Российская сервисная платформа виртуализации РУСТЭК

Общее описание и архитектура

Релиз 2021.2.5

2023г.

Оглавление

1 Общее описание

Платформа виртуализации РУСТЭК разворачивается на физических серверах на базовой ОС РУСТЭК, входящей в состав дистрибутива, и использует гипервизор KVM, входящий в состав базовой ОС. ОС РУСТЭК является специализированной ОС на основе Linux, предназначенной в основном для работы в составе платформы виртуализации.

Сервисы платформы разворачиваются в кластерном режиме при помощи специализированного инсталлятора платформы РУСТЭК и обеспечивают высокую доступность, как для управляющих сервисов, так и для ВМ, запущенных на платформе.

В качестве оркестратора используется модифицированное ПО OpenStack. Платформа также включает в себя дополнительные сервисы, разработанные компанией РУСТЭК. Управление может осуществляться из консоли при помощи утилиты командной строки, через REST API или используя Web-панель управления.

2 Архитектура

В состав платформы входят следующие сервисы из проекта OpenStack:

1. Nova - управляет жизненным циклом виртуальных машин
2. Glance - хранилище образов
3. Neutron - обеспечивает SDN возможности платформы
4. Cinder - обеспечивает виртуализацию работы с СХД
5. Keystone - обеспечивает авторизацию и управление пользователями, проектами, доменами
6. Placement - обеспечивает управление ресурсами платформы
7. Ceilometer - обеспечивает сбор и хранение метрик как для виртуальной инфраструктуры так и физической
8. Aodh - обеспечивает возможность мониторинга и оповещений
9. Barbican - обеспечивает хранение ключей и других секретов
10. Octavia - обеспечивает сервис сетевых балансировщиков
11. Heat - обеспечивает сервис инфраструктуры как кода
12. Designate - обеспечивает DNS как сервис
13. Magnum - обеспечивает кластеры Kubernetes как сервис
14. Watcher - обеспечивает оптимизацию нагрузки в кластере (балансировку)
15. Mistral - обеспечивает предоставление Workflow как сервис

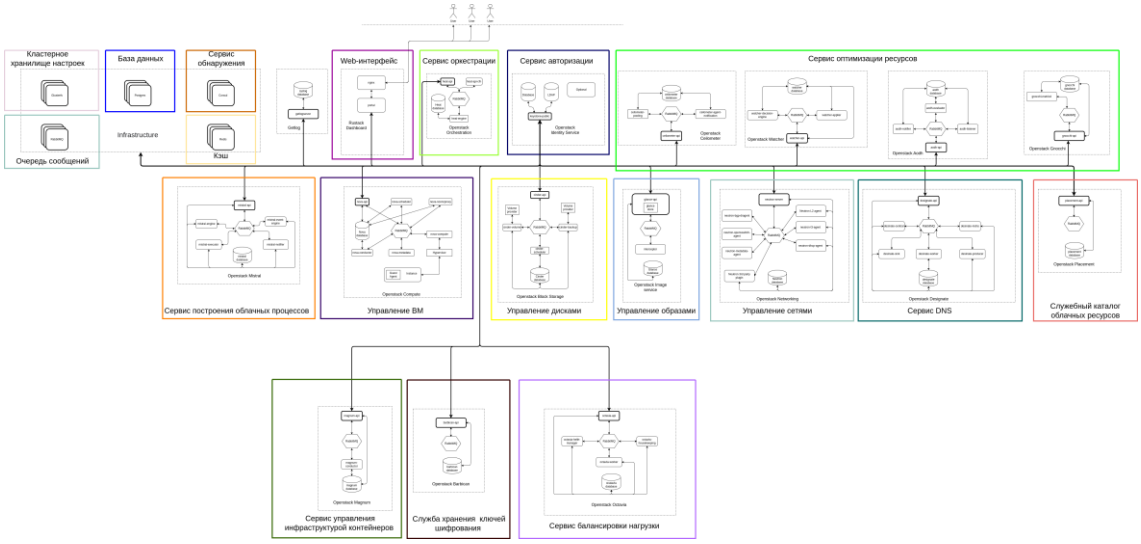
В платформу входят следующие модули, разработанные компанией РУСТЭК:

