирование хранилища

Команда **sr-probe** может использоваться двумя способами:

– нахождение значений неизвестных параметров для использования в создании хранилища;

– возвращение списка уже существующих хранилищ.

В обоих случаях **sr-probe** работает путём определения типа хранилища и одного или нескольких параметров **device**-**config** для этого типа хранилищ. Когда набор параметров неполный, **sr**-**probe** сообщает об ошибке (сообщает о том, что указанные параметры отсутствуют, а также предлагает возможные варианты значений для недостающих параметров). Когда предоставлен полный набор параметров, команда возвращает список существующих хранилищ. Весь вывод sr-probe возвращается в виде XML.

Например, для сканирования определённого места назначения трафика хранения (т. н. «цель») на базе iSCSI следует указать его имя или IP-адрес. В результате сканирования будет возвращён набор IQN, доступных на «цели»:

```
xe sr-probe type=lvmoiscsi device-config:target=192.168.1.10

    Error code: SR_BACKEND_FAILURE_96
    Error parameters: , The request is missing or has an incorrect
target IQN parameter, \
    <?xml version="1.0" ?>
    <iscsi-target-iqns>
        <TGT>
            <Index>
                0
            </Index>
            <IPAddress>
                192.168.1.10
            </IPAddress>
            <TargetIQN>
                iqn.192.168.1.10:filer1
            </TargetIQN>
        </TGT>
    </iscsi-target-iqns>
```

При повторном сканировании цели с заданием как имени/IP-адреса, так и требуемого IQN будет возвращён набор идентификаторов SCSI (точнее, LUN), доступных на target/IQN.

```
xe sr-probe type=lvmoiscsi device-config:target=192.168.1.10 device-
config:targetIQN=iqn.192.168.1.10:filer1

    Error code: SR_BACKEND_FAILURE_107
    Error parameters: , The SCSIid parameter is missing or
incorrect, \
    <?xml version="1.0" ?>
    <iscsi-target>
        <LUN>
            <vendor>
                IET
            </vendor>
            <LUNid>
                0
            </LUNid>
            <size>
                42949672960
            </size>
            <SCSIid>
                14945540000000000000000002000000b70200000f000000
            </SCSIid>
        </LUN>
    </iscsi-target>
```

Повторное сканирование цели с предоставлением всех трёх параметров возвращает список хранилищ, существующих на LUN (если таковые имеются).

```
xe sr-probe type=lvmoiscsi device-config:target=192.168.1.10  \
    device-config:targetIQN=192.168.1.10:filer1 \
    device-
config:SCSIid=149455400000000000000000002000000b70200000f000000

    <?xml version="1.0" ?>
    <SRlist>
        <SR>
            <UUID>
                3f6e1ebd-8687-0315-f9d3-b02ab3adc4a6
            </UUID>
            <Devlist>
                /dev/disk/by-id/scsi-
149455400000000000000000002000000b70200000f000000
            </Devlist>
        </SR>
    </SRlist>
```

Таблица 10 содержит параметры для каждого типа хранилищ с указанием, какие из параметров сканируются.

Таблица 10 – Параметры хранилищ

Тип хранилища	**Параметр конфигурации устройства, в порядке зависимости**	**Может ли быть сканирован**	**Требуется ли для sr-create**
lvmoiscsi	target	Нет	Да
	Chapuser	Нет	Нет
	chappassword	Нет	Нет
	targetIQN	Да	Да
	SCSIid	Да	Да
lvmohba	SCSIid	Да	Да
NetApp	Target	Нет	Да
	Username	Нет	Да
	Password	Нет	Да
	Chapuser	Нет	Нет
	chappassword	Нет	Нет
	Aggregate	Нет*	Да

Тип хранилища	**Параметр конфигурации устройства, в порядке зависимости**	**Может ли быть сканирован**	**Требуется ли для sr-create**
	FlexVols	Нет	Нет
	Allocation	Нет	Нет
	Asis	Нет	Нет
nfs	Server	Нет	Да
	serverpath	Да	Да
lvm	Device	Нет	Да
ext	Device	Нет	Да
EqualLogic	target	Нет	Да
	username	Нет	Да
	password	Нет	Да
	chapuser	Нет	Нет
	chappassword	Нет	Нет
	storagepool	Нет**	Да

* – сканирование Aggregate возможно только во время выполнения sr-create. Должно быть сделано так, чтобы aggregate мог быть определен в точке, в которой было создано хранилище.

** – сканирование пула хранилища возможно только во время sr-create. Должно быть сделано так, чтобы агрегат мог быть определен в точке, в которой было создано хранилище.

6.1.6.3. Удаление хранилища

Хранилище можно удалить различными способами:

– **отсоединить:** разрывает связь между хранилищем и хостом (отключение PBD). SR (и его VDI) становятся недоступными. Содержимое VDI и метаинформация, используемая виртуальными машинами для доступа к VDI, сохраняются. Отсоединение можно использовать, когда необходимо временно отключить SR, например, для обслуживания. Отдельный SR может быть позже присоединен.

– **забыть:** сохраняет содержимое SR на физическом диске, но информация, которая подключает виртуальную машину к ее VDI, навсегда удаляется. Например, позволяет повторно подключить SR к другому серверу Numa vServer, не удаляя содержимое SR.

– **уничтожить:** полностью удаляет содержимое с физического диска и все упоминания о SR.

Для того что бы уничтожить или забыть PBD подключенный к SR, он должен быть отключен от хоста.

– отсоединить PBD для разрыва связи SR с сервером Numa vServer:

```
xe pbd-unplug uuid=<pbd_uuid>
```

– уничтожить хранилище, PBD и записи в базе данных:

```
xe sr-destroy uuid=<sr_uuid>
```

– забыть хранилище, что удалит PBD и записи в базе данных, но оставит содержимое SR на физическом носителе без изменений:

```
xe sr-forget uuid=<sr_uuid>
```

Удаление программного объекта, соответствующего хранилищу, может занять некоторое время, пока «сборщик мусора» не удалит его.

6.1.6.4. Повторный ввод хранилища в работу

Чтобы повторно ввести в работу хранилище, которое ранее было деактивировано («забыто»), необходимо выполнить специальную команду sr-introduce, создать PBD и вручную подключить PBD к надлежащему Numa vServer для активации хранилища.

Следующий пример описывает порядок действий для повторного ввода в работу хранилища типа lvmoiscsi:

– просканировать (*probe*) существующее хранилище для определения его UUID:

```
xe sr-probe type=lvmoiscsi device-config:target=192.168.1.10 device-
config:targetIQN=192.168.1.10:filer1 device-
config:SCSIid=149455400000000000000000002000000b70200000f000000
```

– выполнить **sr-introduce** для хранилища, UUID которого возвращен командой **sr-probe** на первом шаге (команда возвратит UUID нового хранилища):

```
xe sr-introduce content-type=user name-label=<Example Shared LVM over
iSCSI SR> shared=true uuid=<valid_sr_uuid> type=lvmoiscsi
```

– создать для данного хранилища новый PBD (команда возвратит UUID нового PBD):

```
xe pbd-create type=lvmoiscsi host-uuid=<valid_uuid> sr-
uuid=<valid_sr_uuid> device-config:target=<192.168.0.1> device-
config:targetIQN=<192.168.1.10:filer1> device-
config:SCSIid=<149455400000000000000000002000000b70200000f000000>
```

– подключить PBD для присоединения к хранилищу:

```
xe pbd-plug uuid=<pbd_uuid>
```

– проверить статус подключения PBD. Если подключение успешно, то свойство **currently-attached** будет равно true («истина»):

```
xe pbd-list sr-uuid=<sr_uuid>
```

Шаги 3–5 должны быть выполнены для каждого хоста в пуле

6.1.6.5. Расширение LUN

Чтобы увеличить размер LUN, выделенного для сервера Numa vServer необходимо:

– увеличить размер LUN в хранилище;

– на Numa vServer выполнить команду:

```
xe sr-scan sr-uuid=<sr_uuid>
```

Эта команда повторно сканирует SR, что позволяет дополнить его ёмкость.

Уменьшение размера LUN в массиве хранения может привести к потере данных

6.1.6.6. «Живая» миграция VDI

«Живая» миграция (англ. *live migration*) VDI позволяет администратору перемещать VDI виртуальных машин без выключения ВМ. Это делает доступными администратору такие операции:

– перемещение ВМ из низкоскоростных хранилищ в более быстрые, отказоустойчивые, основанные на массиве (*array-backed*) хранилища;

– перемещение ВМ из среды разработки в среду готового продукта;

– перемещение между уровнями системы хранения (*storage tiers*), когда ВМ ограничивается объёмом хранилища;

– выполнение обновлений массива хранения.

6.1.6.6.17. Ограничения и предостережения

«Живая» миграция VDI подвергается следующим ограничениям:

– на целевом хранилище должно быть доступно достаточное дисковое пространство;

– образы VDI, имеющие более одного снимка, не могут быть перемещены.

6.1.6.7. «Холодная» миграция образов VDI между хранилищами (offline-миграция)

VDI, связанные с виртуальной машиной, могут быть скопированы с одного SR на другой для соответствия требованиям обслуживания или многоуровневой конфигурации хранилища.

6.1.6.7.18. Копирование отдельных образов виртуальных дисков на выбранное хранилище

Интерфейс командной строки позволяет копировать отдельные образы виртуальных дисков между хранилищами следующим образом:

– завершить работу ВМ;

– использовать соответствующую команду CLI для идентификации UUID образов VDI, которые предполагается переместить. Если ВМ имеет DVD-привод, то его параметр **vdi-uuid** будет выведен как <not in database> и может быть проигнорирован:

```
xe vbd-list vm-uuid=<valid_vm_uuid>
```

Команда vbd-list выводит на экран идентификаторы UUID VBD и VDI. Последние обязательно следует записать.

