Arenadata™ Database

Версия - v5.8.0-arenadata1

Примечания к выпуску ADB-5.8.0-arenadata1

2019 Arenadata™

Оглавление

1 Новые возможности **3**
- 1.1 Поддержка параллельной записи во внешние JDBC-совместимые СУБД . . . 3
- 1.2 Поддержка проброса условий предикатов при обращении к внешним JDBC-совместимым источникам . . . 4
- 1.3 Поддержка типа компрессии Zstandard (ZSTD) . . . 4
- 1.4 Улучшение оптимизатора GPORCA . . . 4
- 1.5 Limit Slices к запросу . . . 5
- 1.6 Resource Group Global Shared Memory . . . 5
- 1.7 Управление ресурсами PL/Container . . . 5
- 1.8 Проверка подлинности SSL-сертификата . . . 5

2 Измененные функции **7**

3 Экспериментальные функции **8**

4 Ключевые отличия ADB от open-source аналога **9**

5 Поддерживаемые платформы **10**
- 5.1 Поддерживаемые платформы . . . 11

6 Совместимость дистрибутивов Hadoop **12**
- 6.1 Совместимость дистрибутива Hadoop с помощью PXF . . . 12
- 6.2 Совместимость дистрибутива Hadoop с помощью gphdfs . . . 12

7 Обновление до ADB 5.8.0_arenadata1 **13**
- 7.1 Подготовка к обновлению . . . 13
- 7.2 Обновление ADB с 5.x до 5.8.0_arenadata1 . . . 14
- 7.3 Устранение неполадок при неудачном обновлении . . . 15

8 Решенные проблемы **16**

9 Известные проблемы и ограничения **19**
- 9.1 Обновление gp_toolkit.gp_bloat_diag . . . 22

Arenadata DB (ADB) – это сервер базы данных с массовой параллельной обработкой (MPP), который поддерживает хранилище данных следующего поколения и крупномасштабную аналитическую обработку. Благодаря автоматическому секционированию данных и выполнению параллельных запросов кластер может работать как один суперкомпьютер базы данных, работающий в десятки или сотни раз быстрее, чем традиционная база данных. ADB поддерживает SQL, параллельную обработку MapReduce и объемы данных от сотен гигабайт до сотен терабайт.

Important: Данная документация содержит информацию о выпуске Arenadata DB 5.8.0_arenadata1. Для предыдущих версий продукта необходимо перейти к соответствующей документации

ADB 5.8.0_arenadata1 – это первый выпуск Arenadata 5.8.0_arenadata1 с добавлением и изменением некоторых функций.

Выпуск устраняет проблему часовых поясов, в документации разъясняется поведение базы данных ADB при выборе часового пояса и при ненастроенной конфигурации. Arenadata рекомендует настраивать фиксированное значение часового пояса, совместимое с ADB и хост-системой. Это предотвращает пересчет часового пояса “наилучшего соответствия” при каждом повторном запуске кластера. При ненастроенном параметре часового пояса после обновления до версии 5.8.0_arenadata1 необходимо убедиться, что выбранный часовой пояс является приемлемым для развертывания.

Important: Контактная информация службы поддержки – e-mail: info@arenadata.io

Глава 1

Новые возможности

ADB 5.8.0_arenadata1 включает в себя следующие новые функции:

- *Поддержка параллельной записи во внешние JDBC-совместимые СУБД*
- *Поддержка проброса условий предикатов при обращении к внешним JDBC-совместимым источникам*
- *Поддержка типа компрессии Zstandard (ZSTD)*
- *Улучшение оптимизатора GPORCA*
- *Limit Slices к запросу*
- *Resource Group Global Shared Memory*
- *Управление ресурсами PL/Container*
- *Проверка подлинности SSL-сертификата*

1.1 Поддержка параллельной записи во внешние JDBC-совместимые СУБД

С помощью плагина **PXF** возможно записывать данные во внешние СУБД, используя пользовательский JDBC-драйвер, предварительно загруженный в память фреймворка **PXF**. Запись выполняется параллельно каждым сегментом СУБД **ADB**. Поддерживается агрегация данных для записи с помощью параметра *BATCH_SIZE*.

Пример создания внешней таблицы, предназначенной для параллельной записи в СУБД **Oracle DB**:

```
create writable external table oracle_target_write (
id1 int
, text1 text
)
LOCATION ('pxf://ora_schema.ora_table?PROFILE=JDBC&JDBC_DRIVER=oracle.jdbc.driver.OracleDriver&DB_
↪URL=jdbc:oracle:thin:@//ora_host:1521/XE&USER=ora_user&PASS= ora_password&BATCH_SIZE=50000')
FORMAT 'CUSTOM' (FORMATTER='pxfwritable_export')
distributed by (id1);
```

1.2 Поддержка проброса условий предикатов при обращении к внешним JDBC-совместимым источникам

В версии **ADB 5.8.0_arenadata1** при использовании в сложном запросе внешней таблицы, использующей **PXF JDBC** адаптер для чтения данных из внешней системы, **ADB** автоматически определяет возможность передачи предикатов во внешнюю систему и при наличии такой возможности передает их. Это позволяет существенно снизить объем данных, передающихся по сети, при использовании **PXF JDBC**. Опция включается параметром *enable_filter_pushdown= on*.

1.3 Поддержка типа компрессии Zstandard (ZSTD)

Zstandard обеспечивает более высокую производительность и меньшую загрузку CPU при аналогичном **ZLIB** уровне сжатия (Рис.1.1.).

