1.

(a)

T = 60 mins = 3600 s

C = 1267

X = C / T = 1267 / 3600 for the system

CPU:

U(CPU) = B(CPU) / T = 2929 seconds / 3600 s

So:

D(CPU) = U(CPU) / X(system) = B(CPU) / C = 2929 / 1267 = 2.312

Disk:

U(Disk) = B(Disk) / T

D(Disk = U(Disk) / X(system) = B(Disk) / C = 2765 / 1267 = 2.182

(b)

According to X(0) = U(j) / D(j) and U(j) <= 1, we can get

X(0) <= 1 / D(j), should hold for all devices in the system.

So X(0) <= 1 / max(D(j))

Here max(D(j)) = D(CPU)

So X(0) <= 1 / D(CPU) = 1 / 2.312 = 0.432

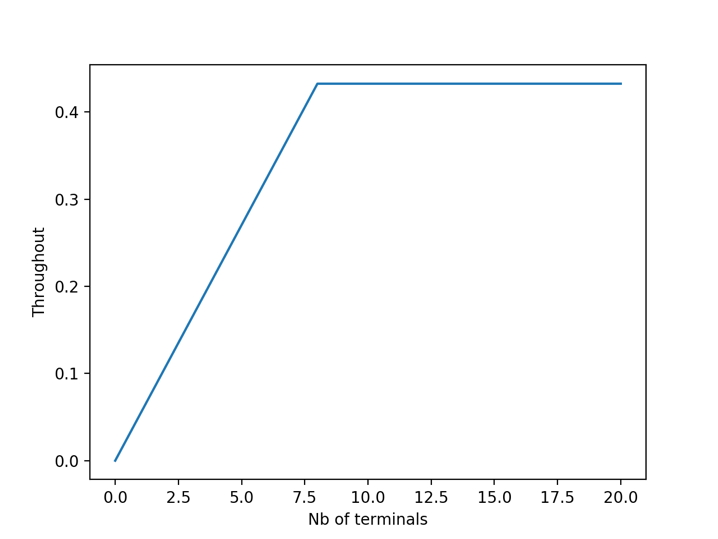
There is another bound for X(0), as described in slide, we get this bound from Little’s Law,

X(0) <= N / (think time + D(CPU) + D(Disk))

Set slop = 1 / (think time + D(CPU) + D(Disk)) = 1 / (14 + 2.312 + 2.182) = 1 / 18.494

Therefore when X(0) == 1 / D(CPU), N = (14 + 2.312 + 2.182) \* 1 / D(CPU) = 7.999 = 8

So we can get the plot:



Question 2

(a)

This is M/M/4/n+4 model.

The exponential inter-arrivals is λ = 15 queries/ hour, the exponential service time is μ = 3 queries / hour.

State 0 = there is zero query in the system.

State 1 = there is 1 query in the system.

State 2 = there is 2 query in the system.

State 3 = there is 3 query in the system.

State 4 = there is 4 query in the system.