– для каждого требуемого VDI выполнить команду **vbd-destroy**;

При использовании команды **vbd-destroy** для отсоединения VDI, заданных идентификаторами UUID, обязательно нужно убедиться, что параметр VBD **other-config:owner** установлен в значение **true**. Если это так, следует установить его в значение **false**, поскольку команда **vbd-destroy** с параметром **other-config:owner=true** уничтожит также связанный с VBD образ VDI.

– скопировать каждый из образов VDI ВМ командой **vdi-copy** для перемещения их в требуемое хранилище:

```
xe vdi-copy uuid=<valid_vdi_uuid> sr-uuid=<valid_sr_uuid>
```

– использовать команду **vbd-create**, указав в ней образы VDI из нового хранилища;

– удалить исходные образы:

```
vdi-destroy uuid=<uuid_of_vdi>
```

6.1.6.8. Преобразование локальных хранилищ на основе Fibre Channel в совместно используемое хранилище

Для подобного преобразования следует выполнить приведённую ниже последовательность действий:

– обеспечить, чтобы всем хостам в пуле зонировали LUN хранилища (см. пп. 6.1.6.2 Сканирование хранилища для получения подробностей по использованию команды **sr-probe** для проверки LUN, имеющихся на хосте).

– преобразовать хранилище в совместно используемое:

```
xe sr-param-set shared=true uuid=<local_fc_sr>
```

6.1.6.9. Автоматическое исправление пространства при удалении снимков состояния

При удалении моментальных снимков с Numa vServer автоматически исправляется всё выделенное место в хранилищах, основанных на LVM, перезагрузка ВМ не требуется (так называемая функция *Online Coalesce*).

Функция Online Coalesce применима только к хранилищам на основе LVM (LVM, LVMoISCSI и LVMoHBA) и не применима к хранилищам EXT или NFS, поведение которых остаётся неизменным

Рекомендуется использовать инструмент Off-Line Coalesce в следующих случаях:

– в условиях значительных объёмов ввода-вывода ВМ;

– когда пространство не остаётся неисправленным даже после истечения некоторого промежутка времени после удаления снимка.

Использование инструмента Off Line Coalesce ведёт к некоторому простою виртуальной машины вследствие выполнения операций приостановки и возобновления

Прежде чем прибегнуть к этому инструменту, необходимо удалить любые снимки и клоны, которые больше не требуются; сценарий исправит как можно больше пространства. Если необходимо исправить всё пространство, следует удалить все снимки и клоны.

Все диски ВМ должны располагаться в общем или локальном хранилище единственного хоста. ВМ с дисками в хранилищах двух этих типов не могут быть объединены.

6.1.6.9.19. Исправление пространства при помощи инструмента Off Line Coalesce

Функция Online Coalesce применима только к хранилищам, основанным на LVM (LVM, LVMoISCSI и LVMoHBA) и не применима к хранилищам на основе EXT или NFS, поведение которых остается неизменным

– открыть консоль на хосте и выполнить следующую команду:

```
xe host-call-plugin host-uuid=<host-uuid> plugin=coalesce-leaf
fn=leaf-coalesce args:vm_uuid=<VM-uuid>
```

Например, если UUID ВМ – «9bad4022-2c2d-dee6-abf5-1b6195b1dad5», а идентификатор UUID хоста – «b8722062-de95-4d95-9baa-a5fe343898ea», необходимо выполнить следующую команду:

```
xe host-call-plugin host-uuid=b8722062-de95-4d95-9baa-a5fe343898ea
plugin=coalesce-leaf fn=leaf-coalesce args:vm_uuid=9bad4022-2c2d-
dee6-abf5-1b6195b1dad5
```

– данная команда приостанавливает ВМ (если она ещё не выключена), инициирует процесс восстановления пространства и затем возобновляет ВМ.

Перед использованием Off Line Coalesce рекомендуется завершить работу или приостановить ВМ вручную посредством CLI. При применении инструмента Coalesce к работающей ВМ она будет автоматически приостанавлена и возобновлена после выполнения требуемых VDI Coalesce действий (по объединению).

Если образы VDI, которые предполагается объединить, находятся на совместно используемом хранилище, необходимо выполнить операцию Off Line Coalesce на хосте, являющемся мастером пула.

Если образы виртуальных дисков, которые предполагается объединить, хранятся локально, необходимо выполнить операцию Off Line Coalesce на сервере, к которому присоединено локальное хранилище.

6.1.6.10. Настройка планировщика дискового ввода-вывода

Для общей производительности планировщик дискового ввода-вывода (англ. *Disk IO Scheduler*, по умолчанию используется планировщик *Noop*) применяется ко всем новым типам хранилищ. Планировщик ввода-вывода Noop обеспечивает самую справедливую производительность для ВМ, конкурирующих за доступ к одному устройству. При использовании

механизма QoS для диска (см. пп 6.1.6.11 Настройки QoS для виртуальных дисков) необходимо переопределить настройку по умолчанию и присвоить параметру **other-config:scheduler** (выбор планировщика для хранилища) значение **cfq**. Соответствующий PBD должен быть отключен и повторно включен для применения новых параметров планировщика. Дисковый планировщик может быть скорректирован с помощью следующей команды:

```
xe sr-param-set other-
config:scheduler=noop|cfq|anticipatory|deadline uuid=<valid_sr_uuid>
```

Эта команда не работает с хранилищами на основе EqualLogic, NetApp или NFS

6.1.6.11. Настройки QoS для виртуальных дисков

Виртуальные диски имеют настройку QoS приоритета ввода-вывода (необязательная опция). В данном пункте описано, как применить эту установку к существующим виртуальным дискам с помощью интерфейса xe CLI.

В случае совместно используемого хранилища, когда множество хостов получает доступ к одному и тому же LUN, установки QoS применяются к виртуальным блочным устройствам, осуществляющим доступ к LUN на том же хосте. QoS не применяется между хостами пула.

Прежде чем конфигурировать любые параметры QoS для VBD, следует обеспечить хранилище соответствующим дисковым планировщиком (см. пп. 6.1.6.10 Настройка планировщика дискового ввода-вывода). Для хранилищ, которым требуется поддержка QoS, параметр планировщика должен быть установлен в значение cfq.

После установки для требуемого хранилища параметра планировщика в значение cfq следует убедиться, что физическое блочное устройство было переподключено, чтобы изменения для планировщика вступили в силу

Первым параметром является **qos_algorithm_type**. Этот параметр должен быть установлен в значение ionice (единственный тип алгоритма QoS, поддерживаемый для виртуальных дисков в данной версии Numa vServer).

Сами параметры QoS задаются парами «ключ/значение», ассоциированными с параметром **qos_algorithm_param**. Для виртуальных дисков **qos_algorithm_param** использует ключ **sched** (тип плана QoS) и, в зависимости от значения, также требует ключ класса **class**.

Возможные значения параметра **qos_algorithm_param:sched**:

– **sched**=rt или **sched**=real-time устанавливает параметр планирования QoS в значение «приоритет реального времени», требующего установки значения параметра класса;

– **sched**=idle соответствует режиму бездействия планировщика QoS (*idle priority*), не требующему установки значения какого-либо параметра класса;

– **sched**=<anything> соответствует приоритету по принципу «лучшей попытки» (*best effort*), требующему установки значения параметра класса.

Возможные значения для класса:

– одно из следующих ключевых слов: highest («самый высокий»), high («высокий»), normal («нормальный»), low («низкий»), lowest («самый низкий»);

– целое число между 0 и 7, где «7» соответствует самому высокому приоритету, а «0» – самому низкому, так, чтобы, например, запрос на ввод-вывод с приоритетом «5» будет обработан раньше запроса с приоритетом «2».

Для включения настроек QoS диска также необходимо установить параметр **other-config:scheduler** в значение cfq и переподключить физические блочные устройства для рассматриваемого хранилища.

Например, следующие команды устанавливают виртуальное блочное устройство виртуального диска для использования приоритета «5» (реального времени):

```
xe vbd-param-set uuid=<vbd_uuid> qos_algorithm_type=ionice
xe vbd-param-set uuid=<vbd_uuid> qos_algorithm_params:sched=rt
xe vbd-param-set uuid=<vbd_uuid> qos_algorithm_params:class=5
xe sr-param-set uuid=<sr_uuid> other-config:scheduler=cfq
xe pbd-plug uuid=<pbd_uuid>
```

6.1.6.12. Многоканальные соединения в системе хранения

Поддержка динамической многоканальности (*multipathing*) доступна для хранилищ на основе Fibre Channel и iSCSI. По умолчанию, Numa vServer применяет циклическую балансировку загрузки *round-robin*, таким образом, оба маршрута имеют активный трафик на них во время нормального функционирования.

Многоканальное соединение можно включить в xe CLI. Однако прежде необходимо убедиться, что на сервере хранения действительно доступно более одного целевого хранилища. Например, бэкэнд хранения iSCSI, запрошенный командой sendtargets на данном портале, должен возвратить более одной цели, как в примере ниже:

```
iscsiadm -m discovery --type sendtargets --portal 192.168.0.161

192.168.0.161:3260,1 iqn.strawberry:litchie
192.168.0.204:3260,2 iqn.strawberry:litchi
```

Только для iSCSI, dom0 имеет IP-адрес в каждой подсети, используемой многопоточным хранилищем. Необходимо убедиться, что для каждого пути к хранилищу есть сетевая карта, а для каждой сетевой карты настроен IP-адрес. Например, если нужно четыре пути к хранилищу, должно быть четыре сетевых адаптера, для каждого из которых настроен IP-адрес.

