

Отечественные системы хранения данных

Весна 2022 года стала значительным событием в эпохе Российского импортозамещения. Из-за введённых санкций практически все импортные производители ИТ-оборудования были вынуждены прекратить официальную деятельность на территории Российской Федерации. Особенно заметно это событие сказалось на рынке систем хранения данных, не смотря на то, что наработки в этом направлении ведутся давно.

У многих архитекторов и заказчиков назрел вопрос о том, какие системы хранения данных доступны сейчас на рынке, возможно ли использование “традиционного” подхода к построению ИТ-инфраструктуры с созданием сети хранения данных и на каких производителей стоит обратить свое пристальное внимание. Отрасль получила большое ускорение в связи с возросшим спросом, освободившимися бюджетами заказчиков с высокими требованиями к системам хранения данных, привыкших к импортным решениям мировых лидеров отрасли.

Самые печальные последствия, которыми обернулись ограничения на поставки зарубежных систем хранения данных различных типов, можно сформулировать следующим образом:

1. **Тип AllFlash / NVMe** в основном был представлен только импортными решениями уровня: Pure FlashArray X//, DELL EMC PowerStore, Netapp AFF и практически полностью прекратил существовать;
2. **Тип HiEnd** - системы, способные работать с Mainframe и обеспечивающие практически 100% доступность, представлены только двумя производителями: EMC и HDS, которые полностью прекратили поставки и обслуживание;
3. **Типы Backup & D2D** - ввиду отсутствия разработок отечественных ленточных библиотек, скорее всего выпущено не будет никогда. По d2d системам и альтернативным хранилищам перспективы сохраняются.

Самый широкий рынок систем хранения данных - “General Purpose Storage”, к счастью, достаточно хорошо наполнен отечественными решениями, но с небольшой оговоркой, что системы попадают в данный класс в основном из-за своей цены и поддерживаемого количества накопителей, но не по функционалу, сравнимому с начальным классом. По моему мнению, это вызвано скорее желанием производителя зависеть возможности масштабирования собственных решений, высокими возможностями современных процессоров в контексте шины данных и неоптимальным программным кодом Российских СХД.

В следующих разделах я постараюсь разобрать каждый тип систем хранения в отдельности, какими отечественными производителями он представлен и какие перспективы можно ожидать в данном направлении.

Тип: Low End

Развитие недорогих двухконтроллерных систем с поддержкой FibreChannel позволило существенно ускорить темпы внедрения технологий виртуализации, высокой доступности в компаниях малого и среднего бизнеса. Традиционно ИТ-инфраструктура таких заказчиков строилась из нескольких серверов, возможно, пары коммутаторов Brocade и одного хранилища уровня HPE MSA. Этот “стартовый набор” позволял создать современную, надежную,

