**CAPITULO 3**

One of the most important performance measures of a data network is the average delay required to deliver a packet from origin to destination. Furthermore, delay considerations strongly influence the choice and performance of network algorithms, such as routing and flow control.

Queueing theory is the primary methodological framework for analyzing network delay.

Packet delay within the communication subnet is the sum of delays on each subnet link traversed by the packet. Each link delay in turn consists of four components.

1. The **processing delay** between the time the packet is correctly received at the head node of the link and the time the packet is assigned to an outgoing link queue
2. The **queueing delay** between the time the packet is assigned to a queue for transmission and the time it starts being transmitted.
3. The **transmission delay** between the times that the first and last bits of the packet are transmitted.
4. The **propagation delay** between the time the last bit is transmitted at the head node of the link and the time the last bit is received at the tail node.

This accounting neglects the possibility that a packet may require retransmission on a link due to transmission errors or various other causes.

**3.1.1 Multiplexing of Traffic on a Communication Link**

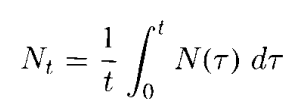
The communication link considered is viewed as a bit pipe over which a given number of bits per second can be transmitted. This number is called the **transmission capacity** of the link.

The manner of allocation of capacity among these traffic streams has a profound effect on packet delay. In the most common scheme, **statistical multiplexing**, the packets of all traffic streams are merged into a single queue and transmitted on a first-come first-serve basis.

In **time-division** (TDM) and **frequency-division multiplexing** (FDM) with m traffic streams, the link capacity is essentially subdivided into m portions – one per traffic stream. In FDM, the channel bandwidth W is subdivided into m channels each with bandwidth W/m. The transmission capacity of each channel is roughly *C /m*, where C is the capacity that would be obtained if the entire bandwidth were allocated to a single channel. The transmission time of a packet that is L bits long is *Lm/C*, or m times larger than in the corresponding statistical multiplexing scheme.

In TDM, allocation is done by dividing the time axis into slots of fixed length. In the case where the slots are short relative to packet length, we may again regard the transmission time of a packet L bits long as *Lm/C*. In the case where the slots are of packet length, the transmission time of an L bit packet is L/C, but there is a wait of (m - 1) packet transmission times between packets of the same stream. Statistical multiplexing has smaller average delay per packet than either TDM or FDM.

**3.2.1 Little’s Theorem**

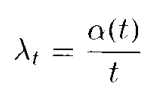


N(t) = Number of customers in the system at time t

A(t) = Number of customers who arrived in the interval [0, t]

time average of N(T) up to time t

Tn = Time spent in the system by the n-th arriving customer

****

N, λ, and T are related by a simple formula that makes it possible to determine one given the other. This result, known as **Little's Theorem.**

time average arrival rate



Little's Theorem expresses the natural idea that crowded systems (large N) are associated with long customer delays (large T) and reversely.

**3.2.3 Applications of Little´s Theorem**

The significance of Little's Theorem is due in large measure to its generality. It holds for almost every queueing system that reaches a steady-state. The system need not consist of just a single queue. Indeed, with appropriate interpretation of the terms N, λ, and T, the theorem holds for many complex arrival-departure systems.

Text, whiteboard

Description automatically generated

A picture containing icon

Description automatically generated**Poisson process -** a Poisson process with average event rate λ, models the probability of a given number of events n occurring in time interval of duration T

Text

Description automatically generated with medium confidence**Poisson distribution** - a Poisson distribution with parameter m gives the probability of a certain number of events n occurring, during a certain time interval, in a Poisson process where the average number of events in that interval is m

3.3 THE M / M /1 QUEUEING SYSTEM

The M / M / 1 queueing system consists of a single queueing station with a single server (in a communication context, a single transmission line).

1. The first letter indicates the nature of the arrival process. M stands for memoryless.
2. The second letter indicates the nature of the probability distribution of the service times. M, G, and D stand for exponential, general, and deterministic distributions, respectively. In all cases, successive interarrival times and service times are assumed to be statistically independent of each other.
3. The last number indicates the number of servers.

**3.3.1 Markov chain formulation**

An important consequence of the memoryless property is that it allows the use of the theory of Markov chains. Once we know the number N(t) of customers in the system at time t, the times at which customers will arrive or complete service in the future are independent of the arrival times of the customers presently in the system and of how much service the customer currently in service (if any) has already received.

Nk = Number of customers in the system at time a\*k (a is a small positive number)

State k - k clients (packets) in the queue

p(i,j) – probability of transition from state i to state j in interval d

Text, letter

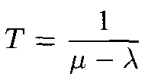
Description automatically generatedWhen d → 0:

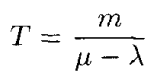
A picture containing text, watch

Description automatically generatedShape

Description automatically generated with medium confidence

**3.3.2 Statistical Multiplexing Compared with Time- and Frequency-Division Multiplexing**

If the streams are merged into a single Poisson stream, with rate A, as in statistical multiplexing, the average delay per packet is

If, instead, the transmission capacity is divided into m equal portions the average delay per packet is, m times larger than for statistical multiplexing:

Text, letter

Description automatically generated

**3.6.1 The Kleinrock Independence Approximation**

A picture containing text

Description automatically generatedWe now formulate a framework for approximation of average delay per packet in data networks. Let xs, in packets/sec, be the arrival rate of the packet stream s. Then the total arrival rate at link (i, j) is:

A picture containing text

Description automatically generatedThe preceding network model is well suited for virtual circuit networks. Assuming that no packets travel in a loop, let xs denote the arrival rate of packet stream s, and let fij(s) denote the fraction of the packets of stream 8 that go through link (I, j). Then the total arrival rate at link (i, j) is:

Kleinrock suggested that merging several packet streams on a transmission line has an effect akin to restoring the independence of interarrival times and packet lengths.

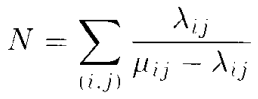
It was concluded that it is often appropriate to adopt an M/M/1 queueing model for each communication link regardless of the interaction of traffic on this link with traffic on other links.

This is known as the **Kleinrock independence approximation** and seems to be a reasonably good approximation for systems involving Poisson stream arrivals at the entry points, packet lengths that are nearly exponentially distributed, a densely connected network, and moderate-to-heavy traffic loads.

Text, letter

Description automatically generatedThe average number of packets in queue or service at (i, j) is:

1/ u(i,j) is the average packet transmission time on link (i, j).

The average number of packets summed over all queues is:

A picture containing clock

Description automatically generated

So by Little's Theorem, the average delay per packet:

Where phi is the total arrival rate in the system.

A picture containing antenna

Description automatically generated

If the average processing and propagation delay d(i,j) at link (i,j) is not negligible, this formula should be adjusted to:

Text, letter

Description automatically generated

Finally, the average delay per packet of a traffic stream traversing a path p is given by:

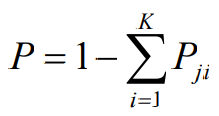
Where the three terms in the sum above represent average waiting time in queue, average transmission time, and processing and propagation delay, respectively.

3.4 Jackson Networks

A picture containing text, clock, watch

Description automatically generatedArrival rate at node j:

Independent routing of packets

When a packet leaves node i it comes to node j with probability P(i,j). Packets can loop inside network. Packet leaves the system at node j with probability.

Text, letter

Description automatically generated

Falta fazer M/M/1/B Queue + M/G/1 Queue