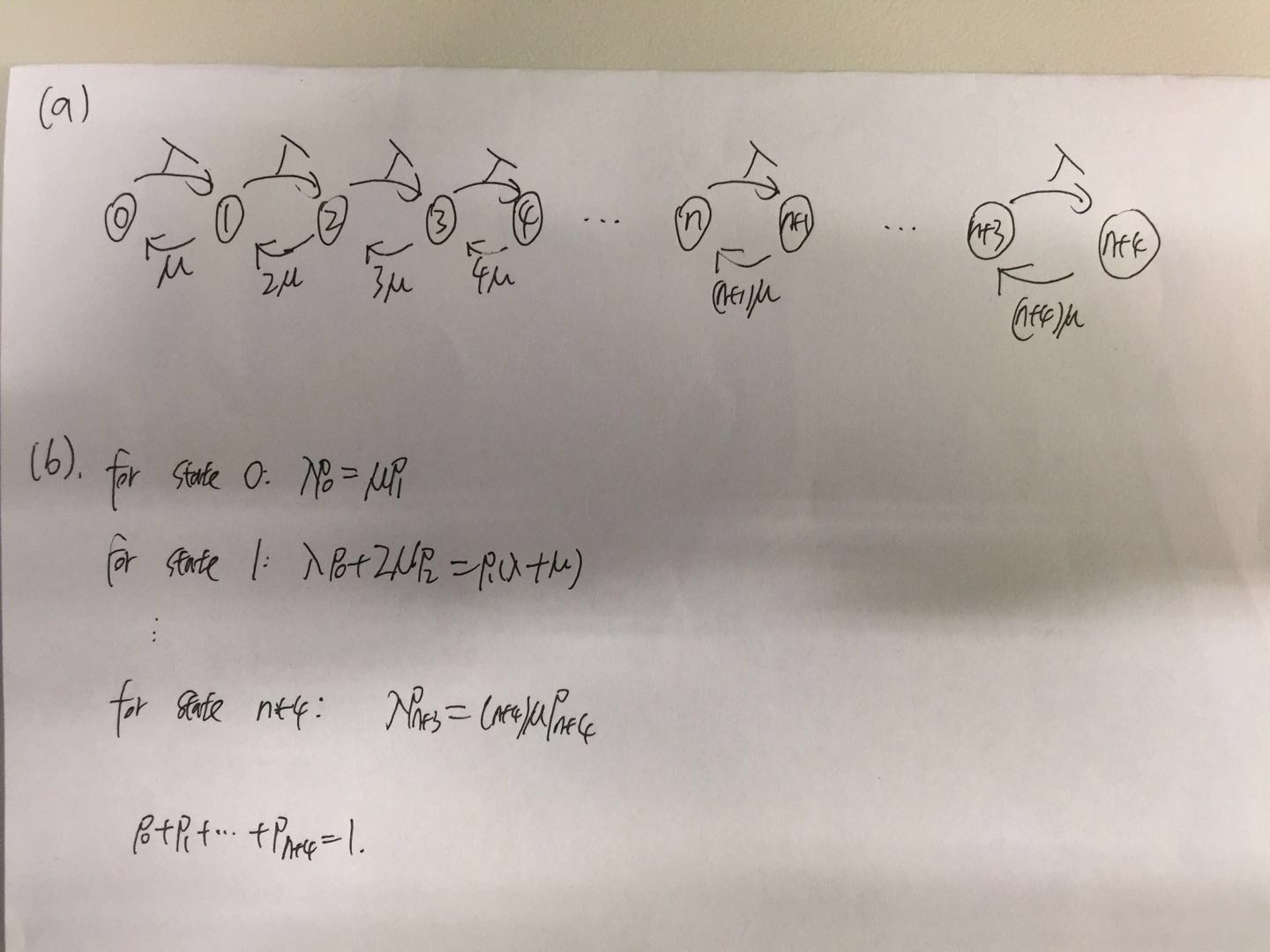
.

.

