Study of Delay and Blocking Probability In M/M/1/B and M/G/1/B Queuing Systems

PACN Mini Project
Project 19

Swarnarup B21140
Joel B21196

Problem Description

In this project we need to estimate analytical / Mathematical model for Blocking probability and Queuing Delay of a M/M/1/B system. Then compare simulation result of Packet Loss % and waiting time with the analytical estimation.

Next we need to simulate M/G/1/B system and check how delay behaves. We need to estimate an approximated mathematical model for estimating delay in M/G/1/B system.

Contents:

M/M/1/B Mathematical Model

Transition States n = number of pkts in system

State change probabilities :

$$P_{0,0} = P[\text{no arrival}] = e^{-\lambda\delta} = (1-\lambda\delta)$$

$P_{n,n}$ for $0<n\leq m$

$$\begin{aligned} P_{n,n} &= P[\text{no arrival}] \times P[\text{no service}] + O(\delta) \\ &= e^{-\lambda\delta} e^{-\mu\delta} + O(\delta) \\ &= 1 - \lambda\delta - \mu\delta + O(\delta) \end{aligned}$$

$$\begin{aligned} P_{n,n+1} &= P[1 \text{ arrival}] \times P[\text{no service}] + O(\delta) \\ &= \lambda\delta e^{-\lambda\delta} \times e^{-\mu\delta} + O(\delta) \\ &= \lambda\delta + O(\delta) \end{aligned}$$

$$\begin{aligned} P_{n,n+1} &= P[1 \text{ arrival}] \times P[\text{no service}] + O(\delta) \\ &= \lambda\delta e^{-\lambda\delta} \times e^{-\mu\delta} + O(\delta) \\ &= \lambda\delta + O(\delta) \end{aligned}$$

$$\begin{aligned} P_{n,n-1} &= P[0 \text{ arrival}] \times P[1 \text{ service}] + O(\delta) \\ &= e^{-\lambda\delta} \mu\delta e^{-\mu\delta} + O(\delta) \\ &= \mu\delta + O(\delta) \end{aligned}$$

Markov chain:

Balance Equation: $\lambda p_{n-1} = \mu p_n$

$$p_n = \rho p_0 \qquad \rho = \frac{\lambda}{\mu}$$

From Normalization

$$\sum_{n=0}^{B} p_n = 1$$

$$\Rightarrow \sum_{n=0}^{B} \rho^n p_0 = 1$$

$$\Rightarrow p_0 \frac{1-\rho^{B+1}}{1-\rho} = 1$$

$$p_0 = \frac{1-\rho}{1-\rho^{B+1}}$$

Blocking Probability

= P[number of PKts in system = B]

$= p_B$

$= \rho^B p_0$

$$= \frac{\rho^B (1-\rho)}{(1-\rho^{B+1})}$$

Average Number of PKts in System ;

$$E[N] = \sum_{n=0}^{B} n p_n$$

$$= \sum_{n=0}^{B} p_0 \, n\rho^n$$

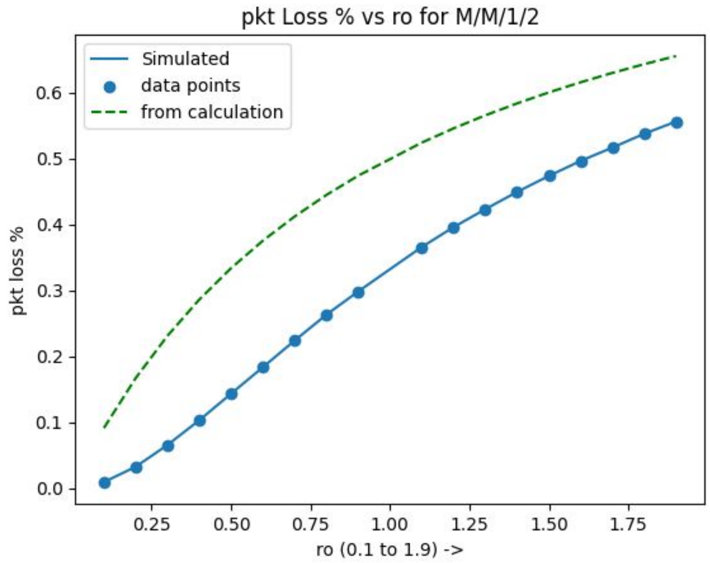
$$= p_0 \rho \left[\frac{1-\rho^{B+1}}{(1-\rho)^2} - \frac{(B+1)\rho^B}{(1-\rho)} \right]$$

