Arenadata™ Database

Версия - v5.5

Примечания к выпуску ADB-5.5.0

2018 Arenadata™

Оглавление

1 Новые возможности **3**
- 1.1 Утилиты gpbackup и gprestore . . . 3
- 1.2 Specify the Temporary Loading Table для утилиты gpload . . . 4
- 1.3 PXF User Impersonation . . . 4
- 1.4 PXF Profile для Parquet Files . . . 5
- 1.5 Запись данных в HDFS с помощью PXF . . . 5
- 1.6 ADB Workload Management Extension . . . 5

2 Измененные функции **6**

3 Экспериментальные функции **7**

4 Устаревшие функции **8**

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

6 Поддерживаемые платформы **10**
- 6.1 Поддерживаемые платформы . . . 11

7 Совместимость дистрибутивов Hadoop **12**
- 7.1 Совместимость дистрибутивов Hadoop с помощью PXF . . . 12
- 7.2 Совместимость дистрибутивов Hadoop с помощью gphdfs . . . 12

8 Обновление до ADB 5.5.0 **13**
- 8.1 Подготовка . . . 13
- 8.2 Обновление ADB с 5.x до 5.5.0 . . . 14
- 8.3 Устранение неполадок при неудачном обновлении . . . 15

9 Решенные проблемы **16**

10 Известные проблемы и ограничения **19**
- 10.1 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 5.5.0. Для предыдущих версий продукта необходимо перейти к соответствующей документации

ADB 5.5.0 – это минорная версия, в которой добавлены и изменены некоторые функции.

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

Глава 1

Новые возможности

ADB 5.5.0 включает в себя следующие новые функции:

- *Утилиты gpbackup и gprestore*
- *Specify the Temporary Loading Table для утилиты gpload*
- *PXF User Impersonation*
- *PXF Profile для Parquet Files*
- *Запись данных в HDFS с помощью PXF*
- *ADB Workload Management Extension*

1.1 Утилиты gpbackup и gprestore

В **ADB 5.5.0** утилиты **gpbackup** и **gprestore** больше не являются экспериментальными.

Утилиты **gpbackup** и **gprestore** предназначены для повышения производительности, функциональности и надежности резервных копий по сравнению с утилитами **gpcrondump** и **gpdbrestore**. Они были введены в качестве экспериментальных в версии *5.2.0*. Утилита **gpbackup** использует блокировку *ACCESS SHARE* на отдельном уровне таблицы (а не блокировку *EXCLUSIVE* в таблице каталога *pg_class*). Это позволяет выполнять команды **DDL** такие, как операции *CREATE*, *ALTER*, *DROP* и *TRUNCATE* в процессе резервного копирования, если эти операции не нацелены на текущий набор резервных копий.

gpbackup и **gprestore** включают следующие улучшения:

- Изменения в уведомлении по электронной почте:
  - Имя файла *YAML* изменено на *gp_email_contacts.yaml*
  - Формат файла *YAML* поддерживает уведомление по электронной почте для операций резервного копирования и восстановления. Электронное письмо можно отправить различным адресатам при успешном завершении операции или при завершении с ошибкой.
- Возвращение кода состояния после завершения работы утилиты:
  - 0 – Резервное копирование или восстановление завершено без проблем
  - 1 – Резервное копирование или восстановление завершено с нефатальными ошибками. Дополнительная информация находится в файле журнала

  - 2 – Резервное копирование или восстановление завершено с фатальной ошибкой. Дополнительная информация находится в файле журнала

- Имя параметра *-backupdir* утилиты **gpbackup** изменено на *-backup-dir*
- **gprestore** поддерживает новые параметры:
  - *-exclude-schema* исключает схему из операции восстановления
  - *-exclude-table* исключает таблицу из операции восстановления
  - *-exclude-table-file* исключает таблицы, которые перечислены в текстовом файле, из операции восстановления
  - *-include-table* восстанавливает таблицу из набора резервных копий
- Имена следующих параметров утилиты **gprestore** изменены:
  - *-backupdir* на *-backup-dir*
  - *-createdb* на *-create-db*
  - *-redirect* на *-redirect-db*
  - *-globals* на *-with-globals*
- Поддерживается восстановление параллельного индекса при указанном параметре *-jobs* утилиты **gprestore**. Ранее параллельно восстанавливались только данные.

