WHITE PAPER

ЛУЧШИЕ ПРАКТИКИ ПО РАБОТЕ С TATLIN.FLEX V.5.2.6

О WHITE PAPERS

Данный документ описывает архитектурные принципы и подходы, а также рекомендации по их применению при построении корпоративных решений на основе продуктов производства компании YADRO.

Документ предназначен для широкой аудитории, но также может включать материалы, требующие специальной профессиональной подготовки, ориентированные на системных архитекторов, администраторов, сервисных инженеров и других IT-специалистов, участвующих в планировании, проектировании и внедрении комплексных решений.

Представленная информация основана на результатах исследований, а также экспертизе специалистов YADRO и носит рекомендательный характер.
Конкретные результаты и эффективность конечных решений зависят от множества факторов, которые могут различаться в зависимости от особенностей инфраструктуры и условий эксплуатации.

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ 4
ОБЗОР ПРИНЦИПОВ РАБОТЫ TATLIN.FLEX 5
РЕКОМЕНДАЦИИ ПО РАБОТЕ С ДИСКОВОЙ ПОДСИСТЕМОЙ И RAID 6
- ОБЩИЕ РЕКОМЕНДАЦИИ 6
- РЕКОМЕНДАЦИИ ПО ИСПОЛЬЗОВАНИЮ GENERIC RAID 6
- РЕКОМЕНДАЦИИ ПО ИСПОЛЬЗОВАНИЮ ERA RAID 6
- ИСПОЛЬЗОВАНИЕ SSD-КЕШИРОВАНИЯ 6
- НАБОРЫ РЕЗЕРВНЫХ ДИСКОВЫХ НАКОПИТЕЛЕЙ (SPARE) 7
- РЕКОМЕНДАЦИИ ПО РАЗМЕРУ RAID 7

БЛОЧНЫЙ ДОСТУП 7
- КОММУТАЦИЯ И ПОДКЛЮЧЕНИЕ СХД К ХОСТАМ 7
- РЕКОМЕНДАЦИИ ДЛЯ БЛОЧНОГО ДОСТУПА ПО ISCSI (ETHERNET) 8
- РЕКОМЕНДАЦИИ ДЛЯ ПРЯМОГО ПОДКЛЮЧЕНИЯ ХОСТОВ 8
- РЕКОМЕНДАЦИИ ПО СОЗДАНИЮ LUN 9
- ОСОБЕННОСТИ МАСКИРОВАНИЯ 9

ФАЙЛОВЫЙ ДОСТУП 9
ДОКУМЕНТАЦИЯ 10

ВВЕДЕНИЕ

Данный документ предназначен для широкого круга читателей: специалистов по предпродаже, сервисных инженеров, системных администраторов, инженеров по сопровождению СХД и др. В документе описаны базовые принципы работы СХД семейства TATLIN.FLEX и общие рекомендации по сценариям их использования. Они применимы как на этапе проектирования инфраструктуры и подготовки к закупке СХД, так и на этапе внедрения.

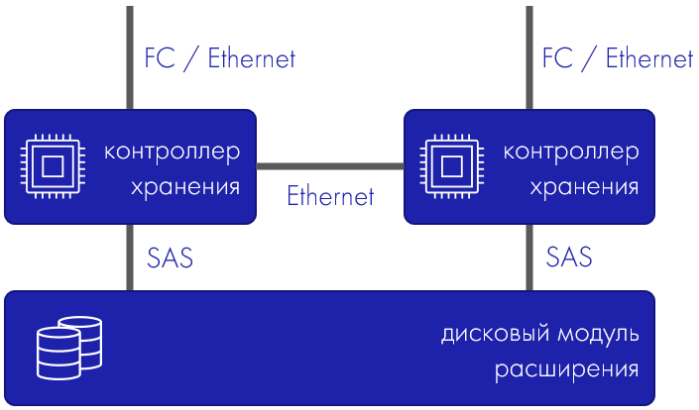
Данный документ не заменяет документацию по продукту, ознакомление с ней также необходимо. Список документов о продуктах семейства TATLIN.FLEX приведен в разделе Документация.

