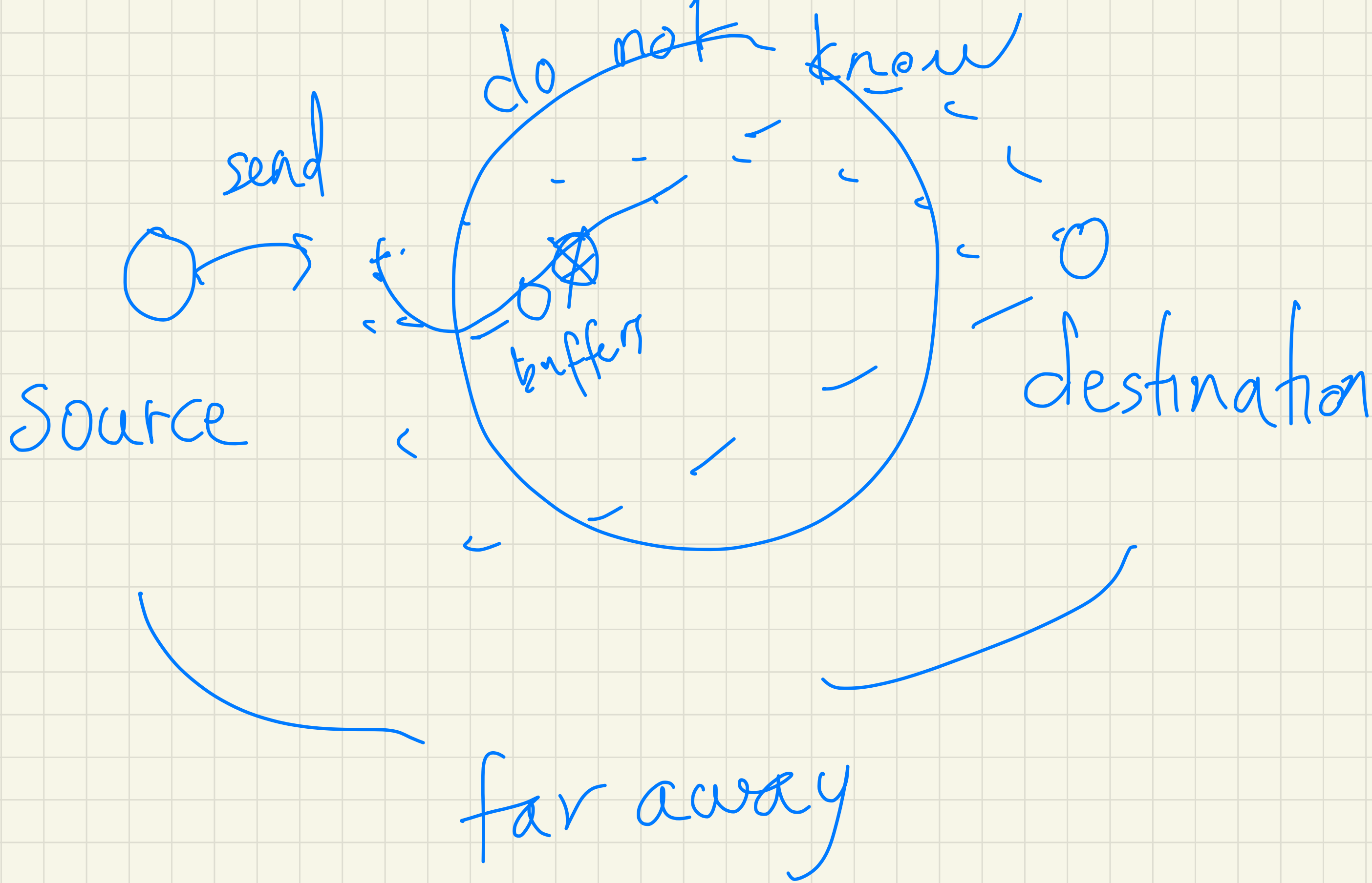
Little's Law

Little is somebody's
name

Little's is invisible in the OSI hierarchy above DLL
We have created the packets and now we must deliver them.
It provides "Control" mechanism

Network Performance

- Development of efficient network algorithms is influenced by transmission delays of packets from source to destination.
  - Which network protocol gives the best delay-throughput characteristics under specified conditions?
  - What size buffers must be employed by a network's users in order to keep the probability of buffer overflow below a particular value?
  - What is the maximum number of voice calls that can be accepted by a network in order to keep the voice packet transfer delay to a minimum?
  - How many users can a satellite link support and still maintain a reasonable response time?

