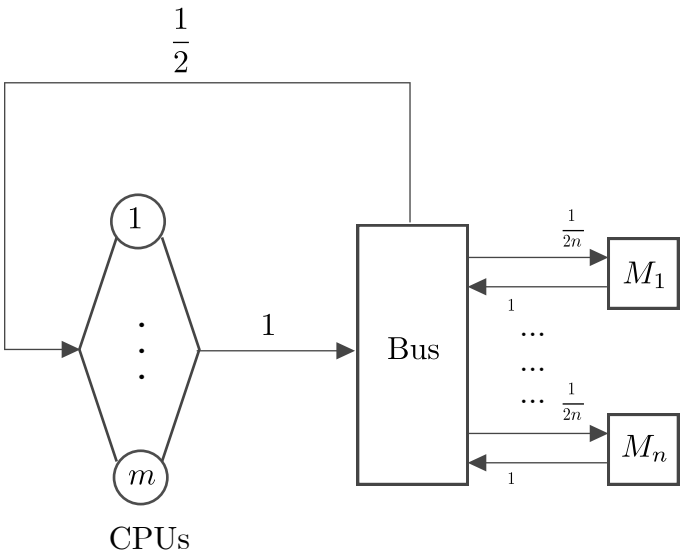
CAS 6EO3
Fall 2019

Project

Nhan Thai
thain1@mcmaster.ca

(a) We can model the system with the following network:



Each request enters and leaves Bus twice. Therefore, in steady state, the probability of a request going from Bus to CPUs equals probability a request going from Bus to memory modules. Thus the numbers $\frac{1}{2}$ and $\frac{1}{2n}$.

With n = 4.

$$Z = 25, \mu_{bus} = \frac{1}{2}, \mu_1 = \mu_2 = \mu_3 = \mu_4 = \frac{1}{4}$$

Let visit ratio of CPUs V_{cpu} be 1:

$$\begin{cases} V_{cpu} = 1 \\ V_{bus} = V_{cpu} + V_1 + V_2 + V_3 + V_4 \\ V_1 = \frac{1}{2n} V_{cpu} \\ V_2 = \frac{1}{2n} V_{cpu} \\ V_3 = \frac{1}{2n} V_{cpu} \\ V_4 = \frac{1}{2n} V_{cpu} \end{cases} \Rightarrow \begin{cases} V_{cpu} = 1 \\ V_{bus} = 2 \\ V_1 = V_2 = V_3 = V_4 = \frac{1}{4} \end{cases}$$

Using MVA (**mva.m**), and in each iteration:

$E[\text{memory request time}] = 2 * E[T_{bus}] + E[T_1]$ (Visit bus twice and visit a memory module)

$\rho_{bus} = \dfrac{X \times V_{bus}}{\mu_{bus}}$

We get the following result table of bus utilization and mean response time of a memory request as a function of m:

m	1	2	3	4	5	6	7	8	9	10
Mean memory request response time	8	8.6	9.31	10.143	11.126	12.29	13.68	15.33	17.27	19.538
Bus utilization	0.12	0.24	0.35	0.45	0.55	0.64	0.72	0.79	0.85	0.89

The analytical result is very close to the simulation result that I have with (**a.c**).

(b) The answer in (a) should not change because the mean think time is still the same.

Another way to explain this is because the distribution is hyperexponential, we can split the CPUs node to 2 think nodes with exponentially distributed think time, and probability of visiting from Bus to them are 0.95 and 0.05 respectively. Doing this will not change visit ratios as well as MVA analysis.

The simulation I ran with (**b.c**) verifies this conclusion as it is very close to the result of that in (a).

(c) Random routing is in (**c1.c**).

- Response time = 11.78523
- Bus utilization = 0.29

JSQ routing is in (**c2.c**).

- Response time = 8.55359
- Bus utilization = 0.379

(d) Simulation is in (**d.c**).

m	1	2	3	4	5	6	7	8	9	10
Mean memory request response time	7.97	8.73	9.58	10.586	11.718	13.02	14.50	16.25	18.24	20.53
Bus utilization	0.12	0.237	0.347	0.45	0.55	0.63	0.70	0.77	0.832	0.88

As we can see, the response time is slightly longer and bus utilization is slightly smaller.