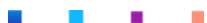




Техническое описание и рекомендации по конфигурации

АЭРОДИСК ENGINE AQ Лайт



Оглавление

Аппаратная платформа СХД ENGINE AQ Лайт	3
Описание функциональности СХД ENGINE AQ Лайт	4
Пул хранения RDG	4
Блочный и файловый доступ	4
Мгновенные снимки и клоны	5
Локальная репликация	5
Тонкое хранилище	6
SSD-кэш для ускорения операции чтения-записи на шпиндельные диски	6
Компрессия	7
Дедупликация	7
Качество обслуживания (QoS)	8
Перестроение RAID групп	8
Миграция LUN	8
Функция автоподдержки	9
Спецификация аппаратных систем, доступных для заказа	10
Рекомендации по логической конфигурации	13
Техническое сопровождение	14

Аппаратная платформа СХД ENGINE AQ Лайт

Система хранения данных (далее СХД) ENGINE AQ Лайт построена на базе аппаратной платформы Аквариус FS440 и оснащена двумя контроллерами. Количество носителей для хранения данных ограничено 24 универсальными отсеками для накопителей 3.5"/2.5" интерфейса NL-SAS\SAS с поддержкой «горячей замены» на передней панели контроллерного шасси. Расширение подсистемы хранения с помощью подключения внешних дополнительных полок не поддерживается.

Для внешней коммутации в базовой конфигурации, стандартно, предусмотрены четыре порта Ethernet 25 Гбит/с с установленными оптическими трансиверами SFP+ 10 Гбит/с – по два порта на каждый контроллер. Опционально поддерживается установка до двух портов 16 Гбит/с Fibre Channel на каждый контроллер системы. Также возможна модернизация с установкой в порты Ethernet оптических трансиверов SFP28 с поддержкой скорости 25 Гбит/с. Таблица с партномерами для заказа этих опций расширения приведена ниже:

AQL08FC	Комплект модификации FC: [АННВ-F16GS2] Front-End адаптер FibreChannel 16G, 2 port – 2 шт. [AQL08FCLIC] Простая (неисключительная) лицензия на ENGINE AQ Лайт, расширение: функционал FC на два контроллера – 1 шт.
B4SFP25	Комплект трансиверов: [AATM-D2LCSR] Трансивер Аэродиск SFP28 для 25GbE, SR – 4 шт.

В системе используются нижеописанные функции управляющего программного обеспечения АЭРОДИСК ВОСТОК, а именно:

- пул хранения RDG;
- блочный и файловый доступ;
- мгновенные снимки и клоны, которые можно делать по расписанию (локальная репликация);
- тонкие тома;
- использование твердотельных накопителей для кэширования данных (только для гибридных конфигураций);
- компрессия и дедупликация;
- возможность ограничения томов по производительности (Quality of service, QoS).

СХД ENGINE AQ Лайт не имеет поддержки функционала удаленной репликации на другую аппаратную систему и метрокластер.

Описание функциональности СХД ENGINE AQ Лайт

Пул хранения RDG

В СХД ENGINE AQ Лайт для организации хранения данных используется метод RAID Distributed Group (RDG), основными особенностями которого являются:

- пул хранения RDG состоит из виртуальных устройств, каждое из которых имеет заданную структуру RAID одного из поддерживаемых уровней – 1/10, 5/50, 6/60, 6P/60P («P» - тройная четность);
- в пуле хранения RDG поддерживаются и файловый, и блочный доступы;
- виртуальные устройства последовательно объединяются в одну виртуальную группу RDG, за счет чего количество дисков в группах (и для данных, и для четности) не ограничено;
- вне зависимости от объема тома или файловой системы все диски в группе участвуют в вводе-выводе для данного тома или файловой системы;
- диски горячей замены являются глобальными;
- для ускорения операций записи на шпиндельные диски используется SSD кэш;
- для ускорения операции чтения на шпиндельные диски есть возможность дополнительно использовать твердотельные накопители (функционал online-tiering);
- миграция логических томов (LUN) «на лету»;
- мгновенные снимки, снапклоны и связанные клоны;
- скорость перестроения RAID можно регулировать политикой перестроения;
- размер пула хранения RDG может быть динамично увеличен (т. е. в онлайн режиме) с помощью добавления новой RDG-группы с ранее используемой в пуле хранения структурой RAID.