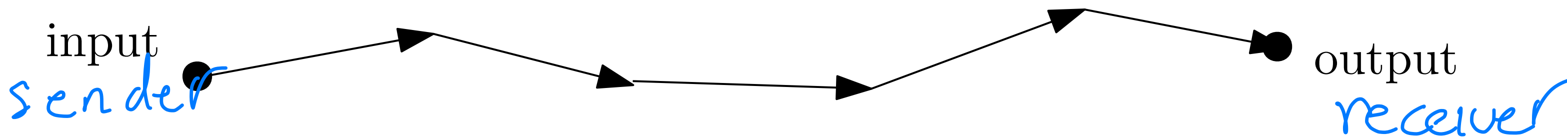

do not know
send
source
buffer
destination
far away

Types of Delay

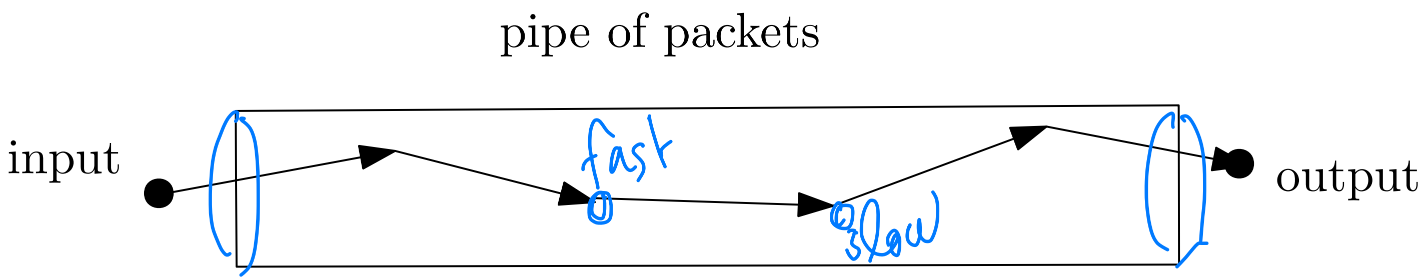
- They are all measured in time units.
  - **Processing:** Delay between time packet is correctly received and the time it is correctly assigned to an outgoing link.
  - **Queuing:** Delay between time packet is assigned to a queue for transmission and the time it starts being transmitted.
  - **Transmission:** Delay between time that the first and the last bits of the packet are transmitted.
  - **Propagation:** Delay between the time that the last bit is transmitted and the bit is received.
- Is there a general principle underlying the various types of delay?

Packet Pipes

- The transmission of packets from source (input) to destination (output) . . .



. . . resembles a pipe of packets



in that you can observe only the input and the output.

- You don't know precisely what is going on inside the pipe!
- Can observations of the input and output teach us something about the performance of the system?

Example

Restaurant Paradigm

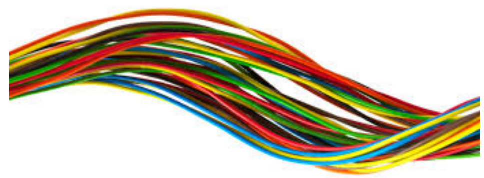
- You are a spy from Burger King trying to figure out how many people are inside MacDonald's.

Buffer Size

- You cannot sit inside MacDonald's all day;
- You must derive the answer based only on observing traffic.

Example

- It's like having a wire with packets entering from the left and exiting from the right.



- You can count how many packets enter the wire in a given time interval: of course the count will be on the average!
- You can count how long a packet stays in the wire before exiting: of course for many customers the measurement will be on the average!
- But you cannot see inside the wire!
- This is observed in many network traffic applications.

Example

- Back to the restaurant example:
- You observe that on the average 40 customers per hour go into the restaurant.
- You observe that on the average a customer stays 15 minutes.
- Any given time there are, on the average, 10 customers inside the restaurant, because

$$40 \text{ customers per hr} \times 1/4 \text{ of an hour} = 10$$

- This sounds like a fundamental principle in networing!.

Modeling Delay

- In typical queueing systems customers (i.e. packets) arrive at random times to obtain service.
- If L = packet length in bits, C = link transmission capacity in bits/sec Then:

$$\text{service time } = \frac{L}{C}$$

- We ignore the distinction between frame and packet.
- We will be interested in estimating:
  - Average # of customers in the system (either waiting in queue or undergoing service);
  - Average delay per customer.
- These will be estimated in terms of customer arrival and service rates.