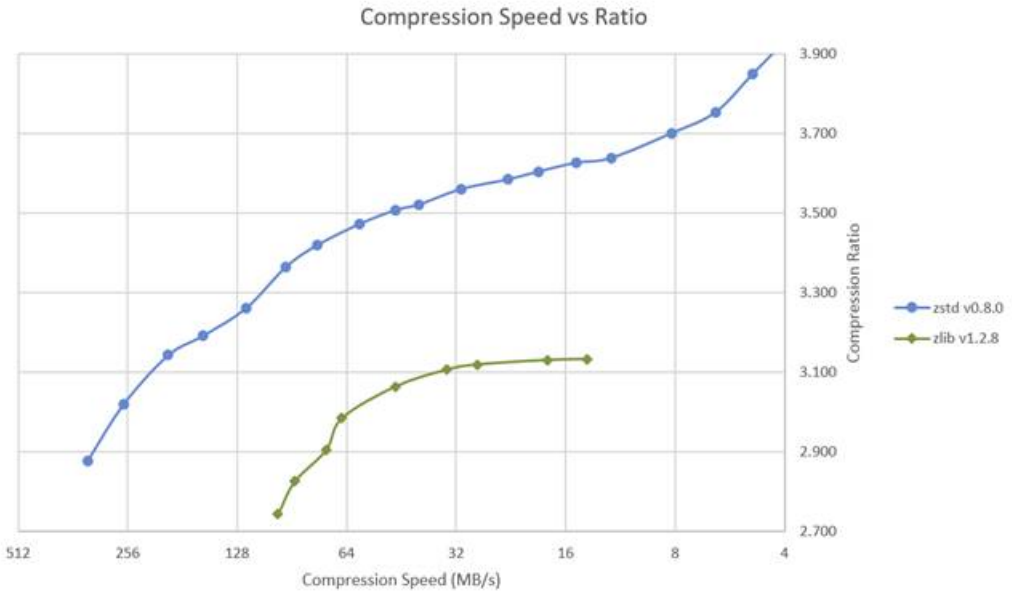


Рис.1.1.: Compression Speed vs Ratio

Zstandard позволяет выбрать уровень сжатия от *1* до *9*.

Пример использования компрессии **ZSTD**:

```
CREATE TABLE sales (txn_id int, qty int, date date)
WITH (appendonly=true, compresstype=zstd, compresslevel=5)
DISTRIBUTED BY (txn_id);
```

1.4 Улучшение оптимизатора GPORCA

Улучшена производительность оптимизатора **GPORCA** для запросов, содержащих предикат, который выполняет сравнение с постоянным массивом большого количества элементов, где элементами являются наборы констант. Например, список *IN* предиката содержит список из двух элементов, которые являются наборами из трех констант.

```
... WHERE (c1, c2, c3) IN ((100,'a','c'), (200,'b','d'))
```

В предыдущих версиях производительность **GPORCA** была хуже и потребляла значительную сумму памяти, пытаясь оптимизировать указанный тип запроса.

Кроме того, на основе анализа эффективности оптимизации запросов значение по умолчанию для параметра конфигурации сервера базы данных **ADB** *optimizer_array_expansion_threshold* было изменено на *100*, что повысило эффективность оптимизации.

1.5 Limit Slices к запросу

Параметр конфигурации сервера базы данных **ADB** *gp_max_slices* указывает максимальное количество срезов (части плана запроса, которые выполняются на инстансах сегмента), которые могут быть сгенерированы запросом. Если запрос генерирует больше срезов, чем заданное количество, кластер возвращает ошибку и не выполняет запрос. Значение по умолчанию – *0*.

Выполнение запроса, генерируемого большое количество срезов, может повлиять на производительность базы данных **ADB**. Например, запрос, содержащий операторы *UNION* или *UNION ALL* над несколькими сложными представлениями, может генерировать большое количество срезов. Для просмотра статистики срезов следует в запросе запустить *EXPLAIN ANALYZE*.

Таблица1.1.: Limit Slices к запросу

Диапазон значений	Значение по умолчанию	Классификация
0 - INT_MAX	0	master/session/reload

1.6 Resource Group Global Shared Memory

В настоящее время база данных **ADB** выделяет незанятую память групп ресурсов в глобальный пул общей памяти. С помощью данной функции кластер выделяет память из глобального общего пула групп ресурсов на отдельные запросы в порядке очереди по принципу “first-come first-served”.

1.7 Управление ресурсами PL/Container

Теперь можно использовать группы ресурсов **ADB** для управления памятью и ресурсами процессора для внешних компонентов. С версией **PL/Container 1.2** можно управлять ресурсами процессора и памятью при запуске инстансов **PL/Container**.

Important: PL/Container 1.2 может использовать новые возможности групп ресурсов, представленные в ADB 5.8.0_arenadata1. При PL/Container 1.1 или более ранней версии необходимо использовать *plcontainer runtime-edit* для удаления любых параметров *resource_group_id* из файла конфигурации PL/Container

1.8 Проверка подлинности SSL-сертификата

Когда внешняя таблица **ADB** определяется протоколом безопасности SSL **gpfdists**, параметр конфигурации сервера базы данных *verify_gpfdists_cert* контролирует, включена ли проверка подлинности SSL-сертификата, чтобы либо считывать данные, либо записывать их во внешний источник. По умолчанию используется значение *true* с включенной аутентификацией SSL.

Значение *false* отключает проверку подлинности SSL-сертификата, за исключением случаев, если:

- SSL-сертификаты, которые использует утилита **gpfdist**, являются самоподписанными;

- Имя хоста, содержащееся в сертификате SSL, не соответствует имени хоста, на котором запущена **gpfdist**.

Значение “false” с целью отключения аутентификации можно установить при тестировании связи между внешней таблицей **ADB** и утилитой **gpfdist**, обслуживающей внешние данные.

Important: Отключение проверки подлинности SSL-сертификата представляет угрозу безопасности

Таблица1.2.: Проверка подлинности SSL-сертификата

Диапазон значений	Значение по умолчанию	Классификация
Boolean	true	master/session/reload

Глава 2

Измененные функции

В **ADB 5.8.0_arenadata1** изменены следующие функции:

- Для параметра конфигурации сервера базы данных *optimizer_array_expansion_threshold* значение по умолчанию изменено на *100*. Когда **GPORCA** оптимизирует запрос, данный параметр ограничивает процесс оптимизации на основе количества констант в массиве. В предыдущих версиях значение по умолчанию было *25*;
- **MADlib 1.14** совместим с **ADB 5.8.0_arenadata1**.

Глава 3

Экспериментальные функции

Поскольку база данных **ADB** основана на open-source аналоге, она включает в себя несколько экспериментальных функций, позволяющих заинтересованным разработчикам экспериментировать с их использованием в системах разработки. Обратная связь помогает стимулировать разработку данных функций, и они могут поддерживаться в будущих версиях продукта.

