

An Analysis of Multi-Tiered Feedback Queueing Networks with Four Servers Motivated by Automated Logistics Systems

Nidhi Sharma, Surender Kumar

Abstract: *In the Automated logistics systems highly efficient material handling is required to manage fluctuating flow capacity and complex routing demands. This queueing network applications with important multi-tiered feedback of logistics networks operating with four different servers, one central server linked with the three parallel servers, on a modern automated sorting facilities and distribution centre like amazon fulfilment centre. We model the system as a queueing network where items traverse within multiple parallel tiers after reaching the central server, incorporating feedback loops to simulate re-routing, sorting exceptions, and multi-server processing tasks. Furthermore, we investigate the impact of server allocation policies and routing probabilities on bottlenecks and total system delay. Numerical simulations validate the analytical model, demonstrating how optimal tier configuration and feedback management can significantly enhance operational efficiency. The results provides practical design guidelines for logistics engineers/managers aiming to minimize congestion and maximize throughput in automated Amazon fulfilment warehousing environments.*

Nidhi Sharma (nidhishrm247@gmail.com) Assistant Professor, Department of Mathematics, G.B.D.C., Rohtak
Surender Kumar (surender210882@gmail.com) Assistant Professor1, 2*, Department of Mathematics, G.B.D.C., Rohtak
Corresponding Author: Surender Kumar

Introduction

The literature on queueing theory offers many reports of extensive effort regarding complex networks. Many scholars have examined queueing models with different service types, multiclass services, or multi-servers including Bouchentouf A.A. and Yahiaoui L. (2017), Gupta, R., and D. Gupta (2018), Bouchentouf, A. A. and Guendouzi (2021), Kumari S. (2022), Saini, A. (2023), Agassi M. (2023), Lee, S. et al (2024) and Agarwal, A. et al (2024). Additionally, it's possible that a customer will need to repeatedly join several service lines before a server completes their task to their satisfaction. These circumstances, which can be seen in manufacturing, healthcare, and other settings, gave rise to feedback in queueing systems and, as a result, enhanced the literature on queueing theory by introducing the idea of researching queueing systems with feedback. However, in modern practical applications, a customer or a data packet frequently requires repetitive routing before a task finishes to satisfaction. These circumstances are highly prevalent in modern Multi-Tiered AI Cloud Computing Networks (where unconfident AI models loop requests back to a central gateway for re-authentication) and Automated Warehouse Sorting Hubs (where packages failing weight validation are pushed back to the main conveyor). This physical reality gave rise to feedback queueing systems. Kamal et al. (2023) worked on the feedback queueing model with four servers, one linked centrally with the other three servers having revisit at most once but did not discuss the variations in the mean queue length with respect to the other parameters such as mean arrival rate λ , μ_1 (the mean service rate of the 1st server), the mean service rate of second server μ_2 , q_4 (the probability of going outside the system from 4th server 1st time). To bridge this gap and provide actionable insights for automated logistics i.e. Amazon-style fulfilment warehouse, this paper evaluates the impact of structural feedback revisit probabilities and first-pass completion efficiencies (b_1 , c_1 , d_1) on system congestion.

Notations

- λ : mean arrival rate at 1st server (S_1)
- μ_1 : mean service rate of 1st server (S_1)
- μ_2 : mean service rate of 2nd server.
- μ_3 : mean service rate of 3rd server.
- μ_4 : mean service rate of 4th server
- a_1 : the probability of customer leaving 1st server 1st time.
- a_2 : the probability of customer leaving 1st server 2nd time.
- b_1 : the probability of customer leaving 2nd server 1st time.
- b_2 : the probability of customer leaving 2nd server 2nd time.
- c_1 : the probability of customer leaving 3rd server 1st time.
- c_2 : the probability of customer leaving 3rd server 2nd time.
- d_1 : the probability of customer leaving 4th server 1st time
- d_2 : the probability of customer leaving 4th server 2nd time
- q_{12} : the probability of customer going from 1st to 2nd server 1st time.

q_{12} : the probability of customer going from 1st to 2nd server 2nd time.
 q_2 : the probability of exit of customer from 2nd server 1st time.
 q_{23} : the probability of customer going from 2nd to 3rd server 1st time.
 q_{21} : the probability of customer going from 2nd to 1st server 1st time.
 q'_2 : the probability of exit of customer from 2nd server 2nd time.
 q'_{23} : the probability of customer going from 2nd to 3rd server 2nd time.
 q'_{21} : the probability of customer going from 2nd to 1st server 2nd time.
 q_3 : the probability of exit of customer from 3rd server 1st time.
 q_{31} : the probability of customer going from 3rd to 1st server 1st time.
 q_{32} : the probability of customer going from 3rd to 2nd server 1st time.
 q_{34} : the probability of customer going from 3rd to 4th server 1st time
 q'_3 : the probability of exit of customer from 3rd server 2nd time.
 q'_{31} : the probability of customer going from 3rd to 1st server 2nd time.
 q'_{32} : the probability of customer going from 3rd to 2nd server 2nd time.
 q'_{34} : the probability of customer going from 3rd to 4th server 2nd time.
 q_4 : the probability of exit of customer from 4th server 1st time.
 q'_4 : the probability of exit of customer from 4th server 2nd time.
 q_{41} : the probability of exit of customer from 4th to 1st server 1st time.
 q_{42} : the probability of customer going from 4th to 2nd server 1st time.
 q_{43} : the probability of customer going from 4th to 3rd server 1st time.
 q'_{42} : the probability of customer going from 4th to 2nd server 2nd time
 q'_{43} : the probability of customer going from 4th to 3rd server 2nd time.

In an Amazon-style fulfilment warehouse, where boxes move through multiple processing stages the parameters used in the model can be translated as:

S₁ (or Node 1): Unloading & De-palletization. Incoming bulk inventory is unloaded from trailers and placed onto the main conveyor belts.

S₂ (or Node 2): Primary Picking & Packing. Items are picked from inventory shelves, scanned, and placed into shipping boxes.

S₃ (or Node 3): Downstream Sealing & Labeling. Packed boxes are automatically taped, weighed, and stamped with shipping labels.

S₄ (or Node 4): Final Outbound Sorting. Packages are sorted by zip code into specific lanes to be loaded into delivery trucks.

λ : Inbound Freight Volume. The number of items arriving at the warehouse dock per hour from vendor trucks.