Little's Theorem

- Little's theorem concerns time averages in the limit.
- Suppose we observe a sample history of a system from the starting time $t = 0$,

MacDonalds

$$20 \qquad \qquad \qquad \qquad 10$$

$$\boxed{\alpha(t)} \Rightarrow \overbrace{- - - - - - - - - - - - -}^{N(t)} \Rightarrow \boxed{\beta(t)}$$

$$10 \qquad \qquad \qquad \qquad 20$$

- Consider the quantities:

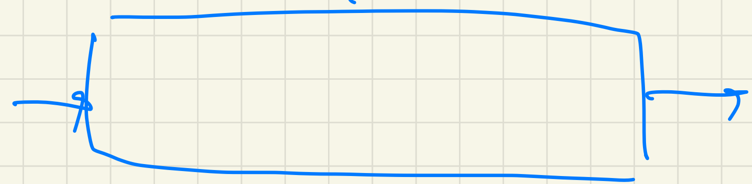
$$N(t) = \#\text{ of customers in the system at time } t;$$
$$\alpha(t) = \#\text{ of customers who arrived in the interval } [0, t];$$
$$\beta(t) = \#\text{ of customers who departed in the interval } [0, t];$$
$$T(i) = \text{time spent in the system by } i\text{-th customer.}$$

$$\frac{\alpha(t)}{t} \qquad \frac{\beta(t)}{t}$$

- How are these quantities related?

On the average there is equilibrium

$\alpha(t)$ arrivals



$\beta(t)$ depart

$$\begin{array}{r} 11 \\ 9 \\ 20 \\ 6 \\ \hline 46 \end{array} \qquad \frac{46}{4}$$

$$\begin{array}{r} 12 \\ 10 \\ 15 \\ 10 \\ \hline 47 \end{array} \qquad \frac{47}{4}$$