Включение многоканального соединение для системы хранения с помощью xe CLI включает следующие шаги:

– отключить все PBD на хосте:

```
xe pbd-unplug uuid=<pbd_uuid>
```

– установить для хоста параметр **other-config:multipathing**:

```
xe host-param-set other-config:multipathing=true uuid=<host_uuid>
```

– установить для хоста параметр other-config:multipathhandle в значение dmp:

```
xe host-param-set other-config:multipathhandle=dmp uuid=<host_uuid>
```

– если существующие на хосте хранилища работают в одноканальном режиме, но имеют возможность многоканального использования:

перенести или приостановить любые работающие ВМ с виртуальными дисками в затронутых SR;

отключить и повторно включить все влияющие на хранилища PBD, чтобы повторно подключить их, используя многоканальное соединение:

```
xe pbd-unplug uuid=<pbd_uuid>
xe pbd-plug uuid=<pbd_uuid>
```

Для отключения многоканального соединения следует сначала отключить VBD, установив на хосте параметр **other-config:multipathing** в значение false и затем повторно включить физические блочные устройства, как это описано выше. Параметр **other-config:multipathing** изменять не следует, поскольку это будет сделано автоматически.

Поддержка многоканальности в Numa vServer основывается на наборе компонентов *multipathd components* модуля ядра Linux, называемого *device-mapper*. Активация и деактивация многоканальных подключенных узлов осуществляется автоматически Storage Manager API. В отличие от стандартных инструментов *dm-multipathd tools* в Linux, узлы *device-mapper* не создаются для всех LUN в системе автоматически: новые узлы вводятся в действие лишь в случаях, когда LUN активно используются уровнем управления системы хранения. В связи с этим нет нужды использовать какой-либо инструмент dm-multipath командной строки для запроса или обновления табличных узлов *device-mapper* на хосте Numa vServer. Если необходимо запросить состояние таблиц *device-mapper* вручную или вывести список активных многоканальных узлов *device-mapper* в системе, следует использовать утилиту *mpathutil*:

– mpathutil status (запрос статуса);
– mpathutil list (вывод списка).

Из-за несовместимостей с интегрированной архитектурой управления многоканальности, стандартная утилита командной строки *dm-multipath* **не должна использоваться** с Numa vServer. Следует использовать инструмент командной строки *mpathutil* для того, чтобы запросить состояние узлов на хосте

Поддержка многоканальности в массивах EqualLogic не охватывает многоканальность ввода-вывода системы хранения (*Storage IO multipathing*) в традиционном смысле этого термина. Управление многоканальными соединениями должно происходить на уровне сетей/агрегаций сетевых интерфейсов. См. документацию *** EqualLogic для получения информации о настройках механизмов обеспечения отказоустойчивости сети для хранилищ на основе EqualLogic или LVMoISCSI.

7. АДМИНИСТРИРОВАНИЕ ПОЛЬЗОВАТЕЛЕЙ

Определение пользователей, групп, ролей и разрешений позволяет контролировать доступ к серверным узлам и ресурсным пулам Numa vServer, а также контролировать, какие действия они могут выполнять.

При установке Numa vServer одна учетная запись пользователя добавляется автоматически. Этой учетной записи присваивается роль локального суперпользователя (*Local Super User*, *LSU*).

LSU является особой учетной записью, предназначенной для администрирования системы и имеет все права и полномочия.

LSU проходит аутентификацию/авторизацию только в Numa vServer и не требует внешней службы аутентификации. Если внешний сервис аутентификации выйдет из строя, LSU всё равно сможет войти и управлять системой. Также суперпользователь всегда имеет доступ к физическому серверу Numa vServer через SSH.

Можно создать другие учётные записи пользователей путем добавления аккаунтов во FreeIPA, Active Directory или LDAP, воспользовавшись командами интерфейса «xe» (далее - CLI или xe). Если перечисленные сервисы не используются, доступной остаётся только учетная запись локального суперпользователя.

При создании новых пользователей Numa vServer не назначает автоматически созданным учетным записям роли согласно RBAC. Таким образом, эти учетные записи не будут иметь доступа к пулу до тех пор, пока им не будет назначена соответствующая роль.

7.1. Аутентификация пользователей с использованием Active Directory

При необходимости иметь несколько пользовательских учетных записей на серверном узле или в ресурсном пуле, можно использовать для аутентификации сервер Active Directory. Это позволяет пользователям Numa vServer авторизоваться на серверах пула Numa vServer, используя свои учетные данные домена в Windows.

Пользователи Active Directory могут использовать команды *xe* (используя аргументы -**u** и -**pw**) для подключения к серверу. Аутентификация осуществляется отдельно для каждого пула. Доступ контролируется с использованием, так называемых *субъектов*. *Субъект* в терминологии Numa vServer соответствует записи на сервере каталогов AD/LDAP/FreeIPA (соответствующей, в свою очередь, пользователю или группе). Когда внешняя аутентификация доступна, учетные данные используются для создания сеанса. Вначале проверяются учетные данные локального пользователя (в случае, если сервер каталогов недоступен), а затем список субъектов. Чтобы разрешить доступ к чему-то, необходимо создать запись для пользователя или группы, которым будет предоставлен этот доступ. Для этого используются команды xe, которые описаны далее.

Numa vServer позволяет использовать авторизационные данные Active Directory для учетных записей Numa vServer. Для этого, Numa vServer отправляет учетные данные Active Directory на контроллер домена Active Directory.

При добавлении в Numa vServer, пользователи и группы Active Directory становятся субъектами Numa vServer. Когда субъект зарегистрирован в Numa vServer, пользователи/группы проходят проверку подлинности в Active Directory при входе в систему. Таким образом, отпадает необходимость квалифицировать имя пользователя с именем домена.

По умолчанию, если домен пользователя не был уточнён (например, в виде «MyDomain\MyUser» или «myser@mydomain.com»), Numa vServer попытается осуществить вход пользователей в серверы аутентификации Active Directory, используя домен, подключенный в настоящее время. Исключением из этого правила является учетная запись локального суперпользователя (LSU), которая всегда аутентифицируется сначала локально (то есть на Numa vServer).

Процесс внешней аутентификации происходит следующим образом:

– учетные данные, при подключении к серверу, передаются на контроллер домена Active Directory для аутентификации;

– контроллер домена проверяет полномочия. Если они являются недействительными, аутентификация тут же прекращается;

– если учетные данные действительны, контроллер Active Directory запрашивает получение идентификатора и группы согласно учетным данным;

– если идентификатор субъекта совпадает с одним из хранящихся в Numa vServer, аутентификация успешно завершается.

При подключении к домену разрешается аутентификация в Active Directory для пула. Однако когда пул подключен к домену, только пользователи из этого домена (или домена, с которым он имеет доверенные отношения) могут подключаться к этому пулу.

Ручная настройка конфигурации DNS физического интерфейса сети с настроенным сервисом DHCP может привести к сбоям в интеграции Active Directory и, следовательно, к сбоям в аутентификации пользователей.

7.2. Настройка аутентификации Active Directory

Numa vServer поддерживает использование серверов Active Directory, начиная с версии Windows 2008.

Для аутентификации в Active Directory серверного узла Numa vServer необходимо, чтобы использовались одни и те же DNS как для Active Directory сервера (настроенного для разрешения в совместимости), так и для серверного узла Numa vServer. В некоторых конфигурациях сервер Active Directory может сам предоставить сервис DNS. Это может быть достигнуто либо с помощью DHCP, чтобы предоставить IP-адрес и список DNS-серверов в Numa vServer, либо путем установки значений в объектах физического интерфейса.

Доменные имена серверных узлов Numa vServer должны быть уникальными в течение всего времени нахождения Numa vServer в домене.

Следует обратить внимание на следующие особенности:

– Numa vServer записывает входы в AD с использованием данного имени сервера в базу данных AD. Поэтому, если два сервера узла имеют одинаковое имя и подключены к одному и тому же домену, второй перезапишет вход в AD первого, независимо от того, находятся они в одном или разных пулах, в результате чего аутентификация на первом перестанет работать. Можно использовать одно и то же доменное имя на двух серверных узлах Numa vServer, пока они подключены к разным AD доменам;

– серверные узлы могут находиться в разных часовых поясах. Для обеспечения корректной синхронизации можно использовать одни и те же NTP сервера для пула Numa vServer и для Active Directory сервера;

– пулы со смешанной аутентификацией не поддерживаются (то есть, невозможно иметь пул, где одни сервера настроены для использования Active Directory, а другие-нет);

– интеграция Active Directory в Numa vServer происходит с использованием протокола Kerberos для соединения с серверами Active Directory. Как следствие, Numa vServer не поддерживает взаимодействие с серверами Active Directory без участия Kerberos;

– чтобы внешняя аутентификация с использованием Active Directory проводилась успешно, важно, чтобы системное время на серверных узлах Numa vServer было синхронизировано с системными часами на сервере Active Directory. Это проверяется, когда Numa vServer присоединяется к домену Active Directory, и если разница превышает допустимые значения, то аутентификация завершится ошибкой.

Доменные имена должны состоять исключительно из букв и цифр – всего не более 63 символов, причём не менее одной буквы.

При добавлении в пул нового серверного узла после того, как был включен механизм аутентификации Active Directory, администратору будет предложено настроить на добавляемом сервере Active Directory. При появлении запроса на ввод учетных данных следует ввести учетные данные Active Directory с правами, достаточными для добавления серверов в этом домене.

7.2.1. Интеграция Active Directory

Необходимо убедиться в том, что сетевые порты для исходящего трафика, перечисленные в таблице 11, открыты в настройках межсетевого экрана Numa vServer.

Таблица 11 – Порты, необходимые для интеграции Active Directory

Порт	Протокол	Назначение
53	UDP/TCP	DNS
88	UDP/TCP	Kerberos 5
123	UDP	NTP
137	UDP	NetBIOS Name Service
139	TCP	NetBIOS Session (SMB)
389	UDP/TCP	LDAP
445	TCP	SMB поверх TCP
464	UDP/TCP	Изменение пароля машины
3268	TCP	Global Catalog Search

Для просмотра правил межсетевого экрана используется команда *iptables*:

```
iptables -nL
```

7.2.2. Управление паролем учетной записи компьютера для интеграции AD

Как и на клиентских машинах под управлением Windows, Numa vServer автоматически обновляет пароль учетной записи компьютера. Numa vServer обновляет пароль каждые 30 дней или в соответствии с политикой обновления пароля учетной записи компьютера на сервере AD.