ВЕРСИИ ИСПОЛЬЗУЕМОГО ПО

ПО	ВЕРСИЯ
TATLIN.FLEX	5.2.6

ОБЗОР ПРИНЦИПОВ РАБОТЫ TATLIN.FLEX

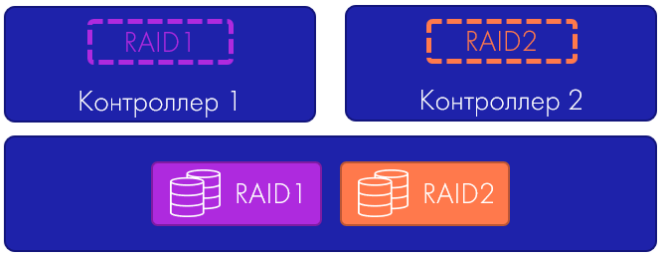
Модели TATLIN.FLEX.PRO и TATLIN.FLEX.TWIN представляют собой классические двухконтроллерные СХД. Они состоят из двух независимых контроллеров хранения фиксированной конфигурации, объединенных высокоскоростным отказоустойчивым соединением Ethernet (для синхронизации кеш-памяти и других взаимодействий) и подключенных к общему набору дисковых или твердотельных накопителей по шине SAS.



TATLIN.FLEX.ONE оснащен одним контроллером хранения и имеет аналогичное устройство.

Дисковые накопители системы объединяются в один или несколько RAID. На основе RAID создаются логические тома, которые могут быть напрямую презентованы хостам по блочным протоколам (FC, iSCSI) или использованы для создания файловой системы. На основе файловой системы создаются общие папки, доступ к которым осуществляется по файловым протоколам (SMB, NFS, FTP).

Каждый RAID принадлежит одному контроллеру хранения. Такой контроллер обозначается как «предпочтительный» для данного RAID. Все логические тома каждого RAID обслуживаются его предпочтительным контроллером хранения.



Режим работы контроллеров хранения в зависимости от типа доступа:

- Доступ к общим папкам (NAS) осуществляется с предпочтительного контроллера хранения;
- Доступ к блочным томам осуществляется с обоих контроллеров:
  - по умолчанию блочный том доступен с предпочтительного контроллера в состоянии Active и с непредпочтительного — в состоянии Standby;
  - при активации режима `lun spare_path` тома доступны в состоянии Active с обоих контроллеров хранения, но с предпочтительного контроллера приоритет доступа выше.

В случае отказа контроллера хранения или отказа путей от контроллера к общим дискам активные на контроллере RAID-ы, затронутые отказом, активируются на соседнем контроллере. При необходимости (например, для балансировки нагрузки) смену предпочтительного контроллера хранения можно произвести вручную с помощью команды «Миграция RAID».

РЕКОМЕНДАЦИИ ПО РАБОТЕ С ДИСКОВОЙ ПОДСИСТЕМОЙ И RAID

ОБЩИЕ РЕКОМЕНДАЦИИ

- Используйте настройки RAID по умолчанию. Изменение настроек по умолчанию допускается либо по указанию сервисной службы, либо по результатам проведенного тестирования, которое выявило положительное влияние альтернативных настроек на работу в специфических условиях конкретного окружения.
- Создавайте несколько RAID и распределяйте их по разным контроллерам хранения (при достаточном количестве накопителей) — это поможет сбалансировать нагрузку на систему.
- В рамках одного RAID необходимо использовать накопители одного типа и емкости.
- Не рекомендуется использовать RAID 5 для накопителей емкостью более 4 ТБ.
- Не используйте RAID 0 для данных, потеря которых недопустима.

РЕКОМЕНДАЦИИ ПО ИСПОЛЬЗОВАНИЮ GENERIC RAID

- Используйте инициализированные RAID во всех случаях.

> **!** Инициализация может занимать несколько десятков часов (около 17 часов для накопителей объемом 10 ТБ). Во время инициализации Generic RAID не доступен для работы.

- СХД семейства TATLIN.FLEX не имеют встроенных батарей, поэтому для защиты кеша оперативной памяти (RAM cache) рекомендуется обеспечить надежное электропитание СХД. В качестве дополнительной меры возможна установка и настройка ИБП.

РЕКОМЕНДАЦИИ ПО ИСПОЛЬЗОВАНИЮ ERA RAID

- В работе ERA RAID не применяется кеш оперативной памяти. Использование ERA RAID с дисковыми накопителями (HDD) не рекомендуется из-за снижения производительности.
- Размер блока логических томов, создаваемых на ERA RAID, не может быть меньше размера блока самого RAID. На RAID с блоком 512 байт можно создать тома с блоком 512 и 4096 байт; на RAID с блоком 4096 байт — тома с блоком 4096 байт.

ИСПОЛЬЗОВАНИЕ SSD-КЕШИРОВАНИЯ

Использование SSD-кеширования может оказать положительное влияние на производительность. Результат зависит от характера подаваемой на СХД нагрузки: локальности, регулярности и др.

