Network Engineering

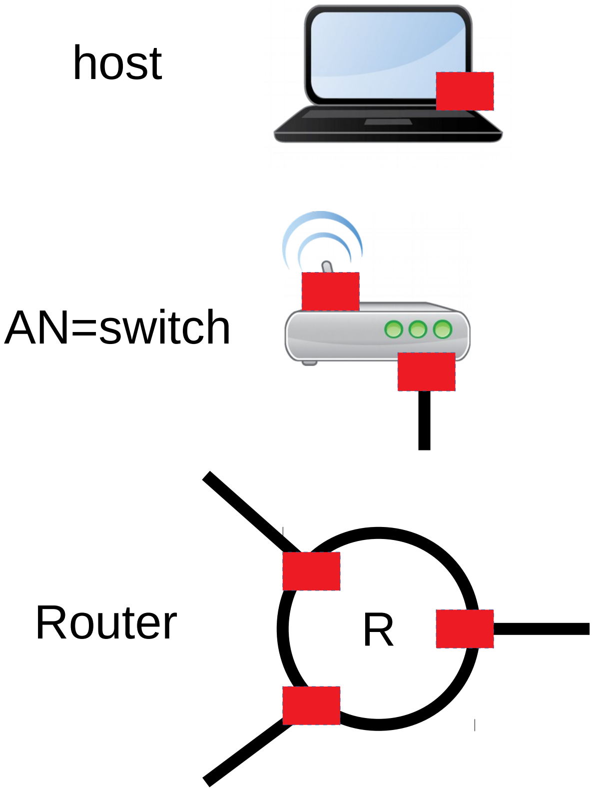
Lecture 6. Modelling a Network Interface 1

