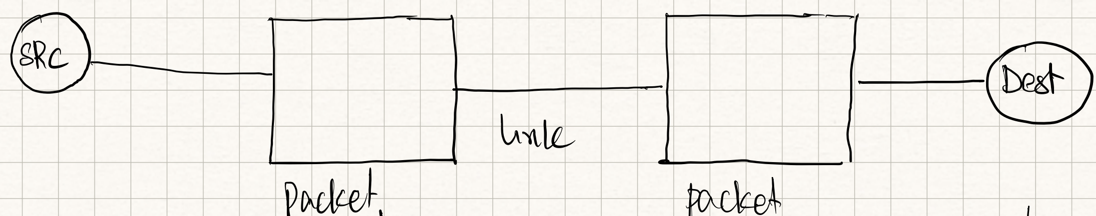
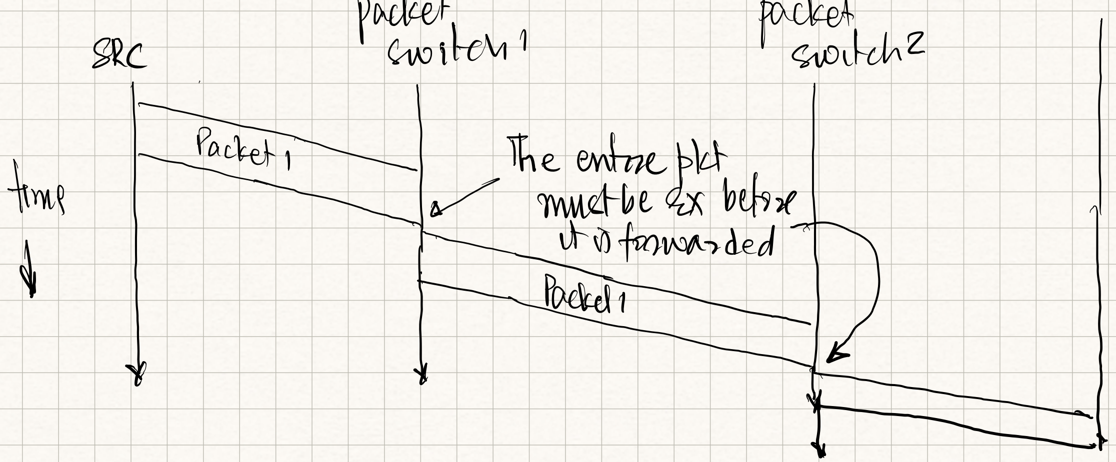
Lecture 3

→ Review of packet switching
→ Queuing delay
→ Simulation

- Packet switch networks are also called store-and-forward networks

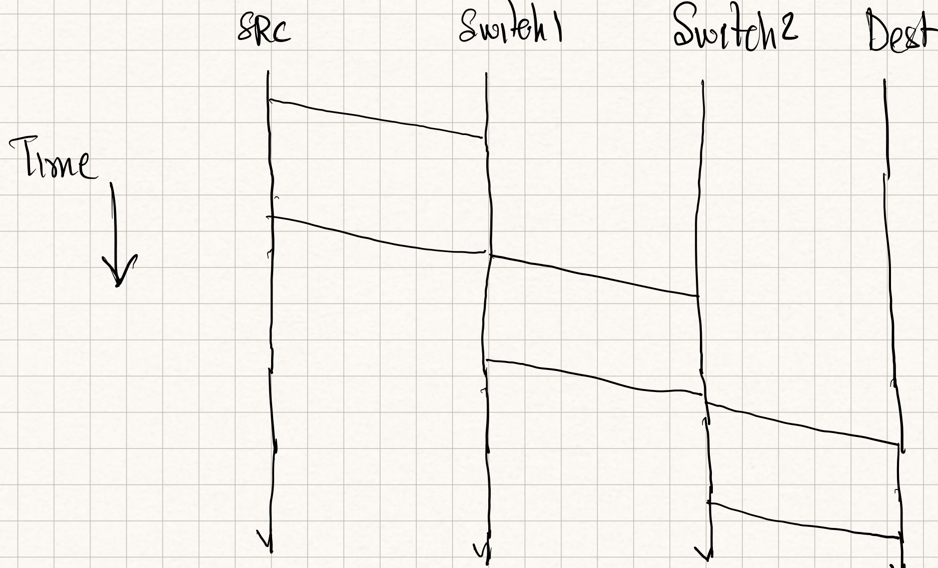
  → At a switch the entire pkt must be received before it is forwarded to the next switch



