

УТВЕРЖДЕН
643.72410666.00067-07 98 01-ЛУ

ЗАЩИЩЕННАЯ СИСТЕМА УПРАВЛЕНИЯ
БАЗАМИ ДАННЫХ «ЈАТОВА»

Руководство по настройке. Часть 3.
Формирование отчетов по журналам СУБД.
Компонент «pgBadger»

643.72410666.00067-07 98 01-03

Листов 30

Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата

АННОТАЦИЯ

В документе приведены сведения, необходимые для формирования аналитической отчетности по заданным параметрам, для выявления потенциальных проблем с системой управления базами данных «Jatoba» (далее – СУБД «Jatoba»). Настоящее руководство предназначено для администратора СУБД «Jatoba».

Администратор СУБД «Jatoba» должен иметь навыки по работе с системами управления базами данных «PostgreSQL» или защищенной СУБД «Jatoba» (ООО «Газинформсервис»).



Все примеры в данном документе приведены для СУБД «Jatoba» версии ядра 4.x, для других версий все шаги выполняются аналогично, разница состоит в именах директорий.

Например, СУБД «Jatoba» версии 6.x по умолчанию устанавливается в директорию:

- ОС Windows – «C:\Program Files\GIS\Jatoba\6\bin»;
- ОС Linux – «/usr/jatoba-6/bin».

Для СУБД «Jatoba» версии ядра 4/5/6/18 используется версия компонента — 13.1.



Важная информация

Для сертифицированной версии СУБД «Jatoba» поддерживается работа только на ОС, указанных в формуляре на поставку!

Степени важности примечаний, применяемые в документе:



Важная информация – указания, требующие особого внимания



Дополнительная информация – указания, позволяющие упростить работу с изделием

№ изменения: _____	Подпись отв. лица: _____	Дата внесения изм: _____
--------------------	--------------------------	--------------------------

СОДЕРЖАНИЕ

1. Назначение компонента.....	4
1.1. Функциональные возможности компонента	4
1.2. Условия применения.....	4
2. Требования по использованию.....	5
3. Установка и удаление	7
3.1. Установка/удаление в ОС GNU/Linux.....	7
3.2. Установка/удаление в ОС Microsoft Windows.....	8
4. Настройка Jatoba/PostgreSQL	10
5. Опции использования компонента.....	12
5.1. Формат значения datetime	16
5.2. SSH опции	16
5.3. Примеры использования	17
6. Журналирование запросов.....	20
7. Параллельная обработка.....	21
8. Инкрементальные отчеты	23
9. Ежемесячные отчеты	24
10. Бинарный формат.....	25
11. Формат JSON	26
12. Отчеты, получаемые с помощью компонента.....	27
Перечень сокращений.....	29

1. НАЗНАЧЕНИЕ КОМПОНЕНТА

«pgBadger» – компонент, обеспечивающий анализ журналов регистрации событий СУБД, в том числе событий безопасности, и формирование статистических отчетов и графиков на их основе.

Компонент «pgBadger» позволяет выполнять обработку в части событий, содержащихся в журналах. С помощью параметров командной строки возможно установить такой режим работы компонента, при котором будет обрабатываться только заданный набор сообщений.

1.1. Функциональные возможности компонента

Компонент «pgBadger» обладает следующими функциональными возможностями:

- сбор общих данных о работе СУБД из файлов регистрации событий, в том числе о сессиях пользователя (выполненные SQL-запросы), и предоставление данных администратору в виде HTML-отчетов;
- выявление проблемных мест в производительности СУБД;
- поиск и анализ запросов, подлежащих оптимизации.

При работе с данным компонентом администратор должен иметь доступ и обладать правами на чтение директории, в которой расположены файлы регистрации событий, и правами на чтение и запись в директорию сохранения отчетов. Допускается использование учетной записи администратора ОС (например, root для Linux) и системной учетной записи postgres.

1.2. Условия применения

Компонент «pgBadger» может использоваться совместно с СУБД «Jatoba» версий 1.x, 4x и выше.

2. ТРЕБОВАНИЯ ПО ИСПОЛЬЗОВАНИЮ

Перед использованием компонента «pgBadger» необходимо выполнить установку динамического языка программирования Perl не ниже версии 5.14, с дополнительными компонентами, приведенными в таблице 2.1.

Таблица 2.1 – Дополнительные компоненты Perl

Компонент	Назначение
Text::CSV_XS	Разбор файлов журналов в формате CSV
JSON::XS	Сериализация/десериализация JSON (для экспорта статистики в файл JSON)



Установка данных компонентов осуществляется только при необходимости разбора файлов журналов в формате CSV и выгрузки статистики в формате JSON.



Названия пакетов могут отличаться в различных дистрибутивах GNU/Linux. Для получения точного названия рекомендуется сверяться с данными подключенных репозиторийев пакетов.

Для установки на Debian-подобную систему используется команда:

```
sudo apt-get install libtext-csv-xs-perl libjson-xs-perl
```

Для установки на RedHat-подобную систему используется команда:

```
sudo yum install perl-Text-CSV_XS perl-JSON-XS
```

Графическое отображение информации осуществляется с помощью библиотеки JavaScript с использованием веб-браузера.