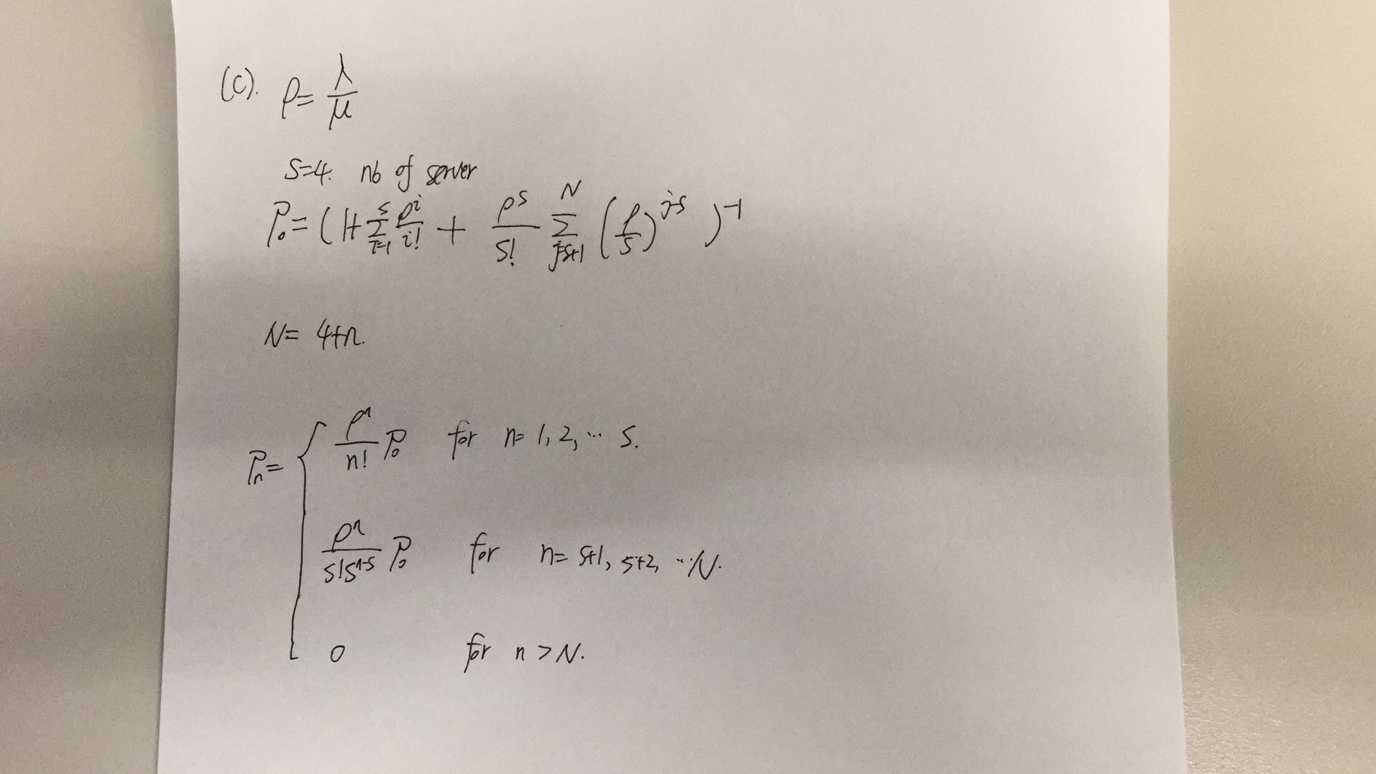
State n = there is 4 query in the system(4 at server, n – 4 at queue)

..

State n + 4 = there is n + 4 query in the system(4 at server, n at queue)



(c)