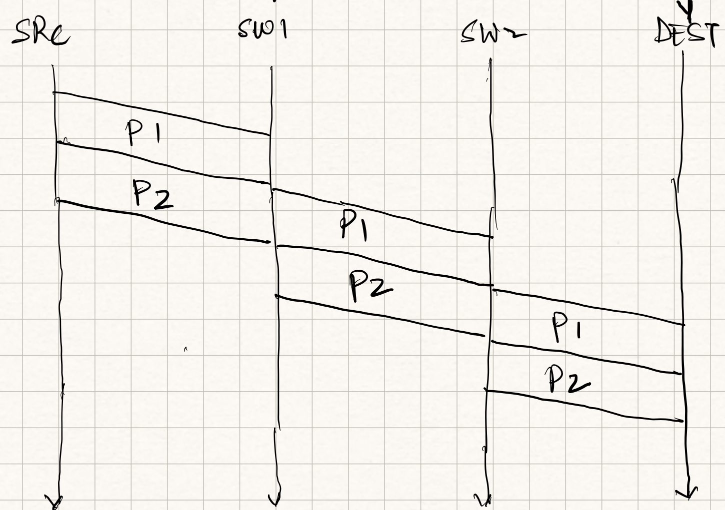


- Breaking the message into smaller pkts is beneficial in multihop networks as sending and receiving in a switch can be pipelined

Consider the same networks as above



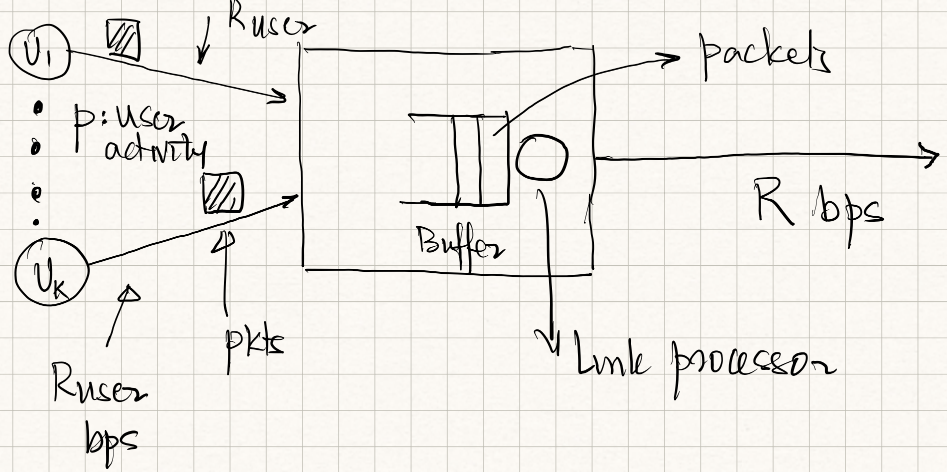
A large message is put into 1 pkt



message divided into 2 pkts

- Since the pkt must have a header, it is counter productive to break the message into too small packets

Queuing Delay

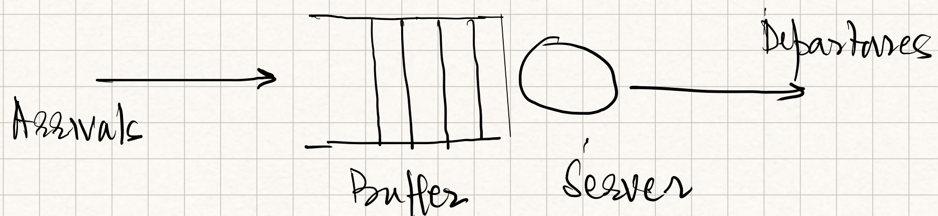


- We can connect K users where $K > \lceil R/R_{user} \rceil$ to the switch
- However, there is some probability of congestion (P_c) which is a function of K and p
- Congestion means that the input rate to the switch is greater than the output rate R bps