Avg. no. of PKts in System, $N = p_0 \rho \left[\frac{1-\rho^{B+1}}{(1-\rho)^2} - \frac{(B+1)\rho^B}{(1-\rho)} \right]$

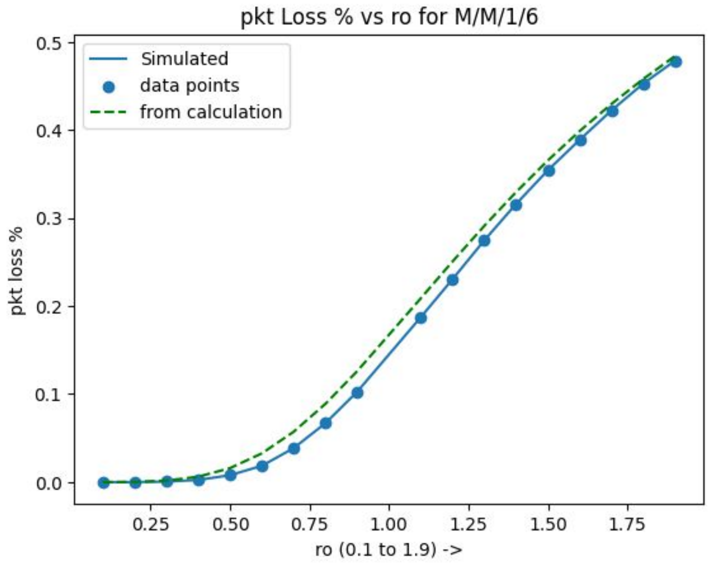
$$\boxed{\text{Avg. waiting time, } W = \frac{p_0 \rho}{\lambda} \left[\frac{1-\rho^{B+1}}{(1-\rho)^2} - \frac{(B+1)\rho^B}{(1-\rho)} \right] - \frac{1}{\mu}}$$

Simulation Result Analysis

Packet Loss % = Incoming pkt found queue/system is full = Blocking Probability vs ρ (keeping queue size fixed)