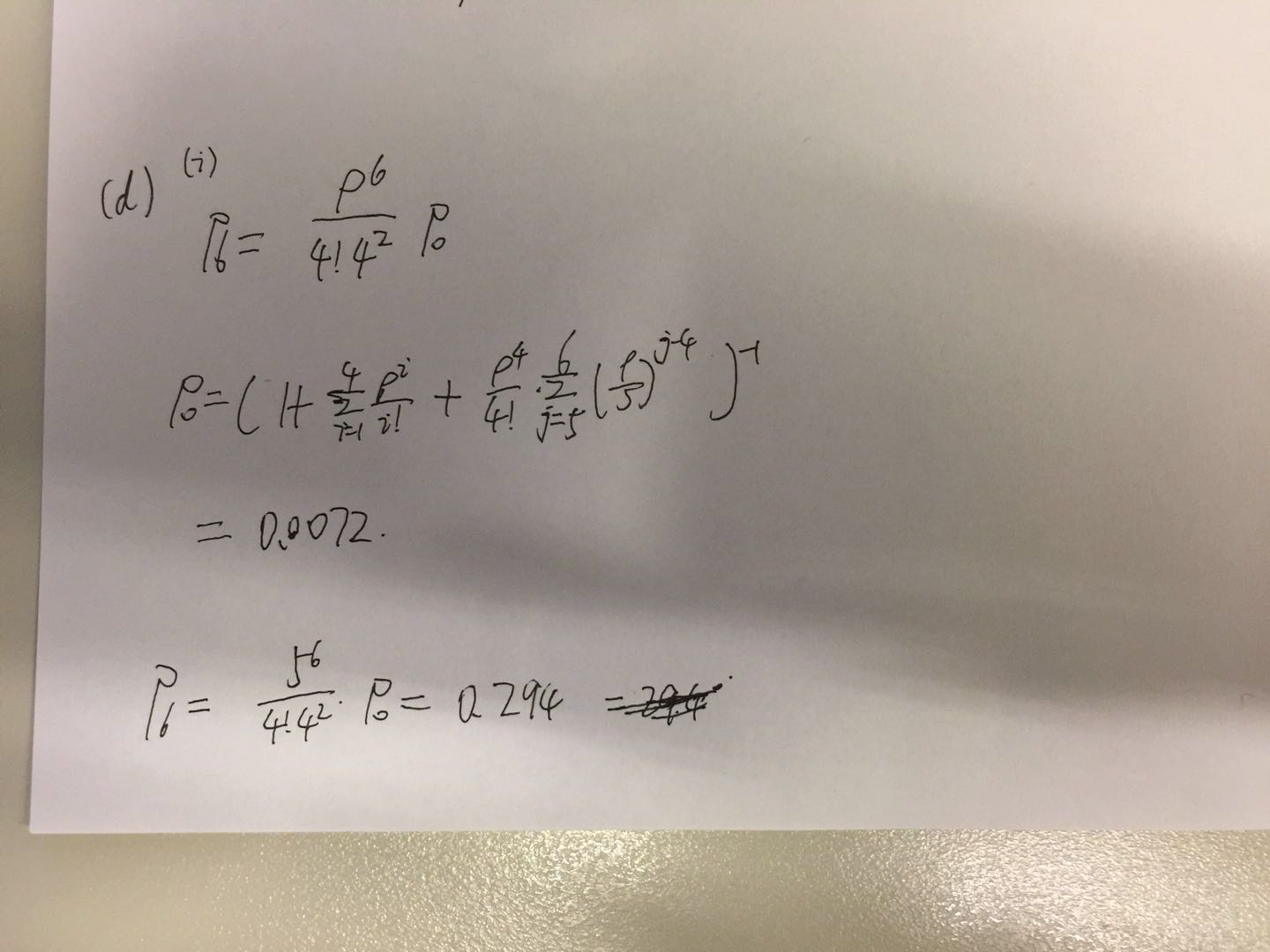


(d)