  ⇒ The buffer can hold pkts during periods of congestion

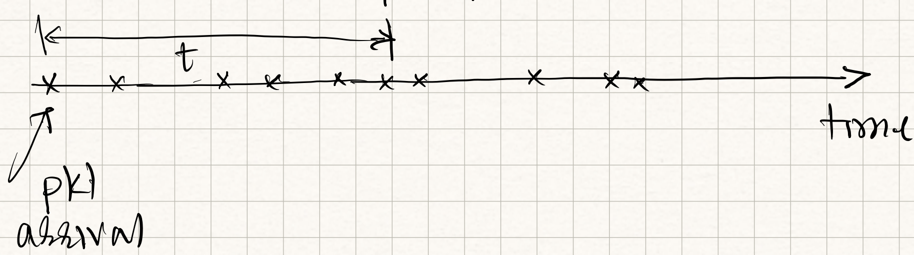
  ⇒ Pkts will experience queuing delay.

A simple queuing system



Assumptions

1. Arrival process follows a Poisson Process with rate λ pkts/sec



If we take an interval of length t

$N(t)$ is the number of pkt arrivals in time t

$N(t)$ is a random variable

$$P(N(t)=k) = \frac{e^{-\lambda t}(\lambda t)^k}{k!} \qquad \lambda > 0, \quad k \in \{0,1,2,\ldots,\infty\}$$

Notes:

- If you take many different intervals of length t, then the average number of arrivals over those different intervals is λt

- λ is the instantaneous arrival rate
  pkts/sec

- The time between pkts is also a random variable. This is the interarrival time. denoted by X

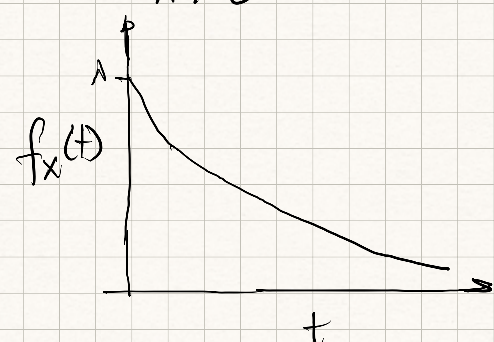
  If the arrivals follow a Poisson process then $X \sim \text{Expo}(\lambda)$

  ↳ negative exponentially distribution

  X is a continuous distribution

  $f_X(t) = \lambda e^{-\lambda t}$ $\qquad t > 0$, $\lambda > 0$

  ↑
  probability density function



  $F_X(t)$: Cummulative Distribution Function

  $$= \int_0^t f_X(y)\,dy = 1 - e^{-\lambda t} \qquad t > 0$$

2. We will assume that the buffer size is infinite. Later we will consider a finite buffer case

3. We will assume that the server is work conserving and services packets in First Come First Served (FCFS) order

Work conserving $\Rightarrow$ the server is not idle when there is work (pkts) in the system

4. We will assume that service time of pkts is a random variable Y and further assume that $Y \sim Expo(\mu)$ where $1/\mu$ sec is the average service time of a pkt.

Notes:

a) Service time is the transmission time of a pkt.

b) Essentially we are saying that pkt size is not fixed but follows an Exponential distribution

Let N denote the number of pkts in the system (buffer + server)

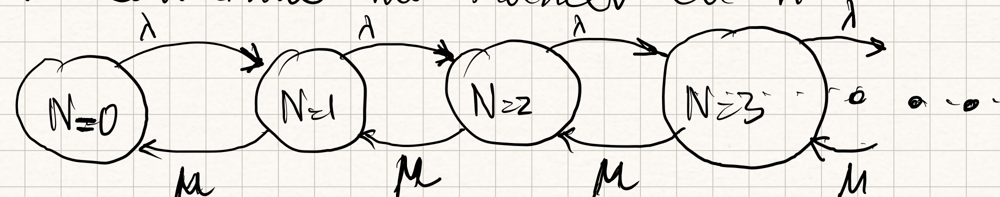
$N = 0 \rightarrow$ the system is empty

$N = 1 \rightarrow$ 1 pkt in the system and it must be in the process of being transmitted

$N = 2 \rightarrow$ 2 pkts in the system
1 being transmitted
1 waiting in the buffer

and so on ...