Important: Экспериментальные функции не рекомендуются и не поддерживаются для развертывания в продуктивной среде. Данные функции могут быть изменены или удалены из будущих версий продукта на основе дальнейшего тестирования и обратной связи. Более того, любые функции, которые могут быть видны в open source code, но при этом не описаны в документации к продукту, должны рассматриваться как экспериментальные и неподдерживаемые для использования в продуктивной среде

В **ADB 5.8.0_arenadata1** входят следующие экспериментальные функции:

- Плагин хранения для **gpbackup** и **gprestore**:
  - Плагин хранения S3. Можно указать опцию *–plugin-config* для хранения резервной копии в местоположении Amazon Web Services S3 и восстановить данные из местоположения S3;
  - Плагин API платформы хранилища. Партнеры, клиенты и разработчики OSS (системы поддержки операций) могут разрабатывать плагины для совместного использования при помощи **gpbackup** и **gprestore**;
- **Recursive WITH Queries** (общие выражения таблицы);
- Группы ресурсов остаются экспериментальной функцией только на платформе **SuSE 11** из-за ограниченной функциональности групп в ядре. **SuSE 12** разрешает проблемы с группой **Linux**, которые вызывают ухудшение производительности, когда включены группы ресурсов **ADB**.

Глава 4

Ключевые отличия ADB от open-source аналога

ADB 5.x имеет ряд ключевых отличий от open-source аналога:

- Скрипт упаковки и установки продукта;
- Поддержка сжатия **ZStandard**;
- Поддержка управления базой данных **ADB** с помощью **Zeppelin management tool**;
- Поддержка мониторинга и управления запросами с помощью **Advanced monitoring stack**;
- Поддержка двустороннего обращения к источникам и приёмникам данных через **JDBC**;
- Поддержка федеративных запросов;
- Поддержка передачи предикатов во внешние системы.

Глава 5

Поддерживаемые платформы

ADB 5.8.0_arenadata1 работает на следующих платформах:

- Red Hat Enterprise Linux 64-bit 7.x
- Red Hat Enterprise Linux 64-bit 6.x
- SuSE Linux Enterprise Server 64-bit 12 SP2 и SP3 с версией ядра выше 4.4.73-5
- SuSE Linux Enterprise Server 64-bit 11 SP4
- CentOS 64-bit 7.x
- CentOS 64-bit 6.x

Important: Для поддерживаемых операционных систем Linux база данных ADB поддерживается на системных хостах с использованием процессоров AMD или Intel на основе архитектуры x86-64. Arenadata рекомендует использовать однородный набор аппаратных средств (системных хостов) в системе базы данных ADB

Значительное ухудшение производительности базы данных **ADB** наблюдалось при включении управления нагрузкой на основе групп ресурсов в системах **Red Hat 6.x**, **CentOS 6.x** и **SuSE 11**. Данная проблема вызвана ошибкой ядра **Linux cgroup**. Ошибка ядра исправлена в системах **CentOS 7.x** и **Red Hat 7.x**.

Если используется **Red Hat 6** и производительность с группами ресурсов приемлема, необходимо обновить ядро до версии *2.6.32-696* или выше, чтобы воспользоваться исправлениями для групп. **SuSE 11** не имеет версии ядра, разрешающего эту проблему; группы ресурсов по-прежнему считаются экспериментальной функцией на данной платформе.

ADB на **SuSE 12** поддерживает группы ресурсов для использования в продуктивной среде. **SuSE 12** разрешает проблемы ядра **Linux cgroup**, которые вызывают ухудшение производительности при включенных группах ресурсов в **ADB**.

Для базы данных **ADB**, установленной на **Red Hat Enterprise Linux 7.x** или **CentOS** версии *7.x - 7.3*, проблема с ОС может привести к зависанию **ADB**, на которой выполняется большая рабочая нагрузка. Данная проблема вызвана ошибками ядра **Linux**, которые исправлены в **RHEL 7.3** и **CentOS 7.3**.

База данных **ADB** на системах **SuSE Linux Enterprise** не поддерживает следующие функции:

- Язык PL/Perl;
- Утилиту gpmapreduce;
- Расширение языка PL/Container;
- Платформу Platform Extension Framework (PXF).

ADB 5.8.0_arenadata1 поддерживает следующие версии **Java**:

- *8.xxx*;
- *7.xxx*.

ПО **ADB 5.8.0_arenadata1**, работающее в системах **Linux**, использует **OpenSSL 1.0.2l** (с **FIPS 2.0.16**), **cURL 7.54**, **OpenLDAP 2.4.44** и **Python 2.7.12**. Клиентское ПО базы данных **ADB**, работающее в системах **Windows** и **AIX**, использует **OpenSSL 0.9.8zg**.

ADB 5.8.0_arenadata1 не поддерживает драйвер **ODBC** для **Cognos Analytics V11**. Подключение к ПО **IBM Cognos** с драйвером **ODBC** не поддерживается. База данных **ADB** поддерживает подключение к ПО **IBM Cognos** с помощью драйвера **DataDirect JDBC**.

5.1 Поддерживаемые платформы

Следующие примечания описывают поддержку платформы для **ADB**:

- Для запуска базы данных **ADB** поддерживается только файловая система **XFS**. Все остальные файловые системы не поддерживаются;
- База данных **ADB** поддерживается на всех подуктовых серверах *1U* и *2U* с локальным хранилищем. Оборудование специального назначения, не являющееся продуктовым, может поддерживаться по собственному усмотрению на основе общего сходства оборудования с продуктовыми серверами;
- База данных **ADB** поддерживается в сетевом или общем хранилище в том случае, если общее хранилище представлено в виде блочного устройства на серверах, на которых запущена база данных **ADB**, и файловая система **XFS** подключена к блочному устройству. Сетевые файловые системы не поддерживаются. При использовании сетевого или общего хранилища зеркальное отображение базы данных **ADB** должно использоваться так же, как и для локального хранилища, при этом никакие изменения не могут быть внесены в схему зеркалирования или схему восстановления сегментов. Другие функции общего хранилища, такие как де-дублирование и/или репликация, напрямую не поддерживаются базой данных **ADB**, но могут использоваться с поддержкой поставщика хранилища, если они не мешают ожидаемой работе **ADB**;
- База данных **ADB** поддерживается при работе в виртуализированных системах, если хранилище представлено в виде блочных устройств, а файловая система **XFS** подключена к хранилищу каталогов сегментов;
- Для конфигурации системы требуется минимум 10-гигабитная сеть;
- Для **Red Hat Enterprise Linux 7.2** или **CentOS 7.2** параметр *systemd*, установленный по умолчанию на *RemoveIPC=yes* удаляет подключения *IPC* при выходе несистемных пользователей. Это приводит к сбою утилиты **gpinitsystem** базы данных **ADB**.