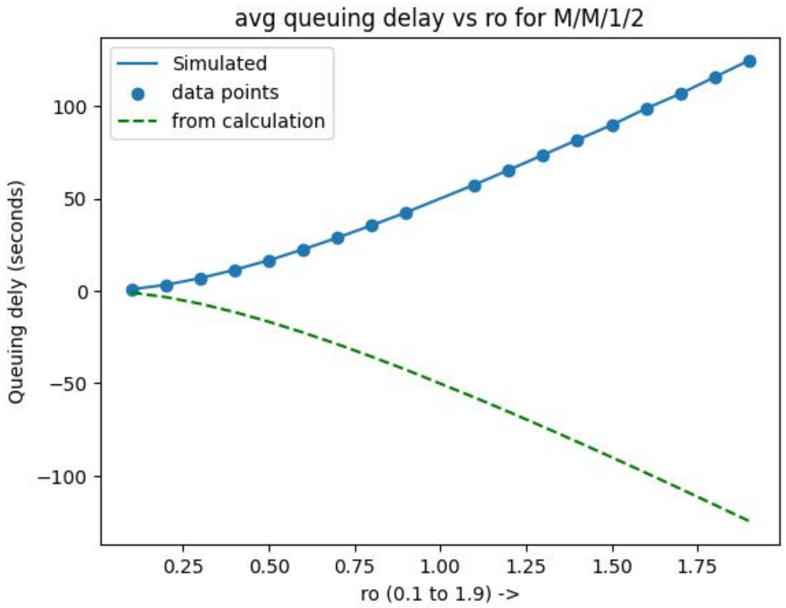
Queue size = 1



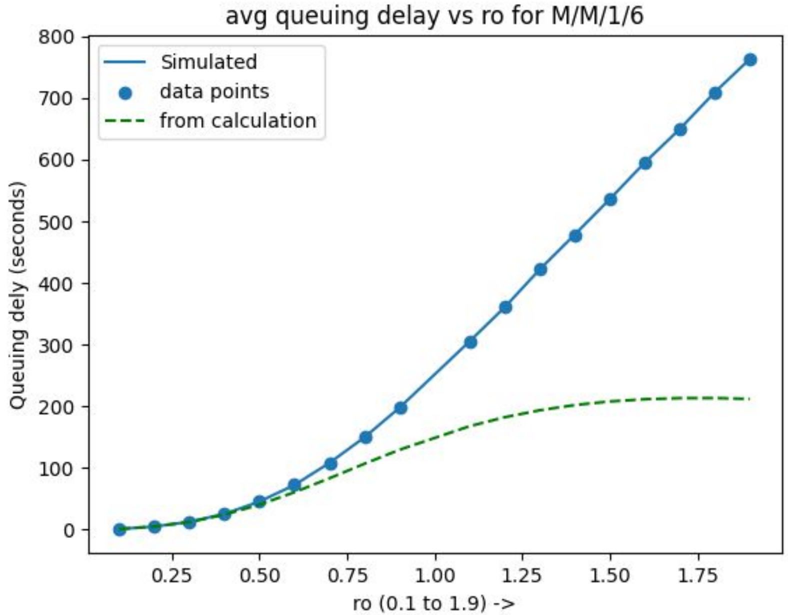
Queue size = 5

Inference : Difference between mathematical model and simulation result reduces when queue size is higher.

M/M/1/B - Waiting time (W) / Queuing Delay vs ρ



Queue size = 1

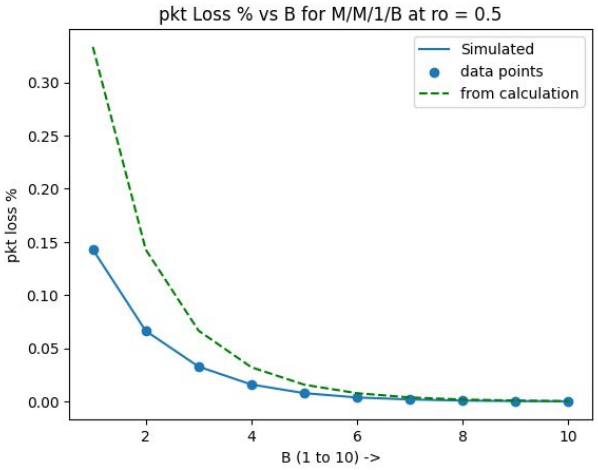


Queue size = 5

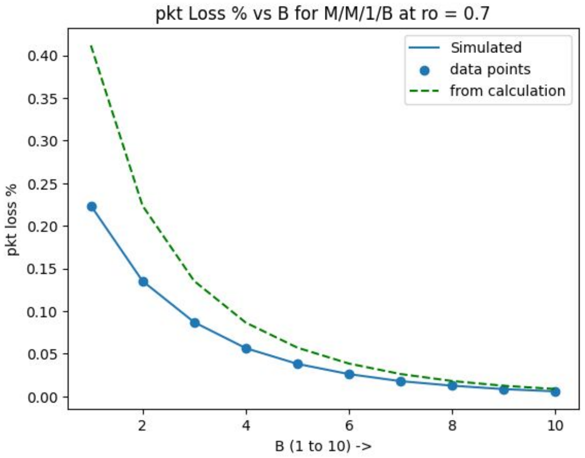
Inference :

- For Higher value of B (queue size), Difference between mathematical model and simulation result reduces
- Mathematical model matches with simulation result where $\rho < 1$
- In simulation delay increases but mathematical model predicts delay to decrease/staying same when ro > 1.

