



АО "НТЦ ИТ РОСА"

**ПЛАТФОРМА ВИРТУАЛИЗАЦИИ
"ROSA VIRTUALIZATION"**

Версия 3.1

Общие сведения о ROSA Virtualization

Листов 14

1 СИСТЕМА ВИРТУАЛИЗАЦИИ ROSA VIRTUALIZATION

Система виртуализации ROSA Virtualization — это платформа виртуализации корпоративного уровня, созданная на основе гипервизоров первого типа, устанавливаемых непосредственно на аппаратную платформу. Технология виртуализации даёт пользователям возможность быстрого создания новых виртуальных серверов и инфраструктуры удаленных рабочих столов (VDI), а также обеспечивает более эффективное использование ресурсов физических серверов.

С помощью системы виртуализации управление всей виртуальной инфраструктурой, включая хосты, виртуальные машины, сети, хранилища и пользователей, производится в централизованном графическом интерфейсе пользователя или посредством REST API, что открывает большие возможности для автоматизации и/или интеграции с другим программным обеспечением.

Ключевые компоненты системы виртуализации перечислены в таблице 1.

Таблица 1 – Ключевые компоненты системы виртуализации

Компонент	Описание
СУСВ (Система управления средой виртуализации)	Служба, предоставляющая графический интерфейс пользователя и REST API для управления ресурсами окружения. СУСВ устанавливается на ВМ посредством графического инсталлятора.
Хосты	Хосты виртуализации (гипервизоры). На хостах используется технология ВМ на базе ядра (KVM). Хосты предоставляют ресурсы для работы виртуальных машин, такие как оперативная память и ядра процессора. Также могут предоставлять ресурсы для программно определяемых хранилищ (гиперконвергентность)
Разделяемое хранилище	Используется для хранения данных виртуальных машин, может быть как программное, так и аппаратное.
Хранилище данных СУСВ	Служба, собирающая конфигурационные данные и данные статистики диспетчера виртуализации.

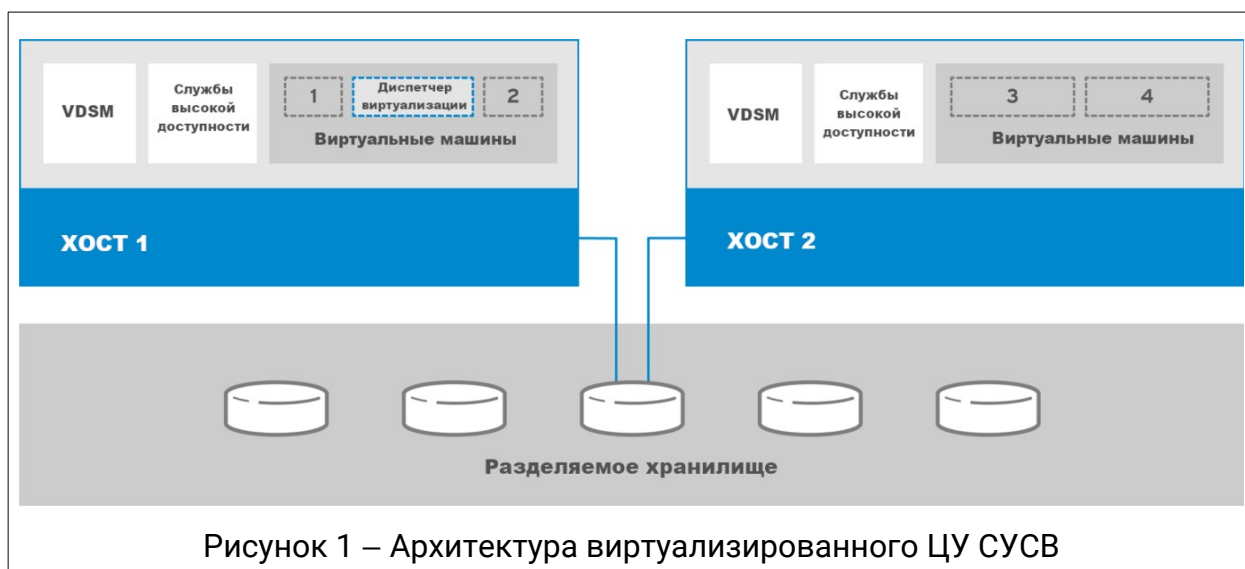
1.1 Архитектура среды виртуализации

СУСВ выполняется в виде ВМ на узлах виртуализированного ЦУ (специализированные хосты). СУСВ является высокодоступной без внешнего управления высокой доступностью.