Блочный и файловый доступ

Блочный доступ обеспечивается путем предоставления блочного устройства (логического тома, LUN) конечному серверу обработки или серверам обработки по протоколу iSCSI с поддержкой iSER и опционально по протоколу Fibre Channel (требуется установка опциональных Front-End адаптеров Fibre Channel). Файловый доступ обеспечивается путем предоставления доступа к файловой системе по протоколам NFS и SMB (CIFS) конечному серверу обработки или серверам обработки. Для SMB (CIFS) может использоваться авторизация пользователей с помощью Active Directory. В рамках одной RDG группы могут функционировать как логические тома, так и файловые системы. Файловый и блочный доступ можно обеспечивать с одного и того же контроллера, достаточно лишь наличия соответствующих Front-End адаптеров Fibre Channel, Ethernet.

Мгновенные снимки и клоны

Мгновенные снимки (*снэпшоты*) и связанные клоны в пуле хранения RDG используют модель переадресации при записи (Redirect-on-Write). Система всегда пишет новые блоки данных в новое место, переставляя на них указатель, а старые блоки данных, на которые уже нет указателя, никогда не стираются, а помечаются системой как освобожденные. Этот механизм позволяет создавать любое количество мгновенных снимков и связанных клонов без какого-либо влияния на производительность. Мгновенные снимки создаются мгновенно и изначально не потребляют дисковое пространство; используемое пространство растёт по мере изменения данных. Связанные клоны создаются мгновенно и сразу могут быть доступны серверу обработки на чтение/запись путем презентации (маппинга).

Снэпклон – это гибрид клона и мгновенного снимка. Снэпклоны создаются быстрее, чем классические клоны и изначально занимают ровно ту полезную емкость, которую занимает источник. Снэпклоны и связанные клоны можно подключить к серверу обработки в виде отдельных LUN или файловых систем. Восстановление данных из мгновенных снимков и связанных клонов можно выполнить двумя способами:

1. Откатить мгновенный снимок/связанный клон, полностью перезаписав данные LUN/файловой системы. Это удобно, когда нужно быстро восстановить LUN/файловую систему полностью.
2. Присоединить связанный клон в виде отдельного LUN/файловой системы к серверу обработки и восстановить данные с этого LUN/файловой системы. Такой способ подходит для ситуаций, когда не нужно восстанавливать LUN/файловую систему целиком, а нужно восстановить только некоторые объекты (файлы).

СХД ENGINE AQ Лайт не имеет ограничений по количеству созданных мгновенных снимков и снэпклонов, за исключением физического ограничения используемого оборудования.

Локальная репликация

Полезной функцией является локальная репликация – создание/удаление мгновенных снимков по расписанию. Это полезно, например, если требуется сохранять резервные копии данных очень часто, что невозможно сделать внешними системами резервного копирования, т.к. в силу их специфики (высокая нагрузка на каналы, долгое время записи и прочее) они резервируют данные обычно не чаще чем раз в сутки. В этом случае есть возможность настроить расписание мгновенных снимков. Например, каждые 3 часа в течение суток со сроком хранения одни сутки. Через сутки мгновенные снимки начнут перезаписываться заново, а данные старше суток уже будут сохранены внешней системой резервного копирования.

Тонкое хранилище

В пуле хранения RDG поддерживается создание тонких томов, что позволяет делать переподписку реально занимаемой дисковой ёмкости. Администратор системы обязан внимательно следить за физически доступным объемом в пуле хранения, чтобы физически занятый объем не достиг предельно возможного. Во избежание переполнения пула хранения рекомендуется не превышать 90% заполняемости пула хранения RDG.