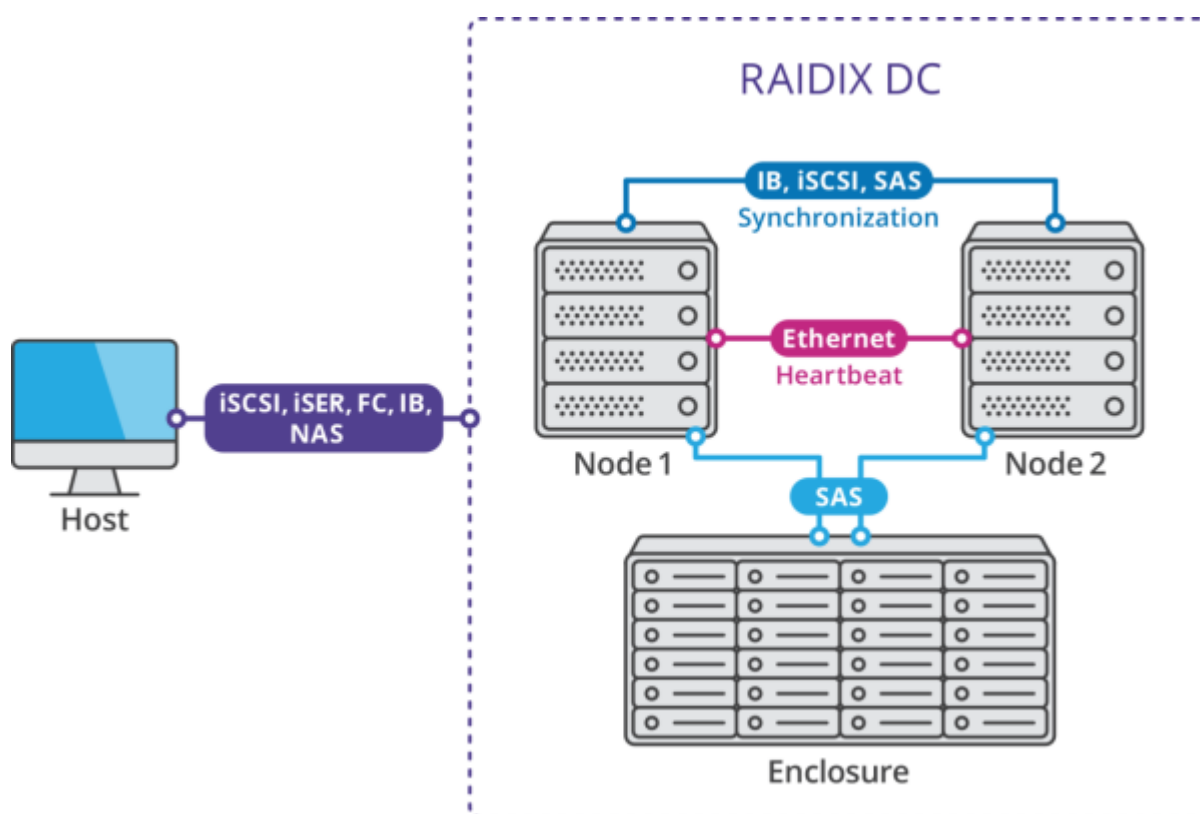
отказоустойчивую инфраструктуру с минимумом вложений в квалифицированный персонал, проектирование и оборудование.

Некоторые заказчики после перехода в класс крупных продолжают следовать данному подходу и используют такие комплексы для решения небольших изолированных задач.

К сожалению, цены отечественных систем хранения совсем не соответствуют ожиданиям прежних заказчиков HPE MSA*, DELL ME*. Аппаратные решения в данном сегменте представлены только, возможно, параллельным импортом, который тоже значительно повышает стоимость решения или системами хранения данных третьего эшелона, например Infortrend.

Для закрытия пробелов в таких проектах необходимо предлагать программно-определяемые решения. Это либо конвергентные среды, либо дешевые SDS, например, RAIDIX.

ПО СХД RAIDIX



Софт для СХД RAIDIX производителем позиционируется в первую очередь для больших потоковых записей, где требуется быстрое перестроение массива и защита от выхода из строя одновременно трёх дисков, например: медиа-архивы, хранилища резервных копий. Но в связи с последними событиями raidix остался один в данном ценовом сегменте из отечественных решений, поэтому компания быстро перестраивает свой маркетинг в направлении СХД общего пользования.

Главным минусом этой системы всегда была неудачная реализация двухконтроллерного режима с неоптимальным механизмом синхронизации кэш-памяти между контроллерами. Хотя многие современные СХД-стартапы сталкиваются с этой проблемой, но каждый решает её по своему. Raidix уже давно не стартап - и наличие проблем с двухконтроллерным режимом может существенно ограничить его применимость в продуктивных средах. Надеюсь, в

современной версии данная проблема решена.

Второе главное неудобство данной СХД в полном отсутствии функционала так называемых пулов хранения данных. СХД позволяет создавать только RAID-группы. Пусть и большие (до 64 дисков на группу), но привязанные к одному контроллеру. Поэтому необходимо либо собирать тома в пулы на серверах, либо использовать Storage DRS в среде виртуализации.

Третья сложность при работе с RAIDIX заключается в том, что наличие внешней дисковой полки является обязательным условием функционирования двухконтроллерной системы. RAIDIX сам не производит хранилища, а только разрабатывает ПО, возможно эта ситуация изменится после поглощения компанией YADRO. Производителей дисковых полок в современных условиях сложно найти на рынке, а компания YADRO до сих пор не предлагает таких устройств как класса.

