

**√**Stack

## Импортозамещение VMware Российская платформа для виртуализации vStack

### Содержание

03	Почему мы решили создать платформу для виртуализации vStack
04	Проектирование
05	Отличия от других решений на основе Open Source
06	Кому и когда будет интересна платформа
07	Как замена VMware
08	Веб-интерфейс vStack
10	Как устроена гиперконвергентная платформа vStack
11	Что такое НСІ
15	Архитектура vStack
21	Как мы развиваем vStack
21	Лицензирование

# Почему мы решили создать платформу для виртуализации vStack

В 2018 году, когда горизонтальное масштабирование практиковалось везде и всюду, нам показалось особенно интересным его применение в инженерной системе. В сущности, необходимость в «золотых» виртуальных машинах или хранилищах с появлением горизонтально масштабируемого ПО отпала. Огромный пласт дорогостоящих обеспечительных мер, которые входили в enterprise-решения, стал не нужен. Это позволило нам всерьез задуматься о своей собственной платформе, построенной по канонам гиперконвергенции.

#### Проектирование

Днем рождения нашего решения можно считать 18 июля 2018 года: тогда мы впервые встретились и начали прикидывать будущий проект на доске. Понимание, что не получится взять и слепить нечто простое, на коленке, и выпустить в прод, было у нас изначально. Нам нужна была система, которая сможет обеспечивать предоставление laaS как в сервисной, так и в on-premise модели.

VMware, являясь стандартом в мире виртуализации, со временем разрослась и превратилась в этакого монстра безумного размера. Задачи построить «убийцу VMware» у нас, разумеется, не было: нам нужно было решение, функциональности которого было бы достаточно для ряда конкретных целей.

Более того, мы не видели смысла в реализации поддержки всякого рода legacy, которым очень сильно переполнен продукт с почти 20-ти летней историей. В качестве стека мы выбрали FreeBSD, ZFS, bhyve.

## Отличия от других решений на основе Open Source

Одним из ключевых отличий нашего продукта является большое количество наших собственных разработок. Другими словами, мы не брали готовое решение, на котором просто меняли вывеску/логотип/шрифты. Мы взяли технологии и объединили их в единое решение своими разработками.

#### Мы сами создали:

- собственную реализацию алгоритма RAFT
- собственный кластерный framework
- собственный слой управления, включающий контроллеры SDC/SDN/SDS и API
- собственную виртуальную сеть, выгодно отличающуюся технологически

#### Кому и когда будет интересна платформа?

## Компаниям с частями инфраструктуры на commodity hardware

Современный бизнес умеет считать деньги, поэтому vStack интересен в первую очередь компаниям, которые понимают логику гиперконвергенции и осознают плюсы использования горизонтального масштабирования, за счет чего отпадает необходимость использования дорогостоящих компонентов. Мы постоянно подчеркиваем этот тезис и делаем на нём фокус: отпадает необходимость, но не возможность. Вы сами решаете, какие именно компоненты должны быть дорогостоящими, а какие нет, на чём именно приемлема экономия.

Один из распространённых случаев с напрашивающейся экономией – тестовые стенды. Итак, в компании есть 20 команд, которые работают над одним продуктом. Каждая из команд «выпросила» для себя отдельный тестовый стенд. И человек, который несет ответственность за бюджет, рано или поздно заметит, что 90% бюджета тратится на обеспечение тестовых стендов. Да, тестовый стенд по канону не должен отличаться от промышленного.

Но когда цена вопроса — 90% бюджета, этот аргумент уже не работает. Поэтому экономическая составляющая — тоже важное преимущество нашей системы. Потребитель самостоятельно формирует контуры и характеристики системы, выбирает компоненты. Если для него важен слой хранения, можно выбрать enterprise-диски. Если нет — можно работать на commodity-дисках и на этом сэкономить.