SSD-кэш для ускорения операции чтения-записи на шпиндельные диски

Для реализации максимальной производительности и гибкости в СХД ENGINE AQ Лайт предусмотрена функция ускорения ввода/вывода для шпиндельных дисков. Данная функция разделяется на три вида:

1. SSD-кэш чтения-записи (SSD-RW), для последовательных нагрузок (бэкап, видеонаблюдение) – два накопителя в RAID1 и один накопитель горячей замены.
2. SSD-кэш чтения-записи и хранение метаданных (SSD-RW + MCACHE), для случайных нагрузок (виртуализация, БД) – два накопителя в RAID1 и один накопитель горячей замены.
3. Online-tiering (SSD online-tiering) – дополнительный кэш на чтение большого размера, накопителя в RAID1/10, для ускорения чтение данных, к которым идут частые обращения.

При создании гибридного хранилища твердотельные накопители SSD добавляются в кэш на запись/чтение в RAID1 – минимум два накопителя. SSD-кэш применяется для всех транзакций. При этом, чтобы исключить переполнение кэша, применяется механизм циклической адаптации (выталкивания) записей из кэша. SSD-RW-кэш является достаточно экономичным вариантом повышения производительности, поскольку не требует дисков большого объема (за счет постоянного выталкивания транзакций). При этом, поскольку данный механизм активно использует запись, это утилизирует ресурс надежности SSD-дисков (DWPD) и для данного типа кэша используются SSD-диски с высоким показателем DWPD равным 3. СХД ENGINE AQ Лайт не имеет ограничения по объему SSD кэша, за исключением физического ограничения используемого оборудования.

Функционал online-tiering позволяет перемещать блоки данных между различными уровнями в зависимости от нагрузки на них, позволяя тем самым размещать более «горячие» данные (т.е. часто используемые) на быстрых твердотельных накопителях, а более «холодные» данные (т.е. редко используемые) на медленных шпиндельных дисках. Перемещение блоков данных между уровнями происходит в онлайн-режиме. Диски добавляются на уровне всей RDG

группы; после добавления дисков в online-tier группа меняет статус на «Быстрый». Минимальное количество дисков на уровень online-tier – 2. В отличие от механизма SSD-кэширования данный функционал хранит данные на SSD-дисках, пока к ним есть обращения. Online-tiering рекомендуется использовать при нагрузке, носящей случайный характер и имеющей ярко выраженные «горячие» области, например, для индексов баз данных. Если в нагрузке преобладают операции записи или нагрузка носит последовательный характер, использование online-tiering не принесет результата.

Возможно одновременное использование технологий SSD-кэш и online-tiering для увеличения производительности и на операциях случайного чтения, и на операциях записи.

Компрессия

СХД ENGINE AQ Лайт поддерживает механизм компрессии транзакции для экономии места. Компрессия транзакций работает в онлайн режиме, то есть данные записываются на носители хранения уже в оптимизированном виде. Так как система оптимизирует размер хранимых данных еще до записи на носители хранения, то в ряде случаев включение этой функции может увеличить количество операций ввода/вывода, так как физических записей/чтений на носители хранения становится меньше. Для выполнения компрессии транзакций «на лету» используются выделенные ядра процессора и процесс компрессии никогда не конкурирует за ресурсы. В случае если количество операций ввода/вывода велико и ресурсов выделенного ядра процессора перестает хватать, система «на лету» автоматически выделяет под процесс компрессии дополнительные выделенные ядра процессора.

Дедупликация

СХД ENGINE AQ Лайт поддерживает функцию дедупликации. Дедупликация это процесс устранения дублей блоков данных при сохранении уникальных блоков для экономии дискового пространства. Применяется онлайн-дедупликация с фиксированным блоком. Она работает и на твердотельных накопителях, и на шпиндельных дисках, и её можно включать на конкретный логический том (LUN). Дедупликация делит входящие данные на равные блоки и устраняет дубли только тогда, когда блоки на 100% совпадают. Процесс дедупликации происходит следующим образом:

- определение данных для дедупликации;
- проверка доступности необходимого объема кэш-памяти (на твердотельных накопителях или ОЗУ);
- наборы данных сохраняются в таблице дедупликации при сохранении их контрольных сумм;

- при создании дубля данных система вместо выделения нового дискового пространства под дубль добавляет в таблицу дедупликации ссылку, которая указывает на реально существующие данные вместо того, чтобы записывать их дубли.

