16.36 Communication Systems Engineering
Spring 2009

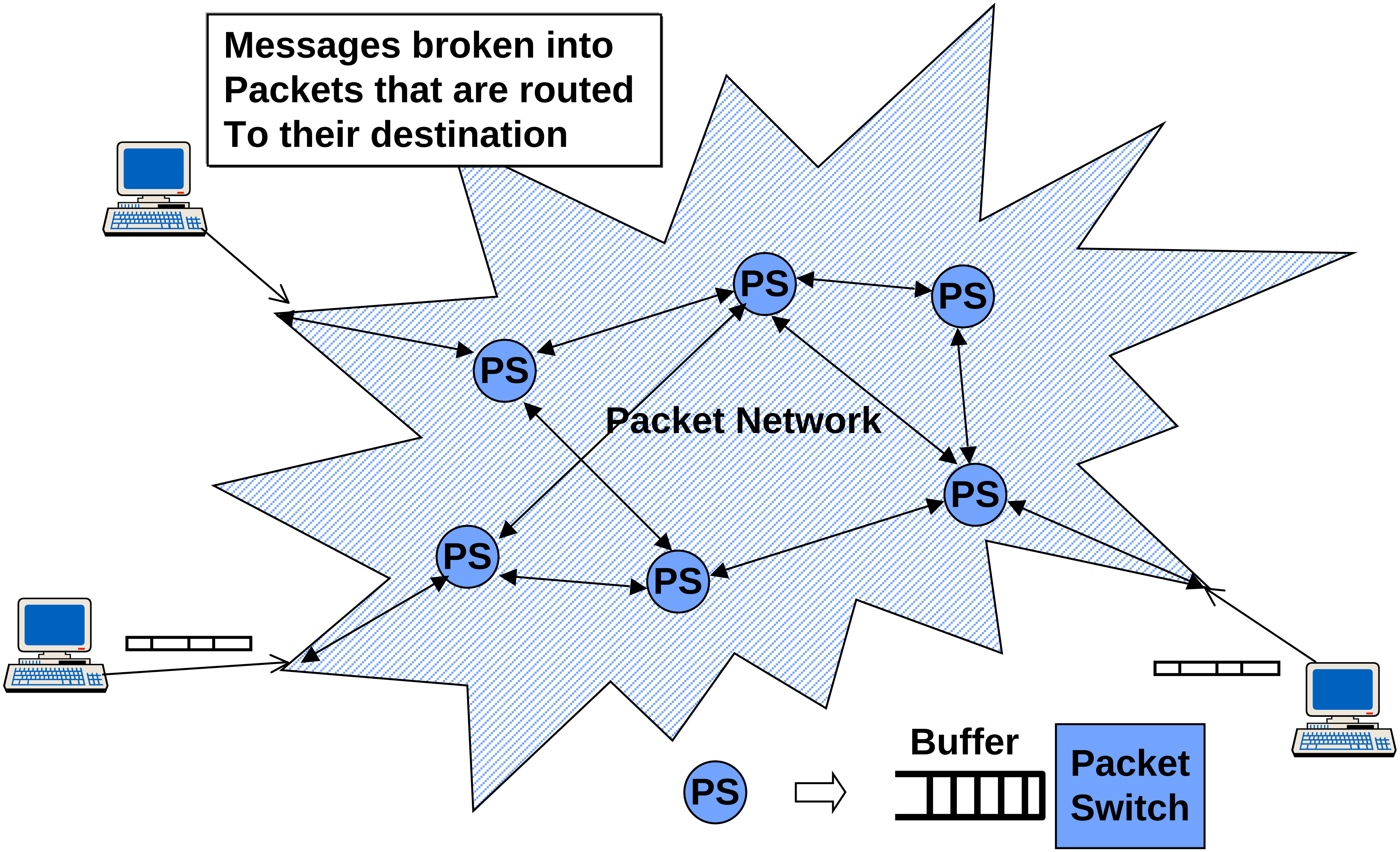
16.36: Communication Systems Engineering

Lecture 19: Delay Models for Data Networks

Part 1: Introduction

Eytan Modiano

Packet Switched Networks



Queueing Systems

- Used for analyzing network performance

- In packet networks, events are random
  - Random packet arrivals
  - Random packet lengths

- While at the physical layer we were concerned with bit-error-rate, at the network layer we care about delays
  - How long does a packet spend waiting in buffers ?
  - How large are the buffers ?

