Probabilistic Modeling of Network Congestion and Queue Delay

Course: CSE400 Fundamentals of Probability in Computing
Done by: Krisha Desai(AU2440103) and Rudra Mehta (AU2440268)

February 19, 2026

Scribe Question 1: Project System and Objective

The probabilistic problem addressed in this project is the modeling of network congestion and queueing delay in a distributed communication network. The project examines how randomness in traffic arrivals and service mechanisms leads to congestion and increased end-to-end packet delay.

The objective here is to focus on developing a concrete probabilistic model for analyzing network congestion and queueing delay. Each network node is modeled as an M/M/1 queue where packets arrive randomly, wait for service, get processed, and move to the next node.

The model uses three probability distributions: Poisson distribution (for packet arrivals), Exponential distribution (for processing times), and Bernoulli distribution (for packet drop events). Here, ρ is the traffic intensity i.e. load at a node:

$$\rho = \frac{\lambda}{\mu}$$

where, λ =packet arrival rate and μ = processing rate

The system is stable only when $\rho < 1$. The total end-to-end delay (W_{total}) is the sum of delays across all nodes:

$$W_{\text{total}} = W_1 + W_2 + \cdots + W_n$$

Scribe Question 2: Key Random Variables, Distributions, and Mathematical Modeling

The random variables used in the system and their key parameters for mathematical modeling are:

Packet arrival count (X): λ = arrival rate (packets/sec)
Time between arrivals (T_a): λ = arrival rate
Packet processing time (T_s): μ = service rate
Queue length (N): $\rho = \frac{\lambda}{\mu}$
Total time at a node (W): λ and μ
Packet drop indicator (D): p = drop probability

To count the number of packets that arrive at a node in a given time$t$, Poisson distribution is used as follows:

$$P(X = k) = \frac{e^{-\lambda t} \cdot (\lambda t)^k}{k!}, \qquad k = 0, 1, 2, 3, \ldots$$

On average, λt packets arrive in time t. Average is used because it can not be said that exactly how many packets would arrive at a time.

Packet processing time and inter-arrival time follow an Exponential distribution:

$$f(x) = \lambda\, e^{-\lambda x}, \qquad x \geq 0$$

The average processing time is $\frac{1}{\mu}$ seconds. Exponential distribution is used here, due to its memoryless property. The time required to finish processing a packet is independent of how long it has already been processed.

Using the above two distributions together in the M/M/1 model, we can figure out how many packets are present at a node at a given time. As we follow Poisson distribution and Exponential distribution, the queue length follows Geometric distribution as follows:

$$P(N = n) = (1 - \rho)\, \rho^n, \qquad n = 0, 1, 2, \ldots$$

Quantities obtained from this are:

- Average queue length: $E[N] = \dfrac{\rho}{1 - \rho}$
- Average waiting time: $E[W_q] = \dfrac{\rho}{\mu(1 - \rho)}$
- Average total time at a node: $E[W] = \dfrac{1}{\mu - \lambda}$

All three quantities grow infinitely without bound as ρ approaches 1, representing mathematical congestion.

For the packet drops, each packet is either dropped or successfully transmitted, so Bernoulli distribution is used as follows:

$$P(D=1)=p, \qquad P(D=0)=1-p, \qquad p\in[0,1]$$

On average, a fraction p of packets are dropped. As ρ increases, p also rises (the busier the node, the more packets it has to reject). We will generate Probability Density Function (PDF) and Cumulative Distribution Function (CDF) plots for all the distributions to better understand how the system behaves under various traffic loads.

Scribe Question 3: Probabilistic Reasoning and Dependencies

Not all the random variables in our model are independent. The queue length and the delay both depend on $\rho=\frac{\lambda}{\mu}$. As λ increases, both $E[N]$ and $E[W]$ grow very quickly. This is the reason why a network that is already busy sometimes suddenly stops working when a little more load is added.

$$P(\text{Arriving packet has to wait})=\rho$$

This result follows from the PASTA property (Poisson Arrivals See Time Averages). This allows steady-state probabilities to be applied directly to arriving packets. The drop probability $P(D=1)$ also increases with ρ, so drops are concentrated in high-traffic periods.

Following are the independence assumptions in our model:

→ Packet arrivals are independent of each other. One packet arriving does not affect when the next packet arrives (justified using Poisson model).

→ Service times are independent of each other and also independent of the arrival process (standard assumption in the M/M/1 model).

→ The delay at one node is treated as independent of the delay at any other node. This helps in simplifying the calculation as we can simply add the per node delays to get the total end to end delay. $W_{\text{total}}=W_1+W_2+\cdots+W_n$.

Scribe Question 4: Model–Implementation Alignment

The mathematical model associates directly with the network system described in the scribe for Milestone 1. Each node is modeled as an M/M/1 queue with Poisson arrivals and exponential service times. The queue length distribution

$$P(N = n) = (1 - \rho)\rho^n$$

and the end-to-end delay formula

$$W_{\text{total}} = \sum_i W_i$$

can be compared against simulation outputs to verify the model. Packet drops are modeled as Bernoulli trials with probability p that increases with ρ.

The performance measures from the previous scribe — waiting time, queue length, and end-to-end delay — are now directly linked to $E[W_q]$, $E[N]$, and $E[W]$. The current assumptions of Poisson arrivals, infinite buffer, exponential service times, and $\rho < 1$ keep the model tractable. Relaxing these assumptions is planned for upcoming phases.

Scribe Question 5: Cross-Milestone Consistency and Change

We have tried to keep the model consistent with the previous scribe. The random variables identified earlier are now formally specified through their distributions, and the qualitative dependencies described before are now expressed through the formulas

$$E[N] = \frac{\rho}{1 - \rho}, \quad E[W] = \frac{1}{\mu - \lambda}.$$

The traffic intensity $\rho = \lambda/\mu$ remains the unifying parameter across the entire model.

In upcoming phases, the model will be extended in the following ways:

- **Finite buffer (M/M/1/K):** The drop probability will be derived as

  $$P(\text{drop}) = \frac{(1 - \rho)\rho^K}{1 - \rho^{K+1}}$$

  rather than being treated as an external parameter.

- **Bursty traffic:** The Poisson arrival assumption will be replaced with distributions that capture more realistic traffic patterns.

- **Numerical validation:** Simulation results will be used to validate theoretical formulas across different values of ρ.

Scribe Question 6: Open Issues and Responsibility Attribution

Several open issues from the Milestone 1 scribe remain under investigation. The effect of bursty or correlated arrivals on queue stability remains unaddressed, as the current Poisson model does not capture burstiness. The drop probability p is still treated as an external parameter rather than being derived from the buffer size K. The assumption that per-node delays are independent needs to be validated through simulation. In addition, realistic values of λ and μ for actual network devices have not yet been determined.

These issues will be addressed in the upcoming phase through model refinement, simulation, and numerical analysis, as outlined in the previous scribe.