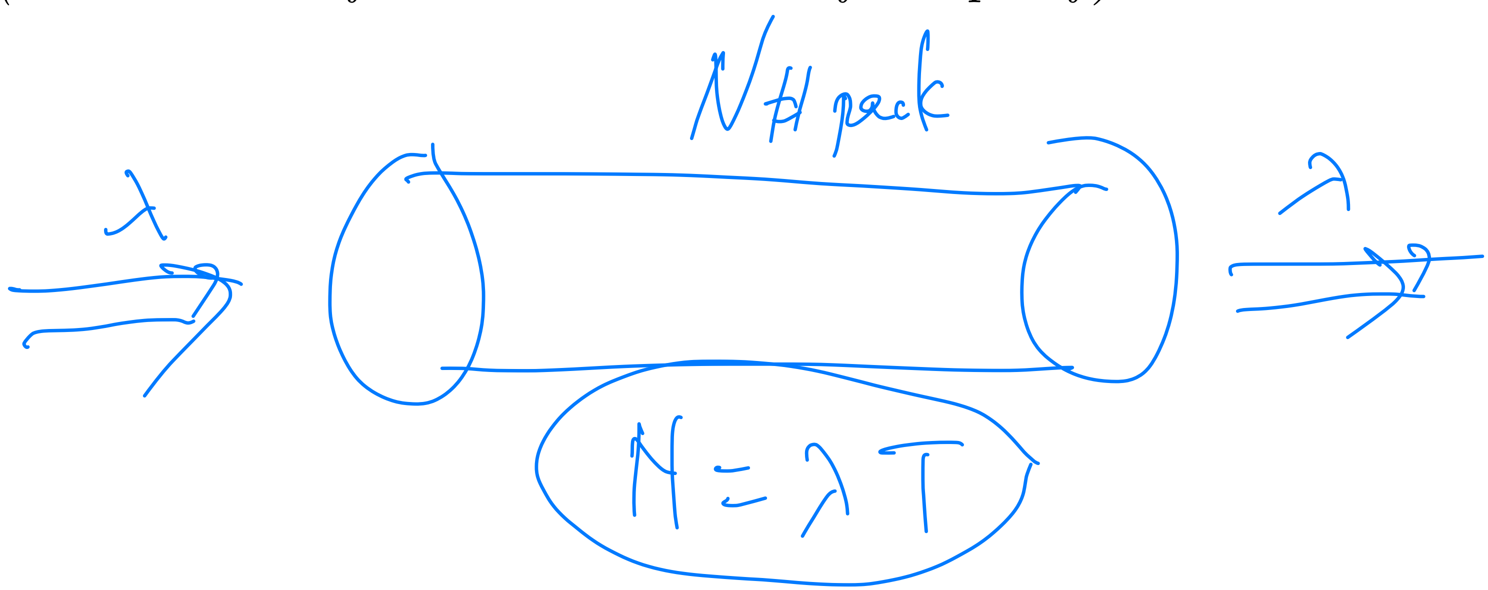
Little's Theorem

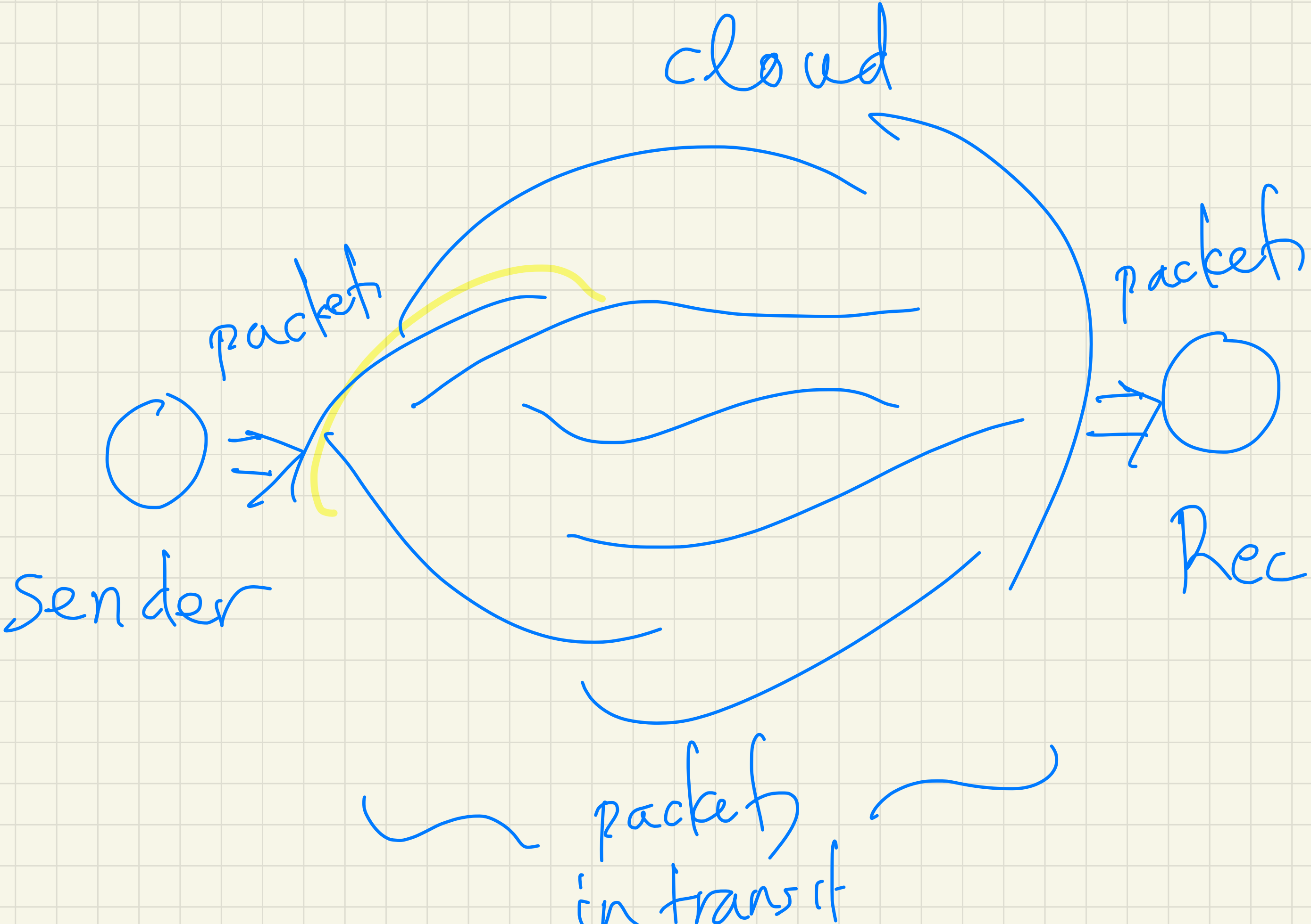
- **Theorem 1 (Little's theorem)** *Assuming a system with steady state behavior, i.e., the rate of arrival and departure are the same (in the limit), we have that*

$$N = \lambda T.$$

they are all averages

where N, T, λ are the averages of the quantities defined before (and will be defined in the course of the proof).




cloud
packet
packet
Sender
Rec.
packet in transit

Proof of Little's Theorem

- We will be interested in finding a relation between these parameters. We define the time average
  - arrival rate over interval $[0, t]$: $\lambda_t = \alpha(t)/t$
  - of the customer delay up to time t:

$$T_t = \frac{1}{\alpha(t)} \sum_{i=1}^{\alpha(t)} T(i)$$

  - of the number of customers up to time t:

$$N_t = \underbrace{\frac{1}{t} \sum_{i=1}^{t} N(i)}_{} \approx \frac{1}{t} \int_0^t N(i)di$$

- In many systems of interest, these quantities tend to a steady state:

$$\Longrightarrow \lambda := \lim_{t\to\infty} \lambda_t, T := \lim_{t\to\infty} T_t, N := \lim_{t\to\infty} N_t$$