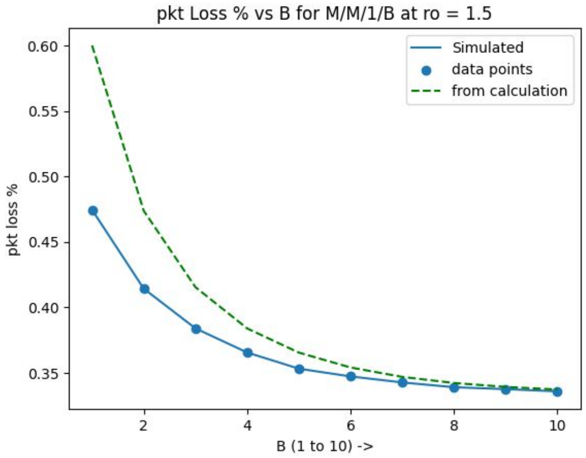
M/M/1/B - Packet Loss % vs B



$\rho = 0.5$



$\rho = 0.7$

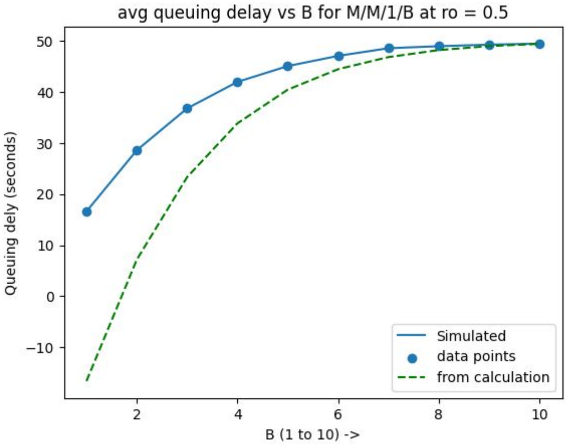


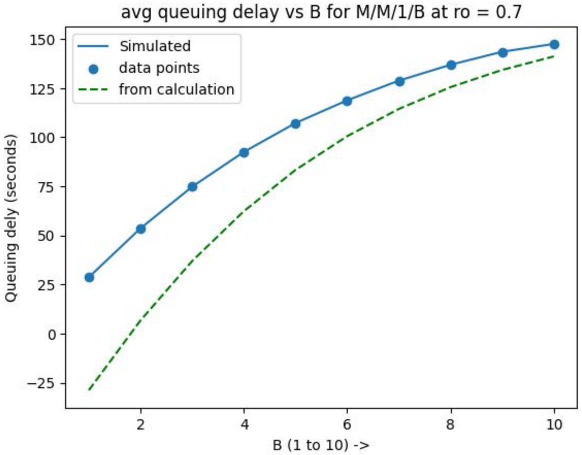
$\rho = 1.5$

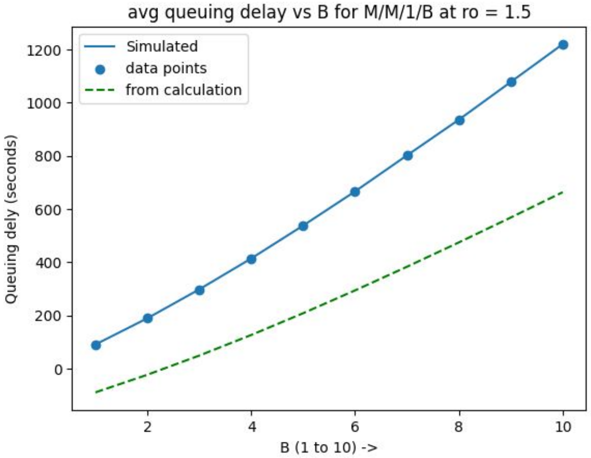
Inferences:

- Mathematical model result and simulation result closes as B increases.
- Blocking Probability is exponentially dependent on B.

M/M/1/B - Queuing Delay vs B







ρ = 0.5 ρ = 0.7 ρ = 1.5

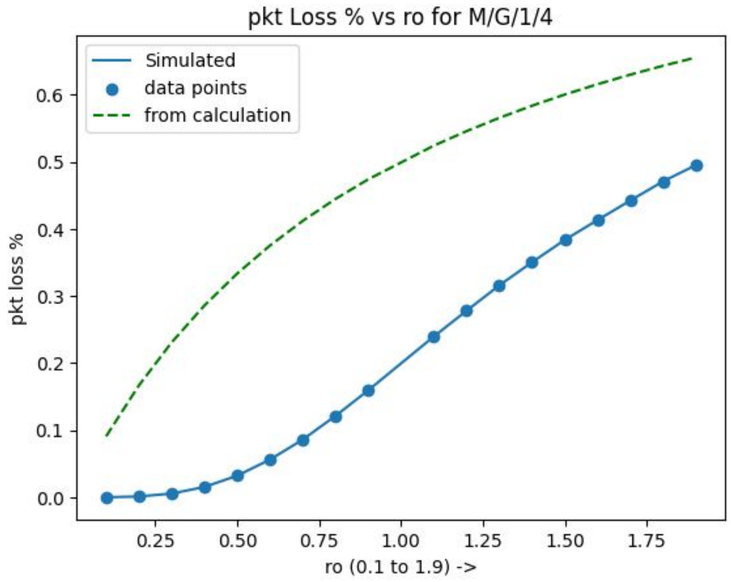
Inferences:

- Mathematical model and simulation result closes for smaller ρ and larger B
- For $\rho < 1$ both simulation and analytical result seem to converge to some asymptotic value but for $\rho > 1$ both analytical result and simulation result increases with increasing rate.
- Have no explanation for why for $\rho > 1$ waiting time isn't converging to a saturation value. Since queue size is fixed an incoming pkt has an upper limit of how much time it's waiting. But both analytical and simulation says waiting time doesn't converge.