μ_1 : De-palletization Throughput. How many items the receiving dock crew can unpack and register per hour.

μ_2 : Primary Packaging Capacity. The average number of boxes an individual packing station can fill and send down the line per hour.

μ_3 : Sorting & Labeling Speed. The number of packages automated scan-and-apply machine can seal and label per hour.

μ_i : Outbound Loading Rate. How fast the shipping dock team can scan and stack packages into delivery trucks.

q_{ij} (Transition Probability): Conveyor Routing Decisions. The percentage of packages that move from station i to station j .

q_{ji} : Quality Defect Probability: The package moves in reverse second time against the main flow of the warehouse. Instead of progressing toward the delivery truck, it is diverted to a correction lane because something went wrong.

a_1 : The probability that a standard tote safely clears the initial conveyor scan checkpoint and moves directly to its designated downstream zone (Picking, Packing, or Shipping) without activating a feedback loop.

a_2 : The probability that an item successfully exits the conveyor loop after initially failing a scanner check, getting stuck, or recirculating because its destination buffer down the line was completely full (blocking).

b_1 : The probability that a picked item correctly leaves the picking zone on its first pass and heads straight to Packing.

b_2 : The probability that an item leaves the pick station *after* a re-pick event occurs. This happens if the first item retrieved was flagged as physically damaged, or if the robotic drive dropped the bin, triggering a safety stall and immediate queue re-injection.

c_1 : The probability that a packed box cleanly leaves the station on its first attempt and heads directly to Shipping.

c_2 : The probability that a box exits Packing *after* triggering an error loop. This occurs if the built-in inline "Slam Scale" detects a weight mismatch (indicating a wrong item or quantity inside), forcing the box into a quality exception lane to be sliced open, corrected, repacked, and rescanned.

d_1 : The probability that a package successfully leaves the warehouse facility on its intended delivery truck on the first attempt.

d_2 : The probability that an item leaves the shipping dock after failing the first attempt. These models reverse logistics scenarios where a truck departs before the late-stage package is loaded (missed dispatch window), or an item is immediately returned by a carrier at the dock due to critical shipping label damage. The item must then be re-routed back upstream to be re-labelled or re-batched for the next shift.

Conceptualization of the Problem

Queue network is made up of four interaction channels of 1st, 2nd, 3rd and 4th; whereas 2nd, 3rd and 4th are commonly linked with the 1st server in series. The presumed customer of arrive at 1st a server external to the system, operating in accordance with a poisson process with mean rate λ and then goes to 2nd, 3rd or 4th servers for required services.

After receiving services from a server for the first time, clients can switch to another server. 2nd, 3rd or 4th server such that $q_{12}+q_{13}+q_{14}=1$.

If a client moves to another service provider after receiving service from the first provider, the connection is severed, or the client returns to the first, third, or fourth service provider. $q_2 + q_{21} + q_{23} + q_{24} = 1$ Similarly, if someone moves from the first server to the third for the first time—whether switching from another server or returning to it—then the results... $q_3 + q_{31} + q_{32} + q_{34} = 1$. Starting from the fourth server, each must exit the system or return to the lowest-level server; this process continues $q_4 + q_{41} + q_{42} + q_{43} = 1$. similarly, equations for re-examine of the customer to any server can be written here. Let Q_{n_1, n_2, n_3, n_4} is the probability of having n_1, n_2, n_3, n_4 customers at server 1st, 2nd, 3rd and 4th at every given time t .

Steady-state equation for different values of n_1, n_2, n_3 , and n_4 are given by:

$$\begin{aligned}
 (\lambda + \mu_1 + \mu_2 + \mu_3 + \mu_4) Q_{n_1, n_2, n_3, n_4} &= \lambda Q_{n_1-1, n_2, n_3, n_4} + \mu_1 (a_1 q_{12} + a_2 q'_{12}) Q_{n_1+1, n_2-1, n_3, n_4} \\
 &+ \mu_1 (a_{14} q_{14} + a_2 q'_{14}) Q_{n_1+1, n_2, n_3, n_4-1} + \mu_1 (a_1 q_{13} + a_2 q'_{13}) Q_{n_1+1, n_2, n_3-1, n_4} \\
 &+ \mu_2 (b_1 q_{21} + b_2 q'_{21}) Q_{n_1-1, n_2+1, n_3, n_4} + \mu_2 (b_1 q_{21} + b_2 q'_{21}) Q_{n_1, n_2+1, n_3, n_4} \\
 &+ \mu_2 (b_1 q_{23} + b_2 q'_{23}) Q_{n_1, n_2+1, n_3-1, n_4} + \mu_2 (b_1 q_{24} + b_2 q'_{24}) Q_{n_1, n_2+1, n_3, n_4-1} \\
 &+ \mu_3 (c_1 q_{31} + c_2 q'_{31}) Q_{n_1, n_2, n_3+1, n_4} + \mu_3 c_1 q_{31} Q_{n_1-1, n_2, n_3+1, n_4} \\
 &+ (c_1 q_{32} + c_2 q'_{32}) \mu_3 Q_{n_1, n_2-1, n_3+1, n_4} + (c_1 q_{32} + c_2 q'_{32}) \mu_3 Q_{n_1, n_2-1, n_3+1, n_4} \\
 &+ (c_1 q_{34} + c_2 q'_{34}) \mu_3 Q_{n_1, n_2, n_3+1, n_4+1} + (d_1 q_4 + d_2 q'_4) \mu_4 Q_{n_1, n_2, n_3, n_4+1} \\
 &+ (d_1 q_{41} + d_2 q'_{41}) \mu_4 Q_{n_1-1, n_2, n_3, n_4+1} + (d_1 q_{42} + d_2 q'_{42}) \mu_4 Q_{n_1, n_2-1, n_3, n_4+1} \\
 &+ (d_1 q_{43} + d_2 q'_{43}) \mu_4 Q_{n_1, n_2, n_3-1, n_4+1} \\
 &\dots (1)
 \end{aligned}$$