1. Конфигуратор/инсталлятор
2. Модуль логирования и работы с логами
3. Панель управления
4. Модуль HA
5. Система диагностики

В платформе также используются служебные сервисы, необходимые для работоспособности платформы:

1. SQL СУБД PostgreSQL
2. NoSQL СУБД Gnocchi
3. Key-value СУБД Redis
4. Сетевая сервисная платформа Consul
5. Распределенная файловая система GlusterFS
6. Очередь сообщений RabbitMQ

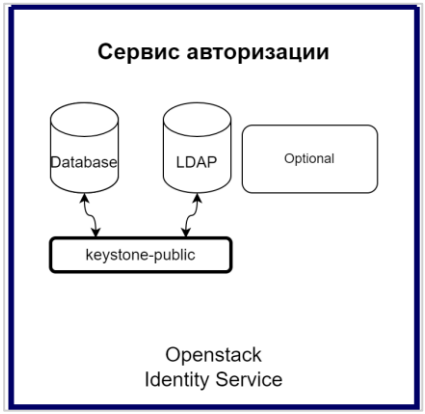
Схема архитектуры платформы РУСТЭК:



3 Описание сервисов

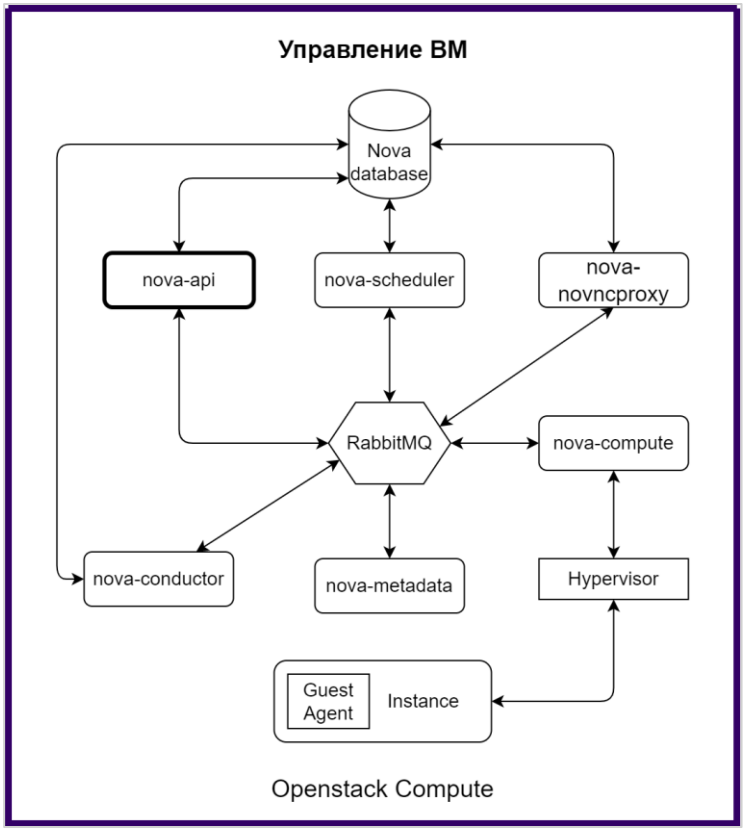
Сервис авторизации

Основным компонентом является сервис идентификации OpenStack Keystone. Обеспечивает аутентификацию и авторизацию пользователей для всех компонентов РУСТЭК. Идентификация поддерживает несколько механизмов аутентификации, включая учетные данные имени пользователя и пароля, а также систему на основе токенов.