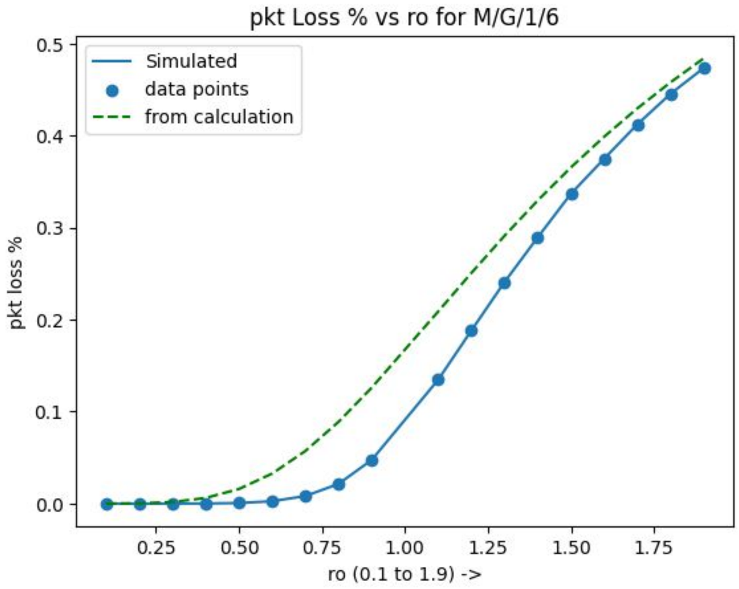
Simulation Result analysis for M/G/1/B

Plotted the packet loss % and delay from simulated result and compared against estimated packet loss % and delay from mathematical model we got for M/M/1/B system.

M/G/1/B - Pkt Loss % vs ρ



Queue size = 3

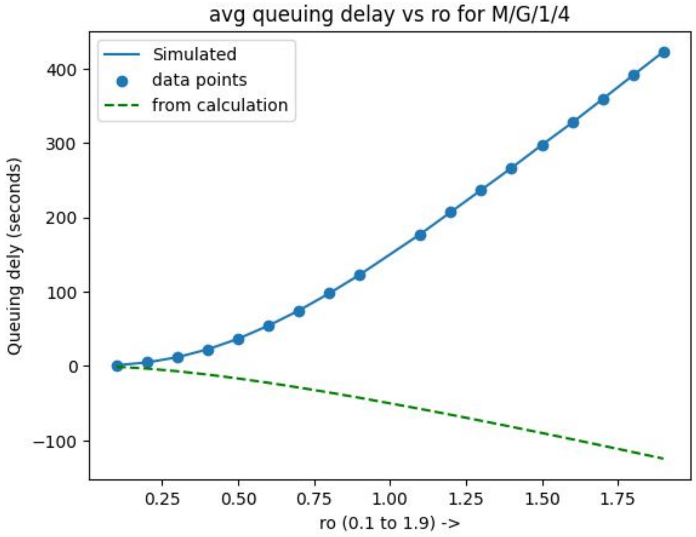


Queue size = 5

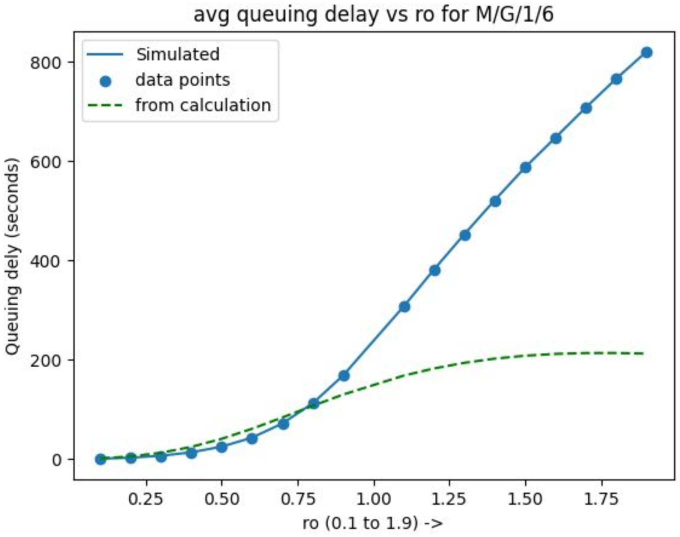
Inference :

- Similar to M/M/1/B system, difference between mathematical model's estimation and simulation result is smaller for higher queue size.