Гиперконвергентные системы и ПО СХД

Для заказчиков, которые планировали купить несколько серверов, коммутатор FC и небольшую двухконтроллерную СХД, самым разумным подходом будет построение гиперконвергентной инфраструктуры, основанной на серверах с установленными дисками.

Эталоном качества в данном классе являются конкурирующие решения от VMware и Nutanix (к сожалению, недоступные сейчас нашему рынку). Но отечественные производители в большинстве своём нашли решение в использовании свободного ПО СХД, такого как GlusterFC или CEPH. Особняком среди них можно выделить компании Росплатформа и Киберпротект с собственной сборкой проприетарной СХД - Virtuozzo Storage.

Таким образом, классифицировать в разрезе СХД современные Российские гиперконвергентные системы можно следующим образом:

1. **Virtuozzo Storage** в вариантах:
 1. Р-хранилище и СКАЛА-Р;
 2. Кибер Инфраструктура (бывш. Акронис);
2. **GlusterFS** - присутствует в большинстве отечественных средств серверной виртуализации:
 1. ECP Veil;
 2. (z/o/red/rosa/host)Virt и другие неизвестные Российские сборки oVirt.
3. **CEPH** - самое функциональное, но и самое сложное свободное ПО СХД, представлено в решениях:
 1. Брест Виртуализация - описывает сценарии использованием CEPH, но данная СХД не встроена в него.
 2. Альт Виртуализация - в режиме PVE обладает отличной интеграцией с CEPH.

Конвергентные решения можно рекомендовать для небольших и средних установок. Главная проблема, которую придется решить, - отсутствие на Российском рынке высокопроизводительных коммутаторов Ethernet с клиентскими портами быстрее 10Гбит. Самое важное для программно-определяемых сред - это наличие высокоскоростных каналов с низкой задержкой.

Например, компания Росплатформа рекомендует в рамках одного кластера СХД ограничиться серверами одной стойки из-за сложности построения высокопроизводительной сети между двумя стойками. Но эти проблемы несложно решить при должном архитектурном подходе,

например, заложив межстоечный СКС в проект.

Тип: Midrange

Данный тип массивов самый обширный и объединяет в себе большинство дисковых массивов общего назначения. Гибридные и Unified системы. Если мы говорим об импортных системах, яркие представители класса это: HPE 3par (Primera / Aletra); IBM Storwize (FlashSystem) 5000, 7000; Hitachi VSP G <900; DELL EMC Unity; Netapp FAS; Huawei OceanStore;

Характерной чертой данных систем является наличие развитого функционала, практически не уступающего системам уровня HiEnd. Следующие функции являются стандартом для таких систем, и заказчики ожидают их увидеть и в отечественных решениях:

- Tiered Storage - многоуровневое хранение а также FlashCache;
- Дедупликация, Компрессия и уплотнение;
- Синхронная и асинхронная репликация;
- Мгновенные снимки, не снижающие производительность;
- Метрокластер;
- “Гибкий RAID” из чанков а не дисков:
 - восстановление диска “много ко многим”,
 - область горячей замены вместо выделенного диска,
 - возможность сочетания нескольких RAID на одном наборе дисков.

Известные отечественные решения которые можно предложить в данный проект включают в себя два решения:

- СХД Аэродиск Engine / Восток;
- СХД YADRO Tatlin Unified;

Вкратце о каждой системе ниже.

Aerodisk



Производитель предлагает систему хранения в двух версиях: сертифицированная(Восток) и обычная (Engine).

Версия Восток отличается наличием станции администрирования и мониторинга, выполненной на процессоре Эльбрус. Хотя на сайте производителя заявлено наличие СХД полностью на контроллерах Эльбрус, в продаже я такую систему не встречал. Возможно, такой вариант предлагается заказчикам, которые не могут в силу объективных причин использовать более высокопроизводительные контроллеры на процессорах Intel.

Система по праву является самой функциональной из отечественных, при этом не все функции, к сожалению, сочетаются между собой. СХД предлагает пользователю

два подхода управления дисками - RDG, назовём их Raid группы и DDP, назовём их пулы. В зависимости от использования того или иного подхода доступны различные функции. Например, пул DDP поддерживает гибкое распределение ёмкости, но не поддерживает файловый доступ.

