

Тихий OOM Killer: как 768 МБ на сессию утопили 16 ГБ памяти

Aurélien LEQUOY · April 2, 2026

MARIADB

OOM-KILLER

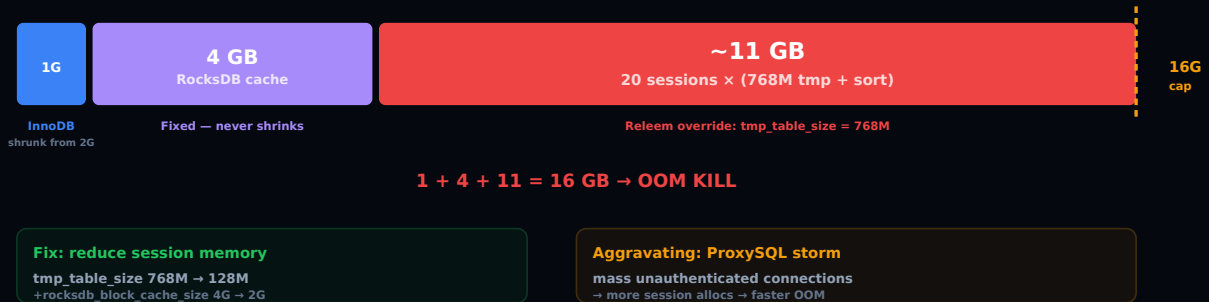
PERFORMANCE-TUNING

SYSTEMD

MEMORY-MANAGEMENT

MEMORY STACK AT OOM KILL

16 GB systemd cap — what fills it up



Calculate worst case, not average — PmaControl memory monitoring

Краткое резюме

MariaDB останавливается не из-за коррупции, не из-за проблемы Galera и не из-за бага SQL. Ядро Linux **убивает процесс mariadbd** за превышение памяти.

Доказательство явно представлено в systemd и логе ядра:

```
mariadb.service: Failed with result 'oom-kill'  
Out of memory: Killed process 1177 (mariadbd) total-vm:22267612kB, anon-rss:16649820kB  
Memory cgroup out of memory: Killed process 1146610 (mariadbd)
```

Окружение

Компонент	Значение
Всего RAM	19.5 GB
Swap	~1 GB
systemd MemoryMax	16 GB

Компонент	Значение
innodb_buffer_pool_size	2 GB (auto-shrink → 1 GB)
rocksdb_block_cache_size	4 GB
tmp_table_size (Releem override)	768 MB
max_heap_table_size (Releem override)	768 MB
sort_buffer_size	32 MB
max_connections	100

Что происходит перед kill

MariaDB обнаруживает давление памяти и пытается защититься, уменьшая buffer pool
InnoDB:

```
Memory pressure event shrunk innodb_buffer_pool_size=1536m from 2048m
→ 1280m → 1152m → 1088m → 1056m → 1040m → 1032m → 1024m
Memory pressure event disregarded; innodb_buffer_pool_size=1024m,
innodb_buffer_pool_size_auto_min=1024m
```

InnoDB уже уменьшил свой buffer pool до минимума (1 ГБ). Но этого недостаточно.
Остальные потребители не отступают.

Расчёт наихудшего случая

При 100 одновременных соединениях наихудший случай потребления памяти на сессию:

```
100 × (768 MB tmp_table + 768 MB heap + 32 MB sort) = ~153 GB
```

Разумеется, не все сессии создают временные таблицы на 768 МБ. Но достаточно **20 сессий**, выполняющих запросы с `GROUP BY` или `ORDER BY` на больших объёмах данных, чтобы превысить 16 ГБ:

```
InnoDB buffer pool : 1 GB (shrunk)
RocksDB cache : 4 GB (фиксирован, не уменьшается)
20 sessions × 768 MB : 15 GB
```

```
Total : 20 GB → kill
```

Усугубляющий фактор: connection storm ProxySQL

Прямо перед OOM лог MariaDB показывает массовые прерванные соединения от 10.68.68.103 (ProxySQL):

```
Aborted connection ... user: 'unauthenticated' host: '10.68.68.103'  
Too many connections
```

Больше соединений = больше сессионной памяти = больше давления.

Исправление

Немедленные действия

1. Уменьшить сессионную память:

```
tmp_table_size = 128M  
max_heap_table_size = 128M  
sort_buffer_size = 8M
```

2. Поднять лимит systemd:

```
MemoryMax=18G
```

3. Проверить кэш RocksDB — 4 ГБ возможно избыточен:

```
rocksdb_block_cache_size = 2G
```

Действия на среднесрочную перспективу

- Удалить файл Releem, который переопределяет значения (`/etc/mysql/releem.conf.d/z_aiops_mysql.cnf`)
- Мониторить `memory_mysqlld` через PmaControl для предупреждения перед kill
- Настроить ProxySQL с `max_connections` на стороне бэкенда ниже, чем `max_connections` MariaDB

Чем это не является

Это **не**:

- ошибка запуска
- сломанное восстановление Galera
- повреждённый datadir
- проблема файловых дескрипторов

MariaDB перезапустилась чисто и немедленно вернулась в состояние `active (running)`.

Заключение

Инструмент автоматической настройки (Releem) установил `tmp_table_size` в 768 МБ — значение, кажущееся разумным в изоляции. Но в сочетании с лимитом `systemd` в 16 ГБ, кэшем RocksDB в 4 ГБ и `connection storms` ProхуSQL оно становится бомбой замедленного действия.

Память сервера MariaDB нужно рассчитывать по наихудшему случаю, а не по среднему.