On solving this equation using generating function technique, we have the function:

$$F(X, Y, Z, R) = \frac{f(X, Y, Z, R)}{g(X, Y, Z, R)}$$

where;

$$f(X, Y, Z, R) = \mu_1 F_0(Y, Z, R)$$

$$\begin{aligned}
 &\left[1 - \frac{1}{X} \{ (a_1 q_{12} + a_2 q'_{12}) Y + (a_1 q_{13} + a_2 q'_{13}) Z + (a_1 q_{14} + a_2 q'_{14}) R \} \right] \\
 &\quad + \mu_2 F_0(X, Z, R) \\
 &\quad \left[1 - \frac{1}{Y} \{ (b_1 q_2 + b_2 q'_2) + (b_1 q_{23} + b_2 q'_{23}) Z + (b_1 q_{24} + b_2 q'_{24}) R + (b_1 q_{21}) X \} \right] \\
 &+ \mu_3 F_0(X, Y, R) \left[1 - \frac{1}{Z} \{ (C_1 q_3 + C_2 q'_3) + (C_1 q_{32} + C_2 q'_{32}) Y + (C_1 q_{34} + C_2 q'_{34}) R + \right. \\
 &\quad \left. (C_1 q_{31}) X \} \right] \\
 &+ \mu_4 F_0(X, Y, Z) \left[1 - \frac{1}{R} \{ (d_1 q_4 + d_2 q'_4) + (d_1 q_{42} + d_2 q'_{42}) Y + (d_1 q_{43} + d_2 q'_{43}) Z + \right. \\
 &\quad \left. d_1 q_{41} X \} \right]
 \end{aligned}$$

and

$$\begin{aligned}
 g(X, Y, Z, R) = & \lambda(1 - X) \\
 & + \mu_1 \left[1 - \frac{1}{X} \{ (a_1 q_{12} + a_2 q'_{12})Y + (a_1 q_{13} + a_2 q'_{13})Z + (a_1 q_{14} + a_2 q'_{14})R \} \right] \\
 & + \mu_2 \left[1 - \frac{1}{Y} \{ (b_1 q_{22} + b_2 q'_{22}) + (b_1 q_{23} + b_2 q'_{23})Z + (b_1 q_{24} + b_2 q'_{24})R + \right. \\
 & \left. (b_1 q_{21})X \} \right] \\
 & + \mu_3 \left[1 - \frac{1}{Z} \{ (C_1 q_{32} + C_2 q'_{32}) + (C_1 q_{33} + C_2 q'_{33})Y + (C_1 q_{34} + C_2 q'_{34})R + \right. \\
 & \left. (C_1 q_{31})X \} \right] \\
 & + \mu_4 \left[1 - \frac{1}{R} \{ (d_1 q_{42} + d_2 q'_{42}) + (d_1 q_{43} + d_2 q'_{43})Y + (d_1 q_{44} + d_2 q'_{44})Z + \right. \\
 & \left. d_1 q_{41}X \} \right]
 \end{aligned}$$

For convenience, let us define:

$$\begin{aligned}
 F_1 &= F_0(Y, Z, R) \\
 F_2 &= F_0(X, Z, R) \\
 F_3 &= F_0(X, Y, R) \\
 F_4 &= F_0(X, Y, Z)
 \end{aligned}$$

For $X = Y = Z = R = 1$ and leveraging the hypothesis/assumption reduces to in determinant form $\left(\frac{0}{0}\right)$. Taking $Y = Z = R = 1$ and limit $X \rightarrow 1$ in (31) we have

$$\begin{aligned}
 -\lambda + \mu_1 - \mu_2 b_1 q_{21} - \mu_3 C_1 q_{31} - \mu_4 d_1 q_{41} &= \mu_1 F_1 + b_1 q_{21} \mu_2 F_2 - C_1 q_{31} \mu_3 F_3 - d_1 q_{41} \mu_4 F_4 \\
 \dots(2)
 \end{aligned}$$

Now, for $X = Z = R = 1$ and taking $Y \rightarrow 1$ in (31):

(Using L' Hospital rule w.r.t y then but $X = Y = Z = R = 1$)

$$\begin{aligned}
 & -\mu_1 (a_1 q_{12} + a_2 q'_{12}) + \mu_2 - \mu_3 (C_1 q_{32} + C_2 q'_{32}) - \mu_4 (d_1 q_{42} + d_2 q'_{42}) \\
 & = -\mu_1 F_1 (a_1 q_{12} + a_2 q'_{12}) + \mu_2 F_2 - \mu_3 (C_1 q_{32} + C_2 q'_{32}) F_3 - F_4 \mu_4 (d_1 q_{42} + d_2 q'_{42}) \\
 \dots(3)
 \end{aligned}$$

For $X = Y = R = 1$ and taking limit $Z \rightarrow 1$ we have :

$$\begin{aligned}
 & -\mu_2 (b_1 q_{23} + b_2 q'_{23}) - \mu_1 (a_1 q_{13} + a_2 q'_{13}) + \mu_3 \\
 & -\mu_4 (d_1 q_{43} + d_2 q'_{43}) = -\mu_1 F_1 (a_1 q_{13} + a_2 q'_{13}) - \mu_2 F_2 (b_1 q_{23} + b_2 q'_{23}) \\
 & + \mu_3 F_3 - \mu_4 F_4 (d_1 q_{43} + d_2 q'_{43}) \\
 \dots(4)
 \end{aligned}$$

For $X = Y = Z = 1$ and taking limit $R \rightarrow 1$ we have :

$$\begin{aligned}
 & -\mu_1 (a_1 q_{14} + a_2 q'_{14}) - \mu_2 (b_1 q_{24} + b_2 q'_{24}) - \mu_3 (C_1 q_{34} + C_2 q'_{34}) + \mu_4 \\
 & = -\mu_1 (a_1 q_{14} + a_2 q'_{14}) F_1 - \mu_2 (b_1 q_{24} + b_2 q'_{24}) F_2 - \mu_3 (C_1 q_{34} + C_2 q'_{34}) F_3 + \mu_4 F_4 \dots(5)
 \end{aligned}$$