We can draw the Markov Ch n



Notes:

Only nearest neighbor transitions because of Poisson process assumption for the arrival process and a single server with exponential service time distribution

Let $p_i = P(N = i)$ [the probability that system has i pkts]

We want to find the steady state or the long run state probabilities

In steady state, the rate into a state must be equal to the rate out of the state

State 0 $\quad p_0\lambda = p_1\mu \quad \Rightarrow \quad p_1 = (\lambda/\mu)p_0$

state 1 $\quad p_1\mu + p_1\lambda = p_0\lambda + p_2\mu \Rightarrow p_2 = (\lambda/\mu)^2 p_0$

In general

$$p_i = (\lambda/\mu)^i p_0 \qquad i = 0, 1, 2, \ldots$$

Since p_i's are probilities

$$\sum_{i=0}^{\infty} p_i = 1 \quad \Rightarrow \quad \sum_{i=0}^{\infty} (\lambda/\mu)^i p_0 = 1$$

$$\Rightarrow \quad p_0 = 1 - \lambda/\mu$$

$$= 1 - \rho$$

where we define $\rho = \lambda/\mu$

Hence,

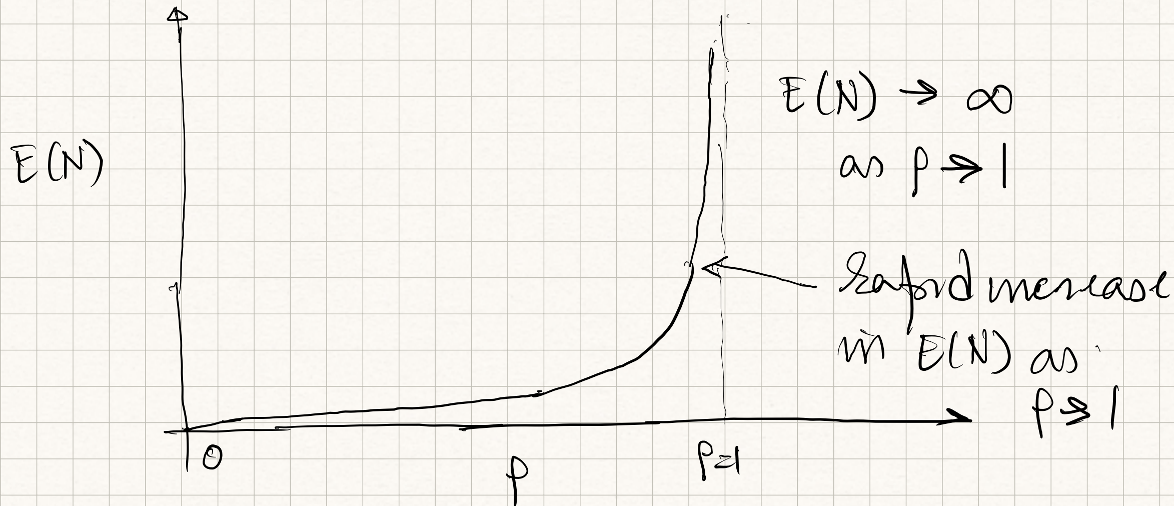
$$p_i = (1-\rho)\rho^i$$

We can find the average number of pkts in the system

$$E(N) = \sum_{i=0}^{\infty} i\, p_i = \sum_{i=0}^{\infty} i(1-\rho)\rho^i$$

$$\boxed{E(N) = \rho/1-\rho}$$

again $\rho = \lambda/\mu$



- ρ is also called the traffic intensity

(avg) arrival rate normalized by the (avg) service rate

$\rho \to 1$ implies $\lambda \to \mu$

- The delay D is related to N

$$E(D) = \frac{1/\mu}{1-\rho}$$

$E(D) \to \infty$ as $\rho \to 1$

i.e., $\lambda \to \mu$