Increasing hps degree for a given workload - CSC148, Prof Mitchell, sept 24, 2018

We demoed an “ideal” homogeneous parallel server for customer cj service requests. In the source code file: ss\_and\_ps\_simpleCompare.gps each cj requests 1 server from the hps “pool” of servers.

The simulated service durations are modeled to be independent of the current number of busy servers or any of the resource contention factors that requests/jobs/users encounter on most “real” systems.

Omitting resource contention means that, in theory, increasing the hps degree should show a measurable decrease in waits and bottlenecks seen by cj. In class, we filled out the first 2 rows of Table 1. A third row, not done in class, has been added to show the effect of hps degree 16. This is an overkill setting, but even degree 16 does not completely eliminate cj waits. Also, 1.25 is the smallest avg(Wc) can ever be (at steady-state).

The conclusion: for ehse hps parameters, service waits do not measurably decrease after degree 8 or so.

**Table 1 – effect of hps degree on cj processing**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **degree** | ****  Held constant | ** = 1/mean = 0.8**  Held constant | **** | **MAX number of cj in S** | **MAX number of cj that waited** | **avg(Wc)** |
| 2 | 1 | 0.8\* | 64% | 18 | 16 | 2.018 |
| 8 | 1 | 0.8 | 15.6% | 9 | 1 | 1.255 |
| 16 | 1 | 0.8 | 7.7% | 9 | 1 | 1.255 |

**Follow-up to Wedn Sept 19th class**

The expression for hps utilization is:

hps degree 

Amazingly, this formula holds for any {ia} and service distributions (not just uniform or exponential distribution), and for any stable system (stable means, analytically, that  < 1).

Consider the following {ia} and service duration mean value settings:

ps\_ia\_mean EQU 1 ; ps Inter-arrival times mean

ps\_service\_mean EQU 1.25 ; Mean ps service duration

With these mean values:  = 1 / (mean service duration) = 1/1.25 = 4/5 = 0.8, and

thus,  1 / (2 \* 0.8) = 5/8 = 0.625.

This  agrees well with the gpssW report value for the hps degree 2:

(1 – UTIL.) = (1 – 0.356) = 0.644. Similarly for the other degrees in Table 1.

**avg(Wc)** **Ideal hps - graph of avg(Wc) vs. {hps degree}**

2.2 | x

|

| Performance knee

|

1.37 | x

1.27 | x

|-------1-------2-------3-------4-------5-------6-------7-------8--------------------------------------------- > **degree**

< -- Linear improvement -- > < -- Degraded: more servers do not decrease avg(Wc) -- >

* Corrected the error in the table noticed in class today