On solving (2), (3), (4) and (5), we have:

$$F_1 = \frac{\begin{aligned} &\mu_1 b_1 (q_{21}(q_{12}(q'_{14}q_{23}-q'_{23})-q_{13}q'_{13}q_{23}-q'_{12}q'_{13}q'_{23})+(-q'_{12}q'_{14}-q_{13})q_{21}^2)+\lambda(q_{13}(\mu_4q_{23}+q'_{21}) \\ &+q'_{14}q_{23}+q'_{12}(q_{23}\mu_4+q'_{14}q'_{21})-q'_{23})+\mu_1(q_{13}(-\mu_4q_{23}-q'_{21})-q'_{14}q_{23}+q'_{12}(-q_{23}\mu_4-q'_{14}q'_{21}) \\ &+q'_{23})+q_{41}\mu_1 d_1(q_{12}(-\mu_4q_{23}-q'_{21})-q'_{13}q_{23}+q_{21}(q'_{12}\mu_4-1)-q'_{12}q'_{13}q'_{21})+q_{31}\mu_1(q_{12} \\ &(-q_{23}\mu_4-q'_{14}q'_{21})+q_{21}(-q_{13}\mu_4-q'_{14})-q'_{13}q'_{23}+q_{13}q'_{13}q'_{21})C_1 \end{aligned}}{\begin{aligned} &\mu_1 b_1 (q_{21}(q_{12}(q'_{14}q_{23}-q'_{23})-q_{13}q'_{13}q_{23}-q'_{12}q'_{13}q'_{23})+(-q'_{12}q'_{14}-q_{13})q_{21}^2)+\mu_1(q_{13}(-\mu_4q_{23} \\ &-q'_{21})-q'_{14}q_{23}+q'_{12}(-q_{23}\mu_4-q'_{14}q'_{21})+q'_{23})+q_{41}\mu_1 d_1(q_{12}(-\mu_4q_{23}-q'_{21})-q'_{13}q_{23} \\ &+q_{21}(q'_{12}\mu_4-1)-q'_{12}q'_{13}q'_{21})+q_{31}\mu_1(q_{12}(-q_{23}\mu_4-q'_{14}q'_{21})+q_{21}(-q_{13}\mu_4-q'_{14}) \\ &-q'_{13}q'_{23}+q_{13}q'_{13}q'_{21})C_1 \end{aligned}} \dots (6)$$

$$F_2 = \frac{\begin{aligned} &\mu_2 b_1 (q_{21}(q_{12}(q'_{14}q_{23}-q'_{23})-q_{13}q'_{13}q_{23}-q'_{12}q'_{13}q'_{23})+(-q'_{12}q'_{14}-q_{13})q_{21}^2) \\ &+\mu_2(q_{13}(-\mu_4q_{23}-q'_{21})-q'_{14}q_{23}+q'_{12}(-q_{23}\mu_4-q'_{14}q'_{21})+q'_{23})+q_{41}\mu_2 \\ &d_1(q_{12}(-\mu_4q_{23}-q'_{21})-q'_{13}q_{23}+q_{21}(q'_{12}\mu_4-1)-q'_{12}q'_{13}q'_{21})+\lambda(q_{12}(q'_{14}q_{23}-q'_{23}) \\ &-q_{13}q'_{13}q_{23}-q'_{12}q'_{13}q'_{23})+\mu_2(-q'_{12}q'_{14}-q_{13})q_{21}+q_{31}\mu_2(q_{12}(-q_{23}\mu_4-q'_{14}q'_{21}) \\ &+q_{21}(-q_{13}\mu_4-q'_{14})-q'_{13}q'_{23}+q_{13}q'_{13}q'_{21})C_1 \end{aligned}}{\begin{aligned} &\mu_2 b_1 (q_{21}(q_{12}(q'_{14}q_{23}-q'_{23})-q_{13}q'_{13}q_{23}-q'_{12}q'_{13}q'_{23})+(-q'_{12}q'_{14}-q_{13})q_{21}^2)+ \\ &\mu_2(q_{13}(-\mu_4q_{23}-q'_{21})-q'_{14}q_{23}+q'_{12}(-q_{23}\mu_4-q'_{14}q'_{21})+q'_{23})+q_{41}\mu_2 d_1 \\ &(q_{12}(-\mu_4q_{23}-q'_{21})-q'_{13}q_{23}+q_{21}(q'_{12}\mu_4-1)-q'_{12}q'_{13}q'_{21})+q_{31}\mu_2 \\ &(q_{12}(-q_{23}\mu_4-q'_{14}q'_{21})+q_{21}(-q_{13}\mu_4-q'_{14})-q'_{13}q'_{23}+q_{13}q'_{13}q'_{21})C_1 \end{aligned}} \dots (7)$$

$$F_3 = \frac{\begin{aligned} &\mu_3 b_1 (q_{21}(q_{12}(q'_{14}q_{23}-q'_{23})-q_{13}q'_{13}q_{23}-q'_{12}q'_{13}q'_{23})+(-q'_{12}q'_{14}-q_{13})q_{21}^2) \\ &+\mu_3(q_{13}(-\mu_4q_{23}-q'_{21})-q'_{14}q_{23}+q'_{12}(-q_{23}\mu_4-q'_{14}q'_{21})+q'_{23})+q_{41}\mu_3 d_1 \\ &(q_{12}(-\mu_4q_{23}-q'_{21})-q'_{13}q_{23}+q_{21}(q'_{12}\mu_4-1)-q'_{12}q'_{13}q'_{21})+q_{31}\mu_3(q_{12} \\ &(-q_{23}\mu_4-q'_{14}q'_{21})+q_{21}(-q_{13}\mu_4-q'_{14})-q'_{13}q'_{23}+q_{13}q'_{13}q'_{21})C_1 \\ &+\lambda(q_{12}(-q_{23}\mu_4-q'_{14}q'_{21})+q_{21}(-q_{13}\mu_4-q'_{14})-q'_{13}q'_{23}+q_{13}q'_{13}q'_{21}) \end{aligned}}{\begin{aligned} &\mu_3 b_1 (q_{21}(q_{12}(q'_{14}q_{23}-q'_{23})-q_{13}q'_{13}q_{23}-q'_{12}q'_{13}q'_{23})+(-q'_{12}q'_{14}-q_{13})q_{21}^2) \\ &+\mu_3(q_{13}(-\mu_4q_{23}-q'_{21})-q'_{14}q_{23}+q'_{12}(-q_{23}\mu_4-q'_{14}q'_{21})+q'_{23})+q_{41}\mu_3 d_1 \\ &(q_{12}(-\mu_4q_{23}-q'_{21})-q'_{13}q_{23}+q_{21}(q'_{12}\mu_4-1)-q'_{12}q'_{13}q'_{21})+q_{31}\mu_3(q_{12} \\ &(-q_{23}\mu_4-q'_{14}q'_{21})+q_{21}(-q_{13}\mu_4-q'_{14})-q'_{13}q'_{23}+q_{13}q'_{13}q'_{21})C_1 \end{aligned}} \dots (8)$$