Important: Если значение параметра *-jobs* больше 1, то в случае создания резервной копии с помощью параметра *-single-data-file* утилиты **gpbackup**, параллельное восстановление не поддерживается

Important: Утилита gprestore в ADB 5.5.0 не поддерживает восстановление резервных копий, созданных с помощью экспериментальных версий gpbackup. Необходимо создавать и восстанавливать резервные копии с помощью одной пары утилит базы данных ADB: либо gpbackup и gprestore, либо gpcrondump и gpdbrestore. То есть gprestore не может восстановить резервную копию, созданную с помощью gpcrondump, и gpdbrestore не сможет восстановить резервную копию, созданную с помощью gpbackup

1.2 Specify the Temporary Loading Table для утилиты gpload

В **ADB 5.5.0** для повышения производительности можно указать внешнюю таблицу, которую утилита **gpload** будет использовать для загрузки данных.

В файле *YAML* файла управления *gpload* можно указать имя таблицы с параметром *STAGING_ TABLE* в секции *PRELOAD*. Если таблицы не существует, утилита ее создает. Если таблица существует, утилита ее использует. Если таблица существует, но схема таблицы не соответствует загружаемой таблице, выдается ошибка.

1.3 PXF User Impersonation

По умолчанию сейчас **PXF** использует роли пользователей для попытки получения доступа к источникам данных (**HDFS**, **Hive**, **HBase**) при помощи идентификатора учетной записи пользователя базы данных **ADB**, которая регистрируется в СУБД и выполняет операцию, используя коннектор профиля **PXF**. В предыдущих релизах все ресурсы были доступны в качестве пользователя *gpadmin*.

1.4 PXF Profile для Parquet Files

PXF теперь включает профиль **HDFS** для чтения файлов, хранящихся в формате *Parquet* (это дополнение к **PXF** коннектору **Hive**, который поддерживает как чтение, так и запись Parquet-данных через **Hive**).

1.5 Запись данных в HDFS с помощью PXF

Запись двоичных файлов формата *SequenceFile* в **HDFS** с использованием платформы **PXF** больше не является экспериментальной функцией.

1.6 ADB Workload Management Extension

В версию добавлено расширение рабочей нагрузки **ADB**. Оно является предварительным условием для предстоящего выпуска **Zeppelin management tool** на платформах **Red Hat** и **CentOS** и еще не включено.

Расширение устанавливается в каталоге *$GPHOME/share/postgresql/extension/gp_wlm*, а файлы библиотеки – в каталоге *$GPHOME/lib/postgresql*. Параметр конфигурации сервера *shared_preload_libraries* в файлах мастреа и сегмента *postgresql.conf* обновляется, чтобы включить имя расшаренной библиотеки расширения – *gp_wlm.so*.

Для перехода на совместимую с двоичным кодом версию базы данных **ADB** раньше версии *5.5.0*, необходимо удалить *gp_wlm.so* из параметра конфигурации *shared_preload_libraries* перед тем, как закрыть текущую версию, и выполнить следующие действия:

1. Проверить текущее значение переменной конфигурации *shared_preload_libraries*:

   ```
   $ gpconfig -s shared_preload_libraries
   ```

2. Если библиотека *gp_wlm.so* является единственной библиотекой в списке, необходимо запустить команду, чтобы закомментировать параметр:

   ```
   $ gpconfig -r shared_preload_libraries
   ```

3. Если в списке есть другие расшаренные библиотеки, необходимо сбросить параметр, не включая *gp_wlm.so* в список:

   ```
   $ gpconfig -c shared_preload_libraries -v '<new-list>'
   ```

Глава 2

Измененные функции

ADB 5.5.0 изменены следующие функции:

- **PXF** Java-сервис теперь использует порт **TCP** по умолчанию *5888* (предыдущие выпуски **PXF** использовали в качестве порта по умолчанию *51200*)
- Утилиты **gpbackup** и **gprestore** включают следующие изменения (подробное описание приведено в разделе *Утилиты gpbackup и gprestore*):
  - Изменения в уведомлении по электронной почте
  - Имя параметра *-backupdir* утилиты **gpbackup** изменено на *-backup-dir*
  - Имена следующих параметров утилиты **gprestore** изменены:
    - *-backupdir* на *-backup-dir*
    - *-createdb* на *-create-db*
    - *-redirect* на *-redirect-db*
    - *-globals* на *-with-globals*
- Утилита **gpdeletesystem** ищет резервные файлы в резервных каталогах (резервные файлы *gpbackup*) и *db_dumps* (резервные файлы *gpcrondump*) мастера и сегмента. В предыдущих выпусках **gpdeletesystem** работала только в каталоге *db_dumps*.

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 поможет стимулировать разработку этих функций, и они могут поддерживаться в будущих версиях продукта.