Поддерживается работа браузеров: Google Chrome, Mozilla Firefox, Microsoft Edge, Internet Explorer, Safari и др.

Формат сжатого файла журнала определяется автоматически по расширению файла. Список расширений архивов и соответствующих им утилит представлен в таблице 2.2.

Таблица 2.2 – Список расширений и соответствующих им утилит

Расширение архива	Используемая утилита
-------------------	----------------------

№ изменения: _____	Подпись отв. лица: _____	Дата внесения изм: _____
--------------------	--------------------------	--------------------------

gz	zcat
bz2	bzcat
lz4	lz4cat
zst	Zstdcat
Zip	unzip
Xz	xz

Если утилиты не найдены в директориях, заданных в переменные окружения PATH, воспользуйтесь опцией командной строки --zcat для изменения этого пути.

Примеры:

```
--zcat="/usr/local/bin/gunzip -c"  
--zcat="/usr/local/bin/bzip2 -dc"  
--zcat="C:\tools\unzip -p"
```

При использовании режима автоопределения формата сжатия файлы gz, bz2, lz4, xz, zip или zstd можно смешивать.



Указание пользовательского значения для опции --zcat отключит возможность смешанного формата сжатия.

3. УСТАНОВКА И УДАЛЕНИЕ

Установка компонента «pgBadger» должна производиться от имени пользователя, обладающего административными привилегиями в системе.

Компонент «pgBadger» может быть установлен штатным образом только с СУБД «Jatoba» (см. документ «Защищенная система управления базами данных «Jatoba». Руководство по установке).

Для использования компонента отдельно от СУБД «Jatoba» потребуется установка вручную (данный режим официально не поддерживается разработчиком).

3.1. Установка/удаление в ОС GNU/Linux

Компонент распространяется в составе СУБД в виде отдельного deb или rpm-пакета. Установка компонента осуществляется средствами пакетного менеджера ОС.

Для разных типов пакетных менеджеров команда установки различается:

– для систем на основе пакетного менеджера APT (к таким системам относятся все ОС семейства Debian, использующие deb-пакеты):

```
apt-get install jatoba4-pgBadger
```

– для систем на основе пакетных менеджеров YUM/DNF (к таким системам относятся все ОС семейства RedHat и вышедшие из нее, использующие rpm-пакеты):

```
yum install jatoba4-pgBadger
```

Установка компонента на ОС ALT Linux и openSUSE имеет следующие особенности:

– ALT Linux использует пакетный менеджер APT, но распространяется в виде rpm-пакетов, команда установки выглядит аналогично Debian:

```
apt-get install jatoba4-pgBadger
```

Установка компонента в составе других версий СУБД «Jatoba» осуществляется аналогично в соответствии с номером версии СУБД, в составе которой он распространяется. Пример имен пакетов приведен в таблице 3.1.

Таблица 3.1 – Имена пакетов для разных версий СУБД

№ изменения: _____	Подпись отв. лица: _____	Дата внесения изм: _____
--------------------	--------------------------	--------------------------

Маркировка частей пакета	Имя пакета	
	jatoba1-pgBadger	jatoba4-pgBadger
Наименование СУБД	jatoba	jatoba
Версия СУБД	1	4
Разделитель	-	-
Наименование компонента	pgBadger	pgBadger

Удаление компонента осуществляется средствами пакетного менеджера ОС. Синтаксис команды удаления аналогичен синтаксису команды установки (install), описанной выше. При этом нужно использовать команду удаления, соответствующую пакетному менеджеру: remove, purge, erase и т.п.

Для получения детальной информации по пакетному менеджеру рекомендуется обратиться к документации по ОС.

3.2. Установка/удаление в ОС Microsoft Windows

Компонент распространяется в составе инсталляционного msi-файла СУБД «Jatoba».

При выборе режима установки «Полный», компонент pgBadger установится автоматически в составе СУБД.

При выборе режима установки «Выборочный», для установки компонента «pgBadger» необходимо в списке компонент выбрать пункт «Анализ журналов сообщений (pgBadger)».

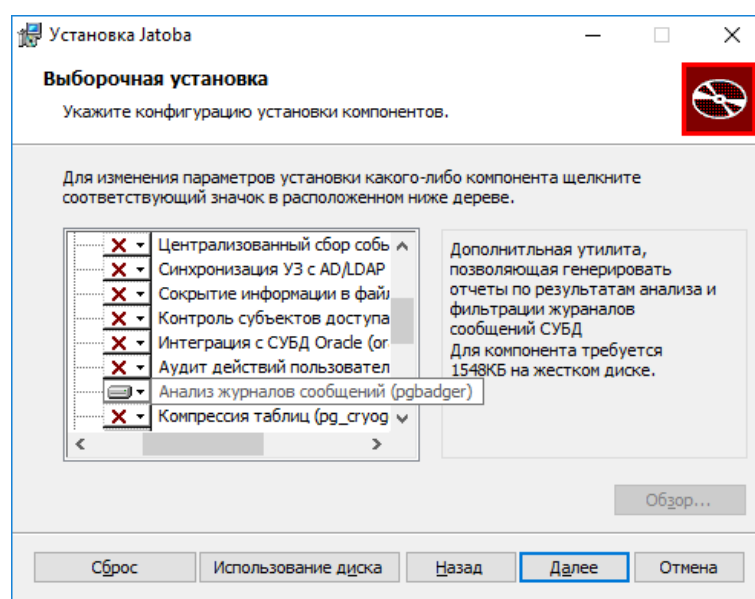


Рисунок 3.1 – Выбор компонента при установке

№ изменения: _____	Подпись отв. лица: _____	Дата внесения изм: _____
--------------------	--------------------------	--------------------------

