# Microsoft Copilot: ваш ИИ-помощник



copilot.microsoft.com/chats/mpKkpH2UueJVNRugQepF5



## Сегодня

du -sh /var/lib/pgsql/14/messenger/base/16709033/23835271\* необходимо подсчитать суммарный реальный размер на диске в байтах

Чтобы определить суммарный реальный размер файлов на диске в байтах, можно использовать команду:

bash

du-cb /var/lib/pgsql/14/messenger/base/16709033/23835271\*

Флаг -с суммирует общий размер, а -b отображает размер в байтах. В конце вывода будет указана строка total, содержащая общий объем занимаемого пространства. Попробуйте и скажите, получилось ли у вас! 🚀

18619056128 байт в Гб, три знака после запятой

Чтобы перевести байты в гигабайты с тремя знаками после запятой, нужно разделить на 102431024^3:

18619056128÷1073741824≈17.338 ГБ 18619056128 \div 1073741824 \approx 17.338 \text{ ГБ}

Таким образом, размер составляет 17.338 ГБ. 🚀

Postgresql тип данных boolean сколько байт занимает с учетом выравнивания

B PostgreSQL тип данных boolean занимает **1 байт**. Однако, с учетом выравнивания, его фактический размер в строке таблицы может увеличиваться.

На 64-битных системах PostgreSQL использует выравнивание на 8 байт для данных в строке. Это означает, что если boolean стоит после типа данных с более строгими требованиями к выравниванию, может добавляться заполняющий байт. Например, если boolean идет после int2 (который требует 2-байтового выравнивания), то общий размер строки увеличится на 3 байта: 1 байт для boolean и 2 байта для int2.

Если boolean используется в одиночной колонке, его размер будет **1 байт**, но в зависимости от расположения в строке таблицы и соседних типов данных, фактический размер может быть больше из-за выравнивания.

Если вам нужно точно определить размер, можно использовать функцию:

sql

```
SELECTpg_column_size(row(true));
```

Это покажет реальный размер хранения boolean в контексте строки. 💉

SELECT pg\_column\_size(row(true)); для какой таблицы?

Функция pg\_column\_size(row(true)) не требует конкретной таблицы, так как row(true) создаёт временный кортеж и измеряет его размер. Это позволяет узнать, сколько байт занимает boolean в контексте строки без необходимости обращаться к реальным таблицам.

Если же вы хотите определить размер boolean в существующей таблице, можно использовать:

sal

```
SELECTpg_column_size(my_boolean_column) FROMmy_table LIMIT 1;
```

### Где:

- my\_boolean\_column название столбца типа boolean.
- my\_table имя таблицы.

Так вы получите реальный размер хранения boolean в контексте строки вашей таблицы. 🖋

SELECT pg\_column\_size(row(true)); = 25 SELECT pg\_column\_size(click\_text) FROM messages LIMIT 1; = 1

Интересные результаты! 🤔

Функция рд\_column\_size(row(true)) = 25 показывает, что при упаковке boolean в кортеж (row()) он занимает больше места из-за служебной информации о строке. Стандартная строка в PostgreSQL содержит заголовок (обычно 24 байта на 64-битных системах), плюс 1 байт на сам boolean, вот и выходит 25 байт.

A вот SELECT pg\_column\_size(click\_text) FROM messages LIMIT 1; = 1 говорит о том, что в вашей таблице messages поле click\_text занимает 1 байт, что может означать, что оно хранит либо NULL, либо очень короткое значение.

Если интересно детальнее разобраться, можно проверить:

sql

```
SELECTpg_column_size(row(click_text)) FROMmessages LIMIT 1;
```

Это покажет, сколько байтов click\_text занимает в контексте строки, аналогично row(true). 💉