#### Как замена VMware

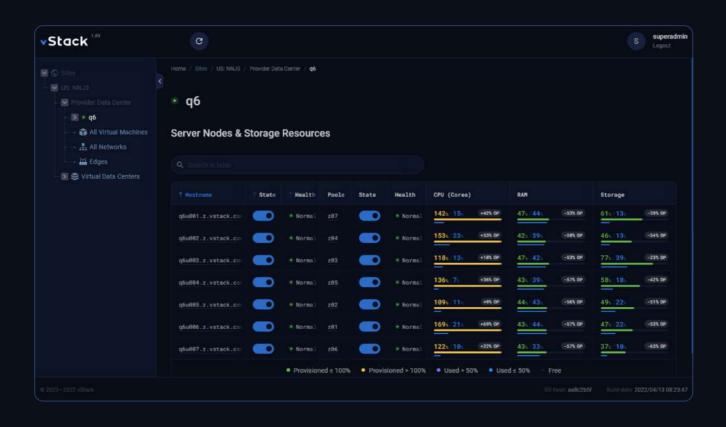
Если вы планируете импортозаместить VMware, можно рассмотреть vStack.

Платформа может вам подойти как альтернатива VMware, если вы:	Наша платформа как замена VMware вряд ли вам подойдёт если:
Задумались о vSAN, но сверхвысокая стоимость вас пугает	Вы хотите «точную копию VMware один-в-один, только не VMware»
Осознаете, что 90% функциональности VMware, которую вы не используете, входит в цену, которую вы платите	Имеющиеся решения с открытым исходным кодом покрывают ваши потребности
Можете с уверенностью сказать, что «вы используете минимум legacy технологий»	Возможность использования привычного инструмента (VMware) для вас важнее любых других факторов
Активно используете ПО с технологиями горизонтального масштабирования	
Понимаете, что текущая инфраструктура усложнилась настолько, что любые изменения в ней происходят с болью, скрипом и за очень большое время	
Не знаете, что делать с этими 12 стойками разношерстного оборудования для тестовых стендов и как на следующем бюджетном комитете защитить статьи на поддержку этого железа без отсылок к Господину Инженеру	

### Веб-интерфейс vStack

Зачастую считается, что веб-интерфейс сложной инженерной системы должен быть максимально ужасным, «лишь бы работал». Мы постарались сделать интерфейс максимально простым и удобным. Все модули и компоненты ИТ-инфраструктуры на базе vStack настраиваются и администрируются в единой панели управления.







## Как устроена гиперконвергентная платформа vStack

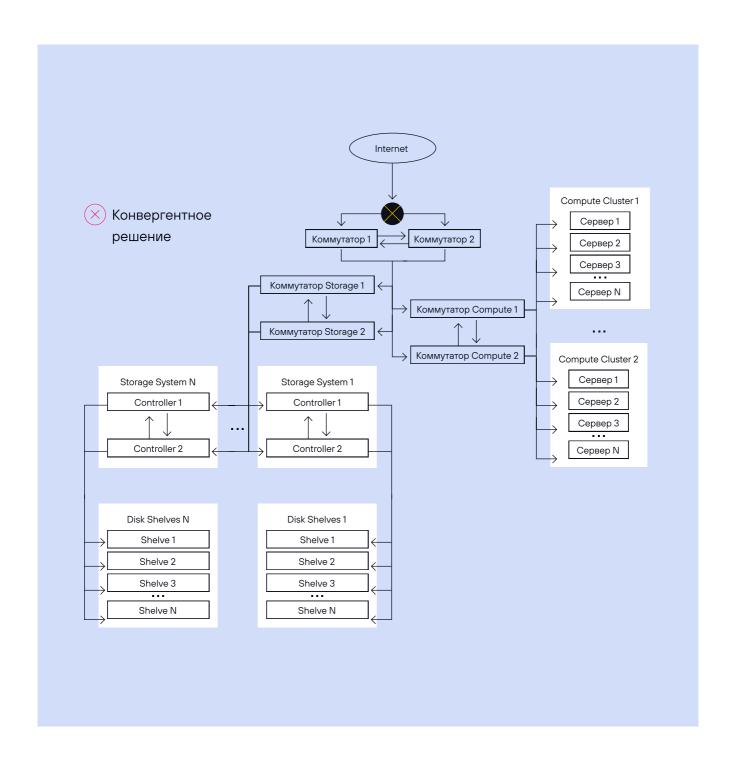
Гиперконвергентные решения, доступные на современном рынке, можно условно разделить на два типа: «полноценные» и интегрированные. Их главное отличие заключается в железе, которое можно использовать при построении инфраструктуры. У настоящей гиперконвергенции (Nutanix HCI и VMware HCI) нет привязки к производителям оборудования: она запросто подружится с аппаратным обеспечением другого вендора. А интегрированные решения работают только на железе вендоров.