Удаление компонента «pgBadger» доступно через инсталлятор СУБД «Jatoba» – компонент удалится при удалении СУБД «Jatoba». Также компонент возможно исключить из состава установленных компонент СУБД через управление приложениями в панели управления Microsoft Windows.

Для получения детальной информации по установке/удалению программ в Microsoft Windows рекомендуется обратиться к документации по ОС.

4. НАСТРОЙКА JATOWA/POSTGRESQL

Перед началом работы необходимо включить журналирование SQL-запросов в файле postgresql.conf:

```
log_min_duration_statement = 0
```

Включение данного параметра позволит осуществлять регистрацию каждого запроса, длительность которого превышает заданное значение в миллисекундах.



На высоко нагруженном сервере, для увеличения производительности СУБД, рекомендуется увеличить данное значение, чтобы регистрировать только длительные запросы.



Компонент pgBadger поддерживает только определенную комбинацию значений параметров log_statement, log_duration и log_min_duration_statement в конфигурационном режиме. Подробное описание приведено в разделе 6 «Журналирование запросов» данного руководства.

Компонент «pgBadger» поддерживает любой пользовательский формат, заданный в параметре log_line_prefix файла postgresql.conf, если в нем указана временная управляющая последовательность (%t, %m или %n) и управляющая последовательность, связанная с процессом СУБД (%p или %c).

Например, для формата журнала 'stderr' префикс log_line_prefix должен принимать следующее минимальное значение:

```
log_line_prefix = '%t [%p]:'
```

Префикс строки журнала может включать имя пользователя, имя базы данных, имя приложения и IP-адрес клиента:

```
log_line_prefix = '%t [%p]: user=%u,db=%d,app=%a,client=%h'
```

Для формата файла журнала syslog используется следующий префикс:

```
log_line_prefix = 'user=%u,db=%d,app=%a,client=%h'
```

Префикс строки журнала для вывода stderr может принимать значение:

```
log_line_prefix = '%t [%p]: db=%d,user=%u,app=%a,  
client=%h'
```

Для вывода журнала syslog используется следующий префикс:

```
log_line_prefix = 'db=%d,user=%u,app=%a,client=%h'
```

Для журналирования детальной информации о работе СУБД необходимо включить дополнительные параметры в файле postgresql.conf:

```
log_checkpoints = on  
log_connections = on  
log_disconnections = on  
log_lock_waits = on  
log_temp_files = 0  
log_autovacuum_min_duration = 0  
log_error_verbosity = default
```




Сообщения журнала должны быть на английском языке. Компонент «pgBadger» не поддерживает обработку и разбор сообщений на других языках:



```
lc_messages='en_US.UTF-8'  
lc_messages='C'
```

5. ОПЦИИ ИСПОЛЬЗОВАНИЯ КОМПОНЕНТА

В данном разделе приведены опции командной строки компонента «pgBadger» (см. таблицу 5.1) и примеры запуска и использования данных опций.

Таблица 5.1 – Опции командной строки «pgBadger»

Опция	Назначение
-a --average min	количество минут для построения усредненных графиков запросов и соединений. <i>По умолчанию – 5 минут</i>
-A --histo-average min	количество минут для построения усредненных гистограмм запросов. <i>По умолчанию – 60 минут</i>
-b --begin datetime	дата / время начала разбора данных в журнале (задается в формате метки времени (включает дату и время), либо только времени без указания даты)
-c --dbclient host	отчет о записях только для данного клиентского узла
-C --nocomment	удаление комментариев типа /* ... */ из запросов
-d --dbname database	сообщение только о записях для данной базы данных
-D --dns-resolv	замена IP-адресов клиентов их DNS-именами  Это может сильно замедлить работу pgBadger.
-e --end datetime	дата / время окончания анализа данных, разбираемых в журнале (формат значения см. п. 5.1)
-E --explode	разбиение общего отчета на части; опция создает по одному отчету для каждой базы данных
f --format logtype	явное указание формата файлов регистрации событий СУБД, когда pgBadger не может определить их формат автоматически. Возможные значения: syslog, syslog2, stderr, jsonlog, cvs, pgbouncer, logplex, rds и redshift
-G --nograph	отключение графики при выводе HTML-отчета. <i>По умолчанию – включено</i>
-h --help	краткий вывод списка опций
-H --html-outdir path	путь к директории, в которую сохраняется HTML-отчет. В инкрементальном режиме бинарные файлы (см. раздел 8 и 10) остаются в каталоге, который определен опцией -O, --outdir
-i --ident name	имя программы, используемое в качестве идентификатора syslog. <i>По умолчанию: postgres</i>
-I --incremental	использование инкрементального режима. Отчеты будут генерироваться по дням в отдельном каталоге, необходимо задать параметр --outdir
-j --jobs number	количество заданий для одновременного выполнения. <i>По умолчанию запускается одно задание</i>
-J --Jobs number	количество файлов журнала для параллельного разбора. <i>По умолчанию обрабатывается один файл</i>

