Forecasting Oracle Performance

Yuri van Buren
Senior Oracle DBA Specialist
End-2-End Performance Engineer
Logica Nederland B.V.
Amstelveen – The Netherlands

Keywords: Performance Management, Service Level Management, Capacity Management

Introduction

Everybody is continuously making predictions. We do this often unnoticed during the day: if we look outside for the weather, drive a car or catch a bal. Forecasting is a natural part of every human being. It is of great importance to be able to predict well. E.g. otherwise you would not catch the thrown ball… This is also very true for predicting the performance of IT systems under different workloads.

In this presentation Yuri will show that you can make accountable forecasts about the Oracle performance in all kinds of practical situations. You could use several forecasting models which are meant to simplify the complexity of the system under analysis. With these methods you can predict when IT systems are running at risk. The risk of overloaded resources (CPU and/or IO) become clear. You learn even to develop strategies to mitigate these risks.

Forecasting Oracle Performance is an important aspect of Service Level Management and Capacity Management. Learn now the methods and skills so that you are able to answer questions like:

- "Can the system handle a 200% increase in workload?"
- "What will happen to performance if we add 2 more CPUs?"
- "What will happen to performance if we add 30% faster CPUs?"
- "Do 25 more GL users significantly impact the response time?"

Learn to manage the service level management risks effectively.
It sure is **"Better than a crystal ball".**

Introduction to Forecasting Models

There are five commonly used forecasting models to forecast Oracle performance

- Simple Math
- Essential Forecasting mathematics
- Queuing Theory
- Ratio Modeling
- Linear Regression Modeling

Simple Math

- Can be used for single component forecasts
- Input can be technical or application metrics
- No queuing effect is involved.
- Appropriate for memory, basic IO predictions and basic networking
- The precision is usually low
- Handy for short-duration projects

Example: If an Oracle client process consumes 10MB of non-shared resident memory. And plans are there to add another 50 users (resulting in 50 additional clients), then the system will require 500MB of additional memory.

Essential Forecasting Mathematics

Response time is the most important term in forecasting and it is also one of the simplest.
Response time is service time (S) plus queue time (W): R = S + W.

Transactions

- Arrival Rate (λ) is the amount of work in a certain time interval
- A Transaction Processor is a "server" that can process transactions
- A Queue is a linked list with a head and tail – First In First Out queues are used
  Transaction Flow – When a business transaction is submitted, it flows throughout the computing system consuming CPU, IO, memory and network resources.

Key variables & Formulas

$$U = \frac{S_t \lambda}{M} \qquad R_{t-cpu} = \frac{S_t}{1-U^M}$$

$$Q = \lambda Q_t \qquad R_{t-io} = \frac{S_t}{1-U}$$

$$R_t = S_t + Q_t$$

Variable	Definition	Example
λ	The arrival rate is how much work enters the system within a period of time.	10 trx/ms
S_t	Service time is how long it takes a server to service a single transaction and is represented as time per transaction.	2 ms/trx
U	Utilization is the busyness of a server and is represented as a simple percentage.	83%
Q	Queue length is simply the length of the queue. For Oracle systems, we typically do not want transactions waiting in the queue.	0 trx
Q_t	Queue time is how long a transaction waits in the queue before it begins to be served. This is represented as time per transaction.	0.02 ms/trx
R_t	Response time is service time plus queue time and is represented as time per transaction.	2.02 ms/trx
M	Number of servers is how many servers are servicing transactions	24 servers

During the presentation Yuri will discuss several essential forecasting mathematics examples.

E.g. DBA John works on a 12 CPU system. He measures an average CPU utilization of 35% during month-end closing. Oracle runs 6.11 user calls per second during that period. Can the system handle a 200% increase in workload?

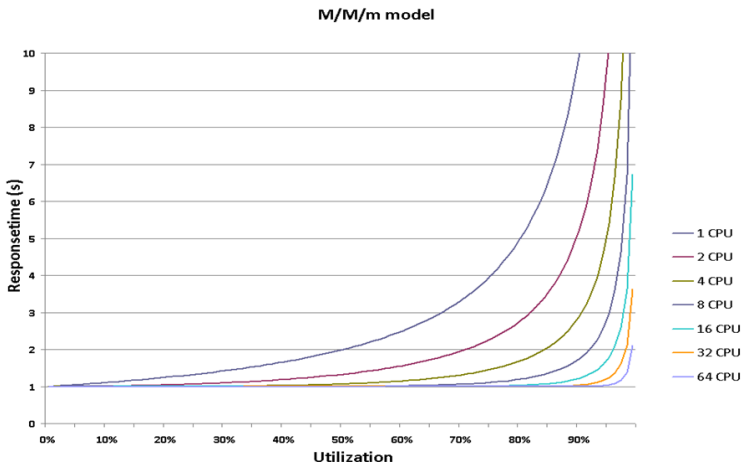
Yuri also demo's the "Interactive Response Time learning Graph.xlsx".