7.2.3. Включение и отключение внешней системы аутентификации с использованием AD

Через интерфейс командной строки внешнюю аутентификацию с помощью Active Directory можно настроить следующей командой:

```
xe pool-enable-external-auth auth-type=AD service-name=<full-
qualified-domain> config:user=<username> config:pass=<password>
```

Указанный пользователь должен иметь привилегии для добавления и удаления компьютеров или рабочих станций, что по умолчанию разрешено администраторам доменов.

Если в сети, используемой Active Directory и серверными узлами Numa vServer, не используется DHCP, можно использовать следующие подходы к настройке DNS:

– настроить порядок поиска суффиксов домена DNS для работы с адресами, не являющимися FQDN:

```
xe pif-param-set uuid=<pif-uuid_in_the_dns_subnetwork> "other-
config:domain=suffix1.com suffix2.com suffix3.com"
```

– назначить DNS-сервер для серверов Numa vServer:

```
xe pif-reconfigure-ip mode=static dns=<dnshost> ip=<ip>
gateway=<gateway> netmask=<netmask> uuid=<uuid>
```

– вручную назначить для интерфейса управления PIF-объект из той же сети, что и DNS-сервер:

```
xe host-management-reconfigure pif-uuid=<pif_in_the_dns_subnetwork>
```

Внешняя аутентификация является параметром, задаваемым отдельно для каждого серверного узла. Тем не менее, рекомендуется включать или отключать это параметр для всего пула – в этом случае при обработке сбоя на любом сервере и выполнении любых требуемых откатов состояний будет гарантироваться, что установленная конфигурация подходит для всего пула.

Для проверки параметров серверного узла и определения состояния внешней аутентификации можно использовать команду **xe host-param-list**, проверяя значения соответствующих полей.

Для отключения внешней аутентификации при помощи командной строки используется команда:

```
xe pool-disable-external-auth
```

7.3. Пользовательская аутентификация

Для разрешения доступа пользователя к серверному узлу Numa vServer, необходимо добавить запись (о субъекте доступа) для этого пользователя, либо для группы, в которой он находится (вложенность групп проверяется в обычном порядке, например, если добавить разрешение для группы A, содержащей группу B, то пользователь из группы B будет иметь это разрешение).

Для управления разрешениями пользователей в Active Directory можно создать единую группу, а затем работать с ней, добавляя и удаляя пользователей. Можно добавлять и удалять отдельных пользователей, или сочетать использование отдельных пользователей и групп. Список субъектов может управляться с помощью интерфейса командной строки, как описано ниже.

При аутентификации пользователя, учетные данные сначала проверяются на соответствие учетной записи локального суперпользователя, что позволяет восстановить систему, если сервер AD вышел из строя. Если учетные данные (имя пользователя, затем пароль) не совпадают, то запрашивается аутентификация на сервере AD – если соединение с сервером AD происходит успешно, информация пользователя будет передана туда и проверен на соответствие списку субъектов, имеющемуся там. Если совпадение не будет найдено, либо не получится установить соединение с сервером AD, в доступе пользователю будет отказано. Проверка по списку субъектов считается успешной, если пользователь или его группа (возможно, в составе вложенной группы) имеются в списке субъектов доступа.

При использовании групп Active Directory для предоставления доступа для пользователей с ролью Администратора пула, которым необходим доступ к хосту по SSH, число пользователей в группе Active Directory не должно превышать 500.

7.3.1. Управление пользовательским доступом к серверному узлу

Для добавления в список AD субъекта доступа к серверному узлу существует команда:

```
xe subject-add subject-name=<entity name>
```

В качестве **entity name** может использоваться имя пользователя или название группы, которому (которой) требуется предоставить доступ. Также можно в необязательном порядке указать домен (например, «<testad\user1>», а не просто «<user1>»), если это требуется для однозначности.

Для запрета пользователю доступа к серверному узлу необходимо выполнить следующую последовательность действий:

– найти идентификатор субъекта пользователя. Это пользователь или группа, содержащая пользователя (удаление группы запрещает доступ всем пользователям этой группы, если они не были также указаны в списке субъектов непосредственно). Для нахождения идентификатора необходимо использовать команду возвращающую список пользователей:

```
xe subject-list
```

Для облегчения поиска имеется возможность задавать фильтры для выводимых результатов поиска. Следующая команда выведет (при условии наличия) информацию о пользователе *user1* в домене *testad*:

```
xe subject-list other-config:subject-name='<testad\user1>'
```

– используя найденный идентификатор (UUID), можно удалить запись о разрешении доступа пользователю при помощи команды:

```
xe subject-remove subject-uuid=<subject-uuid>
```

Для принудительного завершения текущей сессии конкретного пользователя используется команда:

```
xe session-subject-identifier-logout subject-identifier=<subject-id>
```

Для принудительного завершения текущих сессий всех пользователей, работающих в системе в настоящий момент, используется команда:

```
xe session-subject-identifier-logout-all
```

Если принудительно не завершить сессии пользователей, доступ которым был запрещён, они будут иметь доступ к серверному до завершения своей сессии.

7.4. Вывод из домена Active Directory

Когда администратор принимает решение покинуть домен (то есть отключить проверку подлинности Active Directory и отсоединить пул или сервер от этого домена), все пользователи, прошедшие аутентификацию в пуле или на сервере с учетными данными Active Directory, будут отключены.

Чтобы покинуть домен AD, следует выполнить команду:

```
xe pool-disable-external-auth
```

Указав идентификатор UUID пула, если требуется.

Вывод севера из домена не влечёт удаление серверных записей из базы данных Active Directory.

7.5. Управление доступом на основе ролей

Подсистема управления доступом на основе ролей (*Role Based Access Control*, *RBAC*) позволяет назначить пользователей, роли и полномочия для управления доступом к Numa vServer и к действиям, которые смогут выполнять получившие доступ. RBAC связывает пользователя (или группу пользователей) с определенной ролью (именованный набор прав доступа(permissions)). Роли, в свою очередь, имеют соответствующие полномочия для выполнения определенных операций в Numa vServer.

Полномочия не назначаются пользователям напрямую. Пользователи получают права доступа через назначенные им роли. Следовательно, управление полномочиями отдельных пользователей становится вопросом назначения пользователю соответствующей роли, что упрощает общие операции. Numa vServer ведёт список авторизованных пользователей и их ролей.

Подсистема управления доступом на основе ролей позволяет ограничить операции, которые могут выполнять разные группы пользователей, что снижает вероятность сбоя от действий неопытного пользователя.

Также RBAC предоставляет функцию журнала аудита.

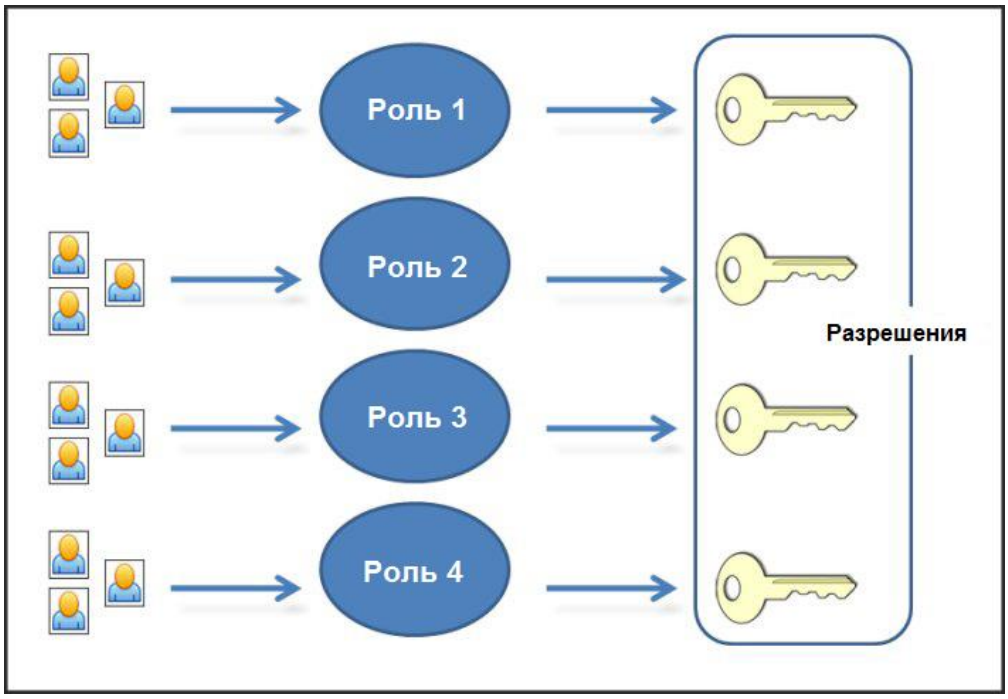

Рисунок 11 – Схема предоставления прав доступа на основе RBAC

RBAC зависит от служб аутентификации Active Directory, LDAP и FreeIPA. В частности, Numa vServer хранит список авторизованных пользователей, основанных на имеющихся в них записях о пользователях и группах. В результате необходимо присоединить ресурсный пул или серверный узел Numa vServer к домену и добавить учетные записи, прежде чем администратор сможет непосредственно назначать роли.

