

Proposal for Automated SQL Execution Advisory in Database

Vaheedbasha Shaik¹

¹Senior Oracle Database Administrator, IT Department, CenturyLink Technologies Pvt. Ltd.

Abstract - This paper explains about adding a new feature to the database for SQL Query executions. Usually, the database would execute the queries which are received from the different ways. Some of them are through applications directed queries, user sessions by directly connecting to the database and other one is internally generated queries which will be triggered within the database by an event.

The Database will execute the queries whatever it reaches to it through session. The current database engine never restricts any SQL from being execute or Propose any suitable timeline to execute it. Let's consider that our Database is running at peak business hours and suddenly some other has triggered highly intensive query, Now the database going to face high load which may leads to outage for the business.

The proposal here is to control such Queries executions depending upon their Priorities and System load.

Key Words: OLTP, OLAP, OPTIMIZER, SQL, PROFILE.

1. INTRODUCTION

The database is a collection and combination of objects. The database binaries come with some system objects and the users will create some user defined objects as per their requirement. The system objects are coming with software binaries to maintain the DB related data. The user defined objects will be used to store user's data after creation in the database.

Let's consider the Oracle database as example though out this paper. The Oracle DB has the schemas to keep a group of objects under a one owner ship. Let's say two applications are using the same database then oracle will keep two schemas to keep objects of each application. We can get the data of one schema's object to another schema's by granting the privileges. By this method the applications can share the data between them. The is how the data can be processed through applications.

In case users want to get the data through SQL session then they must connect to the DB through SQLPLUS session. The retrieved data can be visible in table format but not like GUI mode like in applications.

In some cases, the stored events in the DB will triggered the internally stored SQL queries to pull the data out of DB. However, all these methods of data management will follow the same stages of SQL command processing. Those are Syntax and Symantec Check, Execution Plans

generations and Choosing the Best Plan by Cost and Execution. We are not getting into these stages deeper here.

2. EXISTING SYSTEM

The Applications uses their schemas as container in the Oracle. Let's say the applications continuously required 50% CPU for their SQL Queries processing in the Database server, If all sudden if any user through user session or Inter event triggers high CPU consuming queries then this definitely leads to the slowness.

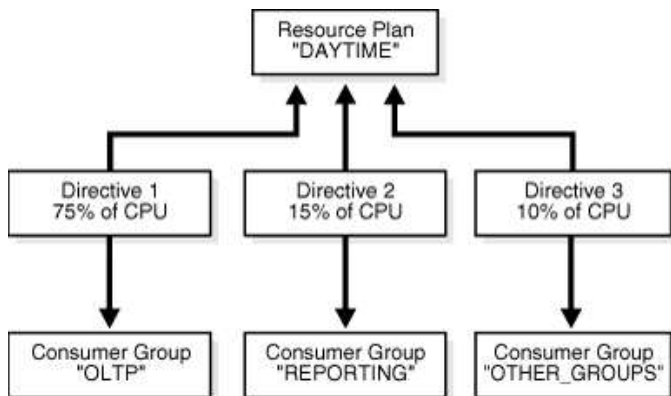
As of now, In Existing system we have Resource management feature at DB Level and CPU allocation in OS level as percentage wise.

As per Oracle Documentation, The Resource Manager allocates resources to consumer groups according to the set of resource plan directives (directives) that belong to the currently active resource plan. There is a parent-child relationship between a resource plan and its resource plan directives. Each directive references one consumer group, and no two directives for the currently active plan can reference the same consumer group.

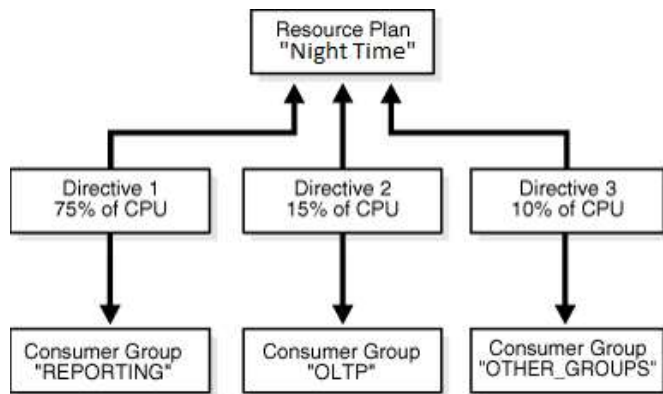
A directive has several ways in which it can limit resource allocation for a consumer group. For example, it can control how much CPU the consumer group gets as a percentage of total CPU, and it can limit the total number of sessions that can be active in the consumer group. See "The Types of Resources Managed by the Resource Manager" for more information.