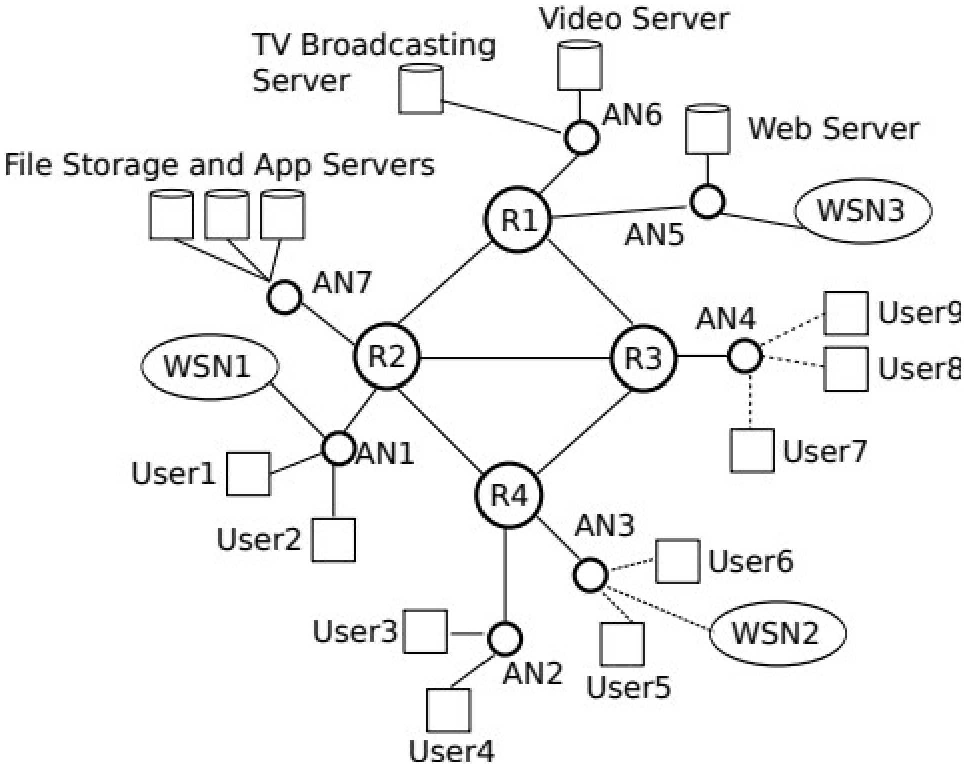
Boris Bellalta

boris.bellalta@upf.edu

WN UPF Wireless Networking
https://www.upf.edu/web/wnrg

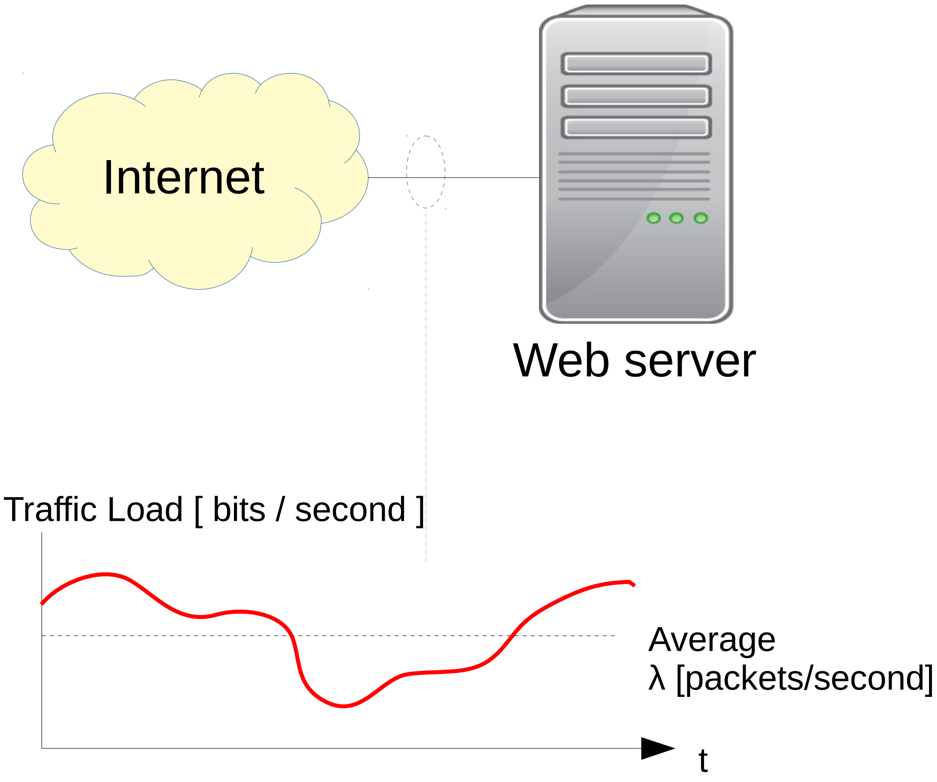
Routers, Access Networks (Switchs/APs), Hosts