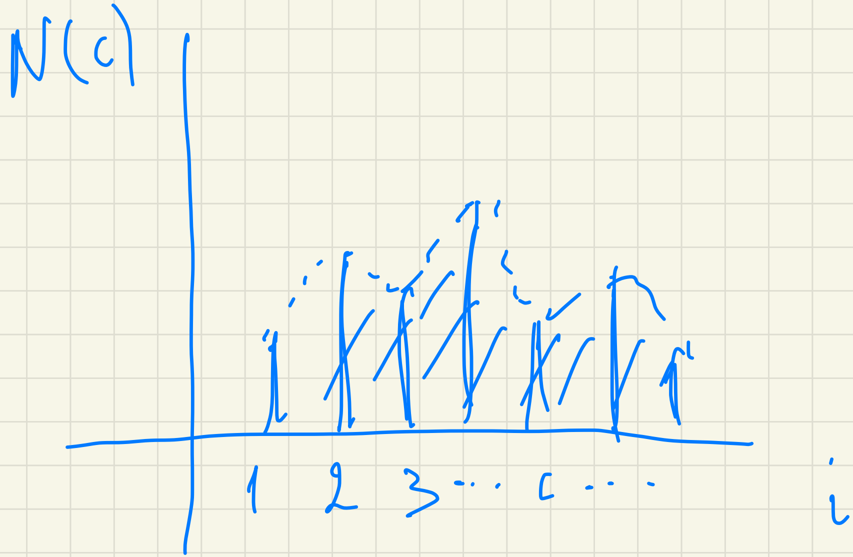
$\alpha(t)$

$\frac{\alpha(t)}{t}$

$[0, t]$

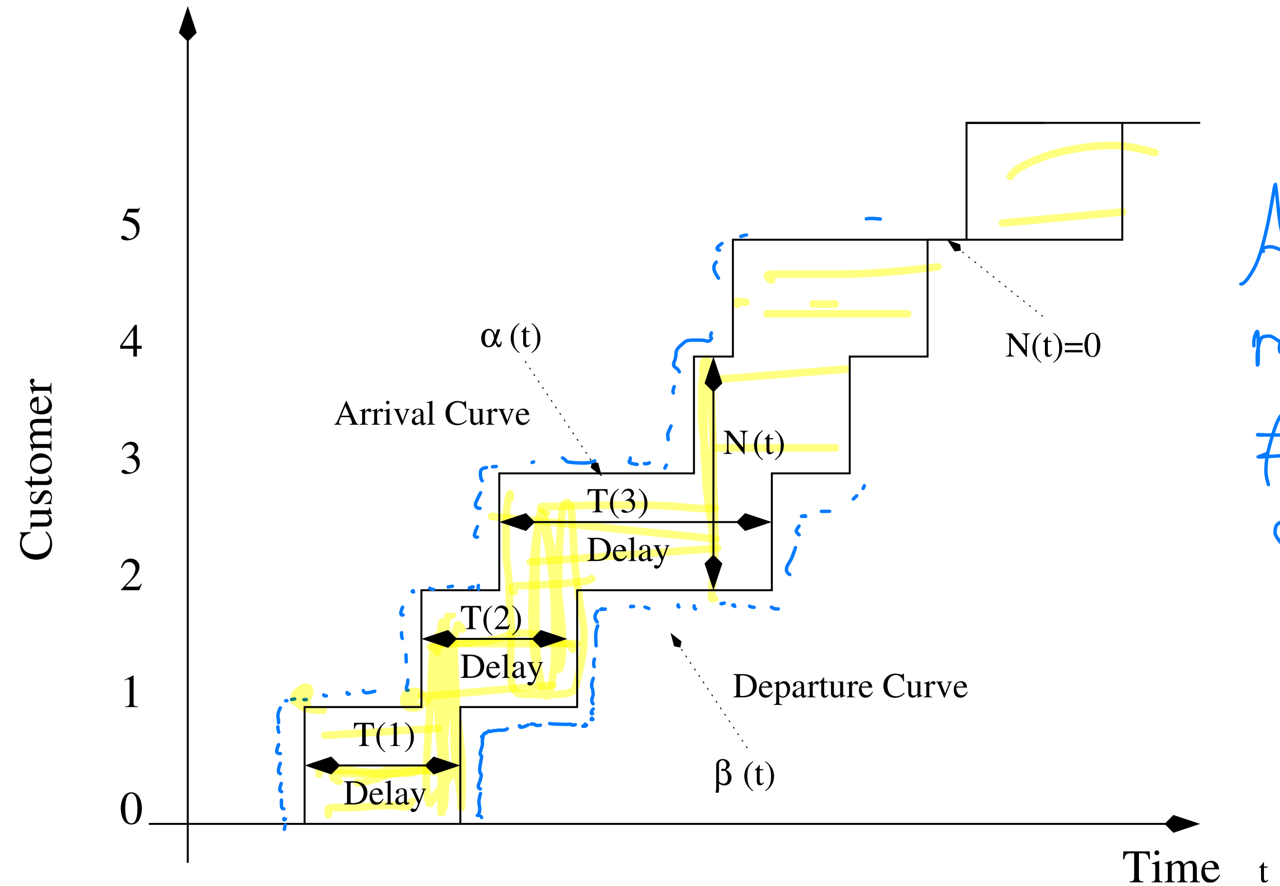
$T(i)$ = time customer i stays in system