Таблица CREATE TABLE IF NOT EXISTS public.messages ( id bigint NOT NULL DEFAULT nextval('messages\_id\_seq'::regclass), profile\_id bigint NOT NULL, template\_id bigint NOT NULL, send\_status boolean, delivery\_status boolean, close\_status boolean, click\_status boolean, read\_status boolean, remove\_status boolean, transaction\_time bigint NOT NULL DEFAULT 0, click\_text boolean DEFAULT false, click\_button2 boolean DEFAULT false, click\_button1 boolean DEFAULT false, click\_button3 boolean DEFAULT false, attributes jsonb DEFAULT '{}`::jsonb, time\_to\_send bigint, created\_at bigint DEFAULT 0, trigger\_uuid text COLLATE pg\_catalog."default" NOT NULL DEFAULT "::text, try\_integration\_send boolean, repeat integer NOT NULL DEFAULT 0, CONSTRAINT messages\_pkey PRIMARY KEY (profile\_id, template\_id, trigger\_uuid, repeat)) WITH (FILLFACTOR = 90, autovacuum\_analyze\_scale\_factor = 0.005, autovacuum\_analyze\_threshold = 50, autovacuum\_vacuum\_scale\_factor = 0.001) TABLESPACE pg\_default; каков расчетный размер одного кортежа?

Чтобы рассчитать размер одного кортежа в PostgreSQL, нужно учесть следующие факторы:

- 1. Заголовок строки: На 64-битных системах занимает 24 байта.
- 2. Размер данных: Сумма размеров всех колонок, с учетом выравнивания.
- 3. Выравнивание: PostgreSQL выравнивает данные для эффективного хранения.
- 4. TOAST: Поля типа jsonb и text могут храниться отдельно.

## Оценка размеров столбцов:

Тип данных	Столбец(-цы)	Размер (байт)
bigint	<pre>id, profile_id, template_id, transaction_time, time_to_send, created_at, repeat</pre>	8 × 7 = 56
boolean	8 столбцов boolean	1 × 8 = 8

Тип данных	Столбец(-цы)	Размер (байт)
jsonb	attributes	Зависит от данных (минимум 1 байт)
text	trigger_uuid	Зависит от данных (минимум 4 байта)
integer	repeat	4 байта
Итого (без выравнивания)		≈ 93+ байта

# Выравнивание:

- bigint требует **8-байтового** выравнивания.
- boolean может занимать больше места из-за выравнивания.
- jsonb и text могут быть вынесены в TOAST, снижая размер.

## Итоговая оценка:

- Базовый размер **≈ 93 байта + выравнивание**.
- Минимально около 100-120 байт на строку (без TOAST).
- Если jsonb и text хранятся отдельно, основной кортеж может быть до 150 байт.

Для точного измерения можно выполнить:

sql

```
SELECTpg_column_size(row(*)) FROMmessages LIMIT 1;
```

Это покажет реальный размер кортежа с учетом выравнивания и хранимых данных. 💉

[42601] ERROR: syntax error at or near "\*"

Ax, небольшая ошибка! В PostgreSQL нельзя использовать row(\*). Вместо этого попробуйте:

```
SELECTpg_column_size(row(id, profile_id, template_id, send_status, delivery_status,
   close_status, click_status, read_status, remove_status, transaction_time, click_text,
   click_button2, click_button1, click_button3, attributes, time_to_send, created_at,
   trigger_uuid, try_integration_send, repeat))
   FROMmessages LIMIT 1;
Здесь я вручную перечислил все столбцы, чтобы создать корректный гом(). Это должно дать точный
размер кортежа. 🖋 Если PostgreSQL все равно ругается, попробуйте просто:
sql
   SELECTpg_column_size(messages.*) FROMmessages LIMIT 1;
Этот вариант измеряет размер всей строки в messages. Попробуйте и расскажите, как получилось! 🔍
       104
Отлично! Это значит, что реальный размер одного кортежа в вашей таблице messages составляет 104
байта. 🎯
Этот результат включает все столбцы, заголовок строки и выравнивание. Возможно, значения jsonb и
text (например, attributes и trigger_uuid) не увеличивают размер строки, так как они хранятся в
TOAST. Если бы эти поля содержали длинные значения, размер кортежа мог бы быть больше.
Если хотите проверить средний размер строк в таблице, можно использовать:
sql
   SELECTpg_relation_size('messages') /pg_table_size('messages') ASavg_tuple_size;
Это даст средний размер строк в вашей таблице, включая возможные TOAST-данные. 🚀 Как вам такие
расчёты? 😊
       SELECT pg relation size('messages') / pg table size('messages') AS avg tuple size; 0
Значение 💿 означает, что рg_relation_size('messages') равно 💿 или гораздо меньше
pg_table_size('messages'), поэтому результат округляется до нуля.
Попробуйте вместо этого:
sql
```

```
SELECTpg_relation_size('messages')::numeric/pg_table_size('messages')::numericASavg_tuple_size;
```