В минимальную установку окружения виртуализированного ЦУ входит:

- одна виртуальная машина с СУСВ, расположенная на узлах виртуализированного ЦУ;
- минимум два узла виртуализированного ЦУ для высокой доступности ВМ. Для упрощения процесса обмена информацией между хостами и СУСВ на всех хостах работает VDSM (агент хоста). Для управления высокой доступностью ВМ СУСВ, на всех узлах виртуализированного ЦУ выполняются службы высокой доступности;
- одно разделяемое хранилище, доступное для всех хостов.

Архитектура виртуализированного ЦУ СУСВ представлена на рисунке 1.



1.2 Терминология системы виртуализации

– **Кластер** – представляет собой группу физических хостов, рассматриваемых как пул ресурсов для виртуальных машин. Хосты в кластере разделяют одну и ту же сетевую инфраструктуру, а также хранилище. Они составляют домен миграции, в пределах которого ВМ могут перемещаться с одного хоста на другой как в ручном режиме, так и в автоматическом.

– **Дата-центр** – это контейнер самого высокого уровня для всех физических и логических ресурсов в управляемом виртуальном окружении. Дата-центр представляет собой набор кластеров, виртуальных машин, доменов хранения и сетей.

– **События** – оповещения, предупреждения и другие уведомления о событиях помогают администратору в наблюдении за информационной безопасностью, производительностью и статусом ресурсов.

– **Службы высокой доступности** – служба ovirt-ha-agent и служба ovirt-ha-broker. Службы высокой доступности выполняются на узлах виртуализированного ЦУ и управляют высокой доступностью ВМ СУСВ.

– **Высокая доступность** – высокая доступность означает, что в случае нарушения работы виртуальная машина будет автоматически перезапущена, либо на исходном хосте либо на другом хосте в кластере. В окружениях высокой доступности допускаются незначительные простои, но эти окружения являются гораздо менее затратными, чем отказоустойчивые окружения, где имеется по две копии каждого ресурса, чтобы в случае сбоя один мог немедленно заменить другой.

– **Хост или гипервизор** – это физический сервер, выполняющий одну или более ВМ. Хосты группируются в кластеры. ВМ могут мигрировать с одного хоста на другой в границах кластера.

– **Диспетчер хранилища хоста** – любой хост в дата-центре, не являющийся диспетчером пула хранилищ (SPM), и который может быть использован для таких действий с данными, как перемещение диска из одного домена хранения в другой. Это предотвращает возникновение узких мест на хосте SPM, который должен использоваться для более кратковременных операций с метаданными.

– **Логическая сеть** – это логическое представление физической сети. В логических сетях концентрируется сетевой трафик и обмен информацией между диспетчером виртуализации, хостами, хранилищем и виртуальными машинами.

– **Удалённый просмотрщик** – графический интерфейс для подключения к консолям ВМ по сети.

– **Узел виртуализированного ЦУ** – хост, настроенный особым образом, чтобы на этом хосте могла разместиться ВМ СУСВ.

– **Снимок (snapshot)** – это сохранение операционной системы ВМ и всех её приложений в определённый момент времени. Снимки можно использовать для сохранения параметров ВМ перед обновлением или перед установкой новых приложений. В случае возникновения проблем, снимок можно использовать для восстановления ВМ до исходного состояния.

– **Домен хранения** – логическая сущность, содержащая данные. Каждый домен хранения используется для хранения виртуальных дисков или образов ISO, а также для импорта или экспорта образов VM.

– **Диспетчер пула хранилищ (SPM)** – роль, присваиваемая одному хосту в дата-центре. Только у хоста SPM имеются полномочия на изменение метаданных дата-центра, например, на создание и удаление виртуальных дисков.