Глава 6

Совместимость дистрибутивов Hadoop

ADB обеспечивает доступ к **HDFS** с помощью **gphdfs** и **PXF**:

- *Совместимость дистрибутива Hadoop с помощью PXF*
- *Совместимость дистрибутива Hadoop с помощью gphdfs*

6.1 Совместимость дистрибутива Hadoop с помощью PXF

PXF поддерживает **Arenadata Hadoop**, **Cloudera**, **Hortonworks Data Platform** и базовые дистрибутивы **Apache Hadoop**.

При необходимости доступа к данным формата *JSON* **PXF** требует версию дистрибутива **Cloudera Hadoop 5.8** или выше, **Arenadata Hadoop 1.4.1** или выше.

6.2 Совместимость дистрибутива Hadoop с помощью gphdfs

Поддерживаемые дистрибутивы **Hadoop**, совместимые с **gphdfs** перечислены в таблице:

Таблица6.1.: Поддерживаемые дистрибутивы Hadoop, совместимые с gphdfs

Дистрибутив Hadoop	Версия	gp_hadoop_target_version
Cloudera	CDH 5.x	cdh
Hortonworks Data Platform	HDP 2.x	hdp
MapR	MapR 4.x, MapR 5.x	mpr
Apache Hadoop	2.x	hadoop
Arenadata Hadoop	1.4.1 и выше	hdp

Important: MapR требует MapR client

Глава 7

Обновление до ADB 5.8.0_arenadata1

В разделе описывается путь обновления от версии **ADB 5.x** до **ADB 5.8.0_arenadata1**:

- *Подготовка к обновлению*
- *Обновление ADB с 5.x до 5.8.0_arenadata1*
- *Устранение неполадок при неудачном обновлении*

7.1 Подготовка к обновлению

Перед началом процесса обновления следует выполнить следующие рекомендации:

- Проверить работоспособность хост-оборудования **ADB** и убедиться, что хосты соответствуют требованиям для работы базы данных. Утилита **gpcheckperf** может помочь в подтверждении требований к хосту. Рекомендуется запускать утилиту **gpcheckcat** за несколько недель до обновления и, в случае обнаружения проблем, будет возможность разрешить их до запланированного обновления.

Утилита находится в *$GPHOME/bin*. Рекомендуется, чтобы **ADB** находилась в ограниченном режиме при запуске утилиты **gpcheckcat**. Если **gpcheckcat** сообщает о несоответствиях каталогов, можно запустить ее с параметром *-g* для создания SQL-скриптов для устранения несоответствий.

После запуска SQL-скриптов необходимо снова запустить **gpcheckcat**. Возможно, придется повторить процесс запуска **gpcheckcat** и создания SQL-скриптов, чтобы убедиться в отсутствии несоответствий. Рекомендуется, чтобы SQL-скрипты, созданные **gpcheckcat**, запускались в статистической системе, так как в ином случае утилита может сообщать ложные предупреждения.

Important: Если утилита gpcheckcat сообщает об ошибках, но не создает SQL-скрипт для их исправления, необходимо обратиться в службу поддержки ADB – e-mail: info@arenadata.io

---

- Во время процесса миграции с **ADB 5.0.0** создается резервная копия некоторых файлов и каталогов в *$MASTER_DATA_DIRECTORY*. Рекомендуется, чтобы файлы и каталоги, которые не использовались в **ADB**, при неободимости были скопированы и затем удалены из файла *$MASTER_DATA_DIRECTORY* перед миграцией.

Если при предыдущей установке **ADB** была настроена платформа **PXF**, необходимо остановить **PXF** и создать резервные копии файлов ее конфигурации до перехода на новую версию **ADB**. В случае если использование **PXF** не планируется и платформа не была настроена ранее, никаких действий не требуется.

7.2 Обновление ADB с 5.x до 5.8.0_arenadata1

Обновление версии с *5.x* до *5.8.0_arenadata1* включает в себя остановку **ADB**, обновление бинарных файлов базы данных, обновление **ADB** и ее перезапуск.

При наличии баз данных, созданных с помощью **ADB 5.3.0** или более ранней версии *5.x*, необходимо обновить функцию *gp_bloat_diag* и визуализацию в схеме *gp_toolkit* (дополнительная информация приведена в разделе Обновление gp_toolkit.gp_bloat_diag).

Если в **ADB** установлена база данных *gpperfmon*, процесс миграции изменяет ключ распределения таблиц *log_alert_** на столбец *logtime*. Перераспределение данных таблицы может занять некоторое время при первом запуске **ADB** после миграции.

Порядок действий для обновления базы данных **ADB** с версии *5.x* до *5.8.0_arenadata1*:

1. Войти в мастер-узел **ADB** от администратора:

   ```
   $ su - gpadmin
   ```

2. Выполнить интеллектуальное завершение работы текущей системы **ADB 5.x** (чтобы не было активных подключений). В примере используется опция *-a* для отключения запросов на подтверждение:

   ```
   $ gpstop -a
   ```

3. Запустить бинарную программу установки для *5.8.0_arenadata1* на мастер-хосте **ADB**. Указать место установки в том же базовом каталоге, что и текущая инсталляция. Например:

   ```
   /usr/lib/gpdb
   ```

Если **ADB** устанавливается с *rpm* (с правами администратора *root*), каталог установки – */usr/lib/gpdb*. Для установки *rpm* необходимо обновить права для новой сборки. Например, запустить команду как *root*, чтобы сменить пользователя и группу устанавливаемых файлов на *gpadmin*:

```
# chown -R gpadmin:gpadmin /usr/lib/gpdb
```

4. Если для развертывания **ADB** используется аутентификация **LDAP**, следует вручную отредактировать файл */usr/lib/gpdb/greenplum_path.sh*, добавив строку:

   ```
   export LDAPCONF=/etc/openldap/ldap.conf
   ```

5. Отредактировать среду суперпользователя (*gpadmin*) **ADB** и убедиться, что для новой установки используется файл *greenplum_path.sh*. Проверить на корректность следующую строку в *.bashrc* или в выбранном файле профиля:

   ```
   source /usr/lib/gpdb/greenplum_path.sh
   ```

Если в файлах профиля используется символическая ссылка (*/usr/lib/gpdb*), необходимо ее обновить и указать новую установленную версию.