Локальный суперпользователь (*LSU*) или *root* – это особая учетная запись пользователя, используемая для системного администрирования, обладающая всеми правами и полномочиями. В Numa vServer учётная запись пользователя root создаётся по умолчанию при установке. Аутентификация с этой учётной записью проверяется только средствами Numa vServer, а не внешней службы аутентификации, поэтому локальный суперпользователь сможет войти и управлять системой даже при сбое соединения с сервером службы аутентификации. Локальный суперпользователь всегда может получить доступ к серверному узлу Numa vServer локально или по SSH (при условии, что доступ по SSH не отключен).

Типовой процесс настройки системы управления доступа на основе ролей в Numa vServer выглядит следующим образом:

– ввод серверного узла или пула в домен (см. пп. 7.2.3 Включение и отключение внешней системы аутентификации с использованием AD).

– добавление пользователя или группы из списков AD в пул (см. пп. 7.6.1 Добавление субъекта в систему RBAC).

– назначение (или модификация) роли субъекта в RBAC (см. пп.7.6.2 Назначение роли созданному субъекту, 7.6.3 Изменение роли субъекта доступа).

7.5.1. Пользовательские роли

Numa vServer поставляется с шестью предустановленными ролями:

– Администратор пула (*Pool Admin*) – та же, что и локальный суперпользователь, роль позволяет выполнять все операции.

– Оператор пула (*Pool Operator*) – роль позволяет делать всё кроме добавления/удаления пользователей и изменения их ролей. Данная роль предназначена, в основном, управление серверными узлами и пулом (создание хранилищ, настройка пулов, управление сетями и т. д.)

– Администратор виртуальных машин с расширенными полномочиями (*VM Power Admin*) – роль создаёт и управляет виртуальными машинами. Эта роль ориентирована на создание и обслуживание виртуальных машин, которые используются администраторами и операторами ВМ.

– Администратор виртуальных машин (*VM Admin*) – роль, схожая с предыдущей, но не позволяющая производить миграцию виртуальных машин или создавать снимки состояния ВМ.

– Оператор виртуальных машин (*VM Operator*) – по аналогии с ролью администратора ВМ, но не имеет полномочий на создание/удаление ВМ, однако позволяет запускать/останавливать операции жизненного цикла ВМ.

– Только для чтения (Read Only) – роль, позволяющая просматривать пул и данные о производительности.

Локальный суперпользователь всегда будет иметь роль «Pool Admin».

В данной версии Numa vServer отсутствует возможность добавления или удаления ролей.

Нельзя назначить роль *Pool Admin* группе служб аутентификации, содержащей более 500 членов, если предполагается, что в дальнейшем эти пользователи будут иметь доступ по SSH.

Для знакомства с полномочиями, доступными для каждой роли, и получения более подробной информации об операциях, доступных для каждого разрешения (см. пп. 7.5.2 Описание ролей и разрешений RBAC).

Всем пользователям Numa vServer должны быть назначены соответствующие им роли. Пользователю может быть назначено множество ролей; в этом случае пользователь будет иметь объединение всех их разрешений.

Роль пользователя может быть изменена двумя способами:

– изменение субъекта и назначение роли (доступно только Администратору пула);

– изменение роли содержащей пользователя группы в AD.

7.5.2. Описание ролей и разрешений RBAC

Таблица 12 суммирует полномочия доступные для каждой роли. Более подробные сведения представлены в таблице 13.

Таблица 12 – Полномочия для базовых ролей

Полномочие	**Роли**					
	Pool Admin	**Pool Operator**	**VM Power Admin**	**VM Admin**	**VM Operator**	**Read Only**
Назначение и изменение ролей пользователей	X					
Резервное копирование/восстановление сервера	X					
Импорт/экспорт OVF-/OVA- контейнеров и образов дисков ВМ	X					

Полномочие	**Роли**					
	Pool Admin	**Pool Operator**	**VM Power Admin**	**VM Admin**	**VM Operator**	**Read Only**
Установка количества ядер на сокет	X	X	X	X		
Преобразование виртуальных машин с помощью диспетчера преобразований	X					
Блокировка портов коммутатора	X	X				
Настройка multipathing	X	X				
Отключение активных пользователей от управления (завершение сеанса работы)	X	X				
Создание и снятие оповещений для пользователей	X	X				
Отмена заданий любого пользователя	X	X				
Управление пулом	X	X				
Живая миграция	X	X	X			
Хранилище для живой миграции	X	X	X			
Расширенные операции по управлению ВМ	X	X	X			
Создание и удаление ВМ	X	X	X	X		
Изменение подключенных CD образов в ВМ	X	X	X	X	X	
Изменение состояния питания ВМ	X	X	X	X	X	
Доступа к консоли ВМ	X	X	X	X	X	
Отмена собственных задач	X	X	X	X	X	X
Чтение журнала аудита	X	X	X	X	X	X
Подключение к пулу и чтение метаданных пула	X	X	X	X	X	X
Настройка vGPU	X	X				
Просмотр конфигурации vGPU	X	X	X	X	X	X
Доступ к конфигурационному диску (только для ВМ CoreOS)	X					
Управление контейнером	X					
Запланированные моментальные снимки	X	X	X			

Полномочие	**Роли**					
	Pool Admin	**Pool Operator**	**VM Power Admin**	**VM Admin**	**VM Operator**	**Read Only**
Запланированные снимки	X	X				
Настройка проверки работоспособности	X	X				
Просмотр результатов и настроек проверки работоспособности	X	X	X	X	X	X
Настройка отслеживания измененных блоков	X	X	X	X		
Просмотр списка измененных блоков	X	X	X	X	X	

Таблица 13 – Дополнительные сведения о полномочиях ролей

Полномочие	**Детали**	**Примечание**
Назначение и изменение ролей пользователей	добавить/удалить пользователей назначить/удалить роль включение и отключение интеграции с AD (присоединение к домену)	Предупреждение. Эта роль позволяет пользователю отключить интеграцию с Active Directory
Доступ к серверной консоли через SSH	доступ с локальной консоли сервера или через SSH	При локальном администрировании администратор может произвольно изменить конфигурацию всей системы, в том числе назначения ролей
Резервное копирование/восстановление сервера	резервное копирование и восстановление серверов резервное копирование и восстановление метаданных пула	Возможность восстановления резервной копии позволяет восстановить изменения конфигурации ролей
Импорт/экспорт OVF-/OVA-контейнеров и образов дисков ВМ	импорт OVF- и OVA-контейнеров импорт образов дисков экспорт виртуальных машин как OVF-/OVA-контейнеров	
Установка количества ядер на сокет	установка количества ядер на сокет для виртуальных процессоров ВМ	Это разрешение позволяет пользователю указать топологию для виртуальных процессоров ВМ

Полномочие	**Детали**	**Примечание**
Преобразование виртуальных машин с помощью диспетчера преобразований	преобразование виртуальных машин VMware в виртуальные машины Numa vServer	Это разрешение позволяет пользователю преобразовывать ВМ из VMware в Numa vServer путем копирования образов виртуальных машин VMware в среду Numa vServer
Блокировка портов коммутатора	контроль трафика в сети	Это разрешение позволяет пользователю по умолчанию блокировать весь трафик в сети или определять конкретные IP-адреса, с которых виртуальной машине разрешено отправлять трафик
Настройка multipathing	включение/отключение multipathing	
Отключение активных пользователей от управления (завершение сеанса работы)	возможность отключения вошедших в систему пользователей	
Создание и снятие оповещений для пользователей	конфигурация предупреждений, когда использование ресурсов пересекает определенные пороги удаление оповещений	Предупреждение. Пользователь с этим разрешением может отклонить оповещения для всего пула Примечание. Возможность просмотра предупреждений является частью подключения к пулу и чтения всех метаданных пула
Отмена заданий любого пользователя	отмена любого запущенного задания пользователя	Это разрешение позволяет пользователю запрашивать у Numa vServer отмену выполняемой задачи, инициированной любым пользователем
Управление пулом	установить свойства пула (название, SR по умолчанию) включить, отключить и настроить механизм HA установить приоритеты	Это разрешение включает в себя все действия, необходимые для поддержки пула Примечание. Если

Полномочие	Детали	Примечание
	перезапуска механизма HA для каждой ВМ конфигурация DR и выполнение DR failover, failback и test failover включить, отключить и настроить балансировку рабочей нагрузки (WLB) добавить и удалить сервер из пула аварийная смена мастера аварийная смена адреса мастера аварийное восстановление подчинённых хостов назначить нового мастера управление пулами настройка свойств сервера настройка ведения журнала на сервере включение и отключение серверов завершение работы, перезагрузка и включение серверов перезапуск набора инструментов отчеты о состоянии системы динамическая миграция всех виртуальных машин на сервере на другой сервер из-за режима обслуживания или высокой доступности настройка интерфейс управления сервером и вторичных интерфейсов отключение управление сервером удаление crashdumps добавление, редактирование и удаление сетей добавление, редактирование и удаление PBD/PIF/VLAN/Bonds/SR добавление, редактирование и	интерфейс управления не работает, никакие пользователи не могут проходить проверку подлинности, кроме локальных пользователей

Полномочие	**Детали**	**Примечание**
	удаление секретов	
Живая миграция	перенос виртуальных машин с одного хоста на другой, когда виртуальные машины находятся в хранилище, совместно используемом обоими хостами	
Хранилище для живой миграции	миграция с одного хоста на другой, если виртуальные машины не находятся в хранилище, совместно используемом двумя хостами перемещение виртуальных дисков (VDI) из одного SR в другой SR	
Расширенные операции по управлению ВМ	настройка памяти ВМ (через динамическое управление памятью) создание снимка виртуальной машины с памятью, создание снимков виртуальной машины и откат виртуальных машин миграция виртуальных машин запуск виртуальных машин, в том числе с указанием физического сервера возобновление работы ВМ	Это разрешение предоставляет уполномоченному достаточно прав для запуска ВМ на другом сервере, если они не удовлетворены выбранным сервером Numa vServer
Создание и удаление ВМ	создание и удаление ВМ клонирование/копирование виртуальных машин добавление, удаление и настройка виртуальных дисков/CD устройств добавление, удаление и настройка виртуальных сетевых устройств импорт/экспорт файлов XVA изменение конфигурации виртуальных машин резервное копирование и восстановление сервера	Примечание. Роль VM Admin может импортировать файлы XVA только в пул с общим SR. Роль VM Admin не имеет достаточных прав для импорта файла XVA на хост или в пул без общего хранилища