Среди функционала систем данного класса Аэродиск закрывает практически все требования, кроме наличия Tiered Storage, подробнее об этом можно прочитать в статье [Описание технологий повышения производительности СХД AERODISK](#).

Трудности, с которыми можно столкнуться при подготовке решения на данной системе, и пути их разрешения приведены ниже:

- Отсутствие Tiered Storage. Данный функционал скорее всего уже никогда не будет реализован, потому что все современные тенденции идут в сторону AllFlash и разделения быстрых СХД и медленных. Остаётся предложить заказчику апгрейд до AllFlash-системы или использование SSD-кэша. Отличие кэширования от многоуровневого хранения только в одном: при кэшировании ёмкость кэша **не добавляется** к пулу, при многоуровневом хранении - **добавляется**. Когда происходит кэширование данные **копируются**, при многоуровневом хранении - **перемещаются**. Таким образом, можно предложить заказчику увеличить ёмкость пула медленными дисками для достижения результата, аналогичного кэшированию. Но стоит отметить, что Аэродиск не рекомендует использовать больше 2 накопителей в кэше на пул.
- Низкая и даже недостаточная производительность по мнению вендора пулов и групп полностью на HDD. Либо это защита от дурака на этапе составления спецификации, либо система крайне плохо работает с HDD, но технический отдел вендора отказывается считать конфигурации без SSD-кэша. Необходимо это иметь ввиду. Так как, возможно, будет выгоднее предлагать AllFlash-конфигурации вместо небольших гибридов.
- Метрокластер и репликация, к сожалению, работают только по протоколу Ethernet. Стоит иметь ввиду, что протокол Ethernet - скорее всего будущее отечественных сетей хранения данных, потому что единственные две компании производящие коммутаторы Fibre Channel, полностью прекратили работу в РФ. Поэтому стоит рассматривать покупку системы хранения данных в первую очередь для сетей Ethernet, а FC устанавливать для обеспечения совместимости с наследуемой инфраструктурой.
- Дедупликация достаточно примитивная, а компрессия не позволяет достичь высоких результатов. Данный функционал не следует брать в расчёт при подготовке решения.

Во всём остальном это неплохая, достаточно функциональная система уровня Midrange, которая отстала от импортных аналогов всего на 5-7 лет. При цене, которую назначают на данную систему, стоит сразу рассматривать AllFlash-конфигурации, обходя стороной небольшие (до 60 дисков) гибриды.

Тип: AFF / HighPerformance

Системы хранения данных, построенные только на твердотельных накопителях (AllFlashSystems), практически полностью вытеснили гибридные СХД в сегменте основного

хранилища предприятия и являются новой базовой единицей построения ИТ-инфраструктуры современных компаний.

Исторически появление такого типа систем можно разделить на два пути:

1. **Усиление гибридных СХД до возможности работы с большим количеством твердотельных дисков.** Среди зарубежных аналогов можно привести: Hitachi VSP F; HPE Zpar 7450, 8450; Netapp AFF; Netapp EF; DELL EMC Unity *F. В основном в целях маркетинга данные системы не поддерживают вращающиеся накопители, что проводит четкую грань между гибридными СХД и AllFlash. У каждой такой СХД как правило есть гибридная альтернатива, поддерживающая накопители всех типов? Функционал данных систем по большей части схож и может отличаться богатством наполнения базовых лицензий, например флеш-версии обычно поставляются по принципу “всё включено” и специфичными функциями, которые востребованы именно на твердотельных системах ввиду высокой стоимости накопителей: виртуальный RAID, компрессия, дедупликация.
2. **Полностью созданные с нуля работающие только с Flash системы.** Данные СХД по большей части являются ИТ-стартапами, которые смогли завоевать свою популярность на волне удешевления твердотельных накопителей. Примерами послужат СХД PureStorage, RAMSAN (IBM) FlashSystem, Violin Memory, Nimble ¹⁾. Большие ИТ-гиганты немного позже, но уже включились в данный класс со своими собственными разработками, такими как: Hitachi VSP E, Dell PowerStore, Huawei OceanStor Dorado. Для систем данной группы характерны уникальные, современные подходы, не работающие для вращающихся накопителей, наличие только базового функционала необходимого для работы с Flash-дисками.