N = 4 + 2 = 6

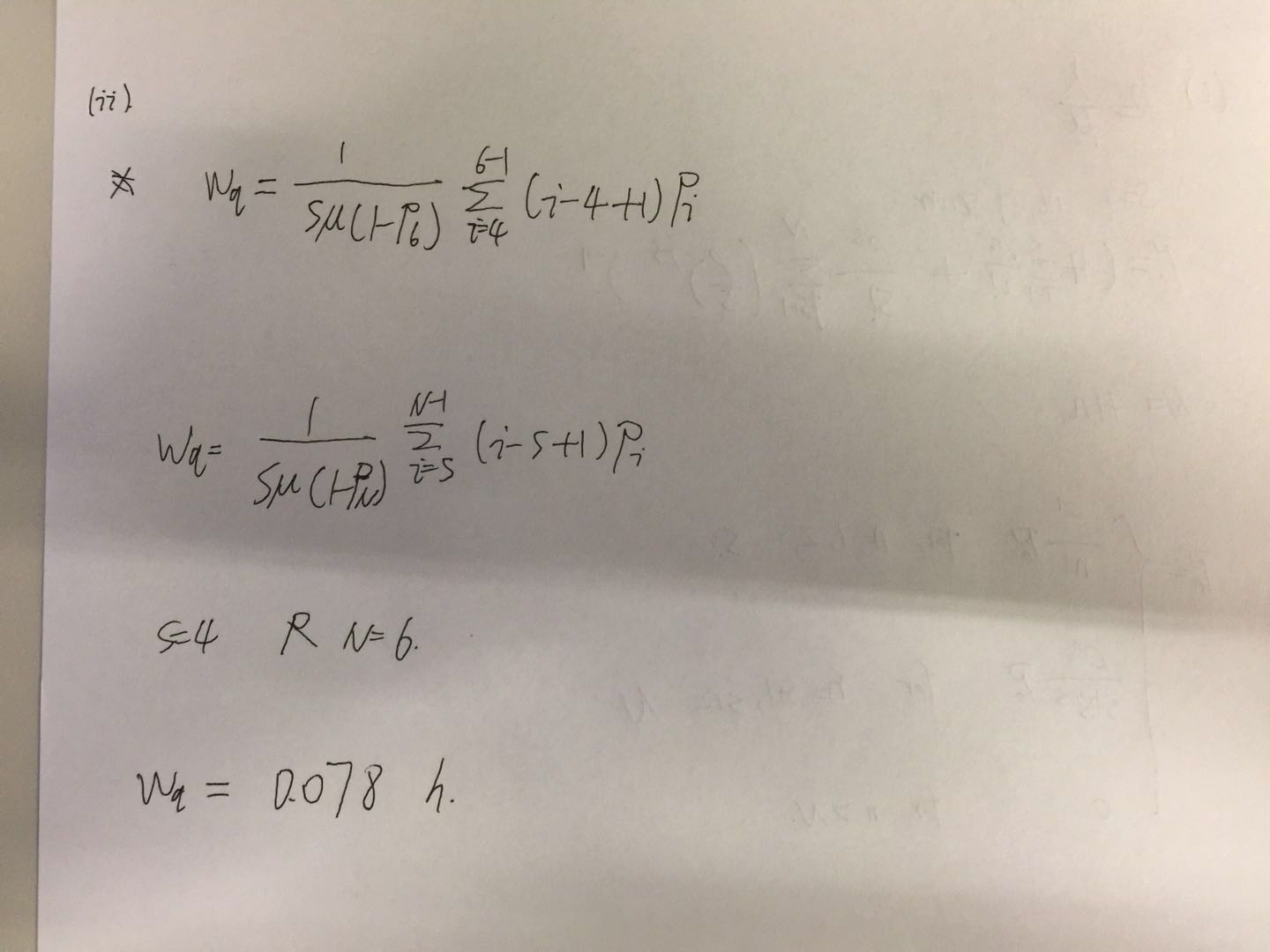
The probability that an arriving query is rejected is the same as the probability that there are 6 queries in the system.

(i)



(ii) mean waiting time = mean response time – mean service time

mean service time = 1 / μ



(e)  
add 5:

0.22331931368078703

add 10:

0.20708581427429248

add 15:

0.20226785993865187

add 20:

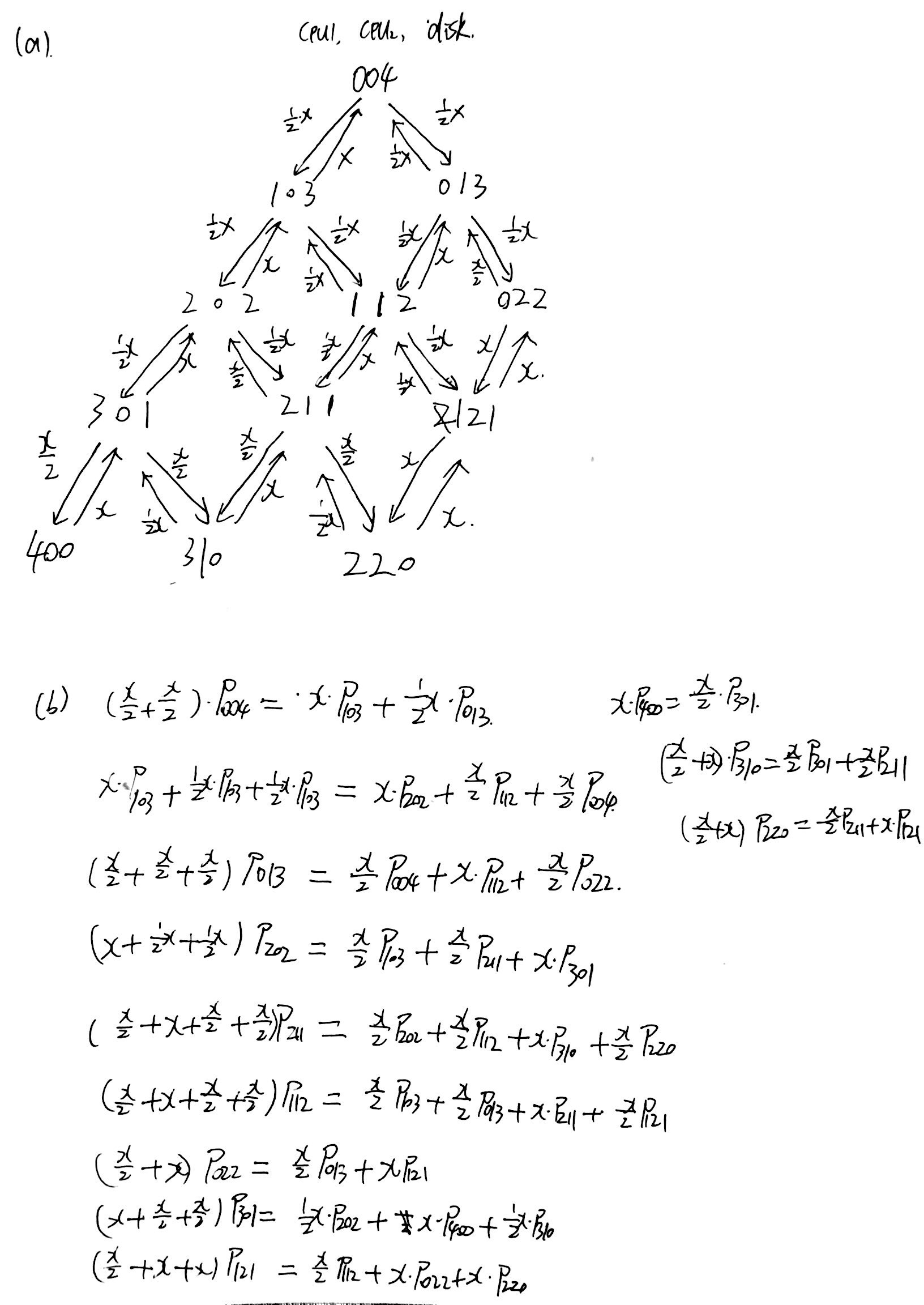
0.2007375098367033

(f)