Important: Экспериментальные функции не рекомендуются и не поддерживаются для развертывания в продуктивной среде. Эти 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 5.5.0** входят следующие экспериментальные функции:

- **Recursive WITH Queries** (общие выражения таблицы);
- **Resource groups** 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

Устаревшие функции

Утилиты **gpcrondump** и **gpdbrestore** устарели и не будут поддерживаться после окончания срока тех.поддержки **ADB 5.x**.

Глава 5

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 6

Поддерживаемые платформы

ADB 5.5.0 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 PL/Container
- Платформу PXF

ADB 5.5.0 поддерживает следующие версии **Java**:

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

ПО **ADB 5.5.0**, 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 5.5.0 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**.

6.1 Поддерживаемые платформы

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

- Единственной файловой системой, поддерживаемой для работы с базой данных **ADB**, является 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 7

Совместимость дистрибутивов Hadoop

ADB обеспечивает доступ к **HDFS** с помощью *gphdfs* и платформы **PXF**.

7.1 Совместимость дистрибутивов 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.

7.2 Совместимость дистрибутивов Hadoop с помощью gphdfs

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

Таблица7.1.: 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. Для MapR 5.x при условии FORMAT команда CREATE EXTERNAL TABLE поддерживается только TEXT и CSV

Глава 8

Обновление до ADB 5.5.0

В разделе описывается путь обновления от версии **ADB 5.x** до **ADB 5.5.0**.

8.1 Подготовка

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 из **ADB 5.0.0** создается резервная копия некоторых файлов и каталогов в *$MASTER_DATA_DIRECTORY*. Рекомендуется, чтобы файлы и каталоги, которые не использовались в **ADB**, были скопированы и, если необходимо, удалены из файла *$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.

8.2 Обновление ADB с 5.x до 5.5.0

Обновление версии с *5.x* до *5.5.0* включает в себя остановку **ADB**, обновление двоичных файлов базы данных, обновление и перезапуск **ADB**.

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.

1. Войти в мастер-узел **ADB** как администратор:

   ```
   $ 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 двоичную программу установки для *5.5.0* на главном хосте **ADB**. При появлении запроса выбрать место установки в том же базовом каталоге, что и текущая установка. 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 установленных 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 условие 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. Только что отредактированный исходный файл:

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

7. Запустить утилиту **gpseginstall** для установки двоичных файлов *5.5.0* на всех узлах сегмента, указанных в *hostfile*. Например:

   ```
   $ gpseginstall -f hostfile
   ```

8. Использовать утилиту *gppkg* для установки расширений **ADB**. Если ранее использовались любые расширения **ADB**, такие как *pgcrypto*, *PL/R*, *PL/Java*, *PL/Perl* и *PostGIS*, следует загрузить соответствующие пакеты и установить их.

9. После того как все хосты сегментов были обновлены, можно войти в систему как пользователь *gpadmin* и перезапустить систему **ADB**:

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

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

8.3 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 5.5.0*);
- Данные журнала из *gpcheckcat* (находится в *~/gpAdminLogs*)

Глава 9

Решенные проблемы

В разделе перечислены проблемы в **ADB 5.5.0** и результат их решения.

- PXF

База данных **ADB 5.4.0** не поддерживала функции **PXF** из-за регрессий, обнаруженных после релиза. Поддержка **PXF** восстановилась в **ADB 5.5.0**.

- 29219 - gpbackup/gprestore

В некоторых случаях сбой утилиты **gplestore** базы данных **ADB** при восстановлении таблиц из резервной копии. Ошибка возникает при создании резервной копии с помощью утилиты **gpbackup** с указанным параметром исключения таблицы *-exclude-table* или *-exclude-table-file*. Резервное копирование не содержит метаданных схемы для таблиц, включенных в резервную копию.

Проблема решена. Резервная копия, созданная с указанными условиями, теперь содержит информацию о схеме.

- 29200 - gpcrondump/gpdbrestore

В некоторых случаях сбой утилиты **gpdbrestore** базы данных **ADB** при операции восстановления. Ошибка вызывается тем, что утилита неправильно поддерживает функции с ключами, которые включали значение параметра *DEFAULT*.

Проблема решена. Теперь утилита **gpcrondump** корректно поддерживает указанный тип функций.

- 29193 - Query Optimizer

Для запроса, содержащего подзапрос в списке проектов, который не может быть декоррелирован, а также содержащего другие столбцы таблицы, **GPORCA** формирует план запроса, который может стать причиной простоя.

Проблема решена. **GPORCA** улучшен для обнаружения и предотвращения возможных простоев для указанного типа запроса.

- 29189, 29043 - Query Optimizer

Для некоторых запросов, которые могут использовать bitmap-индекс, **GPORCA** создает неоптимальный план, который не использует bitmap-индекс.

Проблема решена. **GPORCA** расширен для использования bitmap-индексов в большинстве классов планов запросов.

- 29178, 154194698 - Segment Mirroring

В некоторых ситуациях с включенным зеркалированием сегментов база данных **ADB** не исполняет зеркалирование в основном сегменте. База данных становится непригодной до тех пор, пока не будет выполнено

ручное вмешательство. Проблема возникает, когда процесс **FTS** (**Fault Tolerance Server**) неправильно обрабатывает условие ошибки, вызванное аппаратными проблемами.

Проблема решена. Процесс **FTS** расширен для обработки состояния ошибки.

- 29154 - Vacuum

В некоторых случаях во время операции *VACUUM FULL* в базе данных **ADB** возвращается ошибка, указывающая, что обновленная запись уже *HEAP_MOVED_OFF*. Ошибка происходит, когда операция *VACUUM* неправильно определяет, какие старые записи могут быть удаленными.

Проблема решена. При операции *VACUUM* улучшена проверка на наличие удаленных старых записей.

- 29125 - Query Optimizer

В некоторых случаях **GPORCA** неправильно обрабатывает значение *NULL* при обнаружении противоречий с использованием типов данных текста или символов. Это вызвает *SIGSEGV* для мастера базы данных **ADB**.

Проблема решена. **GPORCA** улучшен при обнаружении противоречия, когда текст или символьные данные имеют значение *NULL*.

- 29062 - Query Optimizer

GPORCA игнорирует модификаторы типа данных, которые были включены во время анализа запроса, и позволяет выполнить запрос на изменение типов данных. Например, столбцы в *CREATE TABLE* определяются символьными (10) и числовыми (5,2) типами данных, которые изменяют типы данных.