Полномочие	**Детали**	**Примечание**
Изменение подключенных CD образов в ВМ	установка/извлечение CD	
Изменение состояния питания ВМ	запуск виртуальных машин выключение виртуальных машин перезагрузка виртуальных машин приостановка виртуальных машин возобновление работы виртуальных машин	Это разрешение не включает start_on, resume_on и migrate, которые являются частью разрешения расширенных операций виртуальной машины
Доступ к консоли ВМ	взаимодействие с консолью виртульных машин	Это разрешение не позволяет пользователю просматривать серверную консоль
Отмена собственных задач	позволяет пользователю отменять собственные задачи	
Чтение журнала аудита	позволяет просматривать журнал аудита	
Подключение к пулу и чтение метаданных пула	подключение к пулу просмотр метаданных пула просмотр данных о производительности просмотр зарегистрированных пользователей просмотр пользователей и ролей просмотр сообщений регистрация и получение событий	
Настройка vGPU	настройка политики размещения в пуле назначение виртуального графического процессора виртуальной машине удаление виртуального графического процессора виртуальной машины изменение разрешенных типов виртуальных графических процессоров создание, удаление и назначение групп GPU	

Полномочие	**Детали**	**Примечание**
Просмотр конфигурации vGPU	просмотр графических процессоров, политик размещения графических процессоров и назначений vGPU	
Доступ к конфигурационному диску (только для ВМ CoreOS)	доступ к драйверу конфигурации виртуальной машины изменение параметров облачной конфигурации	
Управление контейнером	запуск остановка пауза резюме доступ к информации о контейнере	
Запланированные моментальные снимки	создание и удаление моментальных снимков виртуальных машин	
Запланированные снимки	добавление виртуальных машин в расписания снимков удаление виртуальные машины из расписания снимков добавление расписания снимков изменение расписания снимков удаление расписания снимков	
Настройка проверки работоспособности	включение проверки работоспособности отключение проверки работоспособности обновление настроек проверки работоспособности ручная загрузка отчета о состоянии сервера	
Просмотр результатов и настроек проверки работоспособности	просмотр результатов проверки работоспособности просмотр настроек регистрации проверки работоспособности	
Настройка отслеживания	включение отслеживания	

Полномочие	Детали	Примечание
измененных блоков	измененных блоков отключение отслеживания измененных блоков уничтожение данных связанных со снимками и сохранение метаданных получение информации о соединении NBD для VDI	
Просмотр списка измененных блоков	сравнение двух снимков VDI и перечисление блоков, которые изменились	

7.6. Использование RBAC через интерфейс CLI

Для вывода списка доступных ролей в Numa vServer используется команда xe **role-list**. Ниже показан пример такого списка:

```
uuid( RO): 0165f154-ba3e-034e-6b27-5d271af109ba
name ( RO): pool-admin
description ( RO): The Pool Administrator role has full access to
all
features and settings, including accessing Dom0 and managing
subjects,
roles and external authentication

uuid ( RO): b9ce9791-0604-50cd-0649-09b3284c7dfd
name ( RO): pool-operator
description ( RO): The Pool Operator role manages host-  and pool-
wide resources,
including setting up storage, creating resource pools and managing
patches, and
high availability (HA).

uuid( RO): 7955168d-7bec-10ed-105f-c6a7e6e63249
name ( RO): vm-power-admin
description ( RO): The VM Power Administrator role has full access
to VM and
template management and can choose where to start VMs and use the
dynamic memory
control and VM snapshot features

uuid ( RO): aaa00ab5-7340-bfbc-0d1b-7cf342639a6e
name ( RO): vm-admin
description ( RO):  The VM Administrator role can manage VMs and
templates
```

```
uuid ( RO): fb8d4ff9-310c-a959-0613-54101535d3d5
name ( RO): vm-operator
description ( RO):  The VM Operator role can use VMs and interact
with VM consoles

uuid ( RO): 7233b8e3-eacb-d7da-2c95-f2e581cdbf4e
name ( RO): read-only
description ( RO): The Read-Only role can log in with basic read-
only access
```

Список ролей статичен, нельзя добавить новые роли, удалить или изменить старые.

Для отображения списка текущих субъектов доступа используется команда xe **subject-list**, возвращающая список пользователей, их идентификаторы и роли, если такие присутствуют в системе.

7.6.1. Добавление субъекта в систему RBAC

Для того чтобы включить существующих в AD пользователей в систему RBAC, необходимо создать экземпляр субъекта в Numa vServer, либо для пользователя AD непосредственно – либо напрямую для одного из пользователей, прописанных в AD, либо для одной из содержащих его групп:

```
xe subject-add subject-name=<AD user/group>
```

7.6.2. Назначение роли созданному субъекту

Для назначения субъекту роли используется его идентификатор UUID:

```
xe subject-role-add uuid=<subject_uuid> role-uuid=<role_uuid>
```

или имя:

```
xe subject-role-add uuid=<subject_uuid> role-name=<role_name>
```

Пример. Следующая команда назначает пользователю с идентификатором UUID, равным b9b3d03b-3d10-79d3-8ed7-a782c5ea13b4, роль «Pool Administrator»:

```
xe subject-role-add uuid=b9b3d03b-3d10-79d3-8ed7-a782c5ea13b4 role-
name=pool-admin
```

7.6.3. Изменение роли субъекта доступа

Изменение роли пользователя, в частности, необходимо для снятия с него текущей роли (поскольку в любой момент времени хотя бы одна роль должна быть назначена). Используются следующие команды:

```
xe subject-role-remove uuid=<subject_uuid> role-
name=<role_name_to_remove>
xe subject-role-add uuid=<subject_uuid > role-
name=<role_name_to_add>
```

Чтобы убедиться, что новая роль вступила в силу, пользователь должен выйти из системы и снова пройти авторизацию (чтобы иметь возможность сделать это для другого пользователя принудительно, необходимо иметь разрешение на отсоединение активных пользователей (*Logout Active User Connections*), что доступно только Администраторам пула или Операторам пула).

После добавления или удаления субъекта с ролью Администратора пула может возникнуть задержка на несколько секунд для SSH-сессий, связанных с этим субъектом, на всех хостах пула.

7.7. Аудит RBAC

Журнал аудита RBAC записывает все операции, предпринятые вошедшим в систему пользователем.

Запись аудита будет явно содержать идентификатор субъекта и имя пользователя, ассоциированное с сессией, которая вызвала операцию.

В случае успешного выполнения какой-либо операции, фиксируется факт успеха; если операция не удалась, записывается также код ошибки.

Всегда фиксируется факт запроса выполнения операций, для осуществления которых субъект не имеет разрешения.

7.7.1. Команды CLI, связанные с журналом аудита

Следующая команда выгружает в файл все имеющиеся записи файла аудита RBAC пула. Если необязательный параметр **since** присутствует, то выгружаются только записи позднее указанной в этом параметре метки времени/даты.

```
xe audit-log-get [since=<timestamp>] filename=<output_filename>
```

Следующая команда позволяет выгрузить все записи аудита для пула:

```
xe audit-log-get filename=/var/data/auditlog-pool-actions.out
```

Для получения записей журнала аудита пула, датированных позднее точной миллисекундной метки, используется команда:

```
xe audit-log-get since=2019-09-24T17:56:20.530Z
filename=/var/data/auditlog-pool-actions.out
```

Для получения записей журнала аудита пула, датированных позднее метки с точностью до минуты, используется команда:

```
xe audit-log-get since=2019-09-24T17:56Z
filename=/var/data/auditlog-pool-actions.out
```

7.8. Расчёт ролей для сессии в Numa vServer

Роль каждого конкретного субъекта в каждой сессии работы рассчитывается следующим образом.

– субъект проходит аутентификацию через сервер Active Directory для того, чтобы проверить, какие содержащие субъект группы также имеются в списках AD;

– Numa vServer заверяет набор ролей, отведённых субъекту и содержащим его группам;

– субъект, входящий в несколько групп, наследует все доступные им разрешения.

Схема получения окончательного набора ролей для сессии показана на *** рисунке 12