В зависимости от сферы применения и характера записи дедупликация может снизить потребляемый объем дискового пространства. Дедупликация выполняется на блочном уровне, что особенно применимо для больших объемов похожих данных. Например, при дедупликации хранилища виртуальных машин (VM) уникальными, как правило, являются только некоторые блоки данных, а идентичные данные, такие как гостевые ОС, шаблоны VM, клоны VM и прочие являются дублируемыми и, соответственно, при дедупликации не потребляют дополнительного объема. Рекомендуется использовать дедупликацию при наличии твердотельных накопителей в СХД.

Качество обслуживания (QoS)

Настройка качества обслуживания (QoS) позволяет минимизировать эффект «шумного соседа». При правильной настройке этого параметра можно гарантировать, что все потребители ресурсов СХД ENGINE AQ Лайт будут работать так, как ожидает администратор системы. QoS в СХД ENGINE AQ Лайт назначаются на блочные устройства. Параметры качества обслуживания назначаются на уровне каждого конкретного LUN и могут ограничивать его потребление ресурсов в МБ/с и количестве операций ввода-вывода (IOPs).

Перестроение RAID групп

При выходе из строя накопителя в RAID-группе при наличии соответствующего накопителя для горячей замены автоматически начинается процесс перестроения RAID-группы. При перестроении, как правило, может страдать общая производительность, то есть могут страдать конечные потребители ресурсов СХД. Чтобы минимизировать эффект от перестроения, администратор системы может назначить политику перестроения рейдов, в том числе заданную по расписанию. Есть три варианта политики перестроения, чтобы можно было гранулярно управлять скоростью перестроения рейдов – оптимальная, «производительность», «перестроение».

Миграция LUN

СХД ENGINE AQ Лайт поддерживает функционал миграции блочных устройств «на лету» прозрачно для конечного сервера обработки. Функционал может быть полезен для перемещения данных на другой уровень RAID, например, с RAID5 на RAID10 для увеличения

производительности. Также данные могут быть перемещены на другой тип дисков, с NL-SAS дисков на твердотельные накопители. Возможны любые направления перемещения блочных устройств между типами дисков и типами рейдов.

Функция автоподдержки

Для обеспечения максимальной доступности в СХД ENGINE AQ Лайт предусмотрена функция автоматической поддержки. Данная опция обеспечивает постоянный проактивный мониторинг всех компонентов СХД. В случае сбоя происходит автоматическая отправка необходимой диагностической информации в службу технической поддержки АЭРОДИСК для автоматического открытия обращения (тикета) на устранение выявленного сбоя. Открытие обращений производится путем отправки диагностической информации в виде почтовых уведомлений от контроллеров СХД на серверы технической поддержки АЭРОДИСК. После прихода данной информации сообщения автоматически преобразуются в тикет и регистрируются. Далее специалист службы технической поддержки АЭРОДИСК, имея необходимую входную информацию, приступает к работе по устранению сбоя.

Спецификация аппаратных систем доступных для заказа

Фиксированные аппаратные конфигурации и опции расширения подсистемы хранения, доступные для заказа к ним, приведены в таблице ниже.

Гибридные системы ENGINE AQ Лайт	
AQL08	<p>[ДТЛВ.466533.020] ENGINE AQ Лайт 128GB/4x10Gb/3x800GB SSD + 11xHDD 8TB</p> <p><u>Опции расширения подсистемы хранения:</u></p> <p>B5H08: Комплект жестких дисков 5x8TB - [AHHD-S37K08] Жесткий диск 8TB 3.5" NL-SAS 7200 RPM – 5 шт.</p> <p>B5S19: Комплект твердотельных накопителей 5x1.92TB - [AHSD-S21S19] Твердотельный накопитель SSD 1.92TB 1DWPD 2.5" SAS3.0 3D TLC – 5 шт.</p> <p>B5S38: Комплект твердотельных накопителей 5x3.84TB - [AHSD-S21S38] Твердотельный накопитель SSD 3.84TB 1DWPD 2.5" SAS3.0 3D TLC – 5 шт.</p> <p><u>Полезный объем:</u></p> <p>11xHDD 8 ТБ - RAID6 8+2: 52.4 ТиБ; 16xHDD 8 ТБ - RAID6 12+2: 78.6 ТиБ; 21xHDD 8 ТБ - 2xRAID6 8+2: 105 ТиБ.</p>