6. Среда только что отредактированного исходного файла:

   ```
   $ source ~/.bashrc
   ```

7. Запустить утилиту **gpseginstall** для установки бинарных файлов *5.8.0_arenadata1* на всех узлах сегмента, указанных в *hostfile*. Например:

   ```
   $ gpseginstall -f hostfile
   ```

8. Если ранее использовались любые расширения **ADB**, такие как **pgcrypto**, **PL/R**, **PL/Java**, **PL/Perl** и **PostGIS**, следует загрузить соответствующие пакеты и установить их с помощью утилиты **gppkg**.

9. После обновления всех хостов сегмента войти в систему как пользователь *gpadmin* и перезапустить **ADB**:

   ```
   # su - gpadmin
   $ gpstart
   ```

10. Если в предыдущей версии **ADB** была настроена **PXF**, необходимо повторно инициализировать платформу после обновления базы данных.

7.3 Устранение неполадок при неудачном обновлении

При возникновении проблем во время процесса миграции и наличии прав на **ADB**, приобретенных у **Arenadata**, необходимо обратиться в техническую поддержку **Arenadata** по адресу info@arenadata.io и предоставить следующую информацию:

- Завершенный процесс обновления (*Обновление ADB с 5.x до 5.8.0_arenadata1*);
- Данные журнала из *gpcheckcat* (находится в *~/gpAdminLogs*).

Глава 8

Решенные проблемы

В разделе перечислены проблемы в **ADB 5.8.0_arenadata1**, их решения и результат.

- 29368 - Catalog и Metadata

В некоторых случаях при инициализации базы данных **ADB** выдавалась *PANIC* по причине проблем с часовым поясом. Часовой пояс, используемый в **ADB**, выбирается из набора встроенных часовых поясов **PostgreSQL**. База данных выбирает часовой пояс, сопоставляя часовой пояс **PostgreSQL** с указанным пользователем часовым поясом или часовым поясом хост-системы. Реакция *PANIC* вызывалась, когда база данных **ADB** не могла соответствовать часовому поясу хост-системы с внутренним часовым поясом **PostgreSQL**.

Проблема решена. Теперь в указанной ситуации база данных **ADB** возвращает предупреждение и устанавливает часовой пояс кластера как GMT со смещением времени на основе информации из главной системы.

ADB не корректно соответствует системным часовым поясам, которые включают в себя данные leap second. Необходимо вручную настроить часовой пояс для совместимой конфигурации

- 29320 - Catalog и Metadata

Часовой пояс базы данных **ADB** был установлен неправильно и настроен на GMT. **ADB** не использовала значение часового пояса операционной системы, если значение не было установлено в *postgresql.conf*.

Проблема решена. Теперь часовой пояс установлен правильно.

- 29319 - Query Planner

Для запросов, содержащих функцию процентиля, относящихся к партицированной таблице, унаследованный планировщик базы данных выдавал *PANIC*. Планировщик неправильно обрабатывал память при создании многоэтапного плана агрегатов.

Проблема решена. Управление памятью для указанного типа запроса улучшено.

- 29316 - gpbackup/gprestore

Утилита **gprestore** выдавала сбой при попытке восстановления базы данных при установленном языке PL/Java. Утилита **gpbackup** неправильно создавала резервную копию языка PL/Java.

Проблема решена. Утилита **gpbackup** корректно поддерживает установленный язык PL/Java.

- 29301 - gpbackup/gprestore

В некоторых случаях при восстановлении таблиц, содержащих последовательности, утилита **gprestore** возвращала ошибку. Утилита **gpbackup** некорректно выполняла последовательности резервного копирования

при копировании набора таблиц (включая или исключая таблицы) и некоторых из принадлежащих им последовательностей.

Проблема решена. Утилита **gpbackup** правильно поддерживает последовательности таблиц в указанной ситуации.

- 29286 - Analyze

Суперпользователь базы данных **ADB** (*gpadmin*) не мог выполнять операцию анализа на таблице, не принадлежащей суперпользователю кластера. Проблема была вызвана обновлением **PostgreSQL 8.4**, которое было включено в **ADB**.

Проблема решена. Изменение **PostgreSQL 8.4** было учтено для работы с **ADB**.

- 29285 - Query Planner

Для запросов, содержащих *FULL JOIN* и столбцов типа *varchar*, возвращалась ошибка. Проблема была вызвана неправильной обработкой преобразования типов при создании плана запроса.

Проблема решена. Обработка типов преобразований улучшена для заданного типа запросов.

- 29283 - Query Optimizer

Ухудшилась производительность оптимизации **GPORCA** для запросов, содержащих предикат, выполняющий сравнение с постоянным массивом большого количества элементов, где элементами массива являются наборы констант.

Проблема решена. Оптимизация **GPORCA** улучшена для заданного типа запросов.

- 29275 - Query Optimizer

Для запросов с несколькими соединениями **GPORCA** некорректно отключала упорядочение соединения на основе минимальной мощности при достижении порога сопряженности соединения с целью улучшения обработки оптимизации. Это приводило **GPORCA** к генерации субоптимального плана для нескольких запросов соединения.

Проблема решена. **GPORCA** улучшила обработку запросов для мультисоединений.

- 29269 - Internationalization

В случаях когда кодировка **ADB** отличалась от клиентской, клиент получал недопустимую последовательность байтов для кодирования при взаимодействии с базой данных. Вызывалась некорректная функция преобразования кодировки **ADB**.

Проблема решена. Теперь вызывается корректная функция.