$N(i)$



N(c)
1
2
3 ...
c ...
i

Proof of Little's Theorem



Area represent # of customer

Illustrated are the functions $\beta(t), \alpha(t)$ and $\beta(t) \leq \alpha(t)$, $T(i)$, is the delay of customer i.

Proof of Little's Theorem

- $\alpha(t), \beta(t)$ is the number of arrivals and departures up to time t.
- Their difference $\alpha(t) - \beta(t)$ is the number $N(t)$ in the system at time t.

If $\alpha(t) > \beta(t)$ then what can you conclude?

- The area between the arrival and departure curves $\alpha(t), \beta(t)$ is equal to

$$\int_0^t N(\tau) d\tau$$

- If $N(t) = 0$ then the area between the arrival and departure curves is also equal to

$$\sum_{i=1}^{\alpha(t)} T(i).$$

Proof of Little's Theorem

- From the picture $\sum_{i=1}^{\beta(t)} T(i) \leq \underbrace{\int_0^t N(\tau)d\tau \leq \sum_{i=1}^{\alpha(t)} T(i)}$. Therefore

$$
\begin{aligned}
\lambda_t T_t &= \frac{\cancel{\alpha(t)}}{t} \frac{1}{\cancel{\alpha(t)}} \sum_{i=1}^{\alpha(t)} T(i) \\
&= \frac{1}{t} \sum_{i=1}^{\alpha(t)} T(i) \\
&\geq \frac{1}{t} \int_0^t N(\tau)d\tau \\
&\geq \frac{1}{t} \sum_{i=1}^{\beta(t)} T(i) \\
&= \frac{\beta(t)}{t} \frac{1}{\beta(t)} \sum_{i=1}^{\beta(t)} T(i)
\end{aligned}
$$

Proof of Little's Theorem

- Hence:

dep. rate

$$\frac{\beta(t)}{t} \cdot \frac{\sum_{i=1}^{\beta(t)} T(i)}{\beta(t)} \leq N_t \leq \frac{\alpha(t)}{t} \cdot \frac{\sum_{i=1}^{\alpha(t)} T(i)}{\alpha(t)}$$

arrival rate

- Taking the limit we have that

$$\lambda T \leq N \leq \lambda T$$

- Which proves, $N = \lambda T$, i.e., Little's Theorem.

