

Блог доктора Добрянского

Строим кластер на VMware: часть 1 — теория

Давно хотел написать цепочку статей о виртуализированных кластерах. Их архитектуру, принципы построения с инструкциями и примерами установки.

Я не собираюсь писать целые тома с техническими подробностями, хочу лишь поверхностно охватить данную тему. Но даже для поверхностного раскрытия понадобится несколько статей и это первая из них — теоретические основы.

Тех, кто еще не понял, предупреждаю — это статьи для новичков в данной теме, скорее для общего развития чем для приобретения глубокого технического опыта и знаний.

Для начала разберемся, что мы называем кластером. В разрезе **VMware**, кластер — это группа хостов (физических серверов), связанных между собой сетью, которые управляются единым сервисом и совместно выполняют определенные функции, как один большой организм.

На платформе от **VMware** — **vSphere** можно построить 2 вида кластеров: **HA (High-availability)** кластер и **DRS (Distributed Resource Scheduler)** кластер, которые работают на уровне виртуальной машины.

HPC (High-performance computing) кластер, в котором вычислительные мощности всех физических хостов объединяются в одну и все это дело работает как один большой компьютер да еще и производит мега-супер-пупер крутые параллельные вычисления, на VMware построить не получится!

Давайте посмотрим на доступные виды кластеров и на преимущества, которые они нам дают.

Реализация HA-кластера у VMware такова: несколько физических хостов объединяется в кластер и при выходе из строя одного из хостов, виртуальные машинки, которые были на нем, запускаются на других хостах кластера, на которых зарезервированы ресурсы. При этом время простоя виртуальной машины, равняется времени загрузки ее операционной системы.

Как видим, кластером высокой доступности **VMware HA** сложно назвать, т.к. при отказе оборудования мы все равно получаем небольшое время простоя. Для избежания этого простоя, совместно с **VMware HA** обычно используют кластеризацию самого приложения (напр. **MS Exchange** кластер, или **MS SQL** кластер).

Что бы сократить время простоя сервиса (виртуалки) до нуля (или времени очень близкому к нулю), можно также использовать технологию **VMware Fault Tolerance**. Фишка ее заключается в том, что при работе виртуальной машины на одном физическом хосте, ее «призрак» (синхронная реплика) работает на другом хосте кластера и при выходе из строя основного хоста (и соответственно виртуальной машины) происходит мгновенное переключение ввода-вывода и активизация призрака. При этом время простоя виртуальной машины равняется нулю, как говорит маркетинг **VMware**, в реальности все немного не так (на практике замечаются небольшие подтормаживания 1-2 сек и потеря одного «пинга»).

Надо признать, что **Fault Tolerance** на практике используется очень редко из-за своих достаточно серьезных ограничений и требований, но зато в качестве маркетинговой фишки упоминается при любой малейшей возможности :)

Итак, с высокой доступностью разобрались. Теперь посмотрим **что же собой представляет DRS**.

Если не вдаваться в подробности, то **VMware DRS** — достаточно простая штука, которая выбирает на каком физическом хосте запустить виртуальную машину, или на какой хост смигрировать работающую виртуалку, с целью выровнять нагрузку на всех хостах кластера.

DRS объединяет ресурсы кластера в один большой пул и в полу-автоматическом или полностью автоматизированном режиме распределяет виртуальные машины равномерным слоем по физическим хостам кластера.

DRS имеет пару штук настроек и может либо выдвигать рекомендации по запуску, или перемещению виртуальных машин на серверах кластера, либо автоматически мигрировать виртуалки с помощью технологии **vMotion** в зависимости от текущей нагрузки.

Нужно упомянуть о том, что **VMware DRS** как в принципе и **HA**, это не всегда полезно и круто. У данных технологий есть свои нюансы, которые стоит учитывать при разработке проекта.

Перейдем к архитектуре всей этой кухни.

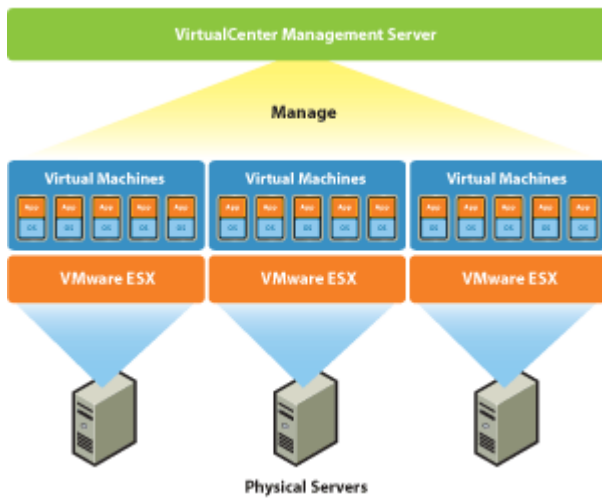
Кластер VMware строится из 2-х или больше (минимум 2, максимум — 32) серверов **ESXi**, которые управляются с помощью **VMware vCenter**.

Для построения кластера нам также понадобится общее хранилище — система хранения данных (далее СХД). СХД может быть **Fibre Channel**, **iSCSI** или **NFS**-ресурс. На СХД хранится один или несколько разделов, в кластерной файловой системе **VMFS**, которые доступны для чтения и записи одновременно всем **ESXi**-серверам кластера. Разделы хранят на себе файлы виртуальных машин (виртуальные диски, файлы конфигурации и т.п.), которые также доступны всем серверам одновременно.