– **Шаблон** – это виртуальная машина-образец с предварительно настроенными параметрами. VM, созданная на основе конкретного шаблона, получает параметры, присутствующие в шаблоне. Использование шаблонов – это самый быстрый способ за один шаг создать множество VM.

– **VDSM** – служба агента хоста, выполняющаяся на хостах, обменивающихся информацией с СУСВ.

– **Виртуальная машина (VM)** – это виртуальная рабочая станция или виртуальный сервер, содержащий операционную систему и набор приложений. Из множества идентичных VM можно создать Пул (VDI).

– **Пул VM** – это группа идентичных VM, доступных по требованию для каждого участника группы. Пулы VM можно настраивать для различных целей.

2 КОМПОНЕНТЫ

2.1 Система Управления Средой Виртуализации

СУСВ предоставляет графический интерфейс пользователя и REST API для управления ресурсами окружения виртуализации. В окружении виртуализированного ЦУ диспетчер виртуализации устанавливается в виде ВМ, расположенной на узлах виртуализированного ЦУ в том же окружении, которым он управляет.

Встроенная высокая доступность диспетчера виртуализации доступна только в окружении виртуализированного ЦУ. Для высокой доступности требуется минимум два узла виртуализированного ЦУ.

2.2 Хосты

Хосты, также известные как гипервизоры, представляют собой физические серверы, на которых выполняются ВМ. Полная виртуализация предоставляется загружаемым модулем ядра Linux, который называется KVM (Kernel-based Virtual Machine, "виртуальная машина на базе ядра"). KVM может одновременно выполнять множество изолированных друг от друга ВМ под управлением разных операционных систем. ВМ выполняются на хосте и удалённо управляются диспетчером виртуализации.

Рекомендуется использовать как минимум два хоста при использовании внешнего разделяемого хранилища или три – при выполнении гиперконвергентной установки. Если использовать только один хост, то будут отсутствовать такие возможности, как миграция и высокая доступность.

2.3 Хранилище

Настройка хранилища и ввод его в окружение системы виртуализации – это предварительное условие, без которого нельзя начать создание виртуальных машин. В системе виртуализации есть три типа доменов хранения:

- **Домен данных** содержит всю информацию, имеющую отношение к виртуальным машинам. Домен данных поддерживает все типы хранилищ, для которых имеется поддержка в системе виртуализации;

- **Домен ISO** является устаревшим типом домена хранения. Эти домены использовались для хранения файлов образов ISO для установки операционных систем или дополнительных приложений, например, гостевых агентов и драйверов Windows. Теперь образы ISO можно загружать в домены данных;

- **Домен экспорта** является устаревшим типом домена хранения. Эти домены использовались в качестве хранилища резервных копий или для перемещения образов между дата-центрами и окружениями системы виртуализации. Теперь это выполняется с помощью домена данных.

Домены ISO и домены экспорта поддерживают только файловые типы хранилищ (NFS, POSIX или GlusterFS). При использовании в дата-центре с локальным хранилищем, домен ISO поддерживает локальное хранилище.

2.4 Хранилище данных СУСВ

В состав СУСВ входит хранилище данных, собирающее данные наблюдения за хостами, ВМ и хранилищем. Хранилище данных, включающее в себя базу данных и службу, должно быть установлено и настроено одновременно с настройкой диспетчера виртуализации.

Во время установки системы виртуализации создаются две базы данных:

- **База данных СУСВ (engine)** является первичным хранилищем данных. В этой базе данных хранится такая информация об окружении, как его состояние, конфигурация и производительность;

- **База данных хранилища данных (ovirt_engine_history)** содержит сведения о конфигурации и данные статистики. Данные конфигурации в базе данных СУСВ проверяются каждую минуту, и изменения дублируются в базе данных хранилища данных. Отслеживание изменений в базе данных предоставляет сведения об объектах в базе данных, что в свою очередь предоставляет возможность для анализа и улучшения производительности окружения системы виртуализации, а также сведения для решения возможных проблем и их прогнозирования.