Опция	Назначение
-l --last-parsed file	разрешение инкрементального разбора журнала, используя последнее разобранное время и строку
-L --logfile-list file	файл, содержащий список файлов журнала для разбора
-m --maxlength size	максимальная длина запроса, которая будет ограничена заданным размером. <i>По умолчанию размер ограничения равен 100000</i>
-M --no-multiline	сокращение размера отчета, опция не собирает данные о многострочных запросах
-n --nohighlight ¹	отключение выделения кода SQL
-N --appname name	составление отчета только по записям для заданного приложения
-o --outfile filename	определение имени файла для вывода. <i>По умолчанию зависит от формата вывода: out.html, out.txt, out.bin, out.json или out.tsung.</i> Опция может быть использована несколько раз для вывода нескольких форматов
-O --outdir path	каталог, в котором будет сохранен файл вывода
-p --prefix string	значение пользовательского префикса log_line_prefix, которое определено в postgresql.conf  Применяется только в случае, если не используется один из стандартных префиксов, указанных в документации pgBadger. Например, если префикс включает дополнительные переменные, такие как IP-клиента или имя приложения.
-P --no-prettyfy	отключение форматирования SQL-запросов
-q --quiet	запрет вывода информации, в том числе строки состояния
-Q --query-numbering	добавление в вывод нумерации запросов при использовании опций --dump-all-queries или --normalized-only
-r --remote-host ip	добавление хоста, на котором будет выполняться команда cat удаленного файла журнала для локального разбора файла
-R --retention N	количество недель для хранения в инкрементальном режиме. <i>По умолчанию 0, отключено</i>  Используется для установки количества недель для хранения данных в выходном каталоге. Более старые недели и дни автоматически удаляются из каталога.
-s --sample number	количество выборок запроса для хранения. <i>По умолчанию: 3</i>
-S --select-only	сообщение только о SELECT-запросах
-t --top number	количество запросов для хранения / отображения. <i>По умолчанию: 20</i>
-T --title string	изменение заголовка отчета HTML-страницы
-u --dbuser username	отчет только по записям для указанного пользователя

¹ данная опция поддерживается только в версии компонента 11.8

Опция	Назначение
-U --exclude-user username	исключение записи для указанного пользователя из отчета. Может быть использовано несколько раз
-v --verbose	включение режима просмотра или отладки. <i>По умолчанию отключен</i>
-V --version	версия pgBadger
-w --watch-mode	сообщение только об ошибках
-W --wide-char	кодирование HTML-вывода запросов в UTF8; позволяет избегать появления сообщения "Wide character in print"
-x --extension	формат вывода. Значения: text, html, bin, json или tsung. <i>По умолчанию: html</i>
-X --extra-files	в инкрементальном режиме позволяет pgBadger записывать CSS и JS файлы в выходной каталог в виде отдельных файлов
-z --zcat exec_path	полный путь к программе zcat. Используется, если zcat, bzcat, или unzip не найдена
-Z --timezone +/-XX	количество часов от GMT для данного часового пояса. Используется для корректировки даты / времени в графиках JavaScript
--pie-limit num	не выводить подписи на круговых диаграммах для значений меньше num (задается в процентах, <i>по умолчанию 2</i>)
--exclude-query regex	запрос, соответствующий заданному регулярному выражению, будет исключен из отчета. Может быть использован несколько раз
--exclude-file filename	путь к файлу, который содержит все регулярные выражения для исключения запросов из отчета. Допускается одно регулярное выражение на строку
--include-query regex	запрос, который не соответствует заданному регулярному выражению, будет исключен из отчета. Может быть использован несколько раз
--include-file filename	путь к файлу, содержащему все регулярные выражения для включения в отчет. Допускается одно регулярное выражение на строку
--disable-error	запрет формирования отчета об ошибках
--disable-hourly	запрет формирования почасового отчета
--disable-type	запрет формирования отчета о запросах по типу, базе данных или пользователю
--disable-query	запрет формирования отчета о запросах (медленные, самые частые, запросы по пользователям, по базам данных и т.п.)
--disable-session	запрет формирования отчета о сессиях
--disable-connection	запрет формирования отчета о соединениях
--disable-lock	запрет формирования отчета о блокировках
--disable-temporary	запрет формирования отчетов о временных событиях
--disable-checkpoint	запрет формирования отчета о контрольной точке / перезагрузке
--disable-autovacuum	запрет формирования отчета об автовакууме
--charset	установка используемой кодировки HTML. По умолчанию: UTF-8
--csv-separator	установка разделителя полей CSV, по умолчанию: ","