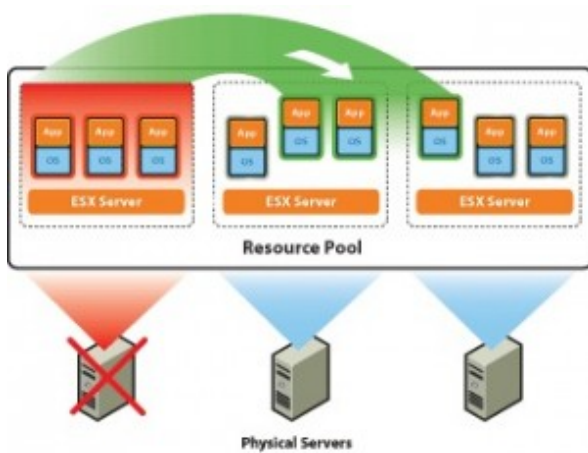
Именно за счет общего хранилища и независимости виртуальной машины от физической платформы, достигается быстрое перемещение/восстановление виртуальных машин.

Ниже **несколько картинок от VMware** для лучшего понимания.

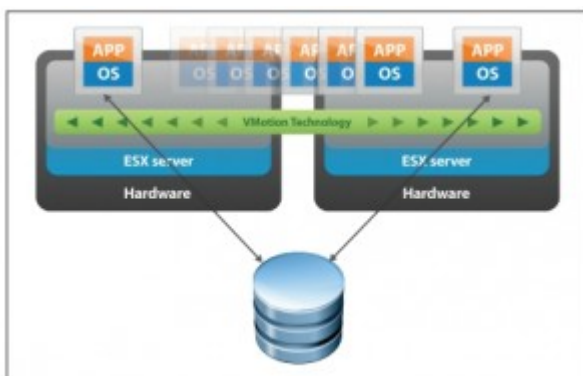
Общая схема платформы VMware vSphere:



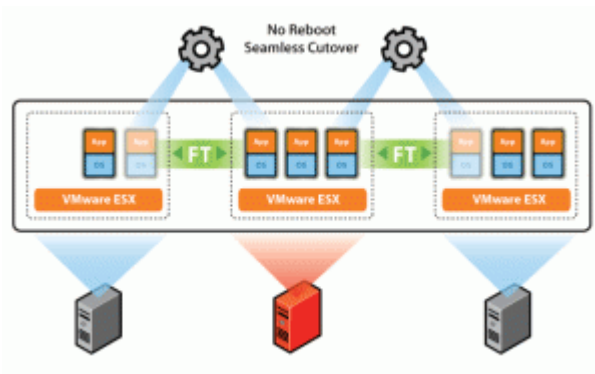
Принцип работы VMware HA + DRS схематически:



vMotion:



И Fault Tolerance:



В данной статье рассматривается **платформа vSphere от VMware**, но у других производителей ПО, например, **Citrix XenServer** или **Microsoft Hyper-V** концепция практически такая же, за исключением нюансов в технологиях и возможностях, а также, разницы в цене решения.

С теорией вроде бы все, переходим к практике. В следующей статье расскажу с картинками об установке и начальному конфигурированию платформы для нашего кластера.

Тем, кто заинтересовался — **продолжение следует... :)**

Поделиться:



Запись опубликована 18 января 2012 [<http://hutpu4.net/puplication/stroim-klaster-na-vmware-chast-1-teoriya.html>] в рубрике Виртуализация, Статьи с метками Citrix, DRS, ESXi, Fault Tolerance, HA, vCenter, vMotion, VMware, vSphere, XenServer.

Строим кластер на VMware: часть 1 — теория: 9 комментариев



Alex

19 июля 2013 в 15:13

А где же часть 2?



Доктор Добрянский

Автор записи

19 июля 2013 в 15:16

Стыдно это признавать, но все никак руки не доходят :(Если есть интерес, то в ближайшее время постараюсь собраться с мыслями и все же написать часть 2, часть 3 и т.п...



Dmitry

30 июля 2013 в 22:54

Интерес есть. Любая статья с доходчивым и грамотным объяснением всегда интересна и нужна. Так что ждем с нетерпением продолжения.

PS: Автору спасибо за уже написанное =)



Александр

1 августа 2013 в 18:16

Очень надеюсь на скорое появление второй части. Желаю удачи и свободного времени! :) Спасибо!



Доктор Добрянский

Автор записи

2 августа 2013 в 15:10

Спасибо за комментарии, они мотивируют на написание нового материала! :) Продолжение статьи в процессе ;)



Вася

4 августа 2013 в 22:02

Ждем продолжение!



Саня

8 августа 2013 в 19:37

С нетерпением жду продолжения



Доктор Добрянский

Автор записи

16 августа 2013 в 11:56

Продолжение здесь — <http://hutpu4.net/puplication/stroim-klaster-na-vmware-chast-2-ustanovka-i-nastrojka-esxi-v-kartinkax.html>



Ольга

13 марта 2014 в 17:02

Очень доступно и понятно
Огромное спасибо автору