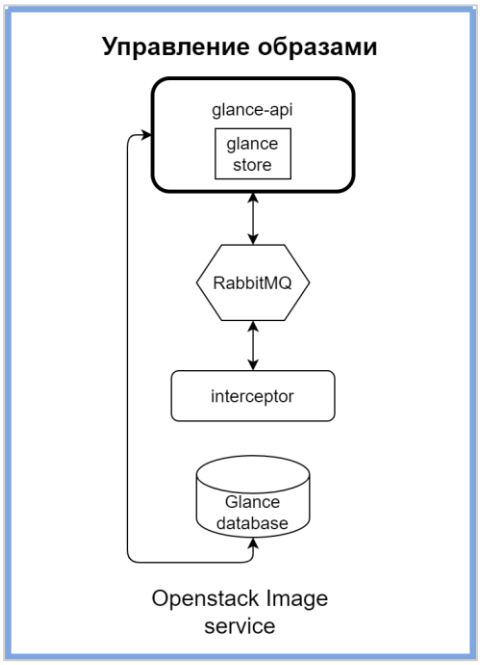
Управление ВМ

Основной составляющей является сервис вычислений OpenStack Nova. Предоставляет способ для развёртывания виртуальных машин. Для выполнения основных функций используются следующие дополнительные сервисы OpenStack: Keystone ("Сервис авторизации"), Glance ("Управление образами"), Cinder ("Управление дисками"), Neutron ("Управление сетями") и Placement ("Служебный каталог облачных ресурсов").



Управление образами

Основным компонентом является сервис образов OpenStack Glance. Создаёт репозиторий образов для создания впоследствии из них виртуальных дисков. Через веб-панель управления платформой РУСТЭК можно добавлять новые образы или делать снимок существующих виртуальных машин для непосредственного хранения. Можно использовать снимки для резервного копирования или в качестве шаблонов для новых виртуальных машин.

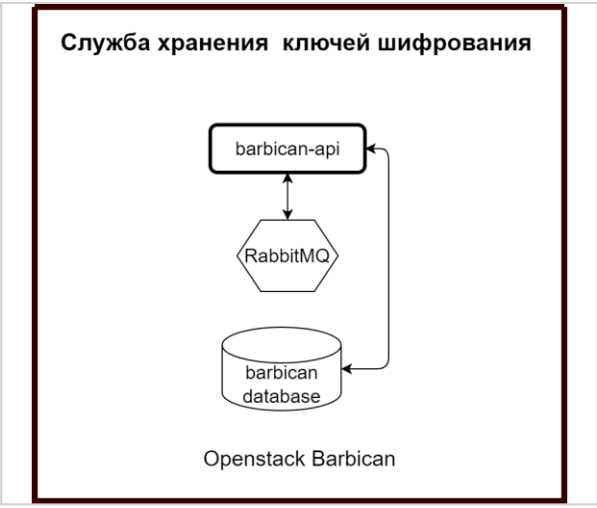


Управление дисками

Основным компонентом является сервис блочного хранения OpenStack Cinder. Обеспечивает управление постоянным блочным хранилищем для виртуальных дисков. Блочное хранилище позволяет пользователю создавать и удалять блочные устройства, а также управлять подключением блочных устройств к виртуальным машинам. Предоставляет механизм для создания и восстановления резервных копий дисков и снимков.

Служба хранения ключей шифрования

Основным компонентом модуля является сервис управления ключами OpenStack Barbican. Функционал модуля обеспечивает безопасное хранение, предоставление и управление секретными данными, такими как симметричные и асимметричные ключи, сертификаты и необработанные двоичные данные.