2.5 Сетевые конфигурации в системе виртуализации

Многие действия, например, работа с хранилищем, управление хостом, обмен информацией между пользователями и виртуальными машинами, зависят от хорошо спланированной и хорошо настроенной сети с оптимальной производительностью. Настройка сетей – это важная процедура для работы окружения системы виртуализации. Система виртуализации разделяет сетевой трафик с помощью логических сетей. Логические сети определяют маршруты, по которым должен проходить трафик определенного типа в сети. Эти сети создаются для изоляции трафика согласно его функционалу или для виртуализации физической топологии.

Логическая сеть `ovirtmgmt` создаётся по умолчанию и получает метку «сеть управления». Логическая сеть `ovirtmgmt` предназначена для управления трафиком между диспетчером виртуализации и хостами. Также можно настроить дополнительные логические сети для отделения:

- общего трафика виртуальных машин;
- трафика хранилищ (например, NFS или iSCSI);
- трафика при миграции ВМ;
- трафика визуализации консоли ВМ;
- трафика хранилищ Gluster.

3 ТЕХНИЧЕСКИЕ ПАРАМЕТРЫ

3.1 Параметры производительности виртуальных машин

3.1.1 Технологические возможности и ограничения ROSA Virtualization

В данном разделе приводятся поддерживаемые и теоретические ограничения платформы.

Поддерживаемая архитектура – x86-64.

Формат нижеизложенных значений:

тестируемое/поддерживаемое [теоретически возможное]

Максимальное число логических ЦП – 768 [5120].

Логический ЦП определяется как любая сущность, допускающая планирование. Таким образом, любое ядро/поток в многоядерном/многопоточном процессоре является логическим ЦП.

Максимальный объём памяти – 12 Тбайт [64 Тбайт].

Архитектурные ограничения основаны на возможностях ядра ROSA Virtualization и физическом аппаратном обеспечении.

Минимальная требуемая память – 1 Гбайт минимум, 1 Гбайт на логический ЦП рекомендуется.

Минимальное требуемое дисковое пространство – 150 Гбайт минимум, 200 Гбайт рекомендуется.

3.1.2 Файловые системы и ограничения хранилища

Ограничения для файловой системы ext4 приведены в таблице 2.

Таблица 2 – Файловая система ext4

Свойство	ROSA Virtualization
Макс. размер файла	16 Тбайт

Свойство	ROSA Virtualization
Макс. размер файловой системы	50 Тбайт [1Эбайт]
Подкаталогов максимально	65000/неогр.
Максимальная глубина симв. ссылки	8
Поддержка ACL	да

Ограничения для файловой системы XFS приведены в таблице 3.

Таблица 3 – Файловая система XFS

Свойство	ROSA Virtualization
Макс. размер файла	500 Тбайт [8 Эбайт]
Макс. размер файловой системы	500 Тбайт [16 Эбайт]
Подкаталогов максимально	неогр.
Максимальная глубина симв. ссылки	8
Поддержка ACL	да

Ограничения для хранилища приведены в таблице 4.

Таблица 4 – Хранилище

Свойство	ROSA Virtualization
Макс. размер загрузочного LUN (BIOS)	2 Тбайт
Макс. размер загрузочного LUN (UEFI)	50 Тбайт
Макс. число путей для устройства (устройства SD)	10,000

3.1.3 Поддерживаемые ограничения ROSA Virtualization 3.1

Для применения ROSA Virtualization 3.1 в крупномасштабных окружениях придерживайтесь нижеследующих рекомендаций.

3.1.3.1 Нормативы для крупномасштабных установок системы виртуализации ROSA Virtualization версии 3.1

Целевое окружение

Для улучшения производительности ROSA Virtualization версии 3.1 в крупномасштабных окружениях должно быть в наличии как минимум:

Крупномасштабные окружения имеют как минимум:

- 80 хостов;
- 1000 виртуальных машин;
- 100 прямых LUN.