In addition to the resource plans that are predefined for each Oracle database, you can create any number of resource plans. However, only one resource plan is active at a time. When a resource plan is active, each of its child resource plan directives controls resource allocation for a different consumer group. Each plan must include a directive that allocates resources to the consumer group named OTHER_GROUPS. OTHER_GROUPS apply to all sessions that belong to a consumer group that is not part of the currently active plan.

A Simple Resource Plan



The above Figure shows a simple resource plan for an organization that runs online transaction processing (OLTP) applications and reporting applications simultaneously during the daytime. The currently active plan, DAYTIME, allocates CPU resources among three resource consumer groups. Specifically, OLTP is allotted 75% of the CPU time, REPORTS is allotted 15%, and OTHER_GROUPS receive the remaining 10%.



But in nighttime, the Plan can switch as mentioned above to support and give high priority to Reporting (OLTA) operations. The currently active plan, Nighttime allocates CPU resources among three resource consumer groups. Specifically, OLTP is allotted 15% of the CPU time, REPORTS is allotted 75%, and OTHER_GROUPS receive the remaining 10%.

At OS level also can be possible the CPU Tool is a simple yet powerful command-line tool for limiting and controlling CPU utilization of any process to a given limit and allows the interruption of process execution if the system load overreach a defined threshold. In order to limit CPU usage, CPU tool sends the SIGSTOP and SIGCONT signals to processes and this is determined by the system load. It relies on the /proc pseudo-filesystem to read PIDs and their CPU usage measures. It may be used to limit the CPU usage or system load influenced by a single process or a group of processes to a given limit and/or suspend processes if the system load goes beyond a threshold.

But The existing system doesn't answer some questions. Let me take you through scenarios.

Scenario 1:

If OLTP operations are already at peak and consumed all allocated 75% CPU and Other groups as well has utilized all allocated CPU, then the Reporting consumer group can be limited to only 15%. In this case the Reporting Consumer group user's operations will be piled up since there is no serialization in execution.

Scenario 2:

All groups have consumed all allocated CPU percentages completely. But we are at peak business hours and our customers are getting slowness issue because we can't increase the CPU more than 75% as per resource plan. In this situation we can't allocate CPU of other's dynamically with immediate effect.

Scenario 3:

The Resource management Plan itself will slow down the database sometimes. We have seen sometimes in wait events this plan management as a cause if you plan complex resource plan with multiple level sub plans.

Scenario 4:

If OLTP operations always needs highest priority without depending upon timeline, then it's difficult to manage dynamically the allocate CPU to OLTP operations.

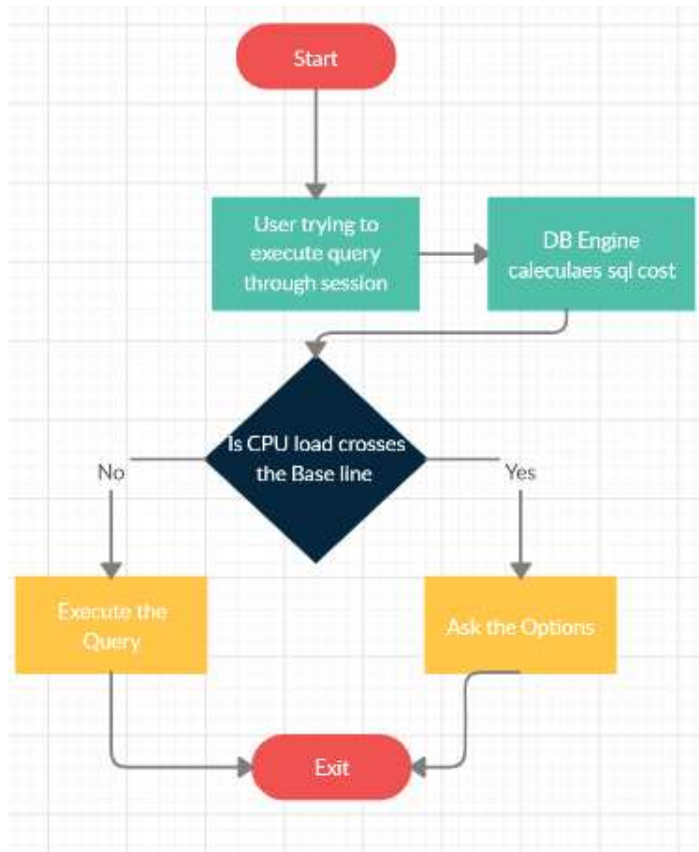
Scenario 5:

Let's Consider the CPU is already 100% utilized and the reporting user wants to execute another report query with parallel to another one. In this scenario the existing system couldn't recommend the user another timeline where the system might be less CPU loaded to execute the report query and automatic schedule of the query.

3. PROPOSED SYSTEM

It's a proposal for enabling a new feature in the database. This feature answers the above 5 scenarios in very effective manner. The applications would connect to the database by using the schema username and password. So, at any cost we shouldn't restrict these operations and need to give full CPU cycles for the OLTP operations. The user defined SQL Queries won't go directly for execution rather than that first it will check and calculate the current system load. If it is less than the actual required load for applications as per their metrics of past and the Query which going to be executed doesn't increase the load, then the DB engine sent the query right away for the execution. In another case if the calculation of CPU load crosses the cpu load baseline by Application's operations or by executing the user queries, Then DB engine proposes for next feasible timeline or

immediate execution for it. The user going to get option wise menu setting for all explicit queries as user interactive. When the CPU load crosses the baseline limits then it starts asking questions to user as mentioned below.

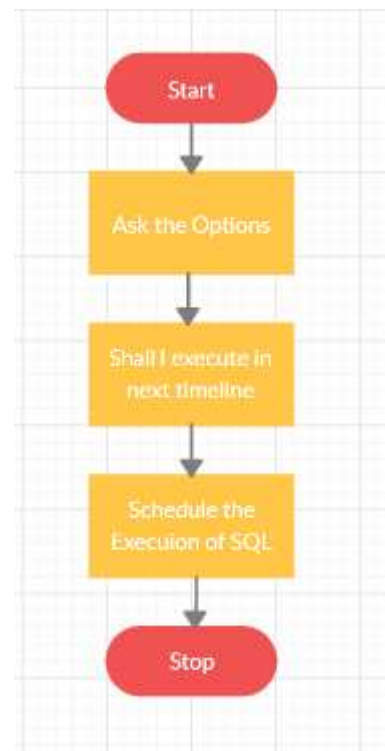


Do you want to execute now?

- If your answer is yes now, it will override and executes it immediately.

Do you want to execute in next available timeline?

- If your answer is yes now, it will show you most suitable timeslot for the execution schedule.



Now let's see how this proposed system answers the scenarios.

Scenario 1:

As mentioned above the schedule of timeslots will be assigned to each report query as they arrive at DB without changing the sequence of queries. All time consuming and high load queries will be serialized. So, No data inconsistency.

Scenario 2:

Here, in oracle databases there will be profile for each and every schema or user. For Application schemas the queries directly routed to execution. But, for high consuming user defined queries of other schemas or users has to go through the Calculations. If their executions don't cross the CPU threshold then they will be directed as normal query. If they don't, then they must go through the option. So, there is no impact on OLTP operations.

Scenario 3:

The Complex Resource management plan will be replaced with this method as optional. No OLTP operations hampered at any level and given highest priority to them.

Scenario 4:

It's dynamically allocates the resources to OLTP operations and gives the choice to users to avoid more contention in the database.

Scenario 5:

It's suggests the timeslot for report query and allocates the automatic schedule of it to execute. It will decrease the contention and load on the Database.

REFERENCES

- [1] https://docs.oracle.com/cd/E11882_01/server.112/e25494/dbrm.htm#ADMIN11860
- [2] <https://app.creately.com/diagram/68p54ntLTHf/edit>

BIOGRAPHIES



Vaheedbasha Shaik has received Master of Technology (M. TECH) in Specialization of Database systems From SRM University and Currently working as Senior Oracle Database Administrator for the IT Department in CenturyLink Technologies Pvt. Ltd.