

## Capacity, Performance & Other Confusing Terms

Author: Mike Ley, UKCMG Committee

A few weeks ago I was phoned by a salesman who was trying to push a “performance measurement tool”. It wasn’t a good time and so I tried my usual brush-off / delaying tactics ~ “now is not the right time, phone me back in a few months”. However this salesman was more persistent than most and he continued on about how the tool could manage the capacity of a system.

Thoroughly confused about what the tool was attempting to do I asked a simple question “what is the difference between capacity and performance?” ...and I got blubber.

Words came across the phone lines, mainly a re-iteration of the latest management jargon, but no clear definition of the difference between the meanings of the terms “capacity” and “performance”.

Later I thought, if this is what specialist companies in the area of capacity and performance are presenting to senior managers, no wonder they are confused about the role of capacity management and the tools that are required.

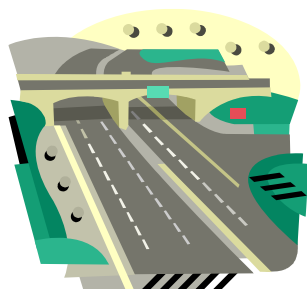
So to set the record straight let me propose the following definition of terms:

**Performance is about the time that an individual transaction or piece of work takes to be carried out**

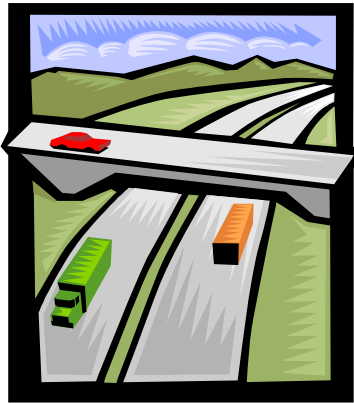
**Capacity is about the amount of work that can be carried out over a period of time**

A simple analogy to remember these definitions relates to road use.

Performance is the time it takes me to drive from my home back to Wales



Capacity is the number of cars per hour the M4 can handle



By extension, utilisation is the number cars per hour travelling on the M4, over the maximum number of cars per hour that the M4 can handle

And high utilisation of the M4 leads to delays because of queuing



As a consequence of these definitions it can be seen that a capacity tool mainly collects utilisation data or technical data (e.g. bytes moved over the last 5 minutes) while a performance tool mainly collects response data. However, there is not always a clean divide between the two and several tools collect both capacity and performance data.

Of course, in my analogy, the measure of capacity does not have to be “cars per hour” it could equally be cars per day. This highlights the fact that it is not meaningful just to say that utilisation is 60%.

Utilisation is a measure over a period of time and this should always be specified. For example, CPU utilisation in the peak hour is 60%. Without this qualification the measure is meaningless as there is a significant difference between a system that has a peak hour utilisation of 60% and one that has a peak day utilisation of 60%, or even 60% utilisation over a week.

So I strongly recommend that you use, remember and apply (URA) these definitions, as they may well save you much money and / or pain. Apart from the saving in time dealing with some salesmen, you may avoid the fate of one company who bought a capacity tool and a performance tool only to find they had bought two tools that collected much the same data, with no discernable extra benefit.