29259 - gprecoverseg

Проблемы производительности **gprecoverseg** при запуске утилиты для восстановления большого количества инстансов сегмента базы данных **ADB**. Некоторые тесты **gprecoverseg** выполнялись последовательно на сегментах.

Проблема решена. Для повышения производительности тесты, выполняемые утилитой, теперь осуществляются параллельно.

- 26529 - Query Optimizer

Для *GORCA* мощность соединения, включая предикат вида *f(a) = b*, оценивалась с использованием постоянного коэффициента. Эта оценка привела к переоценке мощности, а в некоторых случаях – к выбору более низкого плана запросов.

Проблема решена. *GPORCA* улучшена – генерирует оценку мощности указанного типа предиката и приводит к выбору лучшего плана.

- 29206 - Query Planner

Для некоторых запросов, содержащих функции и предикаты, унаследованный планировщик запросов создавал план запроса, который не продвигал предикат под функцию при такой необходимости. Это вызывало регрессию производительности.

Проблема решена. Оценка заданного значения предиката, выполненная планировщиком запросов, улучшена для указанного типа запроса.

Глава 9

Известные проблемы и ограничения

ADB 5.x имеет следующие ограничения:

- Некоторые функции являются незавершенными и считаются экспериментальными. Экспериментальные функции не поддерживаются в продуктивной среде (Экспериментальные функции);
- Пакеты **ADB 4.3.x** не совместимы с **ADB 5.x**.

Далее перечислены основные известные проблемы в **ADB 5.x**.

- **158011506 - Catalog и Metadata**

В некоторых случаях часовой пояс, используемый в базе данных **ADB**, может отличаться от часового пояса хост-системы или заданного пользователем. В некоторых редких случаях время, используемое и отображаемое в базе данных **ADB**, может немного отличаться от времени системы хоста.

Часовой пояс, используемый в базе данных **ADB**, выбирается из набора встроенных часовых поясов **PostgreSQL**. Кластер выбирает часовой пояс, сопоставляя часовой пояс **PostgreSQL** с часовым поясом хост-системы или указанным пользователем. При выборе часового пояса по умолчанию в базе данных **ADB** используется алгоритм выбора часовых поясов **PostgreSQL** на основе таковых в хост-системе. Если в хост-системе в часовом поясе включена информация leap second, кластер **ADB** не может соответствовать системному часовому поясу. Кластер вычисляет наилучшее соответствие с часовым поясом **PostgreSQL** на основе информации из хост-системы.

Временное решение: установить параметры **ADB** и часовых поясов хост-системы в часовой пояс, который поддерживается как базой данных, так и хост-системой. Например, установить часовой пояс **ADB** с помощью утилиты **gpconfig**. Эти команды показывают часовой пояс базы данных и устанавливают часовой пояс для US/Pacific:

```
# gpconfig -s TimeZone
# gpconfig -c TimeZone 'US/Pacific'
```

После изменения часового пояса необходимо перезапустить кластер при помощи команды *gpstop -ra*. В каталоге *pg_timezone_names* представлена информация о часовом поясе базы данных **ADB**.

- **26589 - Storage: Catalog и Metadata**

База данных **ADB** не блокирует схемы при создании таблицы внутри схемы. Одновременные операции *DROP* и *CREATE TABLE* приводят к утечке данных (к несвязанной таблице или потере связей между столбцами) в системном каталоге и, возможно, на диске. Например, отношение, созданное как *CREATE TABLE foobar (col1 int)*, при одновременном *DROP* может привести либо к утечке данных в самой таблице, либо в объекте атрибута *col1*.

Это не влияет на будущие запросы. Несоответствия каталога можно обнаружить с помощью утилиты *gpcheckcat*.

Временное решение: чтобы заблокировать параллельную операцию *DROP*, необходимо нажать и удерживать блокировку *ROW SHARE* в строке таблицы каталога *pg_namespace*, которая соответствует схеме. Например, транзакция получает блокировку *ROW SHARE* в строке таблицы каталога каталога *pg_namespace* для схемы *my_schema*:

```
begin;
SELECT nspname FROM pg_namespace WHERE nspname = 'my_schema' FOR SHARE;
  ...

end;
```

Во время транзакции блокировка предотвращает падение схемы.

- **N/A - PXF**

PXF доступен только для поддерживаемых платформ **Red Hat** и **CentOS**. **PXF** недоступен для платформ **SuSE**.

- **151135629 - Команда COPY**

Когда указано свойство *ON SEGMENT*, команда *COPY* не поддерживает *SELECT* в условии *COPY TO*. Например, следующая команда не поддерживается:

```
COPY (SELECT * FROM testtbl) TO '/tmp/mytst<SEGID>' ON SEGMENT
```

- **29064 - Storage: DLL**

Тип данных *money* принимает значения вне диапазона как отрицательные, и сообщение об ошибке не отображается.

Решение: Использовать только значения в диапазоне для типа данных *money* (64-разрядный для **ADB 5.x**). Или использовать альтернативный тип данных, например *numeric* или *decimal*.

- **3290 - JSON**

Функция *to_json()* не реализована как вызываемая функция. Попытка вызова функции приводит к ошибке. Например:

```
tutorial=# select to_json('Fred said "Hi."'::text);
ERROR: function to_json(text) does not exist
LINE 1: select to_json('Fred said "Hi."'::text);
^
HINT: No function matches the given name and argument types.
You might need to add explicit type casts.
```

Решение: **ADB** вызывает функцию *to_json()* при приведении к типу данных *json*, поэтому следует выполнять листинг. Например: *SELECT '{"foo":"bar"}'::json;* база данных обеспечивает функции *array_to_json()* и *row_to_json()*.