При планировании использования SSD-кеширования рекомендуется учитывать следующие особенности и правила:

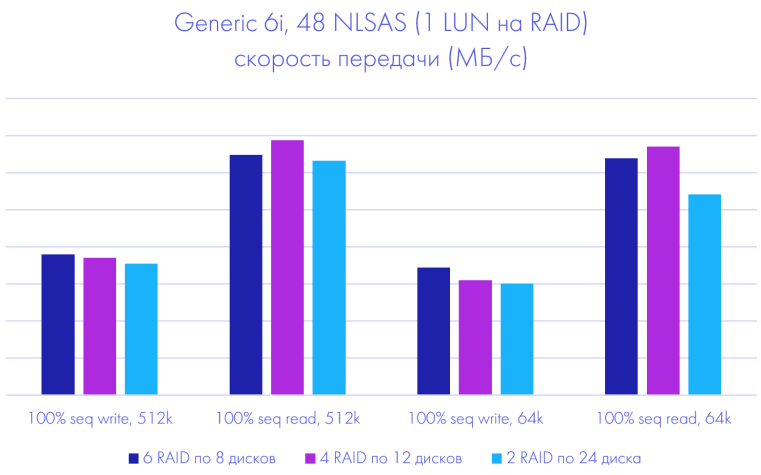
- SSD-кеш является одной из функций Generic RAID, поэтому и том с данными, и кеширующий том должны располагаться на RAID типа Generic.
- SSD-кеш подключается только к томам, расположенным на дисковых накопителях; подключение SSD-кеша к томам твердотельных накопителей не поддерживается.
- Алгоритм SSD-кеширования обрабатывает блоки до 32 КБ включительно, блоки большего размера игнорируются.
- SSD-кеширование требует значительного объема оперативной памяти контроллера хранения (до 54 ГиБ на каждый 1 ТиБ кеширующего тома).
- Кеширующие тома должны быть распределены по обоим контроллерам хранения в соответствии с распределением RAID кеширующих томов.
- К применению рекомендуется конфигурация из четырех SSD-накопителей, на которых создается два независимых RAID1 для использования в кеширующих томах.

НАБОРЫ РЕЗЕРВНЫХ ДИСКОВЫХ НАКОПИТЕЛЕЙ (SPARE)

- Резервные дисковые накопители не имеют предпочтительного контроллера хранения и могут быть использованы для всех RAID (в том числе с разными схемами защиты) в состоянии Active на любом из двух контроллеров.
- Рекомендуется иметь резервные дисковые накопители для каждого используемого в системе типа накопителей.
- Рекомендуемое количество резервных дисковых накопителей:
  - по одному резервному на каждые 30 дисковых накопителей с данными (для накопителей любого типа емкостью до 12 ТБ);
  - по одному резервному на каждые 20 дисковых накопителей (для накопителей любого типа емкостью от 15 ТБ).

РЕКОМЕНДАЦИИ ПО РАЗМЕРУ RAID

При прочих равных условиях и при настройках по умолчанию RAID небольшого и среднего размера показывают более высокую производительность, чем большие RAID.

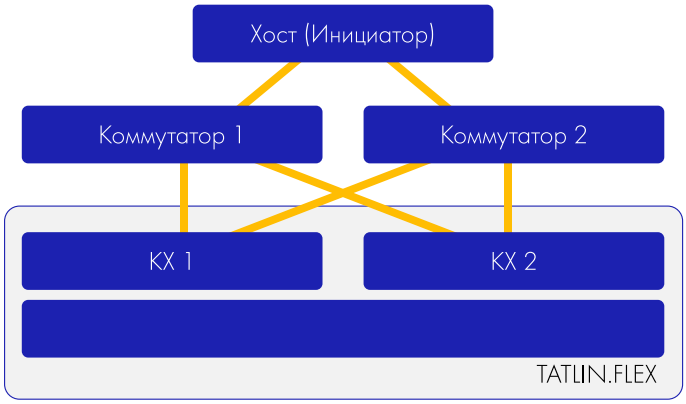


БЛОЧНЫЙ ДОСТУП

КОММУТАЦИЯ И ПОДКЛЮЧЕНИЕ СХД К ХОСТАМ

Использование отказоустойчивой сети хранения данных поможет повысить доступность системы как в штатных (например, плановое обслуживание), так и в нештатных ситуациях (например, сбои):

- Рекомендуется иметь две независимые FC-фабрики или два коммутатора Ethernet;
- Каждый контроллер хранения TATLIN.FLEX должен подключаться к каждой из фабрик или к обоим коммутаторам Ethernet.