Системы, созданные специально для твердотельных дисков, часто не обладают даже базовым функционалом, который для традиционных СХД является стандартом, например:

- **Только один контроллер обрабатывает нагрузку ввода-вывода.** Например, СХД Nibmle работает по принципу Active/Passive, а PureStorage использует второй простаивающий контроллер только для ускорения дедупликации;
- **Функционал метро кластера мало где реализован,** так как для метрокластера является допустимой задержка путешествия пакета порядка 10мс, что совершенно недопустимо для AllFlash-систем;
- **Уровень RAID часто можно изменить только при инициализации системы** без возможности что-либо изменить без полной потери данных;
- **Возможность создания многоуровневых хранилищ не реализована** в СХД такого типа, в виду того, что SSD разного объема работают со сравнимой скоростью и смысла перемещать данные между ними нет.

Следующие функции наоборот являются практически неотъемлемой частью СХД на флеш-дисках:

- **Обслуживание ячеек флеш-памяти на уровне СХД.** В отличие от отдельных SSD или гибридных СХД системы AllFlash можно представить как один большой SSD-диск. Операции очистки “грязных ячеек”, дефрагментации и балансировки износа в таких системах производятся централизованно и не вызывают “Write Cliff” ²⁾
- **Дедупликация и компрессия** ввиду того, что стоимость и производительность таких накопителей очень высока, функционал сжатия позволяет за счет высокой (часто избыточной) скорости дисков увеличить полезную ёмкость, сократив разрыв в стоимости владения системой по сравнению с традиционными СХД.
- **Сетевые протоколы подключения клиентов с поддержкой NVMe.** Данные

протоколы часто применяются как и в самих накопителях, так и для подключения клиентов к СХД. NVMe позволяет достичь более высокого времени отклика и низких задержек, при этом существенно снижает нагрузку на процессор клиентской системы и работает более оптимально при высоких нагрузках.

- **Мгновенные снимки без снижения производительности** также позволяют существенно повысить полезную емкость системы при грамотном использовании. А отсутствие жестких дисков видимо существенно упрощает реализацию данного механизма.
- **Простой интерфейс управления** стал визитной карточкой AllFlash-систем хранения “новой волны”, таким образом производители позиционируют массивы как привлекательные, не сложные в администрировании системы с высокой производительностью, которой хватит практически для любых задач.

Вернемся к отечественным решениям. Что могут предложить наши производители среди систем данного типа?

1. Простых и привлекательных систем среди отечественных разработок нет и в скором времени ожидать их не стоит.
2. Казалось бы, большой простор для развития в классе AllFlash-стартапов, но в данном направлении наработки есть, видимо, только у компании АРГО с разрабатываемой системой “Power IO”. Данная СХД только выходит на рынок про нее пока нечего сказать.
3. Системы Аэродиск и Ядро могут использоваться в формате “только флеш”, но отдельной как-то оптимизированной (пусть даже и маркетинговым способом) версии для работы без жестких дисков, у этих производителей нет, поэтому выделять их в данный класс не стоит.
4. Raidix предлагает продукт программного RAID-массива для NVMe накопителей. Это не готовая СХД, а скорее механизм RAID для локальных дисков.

Подводя промежуточный итог по данному типу систем хранения данных, можно сделать вывод, что самые передовые системы - типа AllFlash в России пока недоступны как класс. При необходимости высокопроизводительного доступа заказчику придется ограничить свой выбор системами общего назначения в старших конфигурациях. Функционал дедупликации и компрессии представлен среди популярных решений только у Аэродиска, но коэффициенты сжатия достаточно скромные, а нагрузка на систему при использовании очень высокая, механизмы требуют доработки.