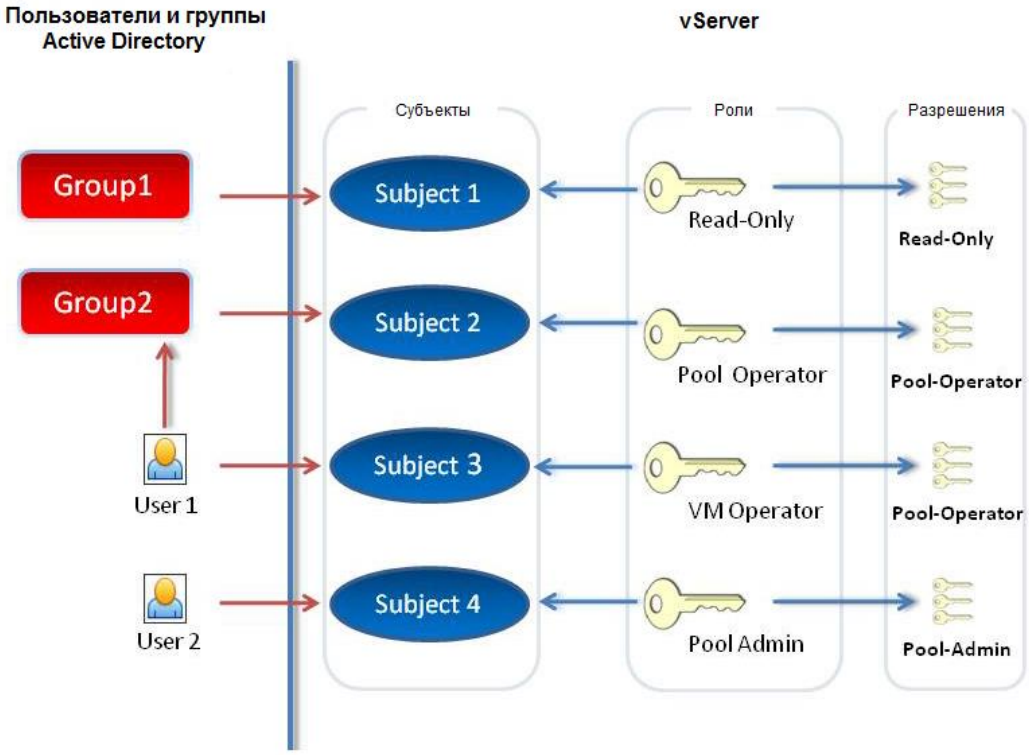

Рисунок 12 – Схема получения окончательного набора ролей для сессии

На этой иллюстрации, поскольку Subject2 (из группы Group2) является Оператором пула и пользователь User1 является членом группы Group2, когда Subject3 (пользователь User1) пытается войти, он наследует роли как Subject3 (роль «Оператор ВМ»), так и группы Group2 (роль «Оператор пула»). Поскольку роль «Оператор пула» выше по числу разрешений, в результате ролью субъекта Subject3 (User1) становится «Оператор пула», а не «Оператор ВМ».

8. НАСТРОЙКА USB PASSTHROUGH

Для настройки сквозного подключения USB устройств в ВМ с помощью Numa vServer необходимо:

1) Подключить USB устройство к аппаратной платформе, на которую установлен Numa vServer.

2) Выполнить команду xe pusb-list для просмотра подключённых USB устройств:

```
[root@localhost:~]# xe pusb-list
uuid ( RO)             : 77b8f209-4171-522d-92a8-3270f171dba5
             path ( RO): 1-10
       vendor-id ( RO): 0781
     vendor-desc ( RO): SanDisk Corp.
      product-id ( RO): 5571
    product-desc ( RO): Cruzer Fit
          serial ( RO): 00017428010822104644
         version ( RO): 2.00
     description ( RO): SanDisk Corp._Cruzer
Fit_00017428010822104644


uuid ( RO)             : 50c69bd0-79d3-dd55-e4e6-85dde8501ee9
             path ( RO): 1-6
       vendor-id ( RO): 058f
     vendor-desc ( RO): Alcor Micro Corp.
      product-id ( RO): 3828
    product-desc ( RO):
          serial ( RO):
         version ( RO): 2.00
     description ( RO): Alcor Micro Corp.


uuid ( RO)             : 730d6957-8ba9-6ba8-ff1d-5b3485fcad10
             path ( RO): 1-7
       vendor-id ( RO): 13fe
     vendor-desc ( RO): Kingston Technology Company Inc.
      product-id ( RO): 4300
    product-desc ( RO):
          serial ( RO): 0708236999B4BE80
         version ( RO): 2.00
     description ( RO): Kingston Technology Company
Inc._0708236999B4BE80
```

3) Выполнить команду для включения сквозной передачи для определенного USB устройства xe pusb-param-set uuid=<pusb_uuid> passthrough-enabled=true, где значение <pusb_uuid> - UUID USB устройства из п.2.

```
[root@localhost:~]# xe pusb-param-set uuid=77b8f209-4171-522d-92a8-
3270f171dba5 passthrough-enabled=true
```

4) Выключить ВМ, для которой необходимо настроить сквозное подключение USB устройства.

Убедитесь, что высокая доступность на целевой ВМ отключена.

5) Подключить USB устройство к ВМ, выполнив команду xe vusb-create usb-group-uuid=<usb-group UUID> vm-uuid=<UUID_VM>, где <UUID_VM> - UUID ВМ, для которой необходимо настроить сквозное подключение USB, <usb-group UUID>- можно узнать выполнив команду xe usb-group-list PUSB-uuids=<PUSB_UUID>

```
[root@localhost:~]# xe usb-group-list PUSB-uuids=77b8f209-4171-522d-
92a8-3270f171dba5
uuid ( RO)                  : 24b42d83-f89e-476b-d5d6-a92877f9fae7
          name-label ( RW): Group of 0781 5571 USBs
    name-description ( RW):

[root@localhost:~]# xe vusb-create usb-group-uuid=24b42d83-f89e-
476b-d5d6-a92877f9fae7 vm-uuid=b5f3055d-c3a8-f8ef-9c56-3c357c3a2ecb
d885b152-eb32-cbc7-0c14-ef51b9ae4b9c
```

6) Включить ВМ, убедиться, что USB устройство работает:

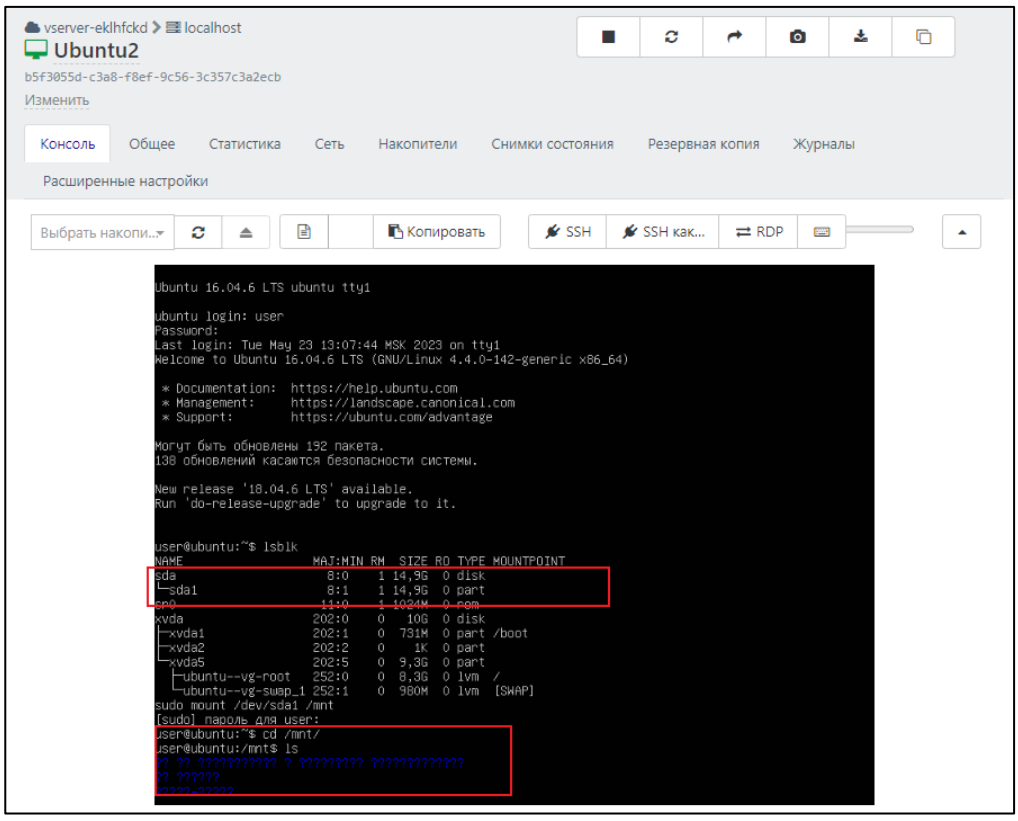

Рисунок 13 – Пример запуска подключенного USB к ВМ

7) Для отключения USB от ВМ необходимо выполнить команду xe vusb-unplug uuid=<vusb_uuid>

8) Для удаления VUSB xe vusb-destroy uuid=<vusb_uuid>

Для дальнейшей корректной работы ВМ необходимо отключить USB от целевой ВМ. В противном случае часть функций, таких как копирование ВМ, снимки состояния и т.п. работать не будут.

9. ПРЕОБРАЗОВАНИЕ И УСТАНОВКА ОБРАЗОВ ВМ

Для преобразования и установки образа ВМ в Numa vServer необходимо выполнить следующие действия:

1) Выполнить подключение к vServer и скопировать необходимый для дальнейшей работы образ ВМ.

2) Преобразовать образ ВМ в формат VHD:

```
qemu-img convert -f <image_format_input> -O vpc
<input_file.image_format_input> <outputfile.vhd>
```

где <image_format_input> указать формат входного файла согласно таблице:

тип формата	<image_format_input>
raw	raw
qcow2	qcow2
VDI	vdi
VMDK	vmdk
VHD	vpc
VHDX	vhdx

<input_file.image_format_input> - наименование входного файла (с расширением)

<outputfile.vhd> - наименование итогового файла(с расширением)

3) Просмотреть информацию о созданном файле:

```
qemu-img info <outputfile.vhd>
```

Пример:

```
[root@vserver-zwjpdzwc:~]#qemu-img convert -f vmdk -O vpc ubuntu-
18.10-server-cloudimg-amd64.vmdk ubuntuKs.vhd

[root@vserver-zwjpdzwc:~]# qemu-img info ubuntuKs.vhd
image: ubuntuKs.vhd
file format: vps
virtual size: 10 GiB (10737893376 bytes)
disk size: 1.12 GiB
cluster_size: 2097152
Child node '/file':
    filename: ubuntuKs.vhd
    protocol type: file
    file length: 1.16 GiB (1250230784 bytes)
    disk size: 1.12 GiB
```

4) Создать ВМ по шаблону импортируемой ВМ:

```
Xe template-list
xe vm-install template=<template-name> new-name-label=<name_VM>
xe vif-create vm-uuid=<uuid_VM> network-uuid=<network-uuid>
mac=random device=0
```

где <network-uuid> можно узнать использовав команду xe network-list.