$$F_4 = 1 + \frac{\lambda(q_{12}(-\mu_4q_{23}-q'_{21})-q'_{13}q_{23}+q_{21}(q'_{12}\mu_4-1)-q'_{12}q'_{13}q'_{21})}{\begin{aligned} &\mu_4 b_1 (q_{21}(q_{12}(q'_{14}q_{23}-q'_{23})-q_{13}q'_{13}q_{23}-q'_{12}q'_{13}q'_{23})+(-q'_{12}q'_{14}-q_{13})q_{21}^2) \\ &+\mu_4(q_{13}(-\mu_4q_{23}-q'_{21})-q'_{14}q_{23}+q'_{12}(-q_{23}\mu_4-q'_{14}q'_{21})+q'_{23})+q_{41}\mu_4 d_1(q_{12}(-\mu_4q_{23} \\ &-q'_{21})-q'_{13}q_{23}+q_{21}(q'_{12}\mu_4-1)-q'_{12}q'_{13}q'_{21})+q_{31}\mu_4(q_{12}(-q_{23}\mu_4-q'_{14}q'_{21}) \\ &+q_{21}(-q_{13}\mu_4-q'_{14})-q'_{13}q'_{23}+q_{13}q'_{13}q'_{21})C_1 \end{aligned}} \dots (9)$$

Let L_{q1} , L_{q2} , L_{q3} and L_{q4} denote the mean queue length at 1st, 2nd, 3rd and 4th server then we have:

$$L_{q1} = \frac{\left(\frac{\partial f}{\partial X}\right)_{(1,1,1,1)} \left(\frac{\partial^2 g}{\partial X^2}\right)_{(1,1,1,1)} + \left(\frac{\partial g}{\partial X}\right)_{(1,1,1,1)} \left(\frac{\partial^2 f}{\partial X^2}\right)_{(1,1,1,1)}}{2 \left[\left(\frac{\partial g}{\partial X}\right)_{(1,1,1,1)} \right]^2} \frac{(\mu_1 F_1 - b_1 q_{21} \mu_2 F_2 - c_1 q_{31} \mu_3 F_3 - d_1 q_{41} \mu_4 F_4)(-2\mu_1) + (-\lambda + \mu_1 - b_1 q_{21} \mu_2 - c_1 q_{31} \mu_3 - d_1 q_{41} \mu_4)(-2\mu_1 F_1)}{2[-\lambda + \mu_1\{(a_1 q_{12} + a_2 q'_{12}) + (a_1 q_{13} + a_2 q'_{13}) + (a_1 q_{14} + a_2 q'_{14})\} - \mu_2 b_1 q_{21} - \mu_3 c_1 q_{31} - \mu_4 d_1 q_{41}]^2} \dots (10)$$

$$L_{q2} = \frac{\left(\frac{\partial f}{\partial Y}\right)_{(1,1,1,1)} \left(\frac{\partial^2 g}{\partial Y^2}\right)_{(1,1,1,1)} + \left(\frac{\partial g}{\partial Y}\right)_{(1,1,1,1)} \left(\frac{\partial^2 f}{\partial Y^2}\right)_{(1,1,1,1)}}{2 \left[\left(\frac{\partial g}{\partial Y}\right)_{(1,1,1,1)} \right]^2}$$

$$= \frac{[-\mu_1 F_1(a_1 q_2 + a_2 q'_{12}) + \mu_2 F_2 - \mu_3 F_3(c_1 q_{32} + c_2 q'_{32}) - \mu_4 F_4(d_1 q_{42} + d_2 q'_{42})](-2\mu_2) + [-\mu_1(a_1 q_{12} + a_2 q'_{12}) + \mu_2 - \mu_3(c_1 q_{32} + c_2 q'_{32}) - \mu_4(d_1 q_{42} + d_2 q'_{42})](2 - \mu_2 F_2)}{2[-\mu_1(a_1 q_{12} + a_2 q'_{12}) + \mu_2 - \mu_3(c_1 q_{32} + c_2 q'_{32}) - \mu_4(d_1 q_{42} + d_2 q'_{42})]^2}$$

...(11)

$$Lq_3 = \frac{\left(\frac{\partial f}{\partial Z}\right)_{(1,1,1,1)} \left(\frac{\partial^2 g}{\partial Z^2}\right)_{(1,1,1,1)} + \left(\frac{\partial g}{\partial Z}\right)_{(1,1,1,1)} \left(\frac{\partial^2 f}{\partial Z^2}\right)_{(1,1,1,1)}}{2 \left[\left(\frac{\partial g}{\partial Z}\right)_{(1,1,1,1)}\right]^2} \frac{[-\mu_1 F_1(a_1 q_{13} + a_2 q'_{13}) - \mu_2 F_2(b_1 q_{23} + b_2 q'_{23}) + \mu_3 F_3 - \mu_4 F_4(d_1 q_{43} + d_2 q'_{43})](-2\mu_3) + [-\mu_1(a_1 q_{13} + a_2 q'_{13}) - \mu_2(b_1 q_{23} + b_2 q'_{23}) + \mu_3 - \mu_4(d_1 q_{43} + d_2 q'_{43})](-2\mu_3 F_3)}{2[-\mu_1(a_1 q_{13} + a_2 q'_{13}) - \mu_2(b_1 q_{23} + b_2 q'_{23}) + \mu_3 - \mu_4(d_1 q_{43} + d_2 q'_{43})]^2}$$

...(12)

$$Lq_4 = \frac{\left(\frac{\partial f}{\partial R}\right)_{(1,1,1,1)} \left(\frac{\partial^2 g}{\partial R^2}\right)_{(1,1,1,1)} + \left(\frac{\partial g}{\partial R}\right)_{(1,1,1,1)} \left(\frac{\partial^2 f}{\partial R^2}\right)_{(1,1,1,1)}}{2 \left[\left(\frac{\partial g}{\partial R}\right)_{(1,1,1,1)}\right]^2} \frac{[-\mu_1 F_1(a_1 q_{14} + a_2 q'_{14}) - \mu_2 F_2(b_1 q_{24} + b_2 q'_{24}) - \mu_3 F_3(c_1 q_{34} + c_2 q'_{34}) + \mu_4 F_4](-2\mu_4) + [-\mu_1(a_1 q_{14} + a_2 q'_{14}) - \mu_2(b_1 q_{24} + b_2 q'_{24}) - \mu_3(c_1 q_{34} + c_2 q'_{34}) + \mu_4(-2\mu_4 F_4)]}{2[-(a_1 q_{14} + a_2 q'_{14})\mu_1 - (b_1 q_{24} + b_2 q'_{24})\mu_2 - (c_1 q_{34} + c_2 q'_{34})\mu_3 + \mu_4]^2}$$