- **148119917 - Resource Groups**

Тестирование функции **Resource Groups** выявило, что при использовании ядра по умолчанию в системе **RHEL/CentOS** может возникнуть ошибка. Проблема возникает из-за проблем в реализации контрольных групп ядра и приводит к ошибке ядра, похожей на следующую:

```
[81375.325947] BUG: unable to handle kernel NULL pointer dereference at 0000000000000010
      [81375.325986] IP: [<ffffffff812f94b1>] rb_next+0x1/0x50 [81375.326014] PGD 0 [81375.326025]
      Oops: 0000 [#1] SMP [81375.326041] Modules linked in: veth ipt_MASQUERADE
```

```
     nf_nat_masquerade_ipv4 iptable_nat nf_conntrack_ipv4 nf_defrag_ipv4 nf_nat_ipv4 xt_addrtype
     iptable_filter xt_conntrack nf_nat nf_conntrack bridge stp llc intel_powerclamp coretemp
     intel_rapl dm_thin_pool dm_persistent_data dm_bio_prison dm_bufio kvm_intel kvm crc32_pclmul
     ghash_clmulni_intel aesni_intel lrw gf128mul glue_helper ablk_helper cryptd iTCO_wdt
     iTCO_vendor_support ses enclosure ipmi_ssif pcspkr lpc_ich sg sb_edac mfd_core edac_core
     mei_me ipmi_si mei wmi ipmi_msghandler shpchp acpi_power_meter acpi_pad ip_tables xfs
     libcrc32c sd_mod crc_t10dif crct10dif_generic mgag200 syscopyarea sysfillrect crct10dif_
↪pclmul
     sysimgblt crct10dif_common crc32c_intel drm_kms_helper ixgbe ttm mdio ahci igb libahci drm␣
↪ptp
     pps_core libata dca i2c_algo_bit [81375.326369]  i2c_core megaraid_sas dm_mirror
     dm_region_hash dm_log dm_mod [81375.326396] CPU: 17 PID: 0 Comm: swapper/17 Not tainted
     3.10.0-327.el7.x86_64 #1 [81375.326422] Hardware name: Cisco Systems Inc
     UCSC-C240-M4L/UCSC-C240-M4L, BIOS C240M4.2.0.8b.0.080620151546 08/06/2015 [81375.326459]␣
↪task:
     ffff88140ecec500 ti: ffff88140ed10000 task.ti: ffff88140ed10000 [81375.326485] RIP:
     0010:[<ffffffff812f94b1>]  [<ffffffff812f94b1>] rb_next+0x1/0x50 [81375.326514] RSP:
     0018:ffff88140ed13e10  EFLAGS: 00010046 [81375.326534] RAX: 0000000000000000 RBX:
     0000000000000000 RCX: 0000000000000000 [81375.326559] RDX: ffff88282f1d4800 RSI:
     ffff88280bc0f140 RDI: 0000000000000010 [81375.326584] RBP: ffff88140ed13e58 R08:
     0000000000000000 R09: 0000000000000001 [81375.326609] R10: 0000000000000000 R11:
     0000000000000001 R12: ffff88280b0e7000 [81375.326634] R13: 0000000000000000 R14:
     0000000000000000 R15: 0000000000b6f979 [81375.326659] FS:  0000000000000000(0000)
     GS:ffff88282f1c0000(0000) knlGS:0000000000000000 [81375.326688] CS:  0010 DS: 0000 ES: 0000
     CR0: 0000000080050033 [81375.326708] CR2: 0000000000000010 CR3: 000000000194a000 CR4:
     00000000001407e0 [81375.326733] DR0: 0000000000000000 DR1: 0000000000000000 DR2:
     0000000000000000 [81375.326758] DR3: 0000000000000000 DR6: 00000000ffff0ff0 DR7:
     0000000000000400 [81375.326783] Stack: [81375.326792]  ffff88140ed13e58 ffffffff810bf539
     ffff88282f1d4780 ffff88282f1d4780 [81375.326826]  ffff88140ececae8 ffff88282f1d4780
     0000000000000011 ffff88140ed10000 [81375.326861]  0000000000000000 ffff88140ed13eb8
     ffffffff8163a10a ffff88140ecec500 [81375.326895] Call Trace: [81375.326912]
     [<ffffffff810bf539>] ? pick_next_task_fair+0x129/0x1d0 [81375.326940]  [<ffffffff8163a10a>]
     __schedule+0x12a/0x900 [81375.326961]  [<ffffffff8163b9e9>] schedule_preempt_disabled+0x29/
↪0x70
     [81375.326987]  [<ffffffff810d6244>] cpu_startup_entry+0x184/0x290 [81375.327011]
     [<ffffffff810475fa>] start_secondary+0x1ba/0x230 [81375.327032] Code: e5 48 85 c0 75 07 eb␣
↪19 66
     90 48 89 d0 48 8b 50 10 48 85 d2 75 f4 48 8b 50 08 48 85 d2 75 eb 5d c3 31 c0 5d c3 0f 1f 44
     00 00 55 <48> 8b 17 48 89 e5 48 39 d7 74 3b 48 8b 47 08 48 85 c0 75 0e eb [81375.327157] RIP
     [<ffffffff812f94b1>] rb_next+0x1/0x50 [81375.327179]  RSP <ffff88140ed13e10> [81375.327192]␣
↪CR2:
     0000000000000010
```

Решение: Обновить ядро **Red Hat** или **CentOS** до последней версии.

- **149789783 - Resource Groups**

Значительное снижение производительности **ADB** наблюдается при включении управления рабочими нагрузками **Resource Groups** на системах **RedHat 6.x**, **CentOS 6.x** и **SuSE 11**. Эта проблема вызвана ошибкой контрольной группы ядра **Linux**. Ошибка была исправлена в **CentOS 7.x** и **Red Hat 7.x**.

Когда **Resource Groups** включены в системах с поврежденным ядром, при запуске транзакции или запроса может быть задержка в *1* секунду или более. Задержка вызвана ошибкой ядра **Linux**, где механизм синхронизации вызывается *synchronize_sched* и выдает ошибку, когда процесс присоединяется к группе. См. http://www.spinics.net/lists/cgroups/msg05708.html и https://lkml.org/lkml/2013/1/14/97 для получения дополнительной информации.

Эта проблема приводит к тому, что операции одиночного присоединения занимают больше времени, а также, что все параллельные вложения выполняются последовательно. Например, одно приложение

процесса может занимать около *0,01* секунды. При одновременном подключении *100* процессов самое быстрое присоединение занимает *0,01* секунды, а самое медленное - около *1* секунды. **ADB** выполняет присоединения процессов при запуске транзакции или запросов. Таким образом, ухудшение производительности зависит от одновременно запущенных транзакций или запросов и не связано с параллельными работающими запросами. Также **ADB** имеет оптимизацию для обхода переписывания, когда **QE** повторно используется несколькими запросами в том же сеансе.