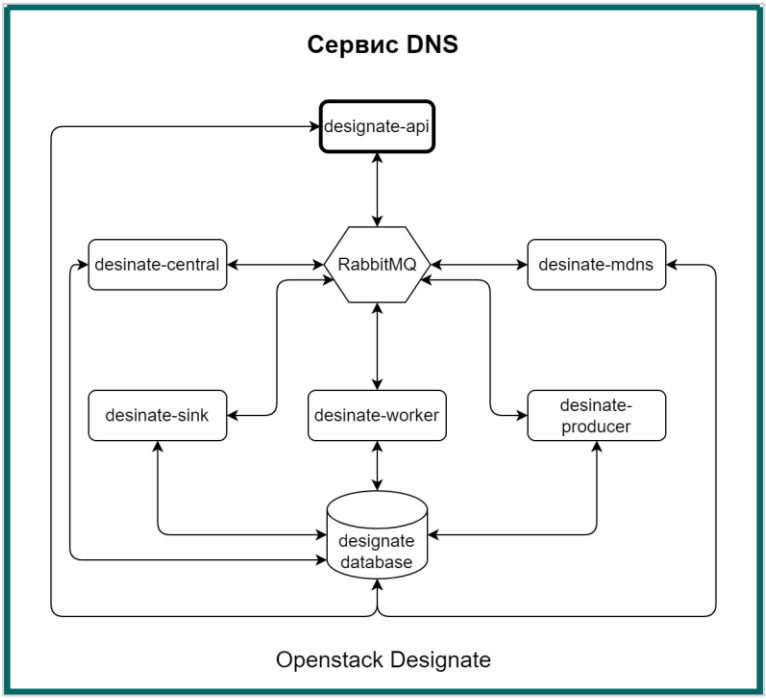


Управление сетями

Основным компонентом является сервис сети OpenStack Neutron. Обеспечивает создание и управление виртуальной сетевой инфраструктурой в облаке. Сетевая подсистема предоставляет абстрактные сети, подсети, порты и роутеры. Каждая из них по функциональности имитирует физические аналоги: сети содержат подсети, роутеры маршрутизируют траффик между разными подсетями. Для создания виртуальных машин в облаке должна быть создана хотя бы одна внутренняя сеть. Эти программно-определяемые сети подключаются к виртуальным машинам. Только виртуальные машины в данной внутренней подсети или в подсетях, подключенных к одному роутеру, имеют прямой доступ к виртуальным машинам в этой сети. Также можно создавать внешние сети. В отличие от остальных сетей внешняя сеть не является полностью виртуальной сетью. Вместо этого она представляет фрагмент физической внешней сети, доступной вне облака. IP-адреса во внешней сети физически доступны всем во внешней сети при соответствующих настройках в профилях безопасности и нижележащем оборудовании.

Сервис DNS

Основным компонентом является сервис OpenStack Designate. Обеспечивает управление доменами и записями в них, а также интегрирован с сервисами OpenStack Nova и OpenStack Neutron для автоматического создания записей в домене. В качестве DNS сервера используется PowedDNS. В качестве кэширующего DNS сервера используется dnsmasq, который устанавливается на все хосты в кластере.



Служебный каталог облачных ресурсов

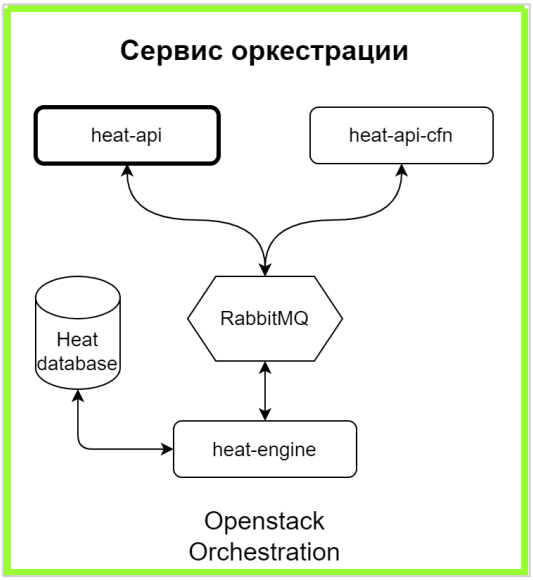
Основным компонентом является сервис Openstack Placement для отслеживания инвентаризации и использования поставщиков ресурсов, а также различных классов ресурсов.

Сервис балансировки нагрузки

Основным компонентом является сервис Openstack Octavia. Позволяет создавать сетевые балансировщики нагрузки (Load Balancer as a Service).

