

Описание функциональных характеристик Arenadata QuickMarts T4 (ADQM T4)

Содержание:

1	Аппаратные требования к серверам	3
1.1.	ClickHouse-сервер	3
1.2.	Zookeeper-сервер	4
2	Аппаратные требования к сети	5
3	Программные требования	7
3.1	Операционная система	7
3.2	Необходимая информация для установки кластера	7
3.3	Особенности установки ADQM	8

1 Аппаратные требования к серверам

Системные требования, приведенные ниже, являются минимальными. Целевой сайзинг необходимо рассчитывать исходя из требований вашей организации.

1.1. ClickHouse-сервер

Минимальные аппаратные требования к ClickHouse-серверам для установки кластера ADQM следующие:

- Физический сервер.
- Хранение: отформатированное в ext4 устройство, примонтированное к файловой системе.
- CPU: от 8 ядер.
- RAM: от 64 ГБ.
- Сеть: сервера должны быть доступны по портам 9000, 8123, 9009 внутри кластера.

Для быстрых витрин (quick marts) подходит массив из SSD-дисков. Для хранения больших таблиц фактов (big tables with facts) предпочтителен RAID 10 из HDD. Рекомендуется использовать программный RAID в Linux (mdadm). Не рекомендуется использовать LVM. При создании RAID 10 необходимо выбрать far-расположение.

На каждом ClickHouse-сервере внутри кластера схема устройства, диски и точка монтирования должны быть одинаковыми.

Примеры требований к ClickHouse-серверам в зависимости от профиля нагрузки кластера ADQM показаны ниже.

- Proof of concept:

Количество узлов	Хранение	CPU	RAM	Сеть
1-4	1 x 500 ГБ HDD для ОС 4 x 1.2 ТБ SAS HDD RAID 10 для данных	8 ядер	96 ГБ	1 x 10 Гбит/с

- Quick marts:

Количество узлов	Хранение	CPU	RAM	Сеть

Количество узлов	Хранение	CPU	RAM	Сеть
4	2 x 500 ГБ SSD RAID 1 для ОС 12 x 1 ТБ SAS12G SSD MixedUse RAID 10 (PCIe Gen3 x8) для данных	24 ядра (2 x 12)	256 ГБ (8 x 32)	4 x 25 Гбит/с (2 адаптера PCIe Gen4 x4)

- Big tables with facts:

Количество узлов	Хранение	CPU	RAM	Сеть
4	2 x 500 ГБ SSD RAID 1 для ОС	2 x 500 ГБ SSD RAID 1 для ОС	2 x 500 ГБ SSD RAID 1 для ОС	2 x 500 ГБ SSD RAID 1 для ОС

1.2. Zookeeper-сервер

Минимальные аппаратные требования к серверам с Zookeeper кластера ADQM следующие:

- Физический или виртуальный сервер. Для пилотных инсталляций можно использовать те же сервера, что и для ClickHouse. При больших инсталляциях — минимум три железных сервера.
- Хранение: от 100 ГБ; для больших кластеров предпочтительны быстрый SSD-диск или RAID 1.
- CPU: от 4 ядер.
- RAM: от 4 ГБ.

В таблице ниже представлен пример конфигурации Zookeeper-сервера.

Количество узлов	Хранение	CPU	RAM	Сеть
3 и более	2 x 500 ГБ SSD RAID 1 для ОС 2 x 2.4 ТБ SSD RAID 1 для данных	4 ядра (P-cores)	8 ГБ	2 x 10 Гбит/с

Для обеспечения HA (High Availability) необходимо, чтобы количество экземпляров Zookeeper-серверов было нечетным.