Тип: Disk Backup

Следующий тип систем хранения данных - дисковые массивы, предназначенные для хранения резервных копий. Данные системы на мировом рынке представлены продуктами линейки HPE StoreOnce, DELL EMC DataDomain, Quantum DXi, Infinidat InfiniGuard, Falconstor. Назначение такого типа систем - хранение резервных копий, поэтому основными характеристиками данного типа продуктов являются:

- Низкая стоимость хранения данных, например, за счет:
 - Встроенной дедупликации;
 - Выгрузки неиспользуемых данных на внешнее S3 хранилище, например в публичное облако;
- Интеграция с наследуемыми ПО резервного копирования по открытым протоколам:
 - SMB, NFS;

- эмуляция ленточной библиотеки;
- Высокая скорость копирования и восстановления за счет:
 - Технологий распределенной дедупликации на стороне клиента системы;
 - Естественной многопоточности дискового массива и низкого времени поиска данных;
- Высокая доступность за счет двухконтроллерных конфигураций и встроенных механизмов репликации данных;

По аналогии с рынком AllFlash-систем в данном сегменте наметился рост популярности стартапов, предлагающих выполнять резервное копирование по-новому. Главным локомотивом прогресса в данном направлении является компания Rubrik, предлагающая ПАК для резервного копирования и восстановления. В состав комплекса входит хранилище с функционалом резервного копирования, таким образом приобретение ПО РК дополнительно не требуется. Основными преимуществами обновленного подхода к резервному копированию озвучиваются:

- Простота развертывания, расширения и администрирования;
- Высокая скорость восстановления.

Среди отечественных решений нет аналогов StoreOnce или DataDomain как и альтернативы Rubrik. Основное решение, которое можно использовать для хранения резервных копий, это сборки на базе ПО СХД RAIDIX, которое поддерживает популярные файловые протоколы. На [странице посвященной резервному копированию RAIDIX](#) не приведено никаких уникальных преимуществ присущих именно системе хранения резервных копий. Все тезисы - основные функциональные особенности RAIDIX. На мой взгляд, преимуществами RAIDIX перед массивами общего назначения будут являться:

- **Низкая стоимость владения** за счет возможности использования общедоступного оборудования и одноконтроллерных конфигураций. Не всегда хранилищу резервных копий требуется отказоустойчивость - это можно использовать для существенного сокращения затрат на построение инфраструктуры. Поддержка общедоступных компонентов позволяет избежать "Vendor Lock" и производить расширение и ремонт системы собственными силами.
- **Высокая производительность последовательной записи** - неотъемлемое преимущество системы на базе ПО RAIDIX, позволит сократить время резервного копирования.
- **Поддержка кластерных файловых систем** позволяет создать сложное распределенное хранилище на базе нескольких СХД.

Также стоит упомянуть про отечественного производителя [СХД BAUM](#). Данные системы не широко популярны, но все еще может стремительно измениться. Особенностью данной СХД является вариант с предустановленным сервером управления резервным копированием - Acronis CyberProtect. Я склонен полагать, что это больше маркетинговый ход, позволивший занять пустующую нишу "Backup Appliance" за счет базового функционала запуска контейнеров или виртуальных машин на контроллерах СХД.

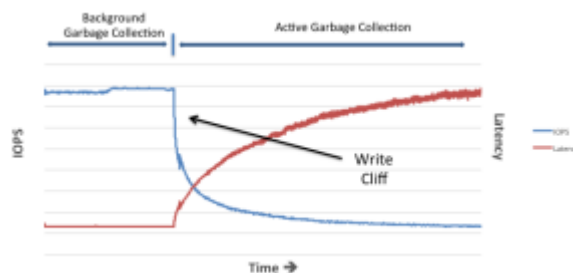
В качестве запасного плана практически во всех типах хранилищ я привожу системы общего назначения, так и здесь в определенных конфигурациях заказчик может размещать данные на СХД Аэродиск или Ядро Татлин. Но такой проект стоит рассматривать только тогда, когда ёмкости или производительности RAIDIX недостаточно, так как данные системы обладают значительно большей стоимостью, а СХД YADRO не имеет "дешевых начальных конфигураций".

1)

Данная СХД немного выбивается из классификации, так как поддерживает и гибридные варианты. Но без SSD работать не может, так как это всё же AllFlash-массив с поддержкой HDD, а не наоборот

2)

Write Cliff - название эффекта резкого падения скорости записи SSD-дисков и СХД, связанное с тем, что чистые ячейки для записи новых данных закончились, и системе требуется провести дефрагментацию в первоочередном порядке. На графике записи данный эффект приобретает форму обрыва.



From:

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

Permanent link:

https://micronode.ru/domestic/alternative_to/storageLast update: **2025/02/20 14:41**