.....(13)

If we denote the average queue length in the queuing system as Lq, we obtain the following:

$$Lq = Lq_1 + Lq_2 + Lq_3 + Lq_4$$

$$Lq = \frac{(\mu_1 F_1 - b_1 q_{21} \mu_2 F_2 - c_1 q_{31} \mu_3 F_3 - d_1 q_{41} \mu_4 F_4)(-2\mu_1) + (-\lambda + \mu_1 - b_1 q_{21} \mu_2 - c_1 q_{31} \mu_3 - d_1 q_{41} \mu_4)(-2\mu_1 F_1)}{2[-\lambda + \mu_1\{(a_1 q_{12} + a_2 q'_{12}) + (a_1 q_{13} + a_2 q'_{13}) + (a_1 q_{14} + a_2 q'_{14})\} - \mu_2 b_1 q_{21} - \mu_3 c_1 q_{31} - \mu_4 d_1 q_{41}]^2} + \frac{[-\mu_1 F_1(a_1 q_2 + a_2 q'_{12}) + \mu_2 F_2 - \mu_3 F_3(c_1 q_{32} + c_2 q'_{32}) - \mu_4 F_4(d_1 q_{42} + d_2 q'_{42})](-2\mu_2) + [-\mu_1(a_1 q_{12} + a_2 q'_{12}) + \mu_2 - \mu_3(c_1 q_{32} + c_2 q'_{32}) - \mu_4(d_1 q_{42} + d_2 q'_{42})](2 - \mu_2 F_2)}{2[-\mu_1(a_1 q_{12} + a_2 q'_{12}) + \mu_2 - \mu_3(c_1 q_{32} + c_2 q'_{32}) - \mu_4(d_1 q_{42} + d_2 q'_{42})]^2} + \frac{[-\mu_1 F_1(a_1 q_{13} + a_2 q'_{13}) - \mu_2 F_2(b_1 q_{23} + b_2 q'_{23}) + \mu_3 F_3 - \mu_4 F_4(d_1 q_{43} + d_2 q'_{43})](-2\mu_3) + [-\mu_1(a_1 q_{13} + a_2 q'_{13}) - \mu_2(b_1 q_{23} + b_2 q'_{23}) + \mu_3 - \mu_4(d_1 q_{43} + d_2 q'_{43})](-2\mu_3 F_3)}{2[-\mu_1(a_1 q_{13} + a_2 q'_{13}) - \mu_2(b_1 q_{23} + b_2 q'_{23}) + \mu_3 - \mu_4(d_1 q_{43} + d_2 q'_{43})]^2} + \frac{[-\mu_1 F_1(a_1 q_{14} + a_2 q'_{14}) - \mu_2 F_2(b_1 q_{24} + b_2 q'_{24}) - \mu_3 F_3(c_1 q_{34} + c_2 q'_{34}) + \mu_4 F_4](-2\mu_4) + [-\mu_1(a_1 q_{14} + a_2 q'_{14}) - \mu_2(b_1 q_{24} + b_2 q'_{24}) - \mu_3(c_1 q_{34} + c_2 q'_{34}) + \mu_4(-2\mu_4 F_4)]}{2[-(a_1 q_{14} + a_2 q'_{14})\mu_1 - (b_1 q_{24} + b_2 q'_{24})\mu_2 - (c_1 q_{34} + c_2 q'_{34})\mu_3 + \mu_4]^2}$$

...(14)

Results and Discussion about the model in Context of Amazon-style fulfilment warehouse:
4.1 With other parameters held constant, the Pattern of the Average queue length (L_q)—which depends on the service rate of the second unit (μ_2)—is presented in Table 1 and Figure 1 for various values of the parameter μ_3 (service rate of the third unit) remaining parameters constant.

Table 1

$\lambda=10, a_1=0.6, a_2=0.4, \mu_1=3, \mu_4=0.2, b_1=0.7, b_2=0.3,$ $c_1=0.5, c_2=0.5, d_1=0.8, d_2=0.2, q_{13}=0.3,$ $q_{12}=0.2, q_{14}=0.5, q_{13}'=0.2, q_{12}'=0.1, q_{14}'=0.7$ $q_2=0.1, q_{21}=0.4, q_{23}=0.3, q_{24}=0.2, q_2'=0.3,$ $q_{21}'=0.3, q_{23}'=0.1, q_{24}'=0.3, q_3=0.6, q_{31}=0.2,$ $q_{32}=0.15, q_{34}=0.05, q_3'=0.2,$ $q_{32}'=0.3, q_{34}'=0.5, q_4=0.7, q_{41}=0.1, q_{42}=0.15,$ $q_{43}=0.05, q_{44}'=0.8, q_{42}'=0.15, q_{43}'=0.05$			
	$\mu_3 = 5$	$\mu_3 = 6$	$\mu_3 = 7$
μ_2	L_q	L_q	L_q
2	526.2041	431.0576	399.4776
3	702.4662	873.4386	1068.438
4	848.5349	1043.921	1266.269
5	1009.34	1228.926	1476.917
6	1187.978	1432.993	1707.629
7	1385.788	1657.723	1960.291
8	1603.722	1904.249	2236.126
9	1842.511	2173.552	2536.176
10	2102.618	2466.522	2861.409

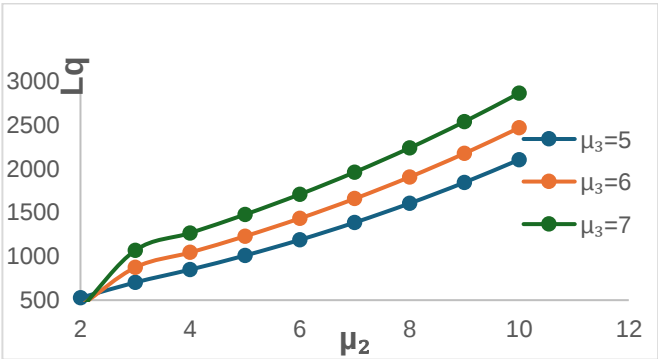


Fig.1

Table 1 and Fig. 1 Interpretation: This data models what happens when we accelerate Primary Picking & Packing (μ_2) under different speeds of Downstream Sealing & Labelling (μ_3).