2 Аппаратные требования к сети

- Требуется одна или несколько выделенных и обособленных одноранговых Interconnect-сетей под цели внутренней коммуникации кластера.
- К Interconnect-сетям должны быть подключены все серверы кластера.
- Все серверы должны иметь полные доменные имена (FQDN) и возможность обмениваться данными друг с другом на заданных скоростях.
- Должен быть известен IP-адрес каждого сервера.
- Должна быть обеспечена доступность Ping любого из серверов (ICMP).
- В случае инсталляции ADQM с доступом в Интернет требуется доступ к развернутому на выделенном сервере Arenadata Cluster Manager.
- В случае инсталляции ADQM без доступа в Интернет (offline) требуется доступ к развернутому на выделенном сервере Arenadata Enterprise Tools.
- На серверах, используемых для установки кластера, должны быть открыты порты, приведенные ниже.

Каждый порт относится к серверу, на который устанавливается конкретный сервис.

Сервис	Порт	Назначение
Общие	22	Порт для подключения к SSH-провайдеру
	81	TCP-порт Arenadata Enterprise Tools (доступ к репозиториям в случае offline-инсталляции)
	8000	TCP-порт для отправки статусов компонентов кластера в ADCM
	2015	TCP-порт для отправки метрик на сервер мониторинга
	2016	UDP-порт для отправки метрик на сервер мониторинга
	80	TCP-порт доступа к web-интерфейсу Graphite
	3000	HTTP-порт доступа к web-интерфейсу Grafana
ClickHouse	8123	HTTP-порт для клиентских или прокси-соединений
	8443	Порт для защищенных клиентских соединений через интерфейс HTTPS

Сервис	Порт	Назначение
	9000	Порт для взаимодействия по протоколу Native (TCP)
	9440	Порт для SSL-защищенного взаимодействия по протоколу Native (TCP)
	9009	Порт для обмена данными между репликами
	9010	Порт для SSL/TLS-защищенного межсерверного взаимодействия
ZooKeeper	2181	Порт доступа к сервису
	2888 3888	Порты межсерверного взаимодействия для компонентов кворума
ClickHouse Keeper	9444	Рекомендуемый порт для соединений <u>RAFT</u> (также используется для безопасного обмена данными, если включена настройка <secure>1</secure>)
	9281	Рекомендуемый порт для SSL-защищенного взаимодействия
JDBC Bridge	9019	Порт JDBC Bridge
Chproxy	9090	HTTP-порт для подключения клиентов

3 Программные требования

3.1 Операционная система

1. ADQM поддерживает только архитектуру x86_64 и следующие операционные системы:
 - Centos 7
 - RHEL 7
 - Alt Linux 8.4 SP (начиная с версии ADQM 22.3.2.2.1.b1)
 - Astra Linux SE 1.7 Орел с Axiom JDK (для Enterprise-версии ADQM, начиная с 22.8.12.45.1.b1)
2. Необходимо подготовить все зависимые репозитории дополнительно к выбранной операционной системе.
3. Со всех машин в кластере должен быть доступ к официальным репозиториям (или созданы локальные зеркала):
 - CentOS Extras
 - CentOS Updates
 - CentOS Base
4. В случае инсталляции ADQM с доступом в Интернет машины должны иметь доступ к репозиториям Arenadata, перечисленным на вкладке Configuration в меню кластера ADQM web-интерфейса ADCM. Для offline-установки ADQM необходим доступ к хосту, на котором развернут кластер Enterprise Tools.

3.2 Необходимая информация для установки кластера

Необходимая информация и условия для установки кластера ADQM:

- IP-адреса и hostnames всех серверов в кластере, сопоставленные с их ролями.
- Доступ к пользователю root или любому пользователю с правами sudo на каждом сервере внутри кластера.
- Правильно работающий DNS. Все hostname и FQDN должны одинаково резолвиться на всех серверах внутри кластера.

3.3 Особенности установки ADQM

1. В процессе установки модифицируются файлы, которые не должны контролироваться системой управления конфигурацией (при ее наличии):
 - системный файл `/etc/hosts`;
 - системный файл `/etc/selinux/config`;
 - системный файл `/etc/sysctl.conf`.
2. В процессе установки создаются файлы в `/usr/lib/systemd/system/`, которые не должны контролироваться системой управления конфигурацией (при ее наличии).
3. Сервисы, устанавливаемые в кластере, не должны контролироваться системой управления конфигурацией (при ее наличии).