#### 0.99972193318036725059

Отлично, теперь мы видим, что средний размер кортежа в вашей таблице messages почти полностью соответствует размеру реальных данных! Значение ≈ 1 говорит о том, что большинство данных хранятся непосредственно в основном отношении таблицы, а не вынесены в TOAST.

Если jsonb или text -поля содержат длинные значения, они могут быть вынесены в TOAST, и тогда средний размер кортежа мог бы быть **меньше 1**. В вашем случае всё компактно хранится, что может быть хорошо для быстрого доступа.

Хотите проверить, есть ли TOAST-записи для messages? Можно выполнить:

sql

```
SELECTpg_relation_size('messages') AStable_size,
    pg_total_relation_size('messages') AStotal_size;
```

table size;total size 18613886976;35679387648

Интересные цифры! 📊

Ваши данные показывают, что:

- Размер основной таблицы ( table\_size ): ~18.6 ГБ
- Общий размер ( total\_size ): ~35.7 ГБ

Разница (~17.1 ГБ) говорит о том, что значительная часть данных **вынесена в TOAST или занята индексами**. Это может означать:

- 1. Поля jsonb (attributes) и text (trigger\_uuid) хранятся в TOAST.
- 2. Индексы занимают значительное пространство.

Если хотите детально изучить TOAST, попробуйте:

```
SELECTrelname, pg_relation_size(oid) ASsize
FROMpg_class
WHERErelname LIKE'pg_toast_%';
```

```
relname;size pg_toast_19499610;0 pg_toast_16709037;0 pg_toast_16709037_index;8192
pg toast 19499610 index;8192 pg toast 16709046;0 pg toast 16709046 index;8192
pg_toast_16709055;0 pg_toast_16709055_index;8192 pg_toast_17017233;0
pg toast 16709060;0 pg toast 16709060 index;8192 pg toast 17017233 index;8192
pg toast 19499630;0 pg toast 19499630 index;8192 pg toast 19499665;0
pg toast 19499665 index;8192 pg toast 1255;8192 pg toast 1255 index;16384
pg toast 1247;0 pg toast 1247 index;8192 pg toast 2604;0 pg toast 2604 index;8192
pg toast 2606;0 pg toast 2606 index;8192 pg toast 2612;0 pg toast 2612 index;8192
pg toast 2600;0 pg toast 2600 index;8192 pg toast 2619;458752
pg toast 2619 index;49152 pg toast 3381;0 pg toast 3381 index;8192 pg toast 3429;0
pg toast 3429 index;8192 pg toast 2618;540672 pg toast 2618 index;16384
pg toast 2620;0 pg toast 2620 index;8192 pg toast 3466;0 pg toast 3466 index;8192
pg toast 2609;0 pg toast 2609 index;8192 pg toast 2615;0 pg toast 2615 index;8192
pg toast 1262;0 pg toast 1262 index;8192 pg toast 2964;0 pg toast 2964 index;8192
pg_toast_1213;0 pg_toast_1213_index;8192 pg_toast_1260;0 pg_toast_1260_index;8192
pg toast 2396;0 pg toast 2396 index;8192 pg toast 3600;0 pg toast 3600 index;8192
pg toast 3079;0 pg toast 3079 index;8192 pg toast 2328;0 pg toast 2328 index;8192
pg toast 1417;0 pg toast 1417 index;8192 pg toast 1418;0 pg toast 1418 index;8192
pg_toast_3118;0 pg_toast_3118_index;8192 pg_toast_3256;0 pg_toast_3256_index;8192
pg toast 6000;0 pg toast 6000 index;8192 pg toast 826;0 pg toast 826 index;8192
pg toast 3394;0 pg toast 3394 index;8192 pg toast 3596;0 pg toast 3596 index;8192
pg toast 3592;0 pg toast 3592 index;8192 pg toast 3456;0 pg toast 3456 index;8192
pg toast 3350;0 pg toast 3350 index;8192 pg toast 6100;0 pg toast 6100 index;8192
pg_toast_19499693;0 pg_toast_17017259;0 pg_toast_19499693_index;8192
pg toast 17017259 index;8192 pg toast 17017297 index;8192
pg toast 17017240 index;8192 pg toast 17017297;0 pg toast 17017271;0
pg toast 17017271 index;8192 pg toast 17017240;0 pg toast 20860887;0
pg_toast_20860887_index;8192 pg_toast_17017291;0 pg_toast_17017291_index;8192
pg_toast_17017333_index;16384 pg_toast_17017333;65536 pg_toast_14172;0
pg toast 14172 index;8192 pg toast 14167;0 pg toast 14167 index;8192 pg toast 14177;0
pg toast 14177 index;8192 pg toast 14182;0 pg toast 14182 index;8192
pg toast 17017342;0 pg toast 17017342 index;8192 pg toast 17017357;0
pg_toast_17017357_index;8192 pg_toast_17017363;0 pg_toast_17017363_index;8192
pg toast 17017392;0 pg toast 17017392 index;8192 pg toast 17017405;0
pg toast 17017405 index;8192 pg toast 17017426;0 pg toast 17017426 index;8192
pg toast 17017454 index;49152 pg toast 17017454;827392 pg toast 17017437;0
pg_toast_17017437_index;8192
```