5) Создать неразмеченный виртуальный диск:

```
xe vdi-create name-label=<name_VDI> virtual-size=<vitrual_size_GiB>
sr-uuid=<uuid_sr>
```

где <name_VDI> - имя виртуального диска

<vitrual_size_GiB> - объем создаваемого диска. Указываемый объем должен быть на 15-20% больше, чем планируемый объем импортируемого диска. Указываются также единицы изменения.

<uuid_sr> - UUID хранилища, в которое требуется поместить виртуальный диск

В качестве вывода vServer присвоит созданному VDI UUID.

Пример:

```
[root@vserver-zwjpdzwc:~]#xe vdi-create name-label=ubuntuKs virtual-
size=12GiB sr-uuid=5df7ebd8-cb45-34ee-fef7-5bafbf8fe717
df1feba5-6984-4623-8e56-8e4e5c8d9bb4
```

6) Создать новый VBD для отображения виртуального диска в ВМ:

```
xe vbd-create vm-uuid=<uuid_VM> device=1 vdi-uuid=<uuid_VDI>
bootable=true type=Disk mode=RW
```

где <uuid_VM> - UUID виртуальной машины

<uuid_VDI> - UUID VDI полученный на предыдущей этапе.

Пример:

```
[root@vserver-zwjpdzwc:~]#xe vbd-create vm-uuid=7681fc64-7c32-f046-
eac8-1eae3310cfc5 device=1 vdi-uuid=df1feba5-6984-4623-8e56-
8e4e5c8d9bb4 bootable=true type=Disk mode=RW
f94bd2f9-9e43-bb4e-9e57-74c008df772c
```

В качестве вывода vServer присвоит созданному VDB UUID.

7) Импортировать сконвертированный на 1 этапе VDI

```
xe vdi-import filename=<outputfile.vhd> format=vhd uuid=<uuid_VDI>
```

где <outputfile.vhd> - итоговый файл, полученный на этапе 1.

<uuid_VDI> - UUID VDI, полученный на этапе 5.

Пример:

```
[root@vserver-zwjpdzwc:~]#xe vdi-import filename=ubuntuKs.vhd
format=vhd uuid=df1feba5-6984-4623-8e56-8e4e5c8d9bb4
```

После чего ВМ будет доступна к запуску и дальнейшей работе.

Внимание! Для корректного запуска ОС семейства Red Hat следует выполнить следующие команды в shell ОС до импорта или в shell ОС в rescue-режиме:

```
yum install dracut-config-generic dracut-network
dracut --add-drivers xen-blkfront -f /boot/initramfs-$(uname -
r).img $(uname -r)
```

В случае дальнейшего использование Legacy:

```
dracut --regenerate-all -f && grub2-mkconfig -o /boot/grub2/grub.cfg
```

При использовании UEFI:

```
dracut --regenerate-all -f && grub2-mkconfig -o /boot/efi/EFI/<your
distribution>/grub.cfg
```

10. НАСТРОЙКА SNMP В NUMA VSERVER

Перед использованием протокола SNMP в составе Numa vServer необходимо провести следующие операции:

1) Создать резервную копию snmpd.conf:

```
cp /etc/snmp/snmpd.conf /etc/snmp/snmpd.conf.backup
```

2) В snmpd.conf указать режим доступа, пароль, перечень ip у которых будет доступ к данному серверу. Пример:

```
rocommunity public 192.168.1.0/24 rwcommunity private 192.168.1.1
```

3) Открыть порты 161/udp и 162/udp. Для этого необходимо прописать в "/etc/iptables/iptables" следующие строки:

```
-A vServer-Firewall-0-INPUT -m conntrack --ctstate NEW -m udp -p
udp --dport 161 -j ACCEPT -A vServer-Firewall-0-INPUT -m conntrack
--ctstate NEW -m udp -p udp --dport 162 -j ACCEPT
```

Пример:

```
*filter
:INPUT ACCEPT [0:0]
:FORWARD ACCEPT [0:0]
:OUTPUT ACCEPT [0:0]
:vServer-Firewall-0-INPUT - [0:0]
-A INPUT -j vServer-Firewall-0-INPUT
-A FORWARD -j vServer-Firewall-0-INPUT
-A vServer-Firewall-0-INPUT -i lo -j ACCEPT
-A vServer-Firewall-0-INPUT -p icmp --icmp-type any -j ACCEPT
# DHCP for host internal networks (CA-6996)
-A vServer-Firewall-0-INPUT -p udp -m udp --dport 67 --in-interface xenapi -j ACCEPT
-A vServer-Firewall-0-INPUT -m conntrack --ctstate ESTABLISHED,RELATED -j ACCEPT
# Linux HA hearbeat
-A vServer-Firewall-0-INPUT -m conntrack --ctstate NEW -m udp -p udp --dport 694 -j ACCEPT
-A vServer-Firewall-0-INPUT -m conntrack --ctstate NEW -m tcp -p tcp --dport 22 -j ACCEPT
-A vServer-Firewall-0-INPUT -m conntrack --ctstate NEW -m tcp -p tcp --dport 80 -j ACCEPT
-A vServer-Firewall-0-INPUT -m conntrack --ctstate NEW -m tcp -p tcp --dport 443 -j ACCEPT
-A vServer-Firewall-0-INPUT -m conntrack --ctstate NEW -m udp -p udp --dport 161 -j ACCEPT
-A vServer-Firewall-0-INPUT -m conntrack --ctstate NEW -m udp -p udp --dport 162 -j ACCEPT

-A vServer-Firewall-0-INPUT -j REJECT --reject-with icmp-host-prohibited
COMMIT
```

Рисунок 14 – Пример конфигурационного файла

4) Перезапустить сервис iptables:

```
systemctl restart iptables
```

5) Перезапустить сервис snmpd:

```
systemctl restart snmpd
```

10.1. Подключение сервера с Numa vServer к Zabbix

Для добавления vServer в Zabbix необходимо выполнить следующие действия:

– пройти авторизацию в веб-оболочке Zabbix-server.

– в панели навигации нажать на кнопку «Monitoring» → «Hosts».

– на открывшейся странице нажать на кнопку «Create host».

– в открывшейся форме в поле для ввода «Host name» ввести имя добавляемого хоста.

– в поле для ввода «Templates» ввести «Linux by SNMP» и кликнуть на совпадение в выпадающем меню.

– в поле для ввода «Host groups» ввести «Hypervisors» и кликнуть на совпадение в выпадающем меню.

– в секции «Interfaces» кликнуть на ссылку «Add», во всплывшем меню выбрать «SNMP».

– в поле для ввода «IP address» ввести ip-адрес хоста с vServer, в поле для ввода «SNMP community» ввести пароль, указанный во время настройки SNMP в Numa vServer.

– нажать на кнопку «Add».

Хост с Numa vServer успешно добавлен.

ПРИЛОЖЕНИЕ А. ПЕРЕЧЕНЬ ЖУРНАЛИРУЕМЫХ СОБЫТИЙ БЕЗОПАСНОСТИ

Тип события	Расшифровка события
AUDIT_VIRT_CONTROL: pause	Постановка ВМ на паузу
AUDIT_VIRT_CONTROL: unpause	Снятие ВМ с паузы
AUDIT_VIRT_CONTROL: start	Запуск ВМ
AUDIT_VIRT_CONTROL: hard_shutdown	принудительное выключение ВМ
AUDIT_VIRT_CONTROL: hard_reboot	принудительная перезагрузка ВМ
AUDIT_VIRT_CONTROL: clean_reboot	перезагрузка ВМ
AUDIT_VIRT_CONTROL: clean_shutdown	выключение ВМ
AUDIT_VIRT_CONTROL: suspend	приостановка ВМ
AUDIT_VIRT_CONTROL: resume	возобновление работы ВМ после приостановки
AUDIT_VIRT_INTEGRITY_CHECK	проверка целостности
AUDIT_VIRT_CREATE	создание ВМ
AUDIT_VIRT_DESTROY	уничтожение ВМ
AUDIT_VIRT_MIGRATE_IN	миграция ВМ в хост
AUDIT_VIRT_MIGRATE_OUT	миграция ВМ из хоста
SERVICE_START	Запуск сервисов
SERVICE_STOP	Остановка сервисов
USER_ROLE_CHANGE	Смена роли пользователя
USER_START	Аутентификация пользователя
USER_LOGIN	Идентификация пользователя
VIRT_CREATE	Создание объекта

СПИСОК СОКРАЩЕНИЙ

ALB	adaptive load balancing
ARP	address resolution protocol
CIFS	common internet file system
CLI	command line interface
CSM	compatibility support module
DHCP	dynamic host configuration protocol
DNS	domain name system
HA	high availability
IDE	integrated development environment
IP	internet protocol
ISCSI	internet small computer system interface
LACP	link aggregation control protocol
LUN	logical unit number
MAC	media access control
NFS	network file system
NIC	network interface controllers
NTP	network time protocol
PBD	physical block devices
PIF	physical interface file
QOS	quality of service
RBAC	role based access control
RO	read only
SMB	server message block
SAS	serial attached scsi
SCSI	small computer system interface
SDN	software-defined networking
SQL	structured query language
SR	storage repsitory
SR-IOV	single root i/o virtualization
SSH	security shell
UEFI	unified extensible firmware interface
USB	universal serial bus
UUID	universally unique identifier

VDI	virtual desktop infrastructure
VIF	virtual interface file
VLAN	virtual local area network
ВМ	виртуальная машина
ПО	программное обеспечение
ЦП	центральный процессор

Лист регистрации изменений									
Изм.	Номера листов (страниц)				Всего листов (страниц) в документе	№ докумен-та	Входящий № сопрово-дительного документа и дата	Подпись	Дата
	изме-ненных	заме-ненных	новых	аннулиро-ванных					