- Applications far beyond just communication networks
  - Air transportation systems, air traffic control
  - Manufacturing systems
  - Service centers, phone banks, etc.

Random events

- **Arrival process**
  - **Packets arrive according to a random process**
  - **Typically the arrival process is modeled as Poisson**

- **The Poisson process**
  - **Arrival rate of λ packets per second**
  - **Over a small interval δ,**

    P(exactly one arrival) = $\lambda\delta$
    P(0 arrivals) = 1 - $\lambda\delta$
    P(more than one arrival) = 0

  - **It can be shown that:**

$$\text{P(n arrivals in interval T)} = \frac{(\lambda T)^n e^{-\lambda T}}{n!}$$

The Poisson Process

$$\text{P(n arrivals in interval T)} = \frac{(\lambda T)^n e^{-\lambda T}}{n!}$$

n = number of arrivals in T

It can be shown that,

$E[n] = \lambda T$

$E[n^2] = \lambda T + (\lambda T)^2$

$\sigma^2 = E[(n-E[n])^2] = E[n^2] - E[n]^2 = \lambda T$

Inter-arrival times

- **Time that elapses between arrivals (IA)**

$$P(IA \leq t) = 1 - P(IA > t)$$

$$= 1 - P(0 \text{ arrivals in time } t)$$

$$= 1 - e^{-\lambda t}$$

- **This is known as the Exponential distribution**
  - **Inter-arrival CDF =** $F_{IA}(t) = 1 - e^{-\lambda t}$
  - **Inter-arrival PDF =** $d/dt\, F_{IA}(t) = \lambda e^{-\lambda t}$
- **The Exponential distribution is often used to model the service times (I.e., the packet length distribution)**

Markov property (Memoryless)

$$P(T \le t_0 + t \mid T > t_0) = P(T \le t)$$

$\Pr oof:$

$$P(T \le t_0 + t \mid T > t_0) = \frac{P(t_0 < T \le t_0 + t)}{P(T > t_0)}$$

$$= \frac{\int_{t_0}^{t_0+t} \lambda e^{-\lambda t} dt}{\int_{t_0}^{\infty} \lambda e^{-\lambda t} dt} = \frac{-e^{-\lambda t} \Big|_{t_0}^{t_0+t}}{-e^{-\lambda t} \Big|_{t_0}^{\infty}} = \frac{-e^{-\lambda(t+t_0)} + e^{-\lambda(t_0)}}{e^{-\lambda(t_0)}}$$

$$= 1 - e^{-\lambda t} = P(T \le t)$$

- **Previous history does not help in predicting the future!**
- **Distribution of the time until the next arrival is independent of when the last arrival occurred!**

Example

- Suppose a train arrives at a station according to a Poisson process with average inter-arrival time of 20 minutes

- When a customer arrives at the station the average amount of time until the next arrival is 20 minutes
  – Regardless of when the previous train arrived

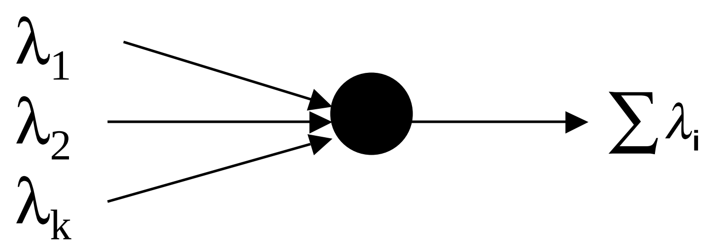
- The average amount of time since the last departure is 20 minutes!

- Paradox: If an average of 20 minutes passed since the last train arrived and an average of 20 minutes until the next train, then an average of 40 minutes will elapse between trains
  – But we assumed an average inter-arrival time of 20 minutes!
  – What happened?