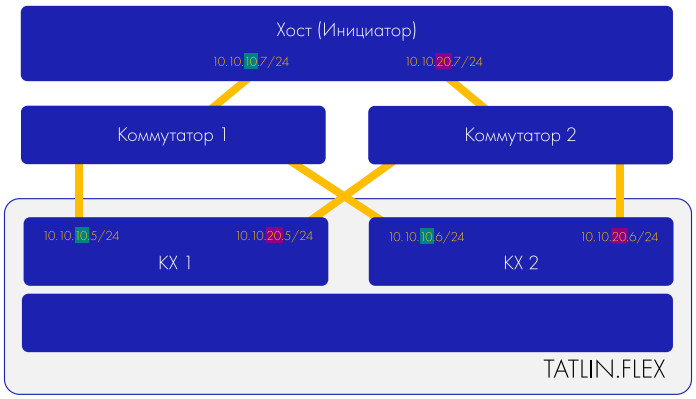


РЕКОМЕНДАЦИИ ДЛЯ БЛОЧНОГО ДОСТУПА ПО ISCSI (ETHERNET)

- Рекомендуется подключать СХД и хост к одному коммутатору. Использование сетей маршрутизации между СХД и хостом не рекомендуется.
- Необходимо использовать многопутевой ввод-вывод (multipathing), а не объединение интерфейсов (bonding).

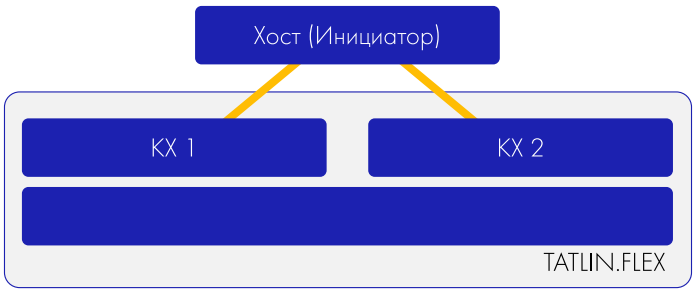
> **!** Рекомендации по параметрам многопутевого ввода-вывода приведены в документе «Настройка ОС и платформ виртуализации».

- Увеличьте стандартный MTU до 9000. Настройка должна быть произведена на всем пути: на интерфейсах TATLIN.FLEX, коммутаторах, интерфейсах хоста, виртуальных коммутаторах, платформах виртуализации.
- Допустимо использовать несколько сессий на каждом подключении для повышения производительности.
- Все сетевые интерфейсы iSCSI на контроллере хранения должны находиться в разных подсетях.
  На схеме ниже представлена корректная настройка адресации.



РЕКОМЕНДАЦИИ ДЛЯ ПРЯМОГО ПОДКЛЮЧЕНИЯ ХОСТОВ

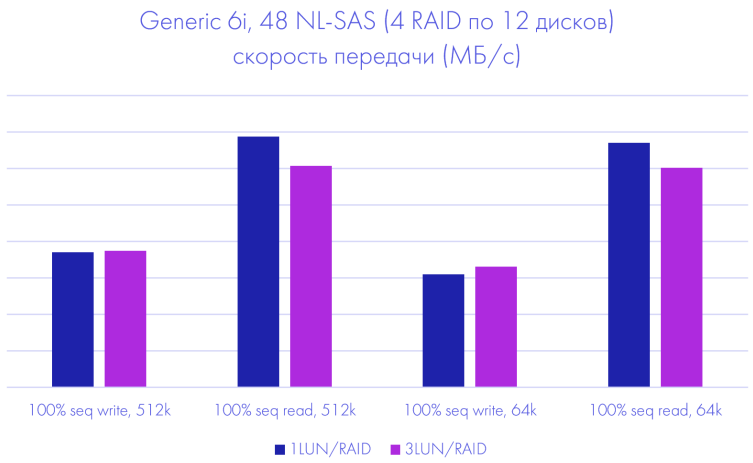
TATLIN.FLEX поддерживает возможность прямого подключения хостов без использования коммутатора при доступе по FC и Ethernet. Типичный сценарий использования прямого подключения — построение небольших сред виртуализации из трех или четырех хостов. При прямом подключении чаще всего каждый хост подключается к каждому контроллеру хранения одним соединением.



> **!** Выход из строя карты ввода-вывода целиком или её отдельного порта не приводит к автоматическому переводу RAID на другой контроллер хранения. Может потребоваться ручное вмешательство для восстановления доступа к данным. Для избежания таких ситуаций в инфраструктурах с прямым подключением хостов допустимо разрешение доступа к блочным томам по резервным путям (spare path).