Эта ошибка не влияет на системы **CentOS 7.x** и **Red Hat 7.x**

Решение: Если используется **RedHat 6** и производительность с **Resource Groups** приемлема, следует обновить ядро до версии *2.6.32-696* или выше, чтобы воспользоваться преимуществами и доработками новой реализации **Resource Groups**.

SuSE 11 не имеет версии ядра, решающей эту проблему; **Resource Groups** по-прежнему считаются экспериментальной функцией на данной платформе.

- **150906510 - Backup и Restore**

Резервные копии базы данных *4.3.15.0* и более поздних версий содержат следующую строку в файлах резервных копий:

```
SET gp_strict_xml_parse = false;
```

Однако, **ADB 5.0.0** не имеет параметр с *gp_strict_xml_parse*. При восстановлении данных из резервного набора версии *4.3* в кластере *5.0.0* может появиться предупреждение:

```
[WARNING]:-gpdbrestore finished but ERRORS were found, please check the restore report file
for details
```

Кроме того, файл отчета может содержать ошибку:

```
ERROR: unrecognized configuration parameter "gp_strict_xml_parse"
```

Данные предупреждения и ошибки не влияют на процедуру восстановления, и могут быть проигнорированы.

9.1 Обновление gp_toolkit.gp_bloat_diag

В **ADB 5.3.0** или более ранней версии **5.x** в некоторых случаях при выполнении запроса в отношении *gp_toolkit.gp_bloat_diag* появлялась ошибка целого числа вне диапазона. Проблема была решена в **ADB 5.4.0**. При обновлении базы данных **ADB** функция *gp_toolkit.gp_bloat_diag* и визуализация должны обновляться в базах данных, созданных с помощью **ADB 5.3.0** или более ранней версии **5.x**. Проблема была решена в **ADB 5.4.0** и более поздних версиях.

Чтобы проверить, требуется ли обновление для функции *gp_toolkit.gp_bloat_diag* и визуализации в базе данных, необходимо запустить **psql** и выполнить команду *df*, чтобы отобразить информацию о функции *gp_toolkit.gp_bloat_diag*:

```
df gp_toolkit.gp_bloat_diag
```

Если тип данных для *btdexppages* является целым числом (“numeric”), требуется обновление. Если тип данных является цифровым (“integer”), обновление не требуется. В примере тип данных *btdexppages* является целым числом и требует обновления:

```
List of functions
-[ RECORD 1 ]-------+-------------------------------------------------------------------------------
↪-------------------
Schema              | gp_toolkit
Name                | gp_bloat_diag
Result data type    | record
```

```
Argument data types | btdrelpages integer, btdexppages integer, aotable boolean, OUT bltidx␣
↪integer, OUT bltdiag text
Type                | normal
```

Необходимо запустить скрипт, чтобы обновить функцию и визуализацию и исправить проблему в каждой базе данных, которая была создана с помощью **ADB 5.3.0** или более ранней версии **5.x**. И в качестве пользователя *gpadmin* выполнить следующие действия:

1. Скопировать скрипт в текстовый файл в базе данных **ADB**.
2. Запустить скрипт в каждой базе данных, требующей обновления.

В примере обновлена функция *gp_toolkit.gp_bloat_diag* и визуализация в базе данных *mytest* и предполагается, что скрипт находится в файле *update_bloat_diag.sql* в доменном каталоге *gpadmin*:

```
psql -f /home/gpadmin/update_bloat_diag.sql -d mytest
```

Запустить скрипт в период низкой активности. Запуск скрипта в период высокой активности не влияет на функциональность базы данных, но может повлиять на производительность.

Скрипт gp_toolkit.gp_bloat_diag для обновления функций и представления

```
BEGIN;
CREATE OR REPLACE FUNCTION gp_toolkit.gp_bloat_diag(btdrelpages int, btdexppages numeric, aotable␣
↪bool,
    OUT bltidx int, OUT bltdiag text)
AS
$$
    SELECT
        bloatidx,
        CASE
            WHEN bloatidx = 0
                THEN 'no bloat detected'::text
            WHEN bloatidx = 1
                THEN 'moderate amount of bloat suspected'::text
            WHEN bloatidx = 2
                THEN 'significant amount of bloat suspected'::text
            WHEN bloatidx = -1
                THEN 'diagnosis inconclusive or no bloat suspected'::text
        END AS bloatdiag
    FROM
    (
        SELECT
            CASE
                WHEN $3 = 't' THEN 0
                WHEN $1 < 10 AND $2 = 0 THEN -1
                WHEN $2 = 0 THEN 2
                WHEN $1 < $2 THEN 0
                WHEN ($1/$2)::numeric > 10 THEN 2
                WHEN ($1/$2)::numeric > 3 THEN 1
                ELSE -1
            END AS bloatidx
    ) AS bloatmapping

$$
LANGUAGE SQL READS SQL DATA;

GRANT EXECUTE ON FUNCTION gp_toolkit.gp_bloat_diag(int, numeric, bool, OUT int, OUT text) TO␣
↪public;
```

```
CREATE OR REPLACE VIEW gp_toolkit.gp_bloat_diag
AS
    SELECT
        btdrelid AS bdirelid,
        fnnspname AS bdinspname,
        fnrelname AS bdirelname,
        btdrelpages AS bdirelpages,
        btdexppages AS bdiexppages,
        bltdiag(bd) AS bdidiag
    FROM
    (
        SELECT
            fn.*, beg.*,
            gp_toolkit.gp_bloat_diag(btdrelpages::int, btdexppages::numeric, iao.iaotype::bool)␣
↪AS bd
        FROM
            gp_toolkit.gp_bloat_expected_pages beg,
            pg_catalog.pg_class pgc,
            gp_toolkit.__gp_fullname fn,
            gp_toolkit.__gp_is_append_only iao

        WHERE beg.btdrelid = pgc.oid
            AND pgc.oid = fn.fnoid
            AND iao.iaooid = pgc.oid
    ) as bloatsummary
    WHERE bltidx(bd) > 0;

GRANT SELECT ON TABLE gp_toolkit.gp_bloat_diag TO public;
COMMIT;
```