- Answer: You tend to arrive during long inter-arrival times
  – If you don't believe me you have not taken the T

Properties of the Poisson process

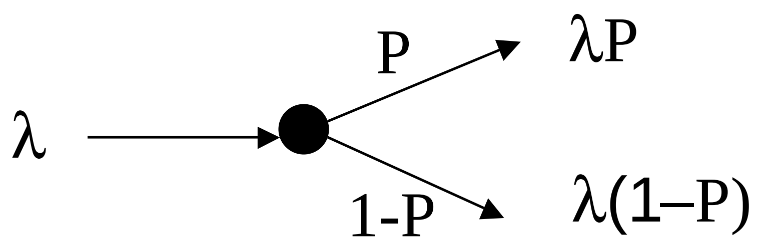
- **Merging Property**



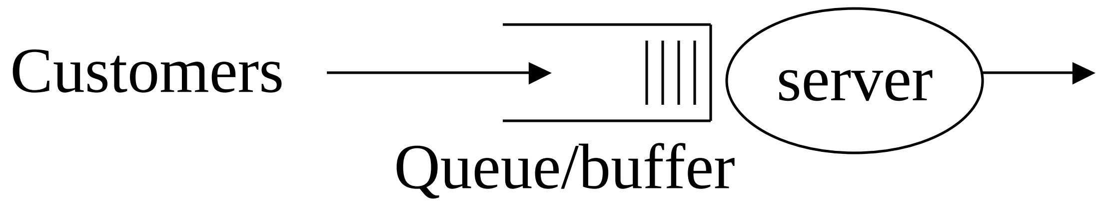
Let A1, A2, … Ak be independent Poisson Processes of rate λ1, λ2, …λk

$$A = \sum A_i \text{ is also Poisson of rate} = \sum \lambda_i$$

- **Splitting property**
  - **Suppose that every arrival is randomly routed with probability P to stream 1 and (1-P) to stream 2**
  - **Streams 1 and 2 are Poisson of rates Pλ and (1-P)λ respectively**



Queueing Models



- **Model for**
  - **Customers waiting in line**
  - **Assembly line**
  - **Packets in a network (transmission line)**

- **Want to know**
  - **Average number of customers in the system**
  - **Average delay experienced by a customer**

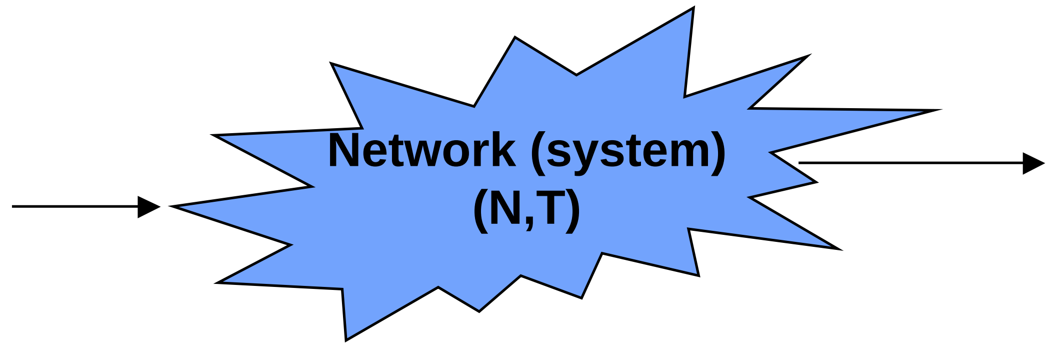
- **Quantities obtained in terms of**
  - **Arrival rate of customers (average number of customers per unit time)**
  - **Service rate (average number of customers that the server can serve per unit time)**

Analyzing delay in networks (queueing theory)

- **Little's theorem**
  - **Relates delay to number of users in the system**
  - **Can be applied to any system**
- **Simple queueing systems (single server)**
  - **M/M/1, M/G/1, M/D/1**
  - **M/M/m/m**
- **Poisson Arrivals ⇒** $P(\text{n arrivals in interval T}) = \frac{(\lambda \text{T})^{\text{n}} e^{-\lambda \text{T}}}{\text{n}!}$
  - **λ = arrival rate in packets/second**
- **Exponential service time ⇒** $P(\text{service time} < \text{T}) = 1 - \text{e}^{-\mu \text{T}}$
  - **μ = service rate in packets/second**