Remarks

- Note in the proof we used the fact that

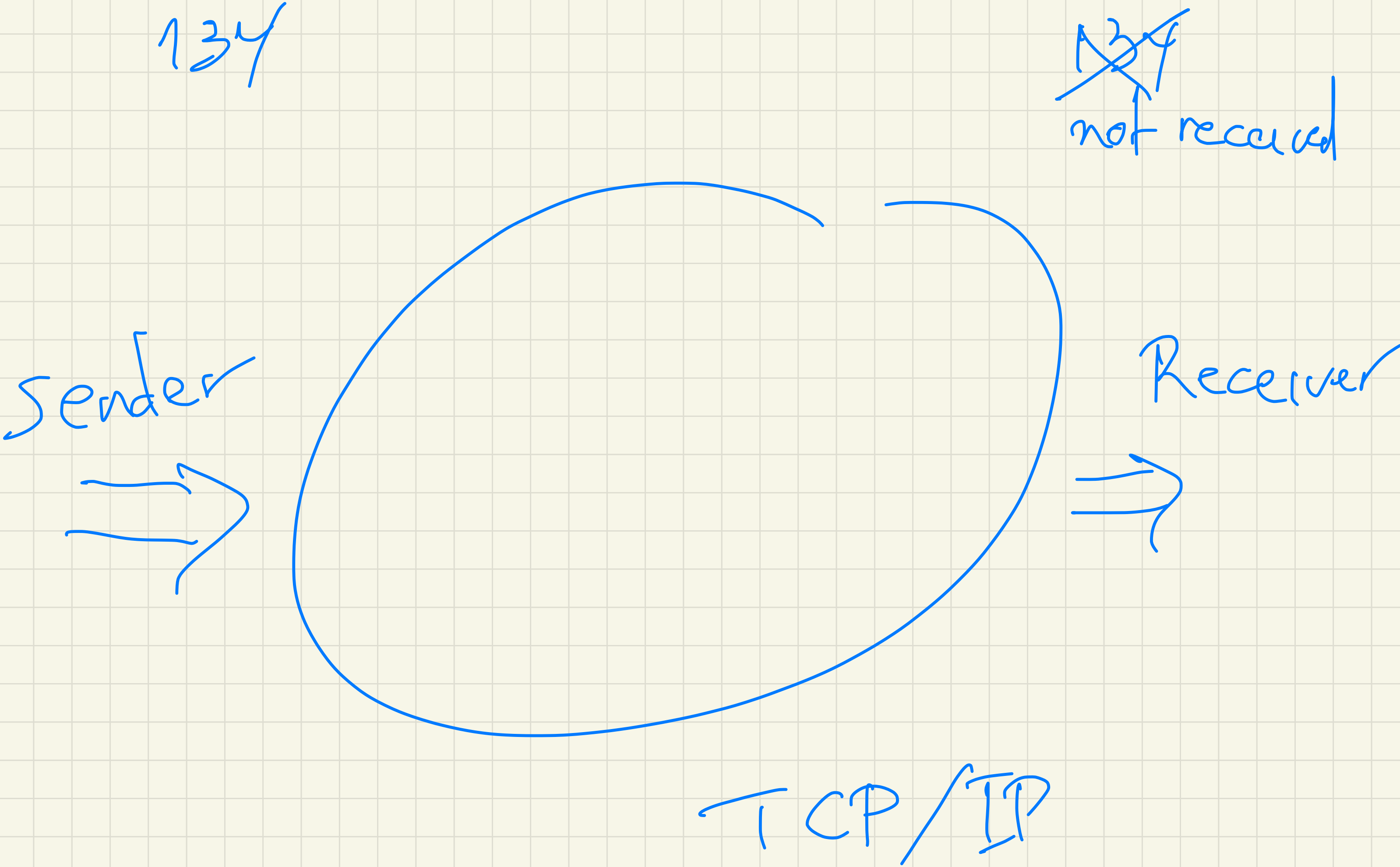
$$\lambda = \lim_{t\to\infty} \frac{\beta(t)}{t} = \lim_{t\to\infty} \frac{\alpha(t)}{t}$$

$$T = \lim_{t\to\infty} \frac{\sum_{i=1}^{\beta(t)} T(i)}{\beta(t)} = \frac{\sum_{i=1}^{\alpha(t)} T(i)}{\alpha(t)}$$

$$N = \lim_{t\to\infty} N_t = \lim_{t\to\infty} \frac{1}{t} \sum_{i=1}^{t} N(i)$$

Systems in steady state satisfies equilibrium conditions

- In a way, these are stability conditions!
- The significance of Little's result is that it holds for any system that reaches a steady state.
- System need not consist of a single queue provided that the terms N, λ, T are properly interpreted.


134
~~134~~
not received
Sender
Receiver
TCP/IP

Example 1

- Suppose we have a closed full system of K servers and N customers, $N \geq K$ (closed means departing customers are always replaced).
- Say average customer service time is $\bar{X}$; we want to find the average customer time T in the system. Apply Little's Theorem on the whole system: $N = \lambda T$.
- Apply Little's Theorem on the service portion: $K = \lambda \overline{X}$ since all K servers are always busy
- If follows that:

$$\frac{N}{T} = \frac{K}{\overline{X}}$$

- Hence:

$$T = N \frac{\overline{X}}{K}$$

Example 2 (Complex)

- Consider now the system under the assumption that customers arrive at a rate λ and are lost (or blocked) if they find the system full.
- In this case the number of busy servers may be less than K. Let $\bar{K}$ be the average number of busy servers, β the proportion of customers that are blocked from entering the system. From Little's theorem we derive that