- **The System Paradox:** As picking and packing get faster (μ_2 goes from 2 to 10), the total backlogged queue L_q dramatically spikes. For instance, at a labelling speed of $\mu_3 = 5$, accelerating our packers pushes the warehouse backlog from 526.2 to 2,102.6 units.
- **The Downstream Flooding Effect:** When we speed up sealing and labelling ($\mu_3 = 5$ to $\mu_3 = 7$), the backlog worsens even faster at higher picking speeds. At $\mu_2 = 10$, increasing the labelling speed pushes the queue up from 2,102.6 to 2,861.4 items.
- **Operational Context:** Speeding up our packing and labelling stations simply drives packages faster toward the Final Outbound Sorting docks (S_4). Because the truck-loading rate is fixed at a crawl ($\mu_4 = 0.2$), accelerating upstream stations creates a massive traffic jam right before the loading docks.

4.2 Table 2 and Figure 2 show how the average queue length in the system (L_q) varies as a function of μ_2 (the average throughput of four servers) for different values of μ_4 (the average throughput of four servers), while keeping the values of the other parameters constant.

Table 2

$\lambda=10, a_1=0.6, a_2=0.4, \mu_1=3, \mu_3=0.1, b_1=0.7, b_2=0.3,$ $c_1=0.5, c_2=0.5, d_1=0.8, d_2=0.2, q_{13}=0.3,$ $q_{12}=0.2, q_{14}=0.5, q_{13}'=0.2, q_{12}'=0.1, q_{14}'=0.7$ $q_2=0.1, q_{21}=0.4, q_{23}=0.3, q_{24}=0.2, q_2'=0.3,$ $q_{21}'=0.3, q_{23}'=0.1, q_{24}'=0.3, q_3=0.6, q_{31}=0.2,$ $q_{32}=0.15, q_{34}=0.05, q_3'=0.2,$ $q_{32}'=0.3, q_{34}'=0.5, q_4=0.7, q_{41}=0.1, q_{42}=0.15,$ $q_{43}=0.05, q_{44}'=0.8, q_{42}'=0.15, q_{43}'=0.05$			
	Lq		
μ_2	$\mu_4=0.2$	$\mu_4=0.3$	$\mu_4=0.4$
2	124.3086	95.34873	70.41079
3	246.5207	209.4175	176.8767
4	327.9664	283.2045	243.6186
5	414.3167	361.109	313.7466
6	512.3118	449.7825	393.8172
7	624.318	551.5481	486.1107
8	751.7753	667.8005	591.9779
9	895.85	799.6543	712.4834
10	1057.606	948.1145	848.5756

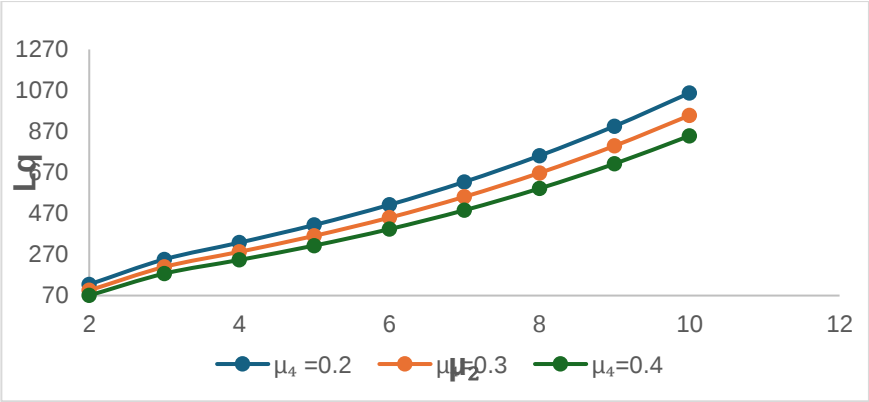


Fig. 2

Table 2 and Fig.2 Interpretation: This table 2 and fig.2 isolates how changing the Outbound Loading Rate (μ_4) affects the backlog as packing speeds change.

- **Positive Relief from the Loading Docks:** Unlike Table 1, increasing the speed at which trailers are loaded (μ_4 moving from 0.2 to 0.4) consistently shrinks the total warehouse backlog.
- **Backlog Reduction Metrics:** When our packers are operating at a moderate pace ($\mu_2 = 4$), upgrading the shipping dock loading speed from $\mu_4 = 0.2$ to $\mu_4 = 0.4$ drops the inventory backlog from 327.9 down to 243.6 boxes. Even at maximum packing speeds ($\mu_2 = 10$), boosting the shipping dock performance reduces the backlog from 1,057.6 to 848.5 units.
- **Operational Context:** This proves that the outbound loading dock is the structural bottleneck of the entire facility. Every increment of efficiency added to the shipping dock directly flushes accumulated work-in-progress (WIP) out of the warehouse doors.

4.3 Table 3 and Figure 3 show how the average queue length in the system (L_q) depends on q_4 (the probability of leaving the system after service by the fourth service unit during the first pass) for various values of λ (the average request arrival rate), while the values of the other parameters remain unchanged.