Наша платформа vStack — это тоже полноценная гиперконвергенция.

#### Что такое HCI

Гиперконвергентная инфраструктура (HCI) — это программно-определяемая инфраструктура, которая объединяет в себе функции вычисления (SDC), хранения данных (SDS) и программной реализации сети (SDN). В отличие от классической инфраструктуры, где эти функции выполняют серверы с выделенными ролями или отдельные сущности вроде СХД и роутеров, в HCI они возложены на каждый сервер-участник инфраструктуры.

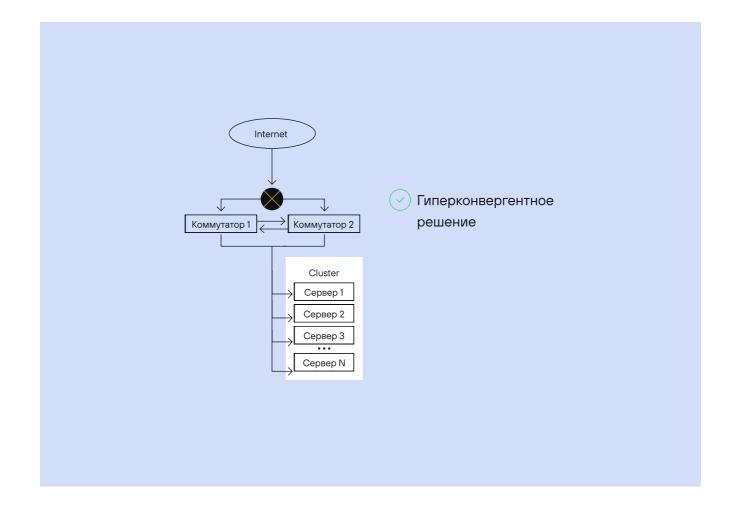
Гиперконвергентное решение. Здесь все функции (SDS, SDC, SDN) выполняются кластером унифицированных x86-серверов с дисками.



Классическая схема организации корпоративной инфраструктуры: роутер, НА-пара коммутаторов ядра, пара коммутаторов сетей общего назначения, непосредственно сами серверы (в том числе резервные хосты), пара коммутаторов SAN, комплекс СХД с зарезервированными контроллерами (возможно, с дополнительными дисковыми полками). В такой ситуации использование оборудования разных вендоров (например, сетевое — от Cisco, СХД — от NetApp и т.д.) — абсолютно нормальная практика.

При этом для поддержки каждого сегмента (вычислительного/сетевого/хранения) требуется отдельный специалист, знакомый с особенностями эксплуатации, взаимодействия с вендором и спецификой оборудования в целом. NAS остаётся правее viewport.

Согласитесь, выглядит намного компактнее? При этом уровень доступности и в первом, и во втором варианте исполнения при прочих равных будет одинаков.



## Ключевые преимущества HCI



#### Более простая инфраструктура

Дискретность гиперконвергентной инфраструктуры существенно ниже даже в приведенном примере, притом, что в случае конвергентной инфраструктуры в слое хранения NAS-составляющая просто не поместилась в экран. Этот фактор напрямую влияет на сложность эксплуатации, так как в ней задействовано немало узких специалистов. Количество людей растет вместе с размером инфраструктуры, так как и управление ею, и она сама дискретны. Каждый из элементов (сетевое оборудование, СХД, SAN, NAS, серверы) имеет собственные интерфейсы управления, концепции, API и т.д. Данная дискретность обуславливает следующий постулат: изменения конвергентной инфраструктуры могут растягиваться на месяцы.

В нашем гиперконвергентном решении такая дискретность отсутствует: управление имеет единые endpoint, интерфейс и API.

Однородность элементов гиперконвергентной инфраструктуры позволяет в случае неисправности (при желании) не тратить драгоценное в XXI веке время на ее локализацию. Достаточно вынуть диски из сервера, вышедшего из строя, и вставить их в ZIP и продолжить работу с прежним (референсным) уровнем резервирования.



#### Нет вендорлока

Подойдет любое доступное на рынке оборудование, что теперь актуально, как никогда раньше. Даже потребительское.



#### Экономия на оборудовании

Можно закупать аппаратное обеспечение потребительского сегмента и таким образом экономить на оборудовании некритичных сред, например, тестовых. Или кардинально снижать стоимость там, где используются горизонтально-масштабируемые решения, не требующие высокой степени резервирования одного элемента. К тому же заменить вышедшее из строя "потребительское" устройство будет несколько проще, чем enterprise.