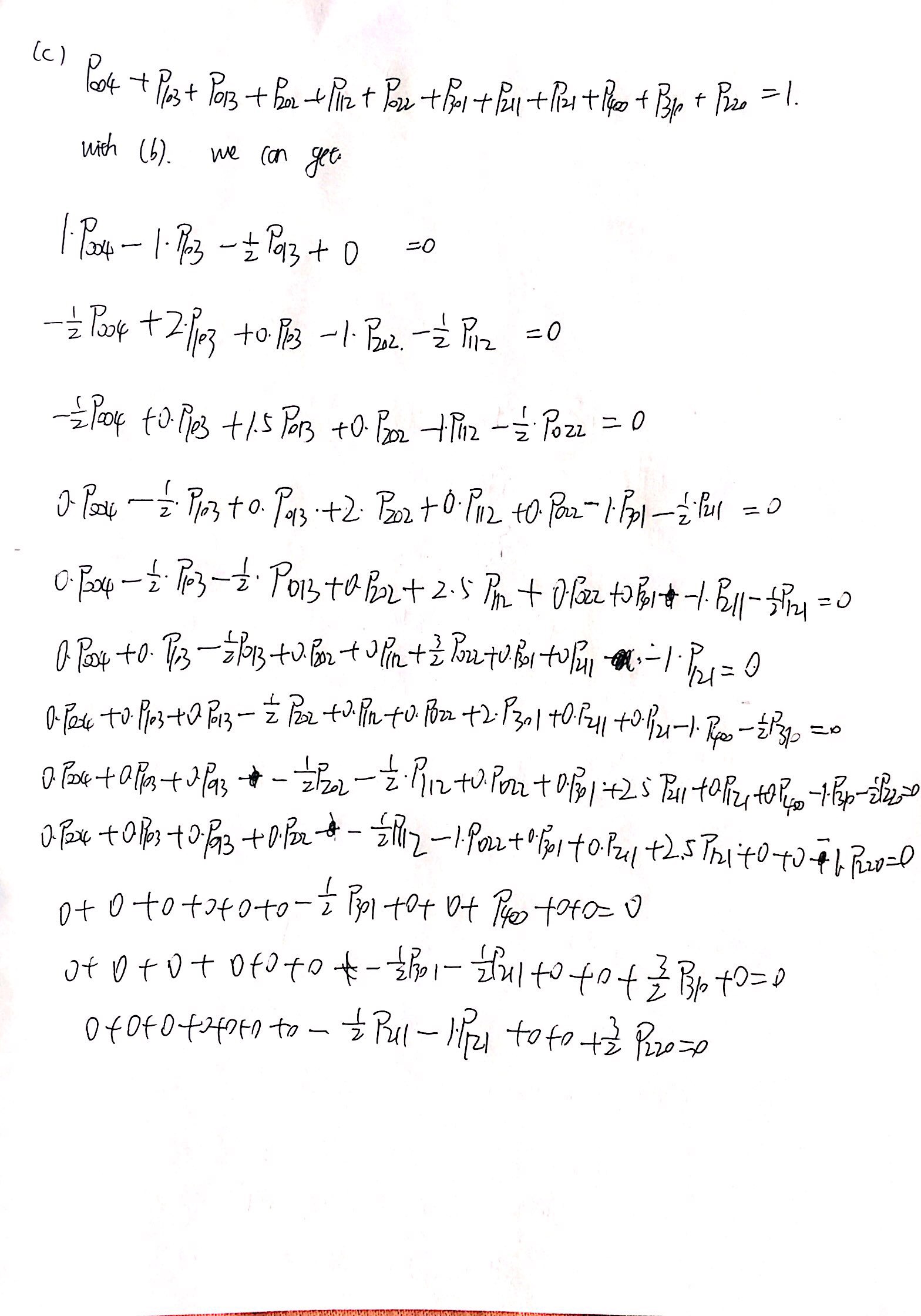
Because after adding 10 waiting slots, the value of P0 decreased as almost the same speed of the value of ρn increased, so there is little drop in the total value of Pn, which is the blocking probability.

Question 3.

Set: jobs complete at Disk at a rate of x = 1 / 0.2 , according to the mean processing times for CPU1 , CPU2 and Disk, the rate for CPU1 is x, the rate for CPU2 is x/2.



(c)



we have result:

[ 0.17107984, 0.09117845, 0.15980279, 0.0500653 , 0.09350334,

0.12132185, 0.02592787, 0.05722703, 0.10208138, 0.01296393,

0.0277183 , 0.08712993]

for:

[P004, P103, P013, P202, P112, P022, P301, P211, P121, P400, P310, P220]

(d)

The throughput of the system is the Disk throughput.

Disk utilization is the sum of probabilities of states where there is at least one job at disk:

U = P004 + P103 + P013 + P202 + P112 + P022 + P301 + P211 + P121 = 0.8721878499999999

S = mean service time = 0.2 s

Throughput = U / S = 4.361 transactions /s

(e)

The mean number of jobs in CPU1 =

1 \* P103 + 2 \* P202 + 1 \* P112 + 3 \* P301 + 2 \* P211 + 1 \* P1212 + 4 \* P400 + 3 \* P310 + 2 \* P 220 = 0.88840192

(f)

CPU1 utilization is the sum of probabilities of states where there is at least one job at CPU1:

U = P103 + P202 + P112 + P 301 + P211 + P121 + P400 + P310 + P220 =

0.54779553

S = 0.2s

X = U / S = 2.73897765

According to Little’s Law:

The mean response time of CPU1 = the mean number of jobs in CPU1/ throughput of CPU1

= 0.88840192 / 4.361 = 0.2037