M/G/1/B - Waiting time (W) / Queuing Delay vs ρ



Queue size = 3

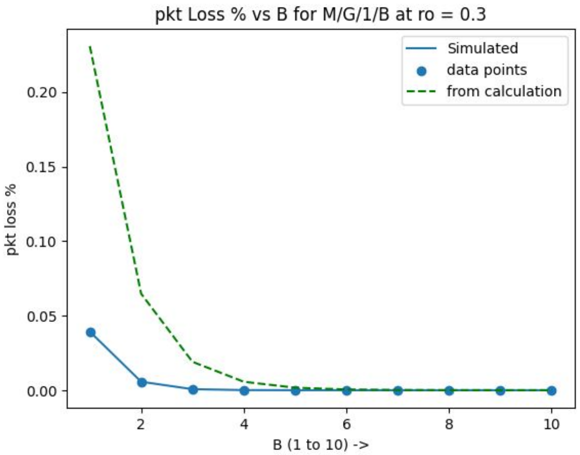


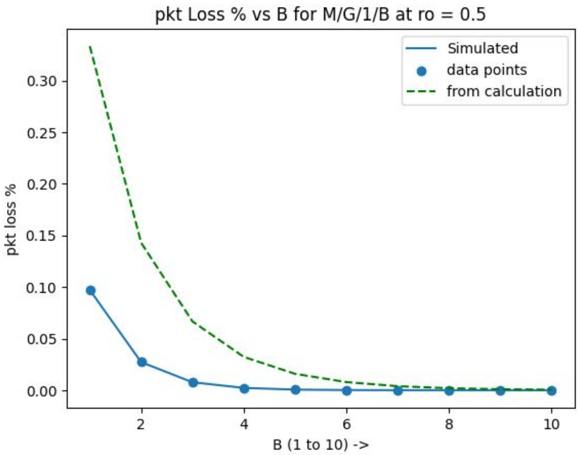
Queue size = 5

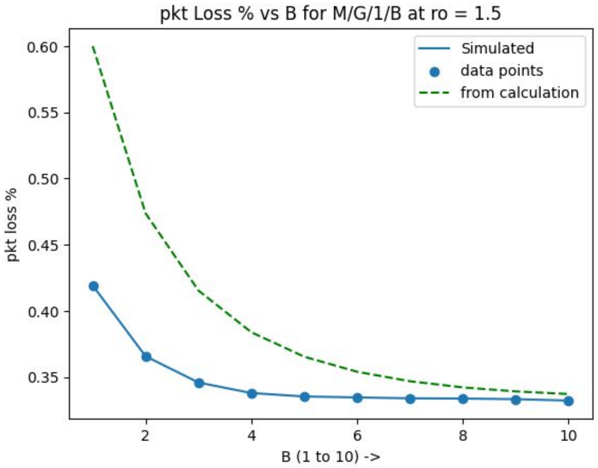
Inference :

- Similar to M/M/1/B for Higher value of B (queue size), Difference between mathematical model and simulation result reduces
- Similar to M/M/1/B Mathematical model matches with simulation result where $\rho < 1$ but simulation result shows that the delay doesn’t decrease or saturates instead it increases with a increasing rate.

M/G/1/B - Packet Loss % vs Queue Size





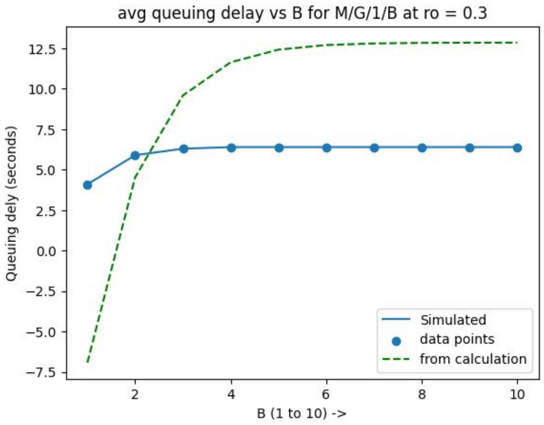


$\rho = 0.3$ $\rho = 0.5$ $\rho = 1.5$

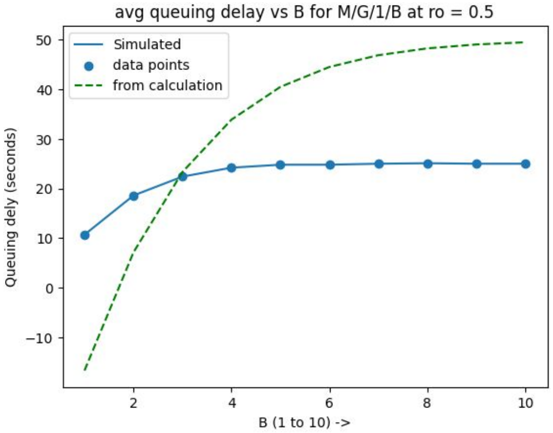
Inferences:

- Similar to M/M/1/B system, Blocking Probability is exponentially dependent on B.