#### Управление всей инфраструктурой через единый интерфейс

Его может обслуживать всего один опытный специалист — вам не придется нанимать отдельных сотрудников под каждый сегмент.

## Архитектура vStack

Платформа создает единое кластерное пространство на базе серверов, одновременно выполняющих сразу 3 функции: Software Defined Storage, Software Defined Networking, Software Defined Computing — хранение, сеть, вычисления. Она объединяет три традиционно разрозненных компонента в одно программно-определяемое решение.

#### Программноопределяемые вычисления (SDC)

Слой SDC работает на базе гипервизора второго типа. Поддерживается спецификация virtio для сетевых портов, дисков и другой периферии (rnd, balloon и т.п.). Кастомизация гостевых ОС в облачных образах реализуется с помощью cloud-init для unix-систем и cloud-base для windows.

#### Облачные образы (cloud images):

- Возможность лимитировать
  производительность сетевого порта (МВрѕ)
  и диска (IOPS, МВрѕ) в реальном времени;
- Лимитирование ресурсов vCPU в реальном времени, в том числе автономное распределение квантов CPU в зависимости от степени overcommit, утилизации узла и нагрузки конкретной ВМ, благодаря чему показатели vCPU overcommit в промышленных кластерах иногда достигают 900%;
- Снимки ВМ, содержащие её конфигурацию (в т.ч. сетевых портов и IP-адресов, а также МАС-адресов на этих сетевых портах).

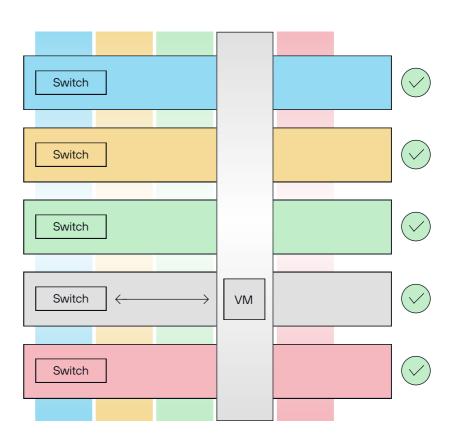
#### Поддерживаемые ОС:

- FreeBSD:
- Linux (OEL/CentOS/Ubuntu/Debian);
- Windows 2019/2022.



Отдельного упоминания заслуживает автономный механизм бюджетирования квантов vCPU, благодаря которому становится возможной экономическая эффективность (efficiency) в слое SDC до 900%. Таким значением может похвастаться далеко не каждое решение.

При создании vStack мы в первую очередь ориентировались на его легковесность. Низкий показатель CPU Overhead (снижение производительности виртуализованного сервера относительно физического) -- один из примеров возможного эффекта этого подхода. Например, overhead y KVM в некоторых нагрузках достигает 15%, а у vStack — всего лишь 2-5%. Таких низких значений удалось достичь благодаря использованию уникальности гипервизора bhyve.



## Программно-определяемое хранилище (SDS)

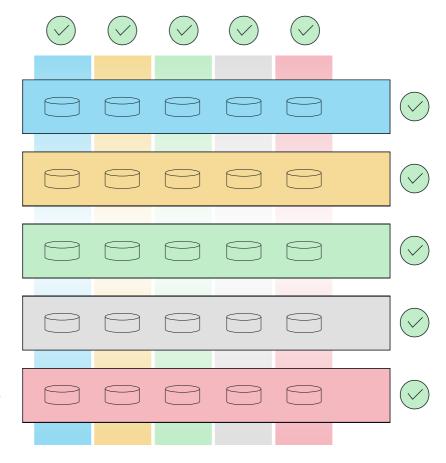
Технологической основой SDS является ZFS.

#### Возможности:

- Компрессия и дедупликация;
- Внутренняя целостность данных;
- Клоны, снимки;
- Самовосстановление данных;
- Транзакционная целостность.

Лимиты (pool/filesystem):

- размер: 1 zettabyte;
- неограниченное количество файловых систем;
- неограниченное количество блочных устройств.