Диспетчер виртуализации

- При развёртывании СУСВ:
 - для производственной среды со средней и высокой занятостью администраторов и параллельными API предоставьте VM с 16 ЦП и 32 Гбайт ОЗУ;
 - для производственной среды с минимальной и средней параллельностью предоставьте VM с 8 ядрами ЦП и 16-32 Гбайт ОЗУ;
- увеличьте число доступных подключений к пулам и базам данных для диспетчера виртуализации:

- добавьте следующие строки в файл /usr/share/ovirt-engine/services/ovirt-engine/ovirt-engine.conf:

```
ENGINE_DB_MIN_CONNECTIONS=1
ENGINE_DB_MAX_CONNECTIONS=200
```

- добавьте следующую строку в файл /var/opt/rh/rh-postgresql10/lib/pgsql/data/postgresql.conf:

```
max_connections = 250
```

Хосты

- Для лучшей производительности добавьте [локальный фильтр LVM](#) (англ.) на всех хостах. Фильтр предотвращает сканирование и активацию логических томов, ненужных хосту;
- если в дата-центре более 250 хостов (максимальное число, поддерживаемое дата-центрами), создавайте дополнительные дата-центры.

Хранилище

- Настройте каждый дата-центр согласно инструкциям поставщика сервера хранилища так, чтобы использовать группу инициатора для изоляции логических томов;

- распределите диски по доменам хранения так, чтобы ни один из доменов не содержал более 1300 дисков.

3.1.4 Ограничения системы виртуализации ROSA Virtualization

Для системы виртуализации ROSA Virtualization применяются следующие ограничения

Максимальные значения для VM:

- максимальное число параллельно выполняющихся VM – без ограничений;
- максимальное число виртуальных ЦП на каждую VM – 710 для Q35 машин, 240 – для PC машин;
- максимальный объём памяти на каждую VM – 16 Тбайт;
- минимальный объём памяти на каждую VM – н/д;
- максимальный размер одного диска на каждую VM – 8 Тбайт.

Примечание – Система виртуализации ROSA Virtualization поддерживает максимальный объём памяти 16 Тбайт на каждую VM хоста, т. е. весь доступный выделяемый виртуальной машине объём. У 32-битных гостей с поддержкой PAE может быть доступ только к 64 Гбайтам, что является ограничением виртуального оборудования.

Обратите внимание, что для виртуализированных гостей размером более 8 ТБ в настоящее время требуется явная конфигурация виртуального NUMA, поскольку максимальный размер виртуального узла NUMA составляет 8 ТБ.

Максимальные значения для хостов применяются к хостам виртуализации и стандартным хостам.

- логических ядер или потоков ЦП – 768;
- ОЗУ – 12 Тбайт;
- динамических миграций одновременно – 2 входящих, 2 исходящих;
- пропускная способность для динамических миграций: по умолчанию – 52 Мбит (~436 Мбайт) на миграцию при использовании устаревшей политики миграции. Другие политики используют адаптивные пропускные значения на базе скорости физического устройства. Политики QoS могут ограничивать пропускную способность для миграций.

Максимальные значения виртуализированного ЦУ :

- число узлов на кластер: 7.

Максимальные значения логических сущностей диспетчера виртуализации:

– Дата-центры:

- максимальное число дата-центров – 400;
- максимальное число хостов – 400 поддерживается, 500 – по данным тестов;
- максимальное число VM – 4000 поддерживается, 5000 – по данным тестов;

– Кластеры:

- максимальное число кластеров – 400;
- максимальное число хостов – 400 поддерживается, 500 – по данным тестов;
- максимальное число VM – 4000 поддерживается, 5000 – по данным тестов;

– Сети:

- Логических сетей /кластер – 300;
- SDN/внешних сетей – 2600 по данным тестов, принудительные ограничения отсутствуют;

– Хранилище:

- максимально доменов – 50 поддерживается, 70 – по данным тестов;
- хостов на домен: без ограничений;
- логических томов на блочный домен: 1500;
- максимальное число LUN – 300;
- максимальный размер диска – 500 ТиБ (ограничение по умолчанию – 8 ТиБ);

Другое:

- параллельных действий диспетчера пула хранилищ (операции с метаданными, динамическая миграция хранилища и т.п.) – 10.

3.1.5 Минимальные требования для развёртывания ROSA Virtualization:

- Процессор, поддерживающий аппаратную виртуализацию
- объем ОЗУ — не менее 16 ГБ;
- количество ядер ЦПУ — не менее 4;
- объем разделяемого хранилища — не менее 500 ГБ;
- скорость сетевого адаптера — не менее 100 Мбит.