Опция	Назначение
--exclude-time regex	любая временная метка, соответствующая заданному регулярному выражению, будет исключена из отчета. Может быть использована несколько раз
--include-time regex	только временная метка, соответствующая заданному регулярному выражению, будет включена в отчет. Может быть использована несколько раз
--exclude-db name	исключение записи журнала для указанной базы данных из отчета. Может быть использована несколько раз
--exclude-appname name	исключение записи журнала для указанного имени приложения из отчета. Может быть использована несколько раз
--exclude-line regex	исключение любой записи журнала, которая соответствует заданному регулярному выражению. Может быть использована несколько раз
--exclude-client name	исключение записи журнала для указанного IP-клиента. Может быть использована несколько раз
--anonymize	скрытие всех фиксированных данных в запросах
--noreport	запрет pgBadger создавать отчеты в инкрементальном режиме
--log-duration	обработка pgBadger записей журнала, созданных как с log_duration = on, так и с log_statement = 'all'
--enable-checksum	добавление md5-суммы под каждым запросом
--journalctl command	добавление к списку файлов регистрации событий СУБД команду вызова системного журнала journalctl для включения его записей в отчет. <i>Например:</i> journalctl -u jatoba-1
--pid-dir path	путь, где должен храниться файл pid. <i>По умолчанию</i> /tmp
--prettify-json	форматирование вывода json
--month-report YYYY-MM	создание кумулятивного HTML-отчета за указанный месяц
--day-report YYYY-MM-DD	создание HTML-отчета за указанный день
--noexplain	не обрабатывать строки, сгенерированные auto_explain
--command CMD	получение записей журнала через стандартный ввод от вывода команды, указанной в параметре
--no-week	отмена еженедельных отчетов в инкрементальном режиме
--explain-url URL	переопределение URL-адреса графического инструмента для EXPLAIN. <i>По умолчанию:</i> http://explain.depesz.com/?is_public=0&is_anon=0&plan=
--tempdir DIR	каталог, в который будут записываться временные файлы. <i>По умолчанию:</i> File::Spec->tmpdir() '/tmp'
--no-process-info	запрет на изменение названия процесса при идентификации процесса pgBadger
--dump-all-queries	сброс всех запросов, найденных в лог-файле
--keep-comments	запрет на удаление комментариев из нормализованных запросов
--no-progressbar	отключение панели прогресса

5.1. Формат значения datetime

Формат datetime может быть задан в формате метки времени (включает дату и время) или в формате времени.

Формат метка времени YYYY-MM-DD hh:mm:ss.SZ.

Формат значения времени hh:mm:ss.SSSZZZ.

YYYY – год, MM – месяц, DD – день, hh – час, mm – минут, ss – секунда, «.S» – миллисекунды, Z – часовой пояс, заданный в формате [+|-] hh[:mm[:ss]]].

Пример.

```
2021-08-27 14:12:45.34567+03.
```

5.2. SSH опции

Компонент pgBadger способен анализировать удаленный файл журнала, используя SSH-соединение. При настройке SSH возможна поддержка аутентификации без пароля, при этом пароль заменяется сгенерированными ключами.

Опция -r или --remote-host задает IP-адрес или имя хоста.

В таблице 5.2 приведены дополнительные опции, позволяющие полностью контролировать SSH-соединение.

Таблица 5.2 – Дополнительные опции для контроля SSH-соединения

Опция	Назначение
--ssh-program ssh	путь к используемой SSH-программе. <i>По умолчанию: ssh</i>
--ssh-port port ssh	порт, используемый для соединения. <i>По умолчанию: 22</i>
--ssh-user name	имя пользователя, логин подключения. <i>По умолчанию: текущий пользователь</i>
--ssh-identity file	путь к используемому файлу идентификации
--ssh-timeout second	время до отказа SSH-соединения. <i>По умолчанию 10 секунд</i>
--ssh-option options	список опций -o, передаваемых SSH-программе. Всегда передаются опции (значение ConnectTimeout берется из параметра --ssh-timeout): -o ConnectTimeout=10 -o PreferredAuthentications=hostbased,publickey

5.3. Примеры использования

В данном разделе приведены общие примеры синтаксиса некоторых опций и специальные примеры, направленные на решение конкретной задачи.

Общие примеры синтаксиса:

```
pgBadger /var/log/jatoba.log
pgBadger /var/log/jatoba.log.2.gz /var/log/jatoba.log.1.gz
/var/log/jatoba.log
pgBadger /var/log/jatoba/jatoba-2021-08-*
pgBadger --exclude-query="^(COPY|COMMIT)" /var/log/jatoba.log
pgBadger -b "2021-09-25 10:56:11" -e "2021-09-25 10:59:11"
/var/log/jatoba.log
cat /var/log/jatoba.log | pgBadger -
pgBadger --prefix '%t [%p]: user=%u,db=%d,client=%h'
/var/lib/jatoba/4/data/log/jatoba-2021-08-21*
pgBadger --prefix '%m %u@d %p %r %a : '
/var/lib/jatoba/4/data/log/jatoba.log
pgBadger --prefix 'user=%u,db=%d,client=%h,appname=%a'
/var/lib/jatoba/4/data/log/jatoba-2021-08-21*
pgBadger -j 8 /var/lib/jatoba/4/data/log/jatoba-10.1-main.log
```

Специальные примеры синтаксиса:

- использование URL для удаленного файла журнала:

```
pgBadger http://10.96.1.50/var/log/jatoba/jatoba-10.1-main.log
pgBadger ftp://username@10.96.1.50/jatoba-10.1-main.log
pgBadger ssh://username@10.96.1.50/var/log/jatoba/jatoba-10.1-
main.log*
```

- совместное использование локальных и удаленных журнальных файлов:

```
pgBadger /var/log/jatoba/jatoba-10.1-main.log ssh
: //username@10.96.1.50/ var/log/ jatoba.log
```

- создание XML-файла в формате Tsung² сессий только с выборочными запросами:

```
pgBadger -S -o sessions.tsung --prefix '%t [%p]: user=%u,db=%d'
' /var/lib/jatoba/4/data/log/jatoba.log
```

² Для использования в системе распределенного нагрузочного тестирования Tsung, <http://tsung.erlang-projects.org/>

- еженедельный отчет об ошибках по заданию cron:

```
30 23 * * * 1 /usr/jatoba-4/bin/pgBadger -q -w  
/var/log/jatoba.log -o /var/reports/jatoba_errors.html
```

- создание еженедельного отчета, используя инкрементальный режим (лог-файл и HTML-отчет обновляются каждую неделю):

```
0 4 * * * 1 /usr/jatoba-4/bin/pgBadger -q `find /var/log/ -  
mtime -7 -name "jatoba.log*" -o /var/reports/jatoba_errors-  
`date +%F`.html -l /var/reports/pgBadger_incremental_file.dat
```

- использование автоматически генерируемых инкрементальных отчетов (генерирует ежедневный и еженедельный отчеты):

```
0 4 * * * /usr/jatoba-4/bin/pgBadger -I -q  
/var/log/jatoba/jatoba.log.1 -O /var/www/html/jatoba_reports/
```

- указание периода в неделях, которые будут храниться в отчетах, в инкрементальном режиме:

```
/usr/jatoba-4/bin/pgBadger --retention 2 -I -q  
/var/log/jatoba/jatoba.log.1 -O /var/www/html/jatoba_reports/
```

- исключение из отчета периода, в который информация не собиралась (если в отчете есть параметр `pg_dump`). Например, `pg_dump` работает в 23:00 и 13:00 каждый день в течение полчаса в сентябре 2021 года. Исключить этот период из отчета возможно следующим образом:

```
pgBadger --exclude-time "2021-09-.* (23|13):.*" jatoba.log
```

Это поможет избежать появления COPY-запросов, генерируемых `pg_dump`, на вершине списка самых медленных запросов.



Параметр `--exclude-appname "pg_dump"` позволяет решить проблему появления COPY-запросов более простым способом.

- вывод `journalctl` может быть разобран, как файл журнала:

```
pgBadger --journalctl 'journalctl -u jatoba-1.service'
```

- вызов journalctl с удаленного хоста:

```
pgBadger -r 10.96.1.50 --journalctl 'journalctl -u jatoba-1.service'
```

- восстановление всех инкрементных HTML-отчетов (обновление всех файлов ресурсов JS и CSS):

```
rm /path/to/reports/*.js  
rm /path/to/reports/*.css  
pgBadger -X -I -O /path/to/reports/ --rebuild
```



Необходимо использовать опцию -E или --explode, если ранее эти отчеты были построены с использованием этой опции.

- для создания кумулятивного отчета за месяц (добавление ссылки в название месяца представления календаря в инкрементальных отчетах) используется команда:

```
pgBadger --month-report 2021-05 -O /path/to/incremental/reports/
```



Необходимо использовать опцию -E или --explode, если ранее инкрементальные отчеты были построены с использованием этой опции.

6. ЖУРНАЛИРОВАНИЕ ЗАПРОСОВ

Журналирование запросов включает в себя параметры – `log_min_duration_statement`, `log_duration` и `log_statement`.

При необходимости включения в статистику запросов продолжительности выполнения всех команд, время работы которых равно или превышает указанное количество миллисекунд, в параметре `log_min_duration_statement` необходимо установить значение равным 0 или больше. По умолчанию отключено (принимает значение «-1»).

Для отображения длительности выполнения всех завершенных команд, нужно установить значение параметра `log_duration = on` в файле `postgresql.conf`.

Значение параметра `log_duration = on` не означает одно и то же, что установка `log_min_duration_statement = 0`. При превышении значения `log_min_duration_statement` в журнал записывается текст запроса, а при включении данного параметра – нет.

При `log_duration = on` и положительном значении `log_min_duration_statement` в журнал записывается длительность всех команд, а текст запросов – только для команд с длительностью, превысившей установленный в параметре предел.

Для добавления отчета о наиболее распространенных запросах, нужно установить `log_min_duration_statement` в значение больше 0, либо включить `log_statement`.

Включение `log_min_duration_statement` добавит отчеты о самых медленных запросах и запросах, которые заняли больше всего времени.



Если для параметра `log_statement` установлено значение 'all', то при включении `log_min_duration_statement` запросы регистрироваться в журнал не будут.



При использовании `log_min_duration_statement` совместно с `log_statement`, текст SQL-запросов будет записываться только один раз (от использования `log_statement`) без дублирования в сообщении о длительности выполнения.