$$\bar{K} = (1-\beta)\lambda\bar{X},$$

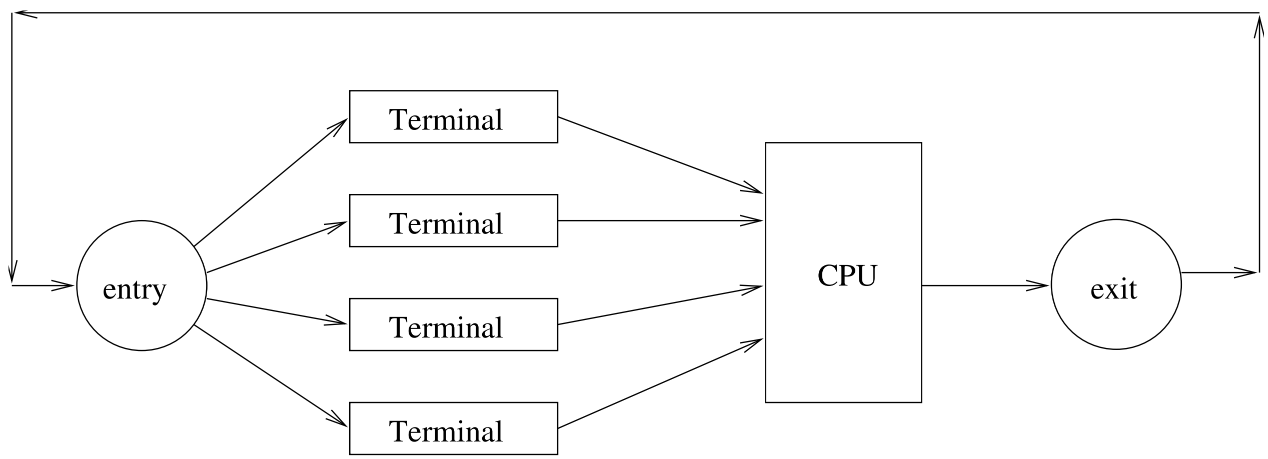
$$\beta = 1 - \frac{\bar{K}}{\lambda\bar{X}}.$$

Since $\bar{K} \leq K$ we obtain the lower bound

$$\beta \geq 1 - \frac{K}{\lambda\bar{X}}.$$

Example 3 (1/3)

- Suppose a system consisting of N terminals connected to a single CPU. Users login through a terminal.



- After reflection R, a user submits a job requiring average processing time P.
- Applying Little's Theorem between entry and exit portion of the system, we have: $N = \lambda T$, where T is average time a user spends in the system, and λ the attainable system throughput.

Example 3 (2/3)

- However, $T = R + D$, where D is the average delay between the time a job is submitted and the time its execution is completed. Clearly D may vary.

$$\underset{\substack{\uparrow \\ \text{case of no other} \\ \text{job submitted}}}{P} \quad \leq \quad D \quad \leq \quad \underset{\substack{\uparrow \\ \text{waiting for other} \\ \text{job to be com-} \\ \text{pleted}}}{NP}$$

$$R + P \quad \leq \quad T \quad \leq \quad R + NP$$

- Hence: $\frac{N}{R+NP} \leq \lambda = \frac{N}{T} \leq \frac{N}{R+P}$
- However, λ is also bounded above by the processing capacity of the computer. Since the CPU can not process more than one terminal per P time units.

Example 3 (3/3)

- We have: $\lambda \le \frac{1}{P}$
- Hence:

$$\begin{array}{ccccc} \frac{N}{R+NP} & \le & \lambda & \le \min & \left\{\frac{1}{P}, \frac{N}{R+P}\right\} \\ \downarrow & & & & \\ (N \to \infty) & & & & \\ \downarrow & & & & \\ \frac{1}{P} & \le & \lambda & \le & \frac{1}{P} \end{array}$$

which means that in the limit $\lambda = 1/P$.

- By using:

$$T = \frac{N}{\lambda}$$

- We obtain:

$$\max\{R+P, NP\} \le T \le R + NP$$

Exercises[a]

1. What happens in Little's theorem if the rate of arrival and departure differ?
   (a) Consider the case: where rate of arrival is less than the rate of departure
   (b) Consider the case: where rate of departure is less than the rate of arrival
2. Packets arrive every k seconds at a regular rate; first packet arrives at time $= 0$. All packets have equal length and require αk seconds for transmission ($\alpha \leq 1$). Suppose delay and propagation time is P seconds. Then:
   (a) What is the arrival rate λ of the packets as a function of k?
   (b) How much time T does a packet spend in the system?
   (c) Now use plug in the formulas for λ and T to Little's

[a] Not to submit

theorem $N = \lambda T$ to give an expression for N.

3. Suppose we have a network of n different nodes $1, 2, \ldots, n$. Suppose the packets arrive at node i at a rate λ_i; let N_i= average number of packets in the system arriving at the node i; and assume T_i = average delay of packets at node i.
   (a) Apply Little's Theorem to each node i. What does the theorem say?
   (b) Now look at it as a whole system, and let N is the average number of packets in the system. Apply Little's theorem to the entire system to derive the average time T a packet is in the system as a function of N and λ_i.
4. Formulate and prove a form of Little's theorem when arriving packets are lost uniformly and independently with probability p.
5. Formulate and prove a form of Little's theorem when departing

packets are lost uniformly and independently with probability p.

6. Formulate and prove a form of Little's theorem when packets are lost inside the wire uniformly and independently with probability p.

7. Formulate and prove a form of Little's theorem when packet faults may occur in any of the arriving and departing packets at the same time.