

Управление памятью в Greenplum

По мотивам [тикета](#) в Яндекс.Облако и встречи с архитектором.

Проблема

У нас достаточно мощные сегмент-хосты (576Гб ОЗУ), но для пользователя dwh_etl выделяется всего 98Мб памяти на запрос. Вопрос:

[blocked URL](#)

Вводные

```
--
select *
from gp_toolkit.gp_resgroup_config;

--
select r.rolname, r.rolconfig, rsg.*
from pg_roles as r
    inner join
        gp_toolkit.gp_resgroup_config as rsg on rsg.groupid = r.rolresgroup
where r.rolname = current_user;

--
show statement_mem;
show gp_resource_group_memory_limit;

--
explain analyze
select 1;
```

Как управлять памятью

Способ 1. Через ресурсные группы

[Документация](#)

Количество памяти, которая выделяется на запрос согласно ресурсной группе, вычисляется по формуле:

*(память на сегмент-хосте * значение параметра gp_resource_group_memory_limit / кол-во сегментов (без зеркал)) * процент памяти ресурсной группы * (1-memory_shared_quota) * memory_spill_ratio*

где:

- память на сегмент-хосте = 576Гб
- show gp_resource_group_memory_limit = 0.8
- кол-во сегментов (без зеркал) = 8
- процент памяти ресурсной группы = 20
- memory_shared_quota = 80
- memory_spill_ratio = 100

Где-то здесь потерялось значение concurrency, но при установленном memory_spill_ratio оно начинает работать, как положено (уменьшаем concurrency - увеличивается память на запрос).

```
select (576 * 0.8 / 8) * 0.2 * 0.2 * 1; -- 2.3 ,      Memory used:  2 326 528 kB
```

При этом важно проверять, что количество памяти, выделяемой на запрос * concurrency не превосходит количество памяти, доступное ресурсной группе.

Количество памяти, доступное ресурсной группе, вычисляется по формуле:

(память на сегмент-хосте * значение параметра `gp_resource_group_memory_limit` / кол-во сегментов (без зеркал)) * процент памяти ресурсной группы

где:

- память на сегмент-хосте = 576Гб
- `show gp_resource_group_memory_limit` = 0.8
- кол-во сегментов (без зеркал) = 8
- процент памяти ресурсной группы = 20

Проверить можно так:

```
select (576 * 0.8 / 8) * 0.2; -- 11.5
```

Важно! Если в настройках ресурсной группы установлено `memory_spill_ratio = 0`, то формула выше по расчету памяти на запрос работать не будет. Вместо этого на запрос буде выделяться память, указанная в параметре `statement_mem`.

Способ 2. Через память для роли

Для роли можно установить значение памяти, которое будет использоваться в запросах. Запросы для этого:

```
alter role <role_name> set statement_mem = '200MB'; --
alter role <role_name> reset statement_mem; --
```

И опять же можно выстрелить в ногу, если количество памяти, выделяемой на запрос * `concurrency` будет превосходить количество памяти ресурсной группы.

Важно! Способ допустим только в исключительных случаях (главная etl-учетка). Запрещено использовать эту настройку для пользовательских учеток. Найденные пользовательские учетки будут баниться