Ваши TOAST-таблицы, связанные с таблицей messages, практически не занимают место — большинство из них имеют размер **0 байт**, а их индексы — по **8192 байта** (8 КБ). Это говорит о том, что длинные значения (например, jsonb и text) не вынесены в TOAST, а хранятся непосредственно в основной таблице.

Единственные TOAST-таблицы с заметным размером:

```
• pg_toast_2619 — 458752 байта (~448 КБ)
```

- pg\_toast\_2618 **540672 байта** (~528 КБ)
- pg\_toast\_17017333 **65536 байт** (~64 КБ)
- pg\_toast\_17017454 **827392 байта** (~808 КБ)

Анализируя это, можно сделать вывод, что:

- В таблице messages хранимые данные в большинстве случаев не превышают порога TOAST (~2 КБ).
- Небольшая часть данных вынесена в TOAST, но их объем незначителен по сравнению с общей структурой таблицы.

Хотите проверить размер индексов отдельно? Можно выполнить:

sql

```
SELECTpg_size_pretty(pg_indexes_size('messages')) ASindex_size;
```

Это покажет, сколько места занимает **индексная структура** вашей таблицы. 🚀 Любопытно, насколько она велика! 😃

Единственные TOAST-таблицы с заметным размером: pg\_toast\_2619 — 458752 байта (~448 КБ) pg\_toast\_2618 — 540672 байта (~528 КБ) pg\_toast\_17017333 — 65536 байт (~64 КБ) pg\_toast\_17017454 — 827392 байта (~808 КБ) какие таблицы используют TOAST-таблицы

Чтобы определить, какие таблицы используют эти TOAST-таблицы, можно выполнить следующий запрос:

sql

```
SELECTrelname, reltoastrelid
FROMpg_class
WHEREreltoastrelid IN(2619, 2618, 17017333, 17017454);
```

Этот запрос ищет таблицы, у которых reltoastrelid соответствует идентификаторам TOAST-таблиц. reltoastrelid — это колонка, которая хранит идентификатор TOAST-таблицы, привязанной к основной таблице.

```
SELECTc.relname AStable_name, t.relname AStoast_table_name
FROMpg_class c
LEFTJOINpg_class t ONc.reltoastrelid =t.oid
WHEREt.relname ISNOTNULL;
```

Этот вариант покажет все таблицы, использующие TOAST, вместе с их TOAST-таблицами. Будет любопытно взглянуть на результат! 😃

```
SELECT relname, reltoastrelid FROM pg_class WHERE reltoastrelid IN (2619, 2618, 17017333, 17017454); HET
```

Если запрос не вернул результатов, это означает, что указанные TOAST-таблицы **не привязаны к каким-либо текущим таблицам** в pg\_class . Возможные причины:

- 1. **Удаленные таблицы**: Если связанные таблицы были удалены, TOAST-таблицы могли остаться, но без привязки.
- 2. **Системные или временные данные**: Некоторые TOAST-таблицы могут быть связаны с системными объектами или временными данными.
- 3. **Ошибка в идентификаторах**: Возможно, стоит проверить oid для TOAST-таблиц заново.