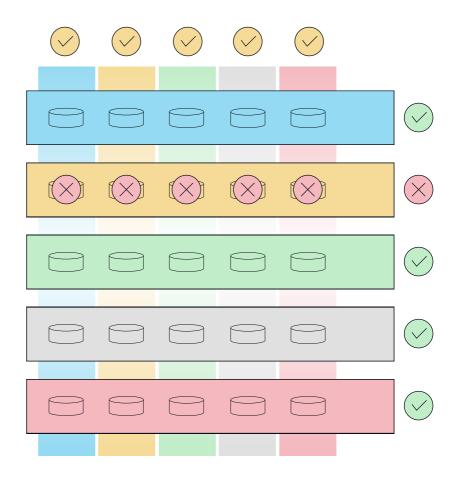
Так выглядит пятиузловой кластер. Вертикальные контейнеры — пулы, горизонтальные — узлы кластера. Каждый пул содержит диск из каждого узла. Redundancy пула всегда равен redundancy кластера.



Для всех промышленных инсталляций мы рекомендуем использовать избыточность N+2 или выше (платформа поддерживает N+1, N+2 и N+3).

Что произойдет, если случится авария?

На картинке ниже желтый хост вышел из строя и за счет механизма fencing был исключен из кластера, а все пулы потеряли по одному диску. Кластер автоматически выполнил процедуру аварийного переключения (failover) ресурсов данного хоста, и желтый пул стал доступен на синем хосте. Redundancy снизился, но пул продолжает работать, причём без потери производительности. Все ВМ, работавшие на желтом хосте, продолжили работать на синем.



## Программно-определяемая сеть (SDN)

Доступно три варианта обеспечения виртуальных сетей:

Vlan

--- Vxlan

GENEVE (собственная имплементация).

Вне зависимости от способа имплементации сеть может быть изолированной или маршрутизированной. Конкретный экземпляр виртуальной сети базируется на распределенном свитче. Реализация самой виртуальной сети, свитча в экосистеме netgraph и модификации в сетевом стеке — наши собственные разработки, прошедшие серьезные испытания и давно (с сентября 2020 г.) работающие в промышленной эксплуатации на большом количестве кластеров.

#### Приемущества: – собственный MTU у каждого Switch экземпляра сети; маршрутизируемые/ изолированные сети; Switch поддержка jumbo frames; поддержка TSO/GSO; поддержка path mtu discovery Switch "из коробки". Лимиты: Switch 65536 сетей; 1 048 576 портов на свитче одного Switch виртуальный порт: 22 GBps / 2.5Mpps.

#### Как мы развиваем vStack

В новом релизе vStack v.2.1 появятся следующие возможности:

- Миграция работающих виртуальных машин.
  Эта функциональность позволит без простоев переносить активные рабочие нагрузки с одного сервера на другой, чтобы пользователи продолжали иметь непрерывный доступ к системам, которые им необходимы.
  Аналогична vMotion в VMware.
- Полноценное клонирование виртуальных машин.
  Мгновенное создание клона виртуальной машины любого объема, ограниченного ресурсами кластера.
- Поддержка сетей на основе GENEVE.
  Открытая технология, имеющая стандарт.

- Поддержка функциональности EDGE
  в графическом интерфейсе. EDGE-роутер это
  программный маршрутизатор, через который
  виртуальные машины внутренних сетей получают
  доступ за её пределами и наоборот.
- Поддержка VM ballooning механизма
   оптимизации работы виртуальных машин
   с оперативной памятью, который позволяет
   возвращать хосту распределённую однажды
   гостевой ОС память, которая была впоследствии
   освобождена.
- Поддержка QinQ (802.1ad).
- Увеличение производительности платформы в слое хранения и вычислений.

#### Лицензирование

vStack предоставляется по двум моделям:

#### Сервисная (ОРЕХ).

Плата за использование лицензий. Ежемесячная тарификация и оплата за points.

В любом варианте доступна возможность Managed vStack — удаленное администрирование от вендора.

#### On-premise (CAPEX).

Покупка лицензий в собственность. Единоразовый платеж плюс ежегодная оплата поддержки. В этой модели доступен overcommit: можно использовать на 33% больше ресурсов, чем выделено по договору.

Если у вас нет сотрудника, который мог бы взять на себя роль администратора, можно делегировать эту задачу нам.

По любым вопросам

info@vstack.com

По вопросам партнеров

partners@vstack.com

Для заказов

sales@vstack.com

Россия, Санкт-Петербург, Литейный проспект, 26

+7 (812) 313-88-15