7. ПАРАЛЛЕЛЬНАЯ ОБРАБОТКА

Компонент pgBadger позволяет выполнять параллельную обработку одного или нескольких файлов журнала с помощью параметров -j N или -J N (например, pgBadger -j 8).

В первом случае (-j N) каждый файл журнала разбивается на N частей и обрабатывается отдельным параллельным процессом. Данный режим рекомендуется для обработки файлов журнала большого объема.

Во втором случае (-J N) каждый файл журнала целиком обрабатывается отдельным параллельным процессом (всего будет создано N параллельных процессов). Данный режим рекомендуется для обработки большого количества файлов, каждый из которых имеет сравнительно небольшой размер.

Указанные режимы можно комбинировать (одновременная обработка сразу нескольких файлов, каждый из которых, в свою очередь, обрабатывается несколькими параллельными процессами).

Значения N этих параметров определяется исходя из количества доступных ядер в системе. Использование значений N при параллельной обработке (в том числе при комбинировании режимов), превышающее количество ядер, приведет к потере скорости параллельной обработки.

Ниже приведен пример параллельной обработки, выполненной разными способами:

Вариант 1 – параллельная обработка, выполненная на сервере с 8 процессорами и одним файлом размером 9,5 ГБ:

Вариант	1 CPU	2 CPU	4 CPU	8 CPU
-j	1h41m18	50m25	25m39	15m58
-J	1h41m18	54m28	41m16	34m45

Вариант 2 – параллельная обработка, выполненная на сервере с 8 процессорами и 200 файлами журналов по 10 МБ каждый и общим объемом 2 ГБ:

Вариант	1 CPU	2 CPU	4 CPU	8 CPU
-j	20m15	9m56	5m20	4m20
-J	20m15	9m49	5m00	2m40



При параллельной обработке компонентом pgBadger осуществляется генерация большого количества временных файлов в каталоге /tmp, которые будут автоматически удалены по завершению работы.

Удалять временные файлы в процессе работы утилиты нельзя.

Для идентификации файлов используется следующий шаблон в наименовании: tmp_pgBadgerXXXX.bin.



Параллельная обработка не может использоваться со сжатыми файлами или CSV-файлами, а также, на платформе Windows.

8. ИНКРЕМЕНТАЛЬНЫЕ ОТЧЕТЫ

Компонент «pgBadger» включает режим автоматического инкрементального отчета с помощью опции -I или --incremental.

При работе в этом режиме «pgBadger» генерирует один отчет в день и суммарный отчет в неделю. Сначала вывод осуществляется в двоичном формате в обязательный выходной каталог (см. опцию -O или --outdir таблицы 5.1), затем в формате HTML для ежедневных и еженедельных отчетов с основным индексным файлом.

В главном индексном файле будет отображено выпадающее меню по неделям со ссылкой на каждый недельный отчет и ссылками на ежедневные отчеты каждой недели.

Если запустить «pgBadger» следующим образом на основе ежедневно чередующихся файлов:

```
0 4 * * * /usr/jatoba-4/bin/pgBadger -I -q  
/var/log/jatoba/jatoba.log.1 -O /var/www/html/jatoba_reports/
```

в результате работы будут получены все ежедневные и еженедельные отчеты за весь период работы.

В этом режиме «pgBadger» создаст автоматический инкрементный файл в выходном каталоге. Для изменения пути к данному файлу необходимо использовать опцию -I. «pgBadger» можно запускать в режиме, при котором каждый день работы будет записываться в файл, предназначенный только для записей этого дня недели.

Для экономии дискового пространства используется опция командной строки -X или --extra-files, чтобы компонент «pgBadger» записывал JavaScript и CSS в отдельные файлы в выходном каталоге. Затем эти ресурсы будут загружены с помощью тегов script и link.

9. ЕЖЕМЕСЯЧНЫЕ ОТЧЕТЫ

По умолчанию «pgBadger» в инкрементальном режиме рассчитывает только ежедневные и еженедельные отчеты.

Для получения ежемесячного сводного отчета необходимо использовать отдельную команду.

Например, чтобы построить отчет за август 2021 года необходимо воспользоваться командой:

```
pgBadger -X --month-report 2021-08  
/var/www/html/jatoba_reports/
```

это добавит ссылку на название месяца в календарный вид инкрементных отчетов для просмотра ежемесячного отчета. Отчет за текущий месяц можно запускать ежедневно, при этом каждый раз он полностью перестраивается.



Ежемесячный отчет не строится по умолчанию, так как занимает много времени из-за большого количества данных.

Если отчеты, на основании которых строится месячный отчет (ежедневные или еженедельные), были построены с опцией -E (--explode), то данная опция должна быть использована при вызове pgBadger для построения ежемесячного отчета:

```
pgBadger -E -X --month-report 2021-08 -O  
/var/www/html/jatoba_reports/
```

То же самое происходит при использовании опции перестройки -R (--rebuild).

10. БИНАРНЫЙ ФОРМАТ

Используя бинарный формат можно создавать пользовательские инкрементальные и кумулятивные отчеты.

Пример 1.