Table 3

b ₁ =0.7, b ₂ =0.3, a ₁ =.6,a ₂ =0.4, μ_1 =3 , μ_2 =2, μ_4 =0.2, μ_3 =1, c ₁ =0.5, c ₂ =0.5, d ₁ =0.8, d ₂ =0.2,q ₁₃ =0.3, q ₁₂ =0.2,q ₁₄ =0.5, q ₁₃ '=0.2,q ₁₂ '=0.1, q ₁₄ '=0.7 q ₂ =0.1 , q ₂₁ =0.4, q ₂₃ =0.3, q ₂₄ =0.2, q ₂ '=0.3, q ₂₁ '=0.3, q ₂₃ '=0.1, q ₂₄ '=0.3, q ₃ =0.6, q ₃₁ =0.2, q ₃₂ =0.15, q ₃₄ =0.05, q ₃ '=0.2, q ₃₂ '=0.3,q ₃₄ '=0.5,q ₄ =0.7,q ₄₂ =0.15, q ₄₃ =0.05, q ₄ '=0.8, q ₄₂ '=0.15, q ₄₃ '=0.05				
q ₄	Lq			
	λ =14	λ =20	λ =30	λ =14.2
0.1	157.1415	199.0751	222.484	161.4348

0.15	157.1024	197.9687	218.3172	161.3289
0.2	157.0662	196.8119	213.9385	161.2225
0.25	157.035	195.603	209.3336	161.1178
0.3	157.0118	194.3409	204.4872	161.0175
0.35	157.0001	193.0247	199.3832	160.9249
0.4	157.0047	191.6543	194.0043	160.8444
0.45	157.0315	190.2304	188.3325	160.7817
0.5	157.0879	188.7552	182.3491	160.7441

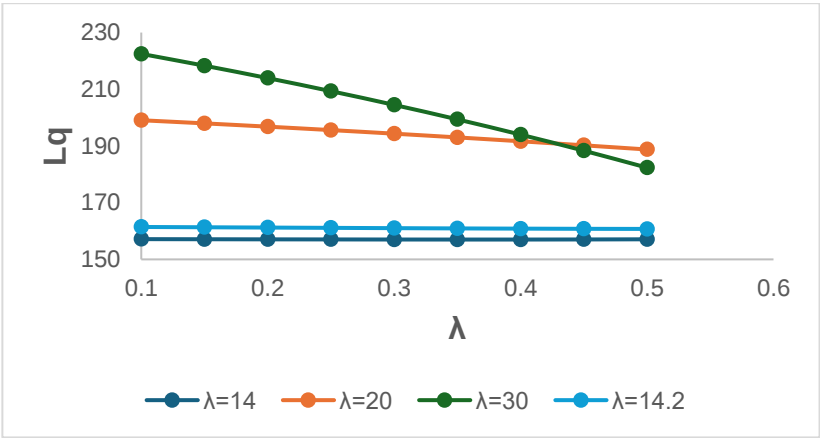


Fig.3

Table 3 and Fig. 3 Interpretation: This table analyzes the probability of a package successfully clearing the shipping dock on its first pass (q_4) against varying Inbound Freight Volumes (λ).

- **Higher Departure Efficiency Lowers Backlog:** As the probability of a package clearing the outbound dock cleanly increases (q_4 from 0.1 to 0.5), the total warehouse queue decreases across almost all volume levels. At a high arrival rate ($\lambda = 30$), raising the exit clearance probability from 0.1 to 0.5 cuts the warehouse backlog down from 222.4 items to 182.3 items.
- **The Inbound Freight Strain:** Higher inbound freight volume (λ) directly scales up the overall queue. Comparing $\lambda = 14$ to $\lambda = 14.2$ shows an immediate, steady increase in backlogged items at every single tier of shipping clearance.
- **The Rework Dip/Anomaly (at $\lambda = 14$):** At lower volumes ($\lambda = 14$), the queue drops until $q_4 = 0.35$ but slightly rises again as q_4 hits 0.5. This indicates an operational sweet spot where clearing first-time packages too quickly can temporarily conflict with the arrival pace of recirculating rework loops entering the same conveyor lines.

Strategic Recommendations for the Warehouse Manager

Stop Optimizing Picking in Isolation: Boosting picking speeds (μ_2) without matching capacity at the loading docks (μ_4) does not ship more orders; it only turns moving items into stagnant conveyor pile-ups.

Invest in Shipping Dock Velocity: Table 2 explicitly shows that the fastest way to lower total warehouse stress metrics L_q is to automate or add labour to the final trailer-loading docks (μ_4).

Prioritize Shipping Accuracy (q_4): Improving label quality, printing alignment, and weight checks to push q_4 from 0.1 toward 0.5 significantly prevents items from entering bad loops, yielding massive queue reductions when the facility is busy ($\lambda = 30$).

References

- Bouchentouf A. A. and Yahiaoui L. On feedback queueing system with reneging and retention of reneged customers, multiple working vacations and Bernoulli schedule vacation interruption. *Arabian Journal of Mathematics*, vol. 6, 2017, pp 1–11. DOI 10.1007/s40065-016-0161-1.
- Gupta, R. and Gupta D. Steady State Analysis of Biserial Server with Bulk Arrival. *International Journal of Mathematics Trends and Technology*, vol. 56 (6), 2018, pp. 430–436.
- Bouchentouf, A. A., and Guendouzi. A. Single Server Batch Arrival Bernoulli Feedback Queueing System with Waiting Server, Variant Vacations and Impatient Customers. *Operations Research Forum*, vol. 2(14), 2021. doi.org/10.1007/s43069-021-00057- 0.
- Kumari S. Waiting Time Analysis of Queueing System Having Combination of Six Servers Centrally Linked with a Common Server. *International Journal of Innovations & Research Analysis*, vol. 2(II), 2022, pp. 233-237.
- Kamal et al. (2023). Mathematical modeling of some feedback queueing systems. Ph. D. Thesis, Baba Mastnath University, Asthal Bohar, Rohtak, Haryana.
- Saini, A., Gupta, D. and Tripathi A. K. Analytical Study of Two Serial Channels with Priority and Reneging. *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 11(6), 2023, pp. 89–93. 10.17762/ijritcc.v11i6.7237.
- Melikov A., Chakravarthy S. R. and Aliyeva S. A retrial queueing model with feedback. *Queueing Models and Service Management*, vol. 6 (1), 2023, pp 63-95.

Lee, S., A. Dudin, and O. Dudina. Analysis of a Priority Queueing System with the Enhanced Fairness of Servers Scheduling. *Journal of Ambient Intelligence and Humanized Computing*, vol. 15, (1), 2024, pp. 465–477. <https://doi.org/10.1007/s12652-022-03903-z>.

Agarwal, A., Agarwal R., and Upadhyaya S. Detection of Optimal Working Vacation Service Rate for Retrial Priority G-Queue with Immediate Bernoulli Feedback. *Results in Control and Optimization*, vol. 14, 2024, p.p 100397. 10.1016/j.rico.2024.100397.