Little's theorem

λ packet per second

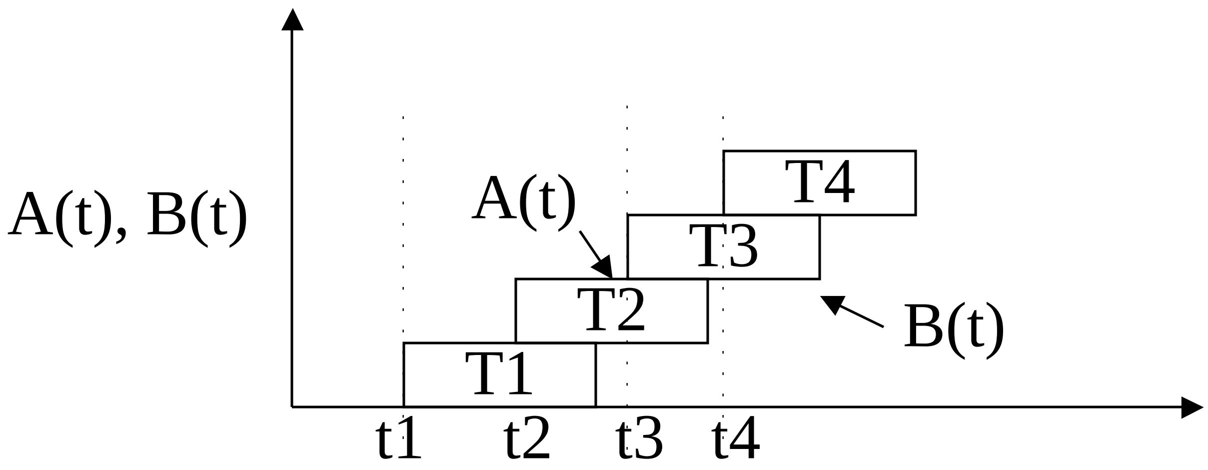


- N = average number of packets in system
- T = average amount of time a packet spends in the system
- λ = arrival rate of packets into the system
  (not necessarily Poisson)

- Little's theorem: $N = \lambda T$
  - Can be applied to entire system or any part of it
  - Crowded system $\leftrightarrow$ long delays

    On a rainy day people drive slowly and roads are more congested!

Proof of Little's Theorem



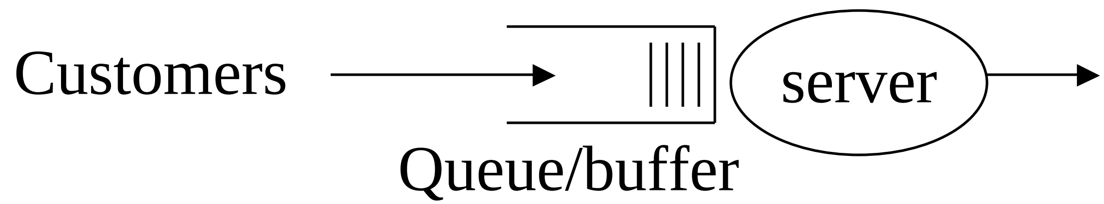
- **A(t) = number of arrivals by time t**
- **B(t) = number of departures by time t**
- **t_i = arrival time of i[th] customer**
- **T_i = amount of time i[th] customer spends in the system**
- **N(t) = number of customers in system at time t = A(t) - B(t)**

$$N = \lim_{t \to \infty} \frac{\sum_{i=1}^{A(t)} T_i}{t}, \quad T = \lim_{t \to \infty} \frac{\sum_{i=1}^{A(t)} T_i}{A(t)} \Rightarrow \sum_{i=1}^{A(t)} T_i = A(t)T$$

$$N = \frac{\sum_{i=1}^{A(t)} T_i}{t} = \left(\frac{A(t)}{t}\right) \frac{\sum_{i=1}^{A(t)} T_i}{A(t)} = \lambda T$$

Application of Little's Theorem

- **Little's Theorem can be applied to almost any system or part of it**
- **Example:**



1) The transmitter: D_{TP} = packet transmission time

- **Average number of packets at transmitter = $\lambda D_{TP} = \rho$ = link utilization**

2) The transmission line: D_p = propagation delay

- **Average number of packets in flight = λD_p**

3) The buffer: D_q = average queueing delay

- **Average number of packets in buffer = $N_q = \lambda D_q$**

4) Transmitter + buffer

- **Average number of packets = $\rho + N_q$**