1. Тест на позицию Junior DB Developer ( SQL )

Таблица T содержит дубли в поле n:

select \* from T

|  |  |  |
| --- | --- | --- |
| N | dt | Msg (clob) |
| 1 | 08.07.2019 |  |
| 1 | 09.07.2019 |  |
| 2 | 08.07.2019 |  |
| 2 | 10.07.2019 |  |
| 3 | 10.07.2019 |  |
| 3 | 10.07.2019 |  |
| … | … | … |

Написать DML запрос, который очистит таблицу от старых дублей, чтобы она имела вид

select \* from T

|  |  |  |
| --- | --- | --- |
| N | dt | Msg(clob) |
| 1 | 09.07.2019 |  |
| 2 | 10.07.2019 |  |
| 3 | 10.07.2019 |  |
| … | .. | … |

1. Тест на позицию Senior DB Developer ( SQL, PL/SQL)

В High-load OLTP системе из-за несовершенства учетной системы и каналов связи в секционированной (PARTITION BY HASH (n)) таблице T из п.1 накопилось за 5 лет 10 млрд. записей c ~30% дублей

Размер секции - 100 млн. записей, размер поля msg = 1Кб

N принимает значения в диапазоне 0..9 999 999 999

Размер табличного пространства для хранения данных в T ограничен 10 Тб.

Написать максимально эффективную хранимую процедуру R, которая очистит таблицу от дублей до вида из п.1, с учетом следующих ограничений ресурсов

1. размер PGA на сервере = 1 Гб и данный ресурс доступен для R монопольно в полном объеме
2. на диске для файлов БД свободно 100 Гб.
3. Объем PGA, выделяемый для WITH и/или вложенных подзапросов = сумме размеров выборок каждого подзапроса, а не размеру результирующей выборки ( внешнего запроса ), т.е. WITH a as (SELECT ‘abc’ as x FROM DUAL) SELECT x FROM a требует 6 байт ОЗУ.
4. Размером памяти на сортировку пренебречь

СУБД – Oracle