РЕКОМЕНДАЦИИ ПО СОЗДАНИЮ LUN

- Размер блока логического тома следует выбирать в соответствии с требованиями со стороны ОС или платформы виртуализации, которая будет использована для работы с логическими томами. При отсутствии явных требований рекомендуется использовать наиболее широко совместимый блок размером 512 байт.
- При использовании накопителей NL-SAS (например, в задачах хранения архивов, резервных копий, видеоархивов и при работе с медиа) рекомендуется создание небольшого количества больших логических томов (по одному LUN на каждый RAID). Использование большого количества томов не гарантирует рост производительности и в некоторых случаях ведет к ее снижению. При использовании Generic RAID не рекомендуется создавать более 60 LUN в одной системе.



ОСОБЕННОСТИ МАСКИРОВАНИЯ

- После создания логические тома по умолчанию не маскируются на таргетах. Данная настройка контролируется параметром `luns_visible_by_default` (см. Справочник CLI).
- При маскировании тома на таргетах он становится доступен всем хостам, имеющим доступ к таргету и не имеющим явно заданных настроек доступа. После настройки доступа к хосту или логическому тому доступ в режиме по умолчанию, когда все тома доступны всем хостам, прекращается.
- Каждый FC-порт является самостоятельным таргетом, поэтому настраивать доступ необходимо на каждом из портов или воспользоваться командой «Маскировать LUN на всех таргетах». Для двухконтроллерных СХД TATLIN.FLEX.TWIN и TATLIN.FLEX.PRO необходимо настроить доступ к логическим томам для таргетов на одном контроллере, затем переключиться на второй контроллер и повторить процедуру настройки или воспользоваться функцией «Таргет удаленного узла».

ФАЙЛОВЫЙ ДОСТУП

- Необходимо организовать отказоустойчивое подключение СХД к сети передачи данных (см. раздел Блочный доступ).
- Для обеспечения отказоустойчивости необходимо использовать объединение интерфейсов (bonding), так как файловые протоколы не поддерживают многопутевой доступ (multipathing).
- Если TATLIN.FLEX подключается к коммутаторам YADRO KORNFELD D1156, следует придерживаться следующих правил:
  - все интерфейсы, участвующие в объединении (bond), должны быть подключены к одному и тому же коммутатору;
  - при использовании LACP на стороне коммутатора необходимо выбирать политику 802.3ad на стороне СХД;
  - при использовании статического LAG на стороне коммутатора необходимо выбирать политику Balance-rr или Balance-xor на стороне СХД.
- Рекомендуется настроить виртуальный IP-адрес (VIP) для сохранения доступности сетевых ресурсов при выходе контроллера хранения из строя:
  - настройку необходимо производить на каждом из контроллеров по отдельности;
  - VIP должен быть активен на том же контроллере, где активен RAID, к общим папкам которого предоставляется доступ.

Не рекомендуется использовать файловый доступ для транзакционных нагрузок: СУБД, виртуальных машин и других задач, требующих непрерывной доступности.

Перевод общих папок между контроллерами хранения может занимать время, превышающее типовые ограничения по времени выполнения операции в подобных средах. После перевода может потребоваться повторное подключение хостов к общей папке.

ДОКУМЕНТАЦИЯ

Перед началом работы с продуктом настоятельно рекомендуется ознакомиться с актуальным набором документации, доступным на сервисном портале YADRO.

Документация для ПО TATLIN.FLEX (единая для всех СХД семейства TATLIN.FLEX):

- **Руководство администратора** — основной документ по работе с СХД. **Обязательно к прочтению.**
- **Характеристики продукта** — поддерживаемые лимиты для актуальной версии ПО СХД, рекомендуется не выходить за них при работе с продуктом.
- **Инструкция по обновлению** — последовательность шагов, необходимых для обновления до актуальной версии ПО СХД.
- **Настройка ОС и платформ виртуализации** — рекомендации по подключению TATLIN.FLEX к различным хостам.
- **Руководство по мониторингу** — настройка интеграции TATLIN.FLEX с системами мониторинга (Zabbix, Syslog, SNMP).
- **Справочник CLI** — описание управления TATLIN.FLEX с помощью командной строки и утилиты rdcli.

Документация, специфичная для конкретной модели СХД TATLIN.FLEX:

- **Руководство по началу работы** — установка, запуск и первоначальная настройка TATLIN.FLEX. **Обязательно к прочтению.**
- **Руководство владельца** — описание и устройство СХД.
- **Руководство по каблированию** — схемы каблирования для СХД.
- **Диагностика и устранение неисправностей** — описание сервисных процедур, доступных для выполнения заказчиком.