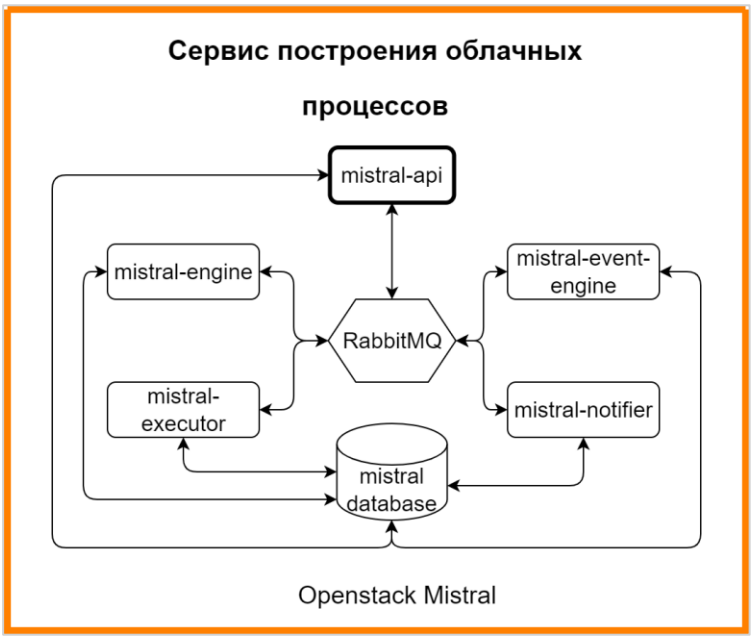
Сервис оркестрации

Основным компонентном является сервис Openstack Heat. Позволяет создавать комплексные среды и управлять ими, используя концепцию инфраструктуры как код.



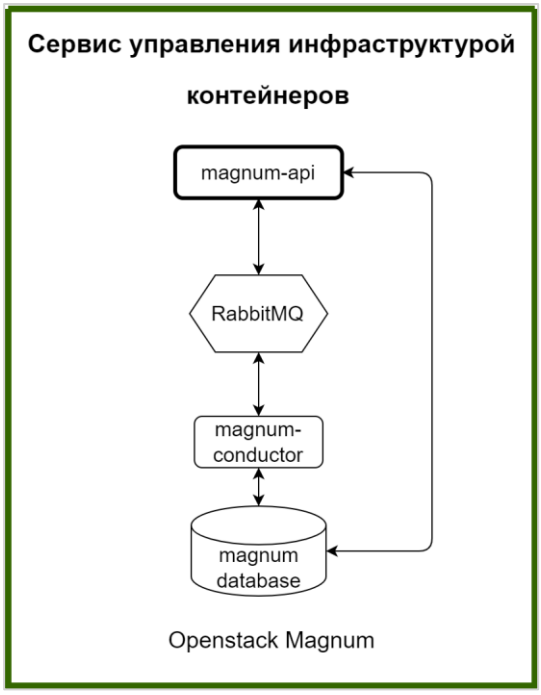
Сервис построения облачных процессов

Основным компонентном является сервис Openstack Mistral. Позволяет создавать всевозможные workflow для выполнения комплексных задач автоматизации.



Сервис управления инфраструктурой контейнеров

Основным компонентом является сервис OpenStack Magnum, предназначенный для создания и управления инфраструктурой контейнеров Kubernetes. Сервис позволяет создавать и настраивать кластер Kubernetes, а после создания кластера добавлять новые ноды Kubernetes в ручном режиме или в режиме автомасштабирования (autoscaling), при котором ноды добавляются по мере увеличения нагрузки в кластере. Кроме того, сервис выполняет проверку доступности компонентов инфраструктуры Kubernetes и запуск процедуры восстановления кластера Kubernetes при включении опции autohealing.

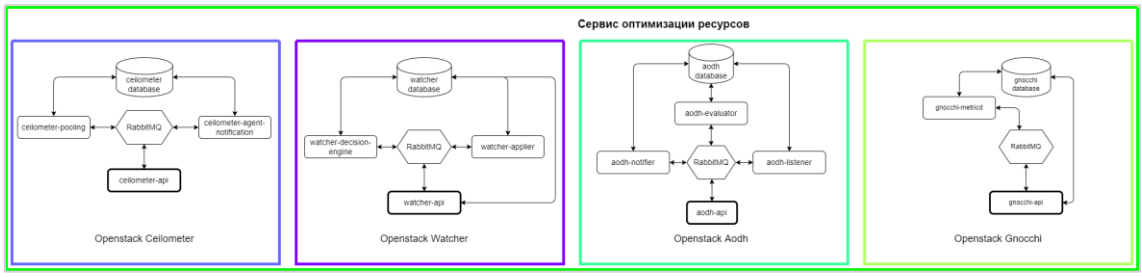


Кластерное хранилище настроек

Основным компонентом является распределённая файловая система GlusterFS. Реализует общее хранилище файлов, используемое различными компонентами РУСТЭК, что обеспечивает отказоустойчивость.