Попробуйте расширенный запрос для поиска всех таблиц с TOAST:

sql

```
SELECTc.relname AStable_name, t.relname AStoast_table_name
FROMpg_class c
LEFTJOINpg_class t ONc.reltoastrelid =t.oid
WHEREt.relname ISNOTNULL;
```

Этот запрос покажет все таблицы, которые **имеют привязку** к TOAST-таблицам. Давайте посмотрим, какие результаты он вернёт! 💉 Будет любопытно разобраться, что именно занимает место в TOAST. 😄

table name; toast table name auth providers; pg toast 19499610 pg statistic; pg toast 2619 pg\_type;pg\_toast\_1247 send\_telegram\_messages\_by\_templates;pg\_toast\_19499665 pg foreign table;pg toast 3118 pg authid;pg toast 1260 telegram bots;pg toast 19499693 telegram subscribers;pg toast 19499630 pg statistic ext data;pg toast 3429 audiences;pg toast 20860887 pg user mapping;pg toast 1418 pg\_subscription;pg\_toast\_6100 pg\_proc;pg\_toast\_1255 pg\_attrdef;pg\_toast\_2604 pg constraint;pg toast 2606 pg language;pg toast 2612 pg aggregate;pg toast 2600 pg statistic ext;pg toast 3381 pg rewrite;pg toast 2618 pg trigger;pg toast 2620 pg\_event\_trigger;pg\_toast\_3466 pg\_description;pg\_toast\_2609 pg\_namespace;pg\_toast\_2615 pg\_database;pg\_toast\_1262 pg\_db\_role\_setting;pg\_toast\_2964 pg\_tablespace;pg\_toast\_1213 pg shdescription;pg toast 2396 pg ts dict;pg toast 3600 pg extension;pg toast 3079 pg\_foreign\_data\_wrapper;pg\_toast\_2328 pg\_foreign\_server;pg\_toast\_1417 pg\_policy;pg\_toast\_3256 pg\_replication\_origin;pg\_toast\_6000 pg\_default\_acl;pg\_toast\_826 pg\_init\_privs;pg\_toast\_3394 pg\_seclabel;pg\_toast\_3596 pg\_shseclabel;pg\_toast\_3592 pg\_collation;pg\_toast\_3456 pg\_partitioned\_table;pg\_toast\_3350 sql\_parts;pg\_toast\_14177 sql\_implementation\_info;pg\_toast\_14172 sql\_features;pg\_toast\_14167 sql\_sizing;pg\_toast\_14182 audiences\_old;pg\_toast\_17017240

```
app_healths;pg_toast_17017233 event_lists;pg_toast_16709037 message_templates;pg_toast_17017297 integration_providers;pg_toast_16709046 integrations;pg_toast_17017271 matching_ttls;pg_toast_17017291 firebase_subscribers;pg_toast_17017259 messages;pg_toast_17017342 message_translates;pg_toast_17017333 nodes;pg_toast_17017357 projects;pg_toast_16709055 route_providers;pg_toast_16709060 send_firebase_messages_by_templates;pg_toast_17017392 statistic_templates;pg_toast_17017405 tags;pg_toast_17017426 profiles;pg_toast_17017363 trigger_events;pg_toast_17017437 workflows;pg_toast_17017454
```

Отличная работа! Теперь у вас есть список таблиц, которые используют TOAST. 🚀

Некоторые интересные наблюдения:

- Системные таблицы: pg\_statistic, pg\_proc, pg\_namespace и другие PostgreSQL использует TOAST для хранения длинных системных метаданных.
- Пользовательские таблицы: messages, message\_templates, audiences, profiles, workflows это говорит о том, что они содержат длинные данные, например jsonb, text или bytea.
- Telegram, Firebase, Analytics: send\_telegram\_messages\_by\_templates, firebase\_subscribers, statistic\_templates TOAST помогает хранить сообщения, подписчиков и статистику.