При необходимости обновления отчета pgBadger каждый час из ежедневного файла журнала СУБД, для генерации инкрементных файлов данных в двоичном формате, следует выполнять следующие команды:

```
pgBadger --last-parsed .pgBadger_last_state_file -o  
sunday/hourX.bin /var/lib/jatoba/4/data/log/jatoba-Sun.log
```

для генерации инкрементных файлов данных в двоичном формате.

Затем сгенерировать HTML-отчет из этого двоичного файла командой:

```
pgBadger sunday/*.bin
```

Пример 2.

При генерации одного файла журнала в час и необходимости перестроения отчетов при каждой ротации файла журнала необходимо выполнить следующие действия:

```
pgBadger -o day1/hour01.bin /var/lib/jatoba/4/data/log/jatoba-  
2021-08-23_10.log  
pgBadger -o day1/hour02.bin /var/lib/jatoba/4/data/log/jatoba-  
2021-08-23_11.log  
pgBadger -o day1/hour03.bin /var/lib/jatoba/4/data/log/jatoba-  
2021-08-23_12.log
```

Для обновления HTML-отчета, например, каждый раз после создания нового двоичного файла нужно ввести следующую команду:

```
pgBadger -o day1_report.html day1/*.bin
```

и настроить команды в соответствии с конкретными потребностями.

11. ФОРМАТ JSON

Формирование отчета возможно в формате JSON через параметр -o | --output. Данный формат хорошо подходит для интеграции «pgBadger» в другие системы мониторинга.

12. ОТЧЕТЫ, ПОЛУЧАЕМЫЕ С ПОМОЩЬЮ КОМПОНЕНТА

Компонент «pgBadger» собирает данные о SQL-запросах:

Overall statistics	Общая статистика
The most frequent waiting queries	Наиболее частые ожидающие запросы
Queries that waited the most	Запросы с наибольшим временем ожидания
Queries generating the most temporary files	Запросы, генерирующие больше всего временных файлов
Queries generating the largest temporary files	Запросы, генерирующие самые большие временные файлы
The slowest queries	Самые медленные запросы
Queries that took up the most time	Запросы, занимающие больше всего времени
The most frequent queries	Самые частые запросы
Histogram of query times	Гистограмма продолжительности запросов
Histogram of sessions times	Гистограмма продолжительности сессий
Users involved in top queries	Пользователи наиболее частых запросов
Applications involved in top queries	Приложения наиболее частых запросов
Queries generating the most cancellation	Запросы, вызывающие наибольшее количество отмен
Queries most cancelled	Наиболее частые отмененные запросы
The most time consuming prepare/bind queries	Самые трудоемкие запросы на подготовку/привязку параметров

Доступны следующие отчеты с почасовыми графиками, разделенными на периоды по пять минут:

SQL queries statistics	Статистика SQL-запросов
Temporary file statistics	Статистика временных файлов
Checkpoints statistics	Статистика контрольных точек
Autovacuum and autoanalyze statistics	Статистика автовакуума и автоанализа

№ изменения: _____	Подпись отв. лица: _____	Дата внесения изм: _____
--------------------	--------------------------	--------------------------

Cancelled queries	Отмененные запросы
Error events (panic, fatal, error and warning)	События ошибок (паника, фатальная ошибка, ошибка и предупреждение)
Error class distribution	Распределение классов ошибок

Круговые диаграммы о распределении:

Locks statistics	Статистика блокировок
Queries by type (select/insert/update/delete)	Запросы по типу (выбрать/вставить/обновить/удалить)
Distribution of queries type per database/application	Распределение типов запросов по базе данных/приложению
Sessions per database/user/client/application	Сеансы для каждой базы данных/пользователей/клиентов/приложений
Connections per database/user/client/application	Количество подключений для каждой базы данных/пользователей/клиентов/приложений
Autovacuum and autoanalyze per table	Автовакуум и автоанализ по таблице
Queries per user and total duration per user	Запросы пользователя и общая продолжительность запросов пользователя

Все диаграммы можно масштабировать и сохранять в формате PNG. Ко всем SQL-запросам автоматически применяется подсветка синтаксиса и выравнивание.

Точность гистограммы настраивается с помощью ключа - A. По умолчанию в гистограммах отображается среднее значение каждого топ-запроса / ошибки, произошедшей за час, но возможно указать гранулярность вплоть до минут.

ПЕРЕЧЕНЬ СОКРАЩЕНИЙ

HTML	–	HyperText Markup Language — стандартизированный язык разметки документов для просмотра веб-страниц в браузере
IP-адрес	–	Internet Protocol — уникальный числовой идентификатор устройства в компьютерной сети, работающей по протоколу TCP/IP
JSON	–	JavaScript Object Notation — текстовый формат обмена данными, основанный на JavaScript
SSH	–	Secure Shell — сетевой протокол прикладного уровня
SQL	–	Structured Query Language — язык структурированных запросов
XML	–	eXtensible Markup Language — расширяемый язык разметки
ОС	–	Операционная система
СУБД	–	Система управления базами данных

Лист регистрации изменений

№ изменения: _____	Подпись отв. лица: _____	Дата внесения изм: _____
--------------------	--------------------------	--------------------------