<p>AQL18</p>	<p>[ДТЛВ.466533.020] ENGINE AQ Лайт 128GB/4x10Gb/3x800GB SSD + 11xHDD 18TB</p> <p><u>Опции расширения подсистемы хранения:</u></p> <p>B5H18: Комплект жестких дисков 5x18TB - [AHHD-S37K18] Жесткий диск 18TB 3.5" NL-SAS 7200 RPM – 5 шт.</p> <p>B5S19: Комплект твердотельных накопителей 5x1.92TB - [AHSD-S21S19] Твердотельный накопитель SSD 1.92TB 1DWPD 2.5" SAS3.0 3D TLC – 5 шт.</p> <p>B5S38: Комплект твердотельных накопителей 5x3.84TB - [AHSD-S21S38] Твердотельный накопитель SSD 3.84TB 1DWPD 2.5" SAS3.0 3D TLC – 5 шт.</p> <p><u>Полезный объем:</u></p> <p>11xHDD 18 ТБ - RAID6 8+2: 117.9 ТиБ; 16xHDD 18 ТБ - RAID6 12+2: 176.9 ТиБ; 21xHDD 18ТБ - 2xRAID6 8+2: 235.9 ТиБ.</p>
	<p>All Flash системы ENGINE AQ Лайт</p>
<p>AQL19</p>	<p>[ДТЛВ.466533.020] ENGINE AQ Лайт 128GB/4x10Gb/12x1.92TB SSD</p> <p><u>Опция расширения подсистемы хранения:</u></p> <p>B4S19: Комплект твердотельных накопителей 4x1.92TB - [AHSD-S21S19] Твердотельный накопитель SSD 1.92TB 1DWPD 2.5" SAS3.0 3D TLC – 4 шт.</p> <p><u>Полезный объем:</u></p> <p>12xSSD 1.92 ТБ - RAID10: 7.86 ТиБ, RAID6: 14.1 ТиБ; 16xSSD 1.92 ТБ - RAID10: 11 ТиБ, RAID6: 20.44 ТиБ; 20xSSD 1.92 ТБ – RAID10: 14.15 ТиБ, RAID6: 26.73 ТиБ; 24xSSD 1.92 ТБ – RAID10: 17.3 ТиБ, RAID6: 31.45 ТиБ.</p>
<p>AQL38</p>	<p>[ДТЛВ.466533.020] ENGINE AQ Лайт 128GB/4x10Gb/12x3.84TB SSD</p>

Опция расширения подсистемы хранения:

B4S38:

Комплект твердотельных накопителей 4x3.84TB - [AHSD-S21S38]

Твердотельный накопитель SSD 3.84TB 1DWPД 2.5" SAS3.0 3D TLC – 4 шт.

Полезный объем:

12xSSD 3.84 ТБ - RAID10: 15.72 ТиБ, RAID6: 28.3 ТиБ;

16xSSD 3.84 ТБ - RAID10: 22 ТиБ, RAID6: 40.88 ТиБ;

20xSSD 3.84 ТБ – RAID10: 28.3 ТиБ, RAID6: 53.46 ТиБ;

24xSSD 3.84 ТБ – RAID10: 34.6 ТиБ, RAID6: 62.9 ТиБ.