Сервис оптимизации ресурсов

Основными компонентами являются сервис оптимизации ресурсов инфраструктуры OpenStack Watcher и сервис сбора измерений OpenStack Ceilometer. Watcher – сервис для оптимизации использования ресурсов платформы, позволяет упростить обслуживание, экономичнее использовать ресурсы и гибко управлять нагрузкой на физические сервера. Ceilometer – сервис сбора данных, который предоставляет возможность нормализовать и преобразовывать данные, отправляемые существующими компонентами РУСТЭК, такими как события от сервиса Compute, или путём опроса ресурсов инфраструктуры, таких как libvirt. В качестве хранилища измерений Ceilometer использует базу данных временных рядов Gnocchi. Это сервис хранения агрегированных измерений, предназначенный для хранения метрик в очень большом масштабе, обеспечивая при этом доступ к метрикам и информации о ресурсах. Эта база данных располагается на общем файловом хранилище Gluster. Также при включении этого модуля устанавливается сервис оповещений телеметрии Aodh, который активирует оповещение, если собранные данные о событиях или измерениях нарушают определённые правила.



Web-службы

Основным компонентом является веб-сервер NGINX. Он используется для балансировки нагрузки и отказоустойчивости компонентов РУСТЭК. Для балансировки используется алгоритм least_conn, когда каждое новое соединение будет передано тому серверу, у которого на данный момент меньше всего активных соединений.

База данных

Основным компонентом является база данных PostgreSQL. База данных PostgreSQL используется компонентами РУСТЭК для хранения информации.

Кэш

Основным компонентом модуля является база данных Redis. Redis – это высокопроизводительная база данных, которая хранит данные в оперативной памяти в виде "ключ-значение". Для обеспечения отказоустойчивости используется решение Redis Sentinel. Механизм кэширования используется сервисом OpenStack Keystone для повышения производительности, так как обычно требуется, чтобы Keystone обрабатывал большое количество запросов на авторизацию и аутентификацию, поскольку Keystone вызывается почти при каждой операции Openstack. Механизм кэширования может хранить токены, идентификаторы пользователей и роли вместо того, чтобы извлекать их из удаленного хранилища. Данный модуль в конфигураторе имеет четыре состояния: "арбитр", "выключено", "дополнительный", "основной". Для обеспечения отказоустойчивости нужно установить модуль "Кэш" на количество хостов больше одного, и чтобы этих хостов было нечётное количество, а также чтобы среди этих хостов как минимум один был основной и один дополнительный.

Web-интерфейс

Основным компонентом является веб-интерфейс для работы с платформой РУСТЭК. Веб-интерфейс представляет собой графический пользовательский интерфейс для работы через браузер.

Хранилище логов

Основным компонентом является сервис управления логами Syslog-ng. Используется для централизованного сбора логов. Логи хранятся в СУБД и доступны для поиска и анализа при помощи сервиса getlog и веб-портала. Клиент Syslog-ng устанавливается на все хосты в кластере.

Сервис времени

Основным компонентом является программа-демон ntpd, которая устанавливает и поддерживает системное время. При включении данного модуля на хосте будет настроен NTP сервер. На хостах с выключенным модулем будет настроен NTP клиент.

Очередь сообщений

Основным компонентом является брокер сообщений RabbitMQ. Это программный брокер сообщений на основе стандарта AMQP, написанный на языке Erlang. РУСТЭК используют очередь сообщений для координации операций и информации о состоянии между сервисами.

Сервис отказоустойчивости

Основным компонентом является сервис VAHA. Он обеспечивает высокую доступность для вычислительных хостов. Определяет доступность хоста на основании сети (ping) и хранилища (файлы heartbeat-ов в корне точек монтирования). В случае недоступности хоста принудительно выключает его через ipmi и эвакуирует виртуальные машины.

Вычислительный узел

Основным компонентом является сервис nova-compute, входящий в сервис вычислений OpenStack Nova. Данный сервис создаёт и удаляет экземпляры виртуальных машин с помощью API-интерфейса libvirt гипервизора KVM.

Сервис обнаружения

Основным компонентом является сервис Consul. Предоставляет функционал обнаружения *сервисов* (service discovery) на основе DNS и проверки их доступности.