Queuing Theory
The most important formula's in queuing theory are:

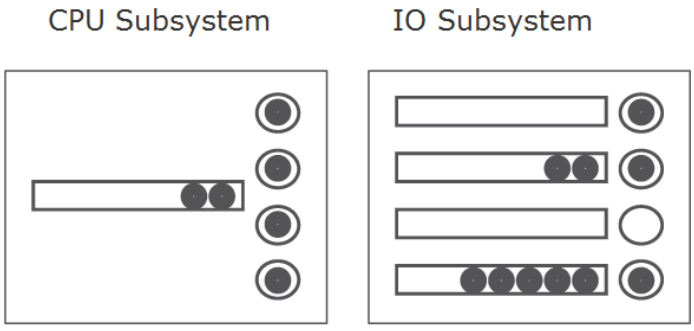
- Little's law
- Kendall's Notation
- Erlang C

Little's Law: $N = \lambda T$
With N the average number (serviced and queued) in a stable system
λ is the average number of arrivals at some interval (that is, the arrival rate)
T is the average time spent in the system (that is the response time).

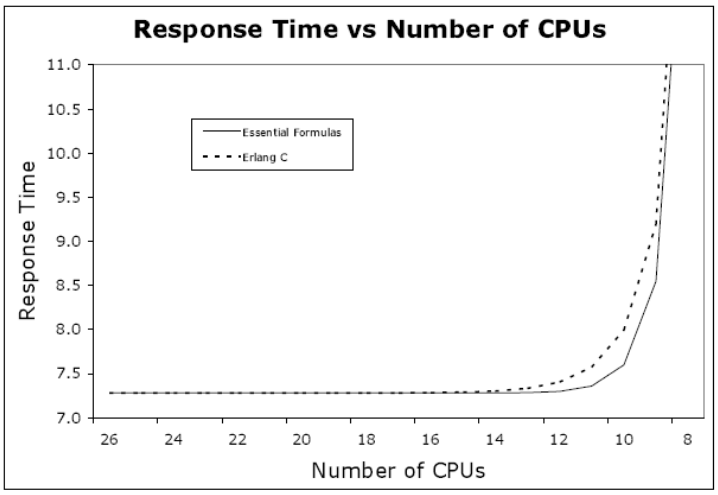
Kendall's Notation: A/B/m or better known as a M/M/m queuing theory model.



With the help of a Queuing Model Data Entry sheet the difference between a 4 CPU and 4 disks IO subsystem is made clear. They are modelled differently as depicted in the next picture.



With the Erlang C formula the essential forecasting mathematics can be improved.
And the corresponding response time curves give more accurate estimates.



Ratio Modeling

$$P = \frac{C_1}{R_1} + \ldots + \frac{C_n}{R_n} \qquad P = UM$$

P is the number of fully utilized CPU's
C is the number of workload category occurrences. (E.g. OLTP and Batch users).
R is the ratio linking a single workload category occurrence to a CPU.
U is the CPU Utilization
M is the number of CPUs (servers).

During the presentation the answer the question "Do 25 more GL users significantly impact the response time?" is given using ratio modelling.

Linear Regression Modeling

To determine how much of some business activity can occur before a system "runs out of gas", linear regression modelling can be used. You can for instance determine the number of orders the system can process within an hour.

Know you know something about mathematical models you still:

- Need a method
- Need to collect data
- Need to know about risk mitigation strategies

Forecasting Method

Step 1: Determine the study question
Step 2: Gather & Characterize workload data
Step 3: Select a forecast model
Step 4: Forecast & Validate
Step 5: Risk Mitigation or "What to tell management".

Summary
With an understanding of the available forecasting models for Oracle based systems and a forecasting method it becomes easy to do service level and capacity management predictions.
It saves you the hassle of looking for answers in a Crystal Ball.

Contact address:

Name	**Yuri van Buren**
	Principal IT Consultant | Senior Oracle DBA Specialist
Company	**Logica Nederland B.V.**
	Logica, now part of CGI
Address	Prof. W.H. Keesomlaan 14
	1183 DJ Amstelveen
	The Netherlands
Phone:	+31 (0)88 5640000
Email	yuri.van.buren@logica.com ; yuri.van.buren@cgi.com
Internet:	www.logica.com; www.cgi.com