Рекомендации по логической конфигурации

Для гибридной системы ENGINE AQ Лайт создаётся единый пул хранения из всех шпиндельных дисков NL-SAS, к которому добавляется SSD-кэш или SSD-кэш + MCACHE в составе двух твердотельных накопителей, собранных в RAID1 (третий твердотельный накопитель в системе установлен для горячего резерва). В системе явно не ограничено количество шпиндельных дисков в RAID-группе, но использование более 14 шпиндельных дисков в одной RAID-группе из дисков большого размера не рекомендуется из-за длительного перестроения в случае выхода из строя шпиндельного диска. Для отказоустойчивости требуется наличие не менее одного шпиндельного диска для горячей замены. Для RDG группы из шпиндельных дисков NL-SAS большого размера рекомендуется использовать уровень RAID6. Оптимальной по производительности формулой является 8+2. Для минимальной конфигурации из 11 шпиндельных дисков рекомендуется создать одну RAID-группу структурой RAID6 по формуле 8+2 и оставить один шпиндельный диск для горячей замены. Полезный объем в таком случае будет равен объему восьми шпиндельных дисков.

При эксплуатации рекомендуется не заполнять систему более чем на 90% из-за деградации параметров производительности. В системах с интенсивной записью для сохранения задержек ввода-вывода (IO latency) на приемлемом уровне рекомендуется не заполнять группу из шпиндельных дисков более чем на 80%. Опционально можно расширить подсистему хранения, добавив один или два комплекта жестких дисков по пять штук в каждом. В системе с 16 шпиндельными дисками рекомендуемой конфигурацией является одна RAID-группа с формулой 12+2 и два диска горячей замены. В системе с 21 шпиндельными дисками рекомендуемой конфигурацией является две RAID-группы с формулой 8+2 и один диск горячей замены.

Опционально в гибридную систему ENGINE AQ Лайт можно добавить дополнительные твердотельные накопители, которые могут использоваться для создания SSD online-tiering или второго пула хранения из твердотельных накопителей. Одновременное использование технологий SSD-кэш и online-tiering увеличивает производительности RDG-группы из шпиндельных дисков и на операциях случайного чтения, и на операциях записи.

Для All Flash системы ENGINE AQ Лайт и второго пула хранения RDG из твердотельных накопителей для гибридной системы ENGINE AQ Лайт рекомендуется использовать уровень RAID10 или RAID5. Оптимальной по производительности формулой для RAID5 является 4+1 или 8+1. Для отказоустойчивости требуется наличие не менее одного твердотельного накопителя для горячей замены.

При создании логического тома на RDG группе можно выбрать размер блока тома – 4К, 16К, 32К, 64К, 128К. Размер блока тома нужно выбирать максимально близким к размеру блока, которым работает конечное приложение, данные которого размещаются на соответствующем томе. Соответствие размера блока приложения размеру блока тома особенно важно при случайном характере нагрузки и для RDG группы из шпиндельных дисков. Для получения максимальной производительности для группы из шпиндельных дисков рекомендуется создавать не менее четырёх томов на группу, а для группы из твердотельных накопителей создавать не менее восьми томов на группу.

Для приложений, создающих интенсивную нагрузку ввода/вывода, рекомендуется разделять её на несколько логических томов. Например, для высоконагруженных баз данных рекомендуется выделять отдельные тома для логов, индексов и размещать файлы базы данных на нескольких логических томах. Для систем видеонаблюдения с большим количеством камер рекомендуется писать видеопотоки с разных групп камер на разные логические тома и не размещать базу каталога системы видеонаблюдения на группе из шпиндельных дисков, выделяя под неё отдельные тома на группе из твердотельных накопителей.

Техническое сопровождение

Поставка СХД ENGINE AQ Лайт осуществляется с уровнем технического сопровождения СТАНДАРТ Плюс сроком на один год. Имеется опциональная возможность увеличения срока сопровождения до трёх лет. Таблица с партномерами для заказа приведена ниже:

APWS-BSN8531	Сертификат на увеличение срока технического сопровождения до 3х лет для AQL08
APWS-BSN8532	Сертификат на увеличение срока технического сопровождения до 3х лет для AQL18
APWS-BSN8533	Сертификат на увеличение срока технического сопровождения до 3х лет для AQL19
APWS-BSN8534	Сертификат на увеличение срока технического сопровождения до 3х лет для AQL38

По отдельному запросу возможна поставка системы с уровнем технического сопровождения БИЗНЕС на один или три года.