```
CREATE TABLE mytest(col1 character(10), col2 numeric(5,2))
```

GPORCA отменяет модификации типов данных во время обработки запросов. Это вызывает проблемы, когда результаты ожидаются как тип данных не по умолчанию.

Проблема решена. **GPORCA** усовершенствован для обработки модификаторов типа данных в столбцах таблицы, скалярных идентификаторах, константах и приведениях длины.

- 27037 - Query Planner, Query Optimizer

В некоторых случаях планировщик **ADB** и **GPORCA** генерируют *SIGSEGV* во время формирования плана запроса. *SIGSEGV* происходит, когда используемая в запросе память для статистики обрабатывается неправильно, и поэтому появляются проблемы со статистикой.

Проблема решена. Теперь управление памятью для статистики улучшено и план сформирован.

- 154732427 - Query Optimizer

Для некоторых запросов с левым внешним соединением, содержащим предикат с условием *IS DISTINCT FROM*, **GPORCA** генерирует план, который приводит к неправильным результатам.

Проблема решена. **GPORCA** расширен, чтобы возвращать правильные результаты для указанного типа запроса.

- 154424446 - Query Optimizer

Для некоторых запросов с большим количеством соединений **GPORCA** генерирует субоптимальный план запроса, в котором используются соединения вложенных циклов вместо хеш-соединений. В предыдущих версиях **GPORCA** учитывала планы запросов, в которых использовались хэш-соединения.

Проблема решена. Теперь **GPORCA** учитывает планы, которые используют хэш-соединения для заданного типа запросов.

- 153757596 - Query Optimizer

GPORCA не оценивает мощность правильно, когда есть сравнение между скалярным идентификатором и массивом, который состоит из констант и не констант.

Проблема решена. Для указанной ситуации была улучшена оценка мощности.

- 153628085 - Interconnect

В некоторых случаях при тяжелых нагрузках база данных **ADB** генерирует *PANIC*, когда запрос не может выделить память, необходимую для установки межсоединения из-за отсутствия системной памяти. *PANIC* возвращается, когда ошибка из памяти не обрабатывается должным образом.

Проблема решена. Теперь в указанной ситуации возвращается ошибка.

- 151934728 - Query Execution

В некоторых случаях при выполнении пользовательской функции запрос завершается с ошибкой *VOID*. Ошибка происходит, когда сопоставление типов не выполняется должным образом во время отправки запроса.

Проблема решена. Улучшено отображение типов во время отправки.

Глава 10

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**.

- **26589 - Хранение: каталог и метаданные**

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 - 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 - Резервное копирование и восстановление**

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.

10.1 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;
```