Если хотите выяснить, какие столбцы используют TOAST в каждой таблице, попробуйте:

sql

```
\label{lem:selectative} SELECTattrelid::regclass \ AStable\_name, \ attname \ AScolumn\_name \\ FROMpg\_attribute \\ WHEREattstorage \ IN('x', 'e');
```

Это покажет **конкретные столбцы** в таблицах, которые используют TOAST для хранения данных. Будет интересно посмотреть, какие именно поля выделяются! 

Если нужно разобрать конкретную таблицу глубже, можно анализировать ресоlumn\_size() на её столбцах.

SELECT attrelid::regclass AS table\_name, attname AS column\_name FROM pg\_attribute WHERE attstorage IN ('x', 'e'); для таблицы messages

Чтобы получить только столбцы из таблицы messages, добавьте условие attrelid = 'messages'::regclass':

```
SELECTattname AScolumn_name
FROMpg_attribute
WHEREattrelid ='messages'::regclass
ANDattstorage IN('x', 'e');
```

Этот запрос покажет **только те столбцы**, которые используют TOAST в <u>messages</u>. 

Попробуйте его и скажите, что получилось! 

Если таблица не содержит TOAST-колонок, результат может быть пустым.

column\_name attributes trigger\_uuid

Отлично, теперь мы знаем, какие столбцы в messages используют TOAST:

- attributes (типа jsonb)
- trigger\_uuid (типа text)

Оба этих типа данных могут быть **довольно длинными**, и PostgreSQL автоматически перемещает их в TOAST, если они превышают определённый порог (обычно ~2 КБ). Это помогает оптимизировать хранение и уменьшить нагрузку на основную таблицу.

Если хотите посмотреть, какие конкретные строки вынесены в TOAST, можно попробовать:

sql

```
SELECTCOUNT(*) FROMmessages WHERELENGTH(attributes::text) >2000;
```

Этот запрос покажет количество строк, где <u>attributes</u> хранит больше 2000 символов, что делает их кандидатами на TOAST.

SELECT COUNT(\*) FROM messages WHERE LENGTH(attributes::text) > 2000; для таблицы messages

Попробуйте выполнить этот запрос, чтобы определить количество строк в messages, где поле attributes содержит длинные JSON-данные, превышающие 2000 символов:

sql

```
SELECTCOUNT(*) FROMmessages WHERELENGTH(attributes::text) >2000;
```

Этот запрос конвертирует jsonb в text и проверяет его длину. Если таблица содержит длинные JSON-объекты, они могут быть вынесены в TOAST. 💉

Если хотите проверить, какие именно значения занимают много места, можно сделать:

```
SELECTId, LENGTH(attributes::text) ASattr_size
FROMmessages
ORDERBYattr_size DESCLIMIT 10;
```

Это покажет **10 самых больших** записей. Попробуйте и скажите, что получилось! **4** Любопытно, насколько велик самый длинный JSON.

column name attributes trigger uuid

Ваш запрос подтвердил, что два столбца в таблице messages — attributes (jsonb) и trigger\_uuid (text) — используют TOAST для хранения длинных данных. Это означает, что если значения в этих колонках превышают определённый размер (обычно 2 КБ), PostgreSQL автоматически перемещает их в TOAST-таблицу для оптимизации хранения.

Если хотите подробнее изучить распределение размеров attributes, попробуйте:

sql

```
SELECTLENGTH(attributes::text) ASattr_size, COUNT(*)
FROMmessages
GROUPBYattr_size
ORDERBYattr_size DESCLIMIT 10;
```

Этот запрос покажет **10 самых часто встречающихся размеров** JSON-данных в таблице. **Я** Любопытно узнать, насколько велики эти объекты!  $\stackrel{\smile}{\hookrightarrow}$