

2023-01-04 Рассказ о том как я спроектировал Ludicrous Speed пул в одном очень большом государственном заказчике

cool-story

Примерно в 2016 году на одном из крупных проектов у нас возникла некоторая проблема финансирования. Изначально инфраструктура хранения была спроектирована с использованием СХД уровня Hi End, (HDS VSP G1000). Но в процессе бюджетирования выяснилось что денег на размещение всей необходимой емкости на данной системе в проекте уже нет.

Системы HiEnd Hitachi Vantara обладают отличным функционалом виртуализации внешних систем хранения данных что позволяет администратору подключить все имеющиеся у него СХД к Hitachi, использовать емкость этих массивов для увеличения собственной емкости СХД Hitachi и производить администрирование дисковой емкости в последствии из единой консоли. Помимо "проброса" емкости система обеспечивает весь функционал идентично как для внешней так и для внутренней емкости.

Основным сценарием использования данного функционала - Universal Volume Manager является миграция с существующих систем и сценарии с внутрисистемной репликацией. Это связано с фундаментальными недостатками подхода виртуализации внешних СХД в принципе:

1. Надежность дискового ресурса будет соответствовать самому ненадежному компоненту цепочки между конечным жестким диском и целевой клиентской системой. К снижению надежности относятся все промежуточные компоненты в этой цепочке:
 1. Виртуализуемый массив, его микрокод, аппаратные компоненты;
 2. Каналы связи между двумя массивами, включая кабели, интерфейсы, коммутаторы;
 3. Если в данном случае выдающей стороной является массив уровня HiEnd о недостатках микрокаода например можно забыть, но при этом он обладает конечными ресурсами;
2. Администрирование усложняется на порядок. Не смотря на то, что внешние системы необходимо грубо говоря настроить только один раз, в случае возникновения НШС ¹⁾ часто приходится быстро вспоминать что где и как было настроено, что куда презентовано и каким образом;
3. Сложность обслуживания увеличивается в случае если оборудование расположено в различных шкафах или территориально разнесено друг с другом. Например в случае необходимости обслуживания одной из стоек с внешним массивом необходимо либо останавливать всю инфраструктуру хранения либо иметь запас свободной емкости для вывода указанного массива из работы.

Не смотря на все недостатки нами было принято решение изъять часть емкости из системы VSP G1000 и заменить её парой G400, одной на 2,5" дисках а другой на 3.5". Это наименьшее из всех зол среди которого можно было выбрать. В проект конечно же вошли совсем другие строки:

``Точкой подключения клиентов к дисковым ресурсам являются системы Hitachi Data Systems - Virtual Storage Platform и Hitachi Data Systems - Virtual Storage Platform G1000. Дисковая емкость систем G400. Данное решение увеличивает гибкость всей подсистемы хранения данных путем создания «единой точки входа» серверов к дисковым ресурсам, позволяет повысить производительность за счет применения твердотельных накопителей и увеличения объема кэш-памяти дисковой подсистемы хранения данных, упрощает администрирование, сокращая число систем требующих постоянной настройки. ``

Нельзя назвать указанное совершенным враньем но в данной ситуации использование механизма виртуализации на система VSP стало нашим спасительным козырем на много лет вперед так как расширение системы теперь можно было проводить за счет добавления более дешевых дисковых массивов и при необходимости увеличения производительности контроллерного модуля G1000. Передо мной стал серьёзный вопрос об объединении дисковой емкости из полученной конфигурации. Дело в том что диски на внешней системе и диски на втыренной системе по мнению СХД и в реальной жизни обладают совершенно отличающимися показателями производительности.

В случае использования механизма многоуровневого хранения Hitachi Dynamic Tiering все внешние ресурсы всегда будут иметь приоритет меньший чем внутренние. Таким образом родилась следующая схема создания пулов внутри СХД. Я точно не могу сказать почему именно мне пришлось использовать название Ludicrous Speed для самого быстрого пула, но к моему недоумению данный пул прошел все проверки и согласования во всех инстанциях заказчика и я не удивлюсь если данное наименование пула эксплуатируется по сей день.

#	Имя	Тип дисков	Размещение
0	DP0_ludicrous_speed	3.2TB Flash Module Drive	G1000 Internal
		1.2TB 10K rpm LFF Disk Drive	G400 UVM
1	DP1_light_speed	600GB 10K rpm SFF Disk Drive	G1000 Internal
		4TB 7.2K rpm LFF Disk Drive	G400 UVM

Таким образом после создания указанных пулов система предоставляла конечному потребителю ресурсы следующих шести уровней. Как минимум три дня я потратил на выяснение наиболее драгоценных металлов чтобы “наскрести” на 6 уровней. Но в итоге решил ввести уровни titanium и diamond выше platinum. Сейчас можно долго спорить в какой последовательности необходимо было распределить данные ископаемые, но в проект ушла следующая схема:

№	Название	Пул	Политика	Тип носителя	Раздел кэш-памяти
1	Diamond	DP0	Level 1 (1) (only tier 1)	3.2TB Flash Module Drive	Diamond
2	Titanium	DP0	Level 2 (2) (tier 1 and tier 2)	3.2TB Flash Module Drive и 1.2TB 10K rpm LFF Disk Drive	Titanium
3	Platinum	DP0	Level 3 (3) (only tier 2)	1.2TB 10K rpm LFF Disk Drive	Platinum
4	Gold	DP1	Level 1 (1) (only tier 1)	600GB 10K rpm SFF Disk Drive	Gold
5	Silver	DP1	Level 2 (2) (tier 1 and tier 2)	600GB 10K rpm SFF Disk Drive и 4TB 7.2K rpm LFF Disk Drive	Silver
6	Copper	DP1	Level 3 (3) (only tier 2)	4TB 7.2K rpm LFF Disk Drive	Copper



1)

Нештатной ситуации

From:
<https://micronode.ru/> - **micronode.ru**



Permanent link:
<https://micronode.ru/blog/2023/01/04>

Last update: **2023/01/08 18:19**