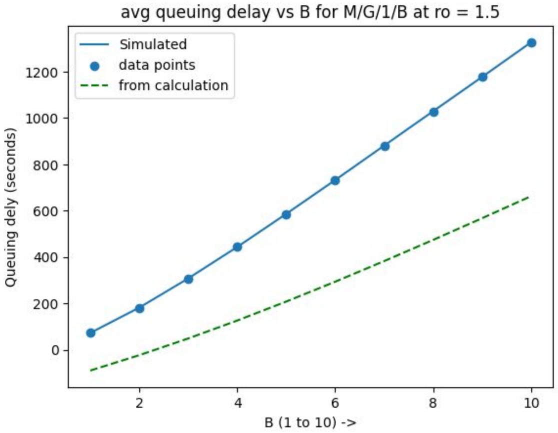
M/G/1/B - Wait Time / Queuing Delay vs Queue Size



ρ = 0.3



ρ = 0.5



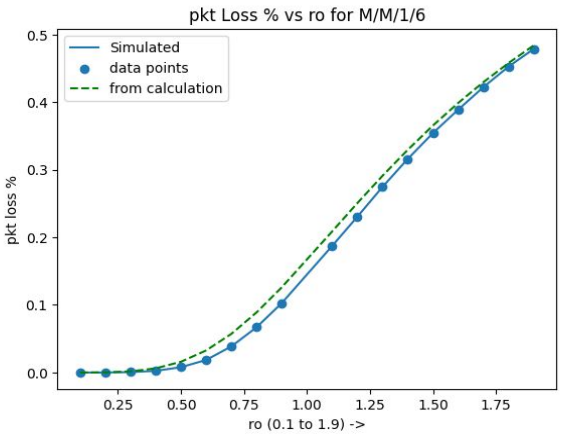
ρ = 1.5

Inferences:

- Similar to M/M/1/B system, when ρ < 1 there is a saturation in delay but when ρ > 1 delay is increasing with a increasing rate.
- There is an offset with the simulated result compared to calculated delay from analytical result of M/M/1/B system; This can be used to get approximate analytical model for M/G/1/B system

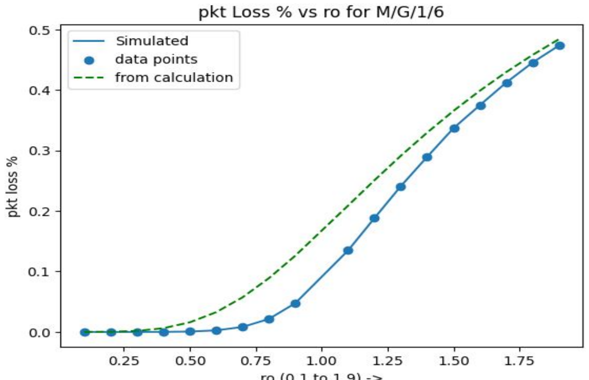
Comparing M/M/1/B and M/G/1/B system

M/M/1/B



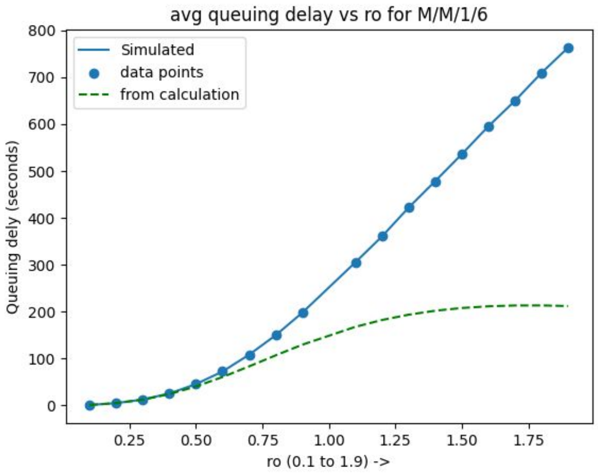
Queue size = 5

M/G/1/B



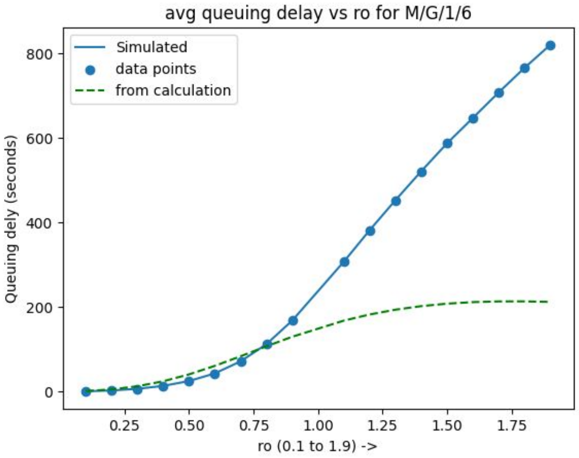
M/G/1/B system retains small pkt loss % till higher ρ value and then steeply increases the pkt loss % and comes close to M/M/1/B system.

M/M/1/B



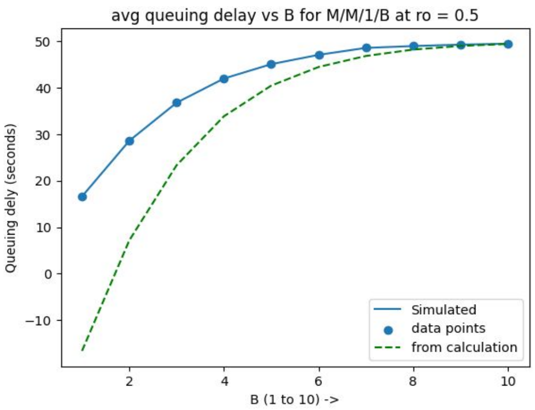
Queue size = 5

M/G/1/B

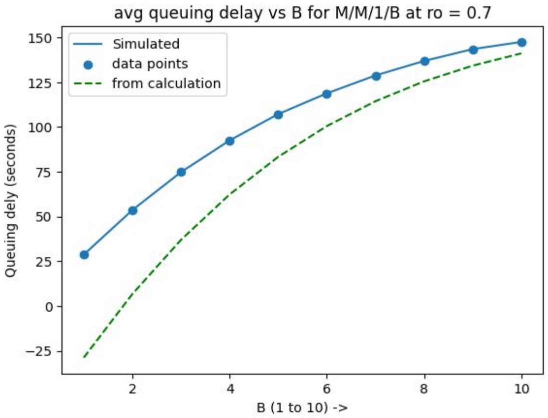


M/G/1/B system delay is smaller than M/M/1/B system till ρ = 0.75 then it increases

M/M/1/B

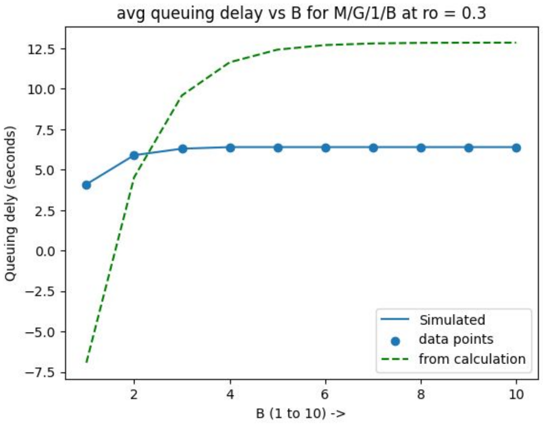


$\rho = 0.5$

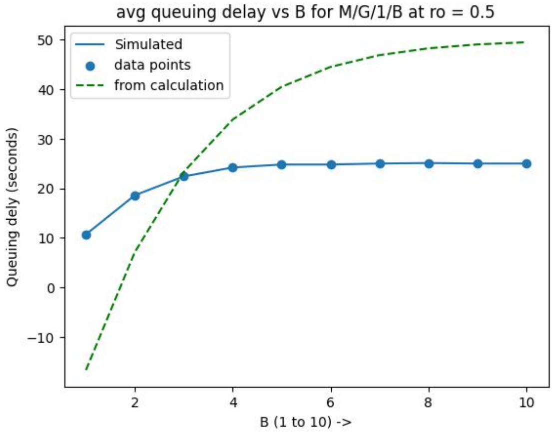


$\rho = 0.7$

M/G/1/B



$\rho = 0.3$



$\rho = 0.5$

Inferences:

- In M/G/1/B system the delay is smaller than M/M/1/B system. (check rho = 0.5 of both system)
- For smaller rho saturation is achieved at smaller value of B
- In M/G/1/B system saturation is achieved at smaller B compared to M/M/1/B system.

Analytical derivation of M/G/1/B

One way to approximate W for M/G/1/B would be to subtract some part from average waiting time for M/G/1/inf (say W_{mg1}) because since there is a limit to how long the queue can be the queuing delay should be reduced by that factor.

Now, queuing delay for M/G/1/B system is,

$W = W_{mg1}$ - ExcessTime $= \lambda * E[x^2]/2(1-\rho)$ - ExcessTime

Work Distribution

Swarnarup - M/M/1/B analytical result, recording data from simulation

Joel - Analysis of simulation results, Approximated mathematical model for M/G/1/B delay.