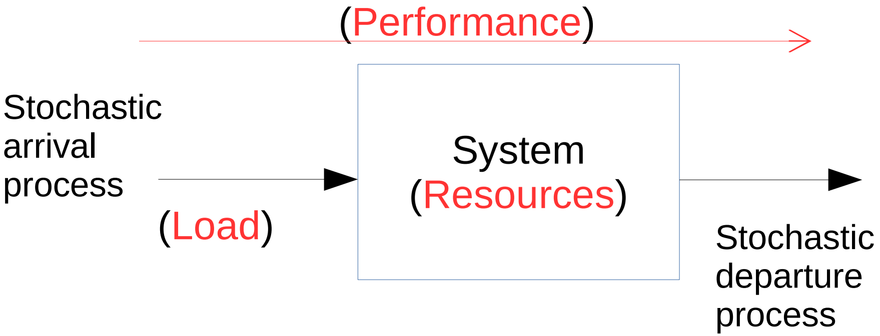


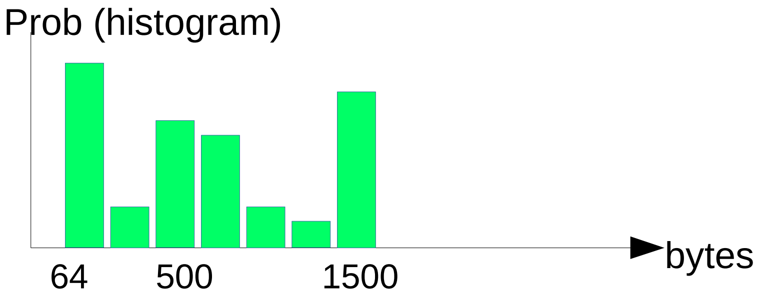


System

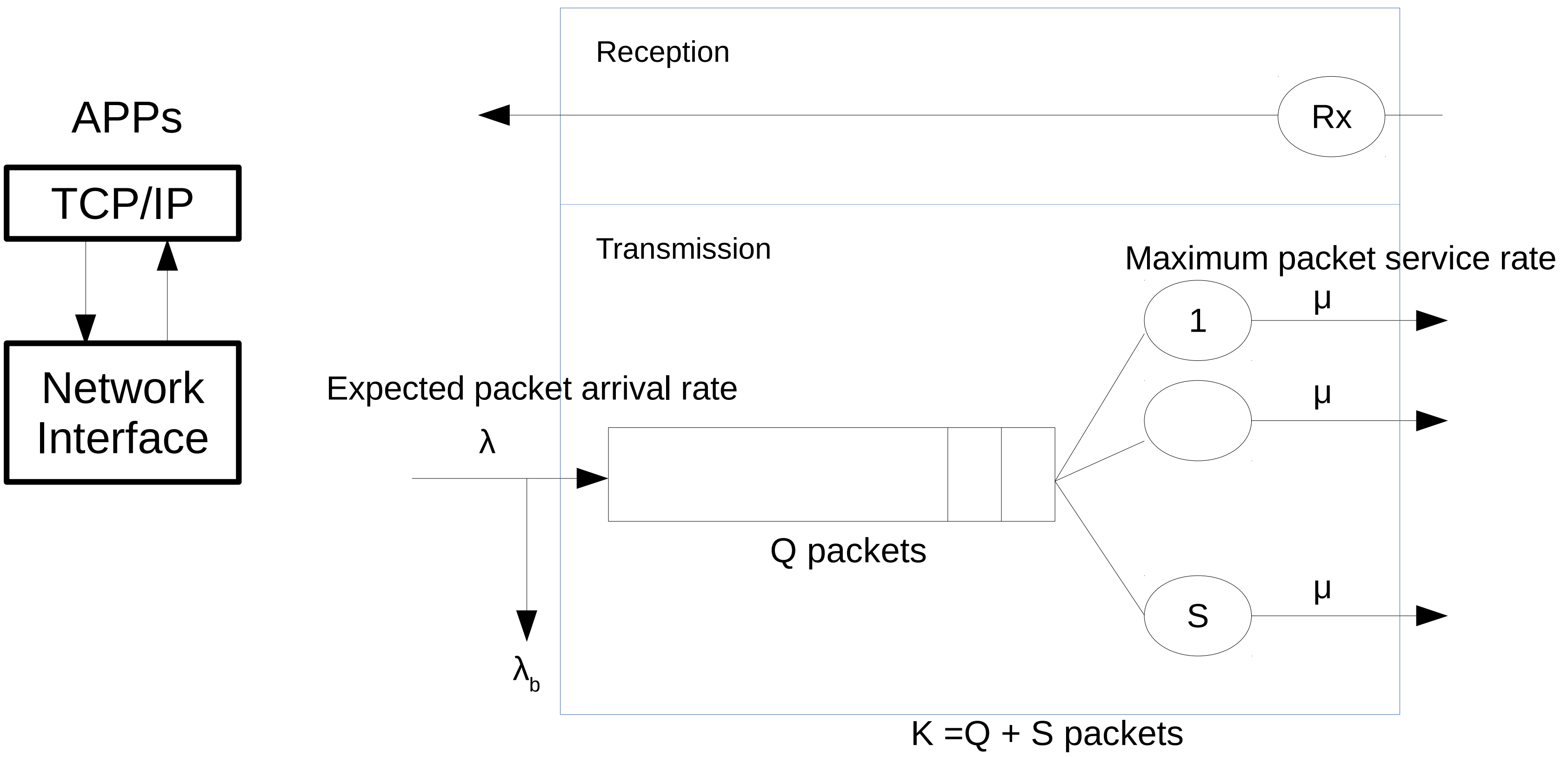
Relationship between ‘Load’, ‘Resources’ and ‘Performance’



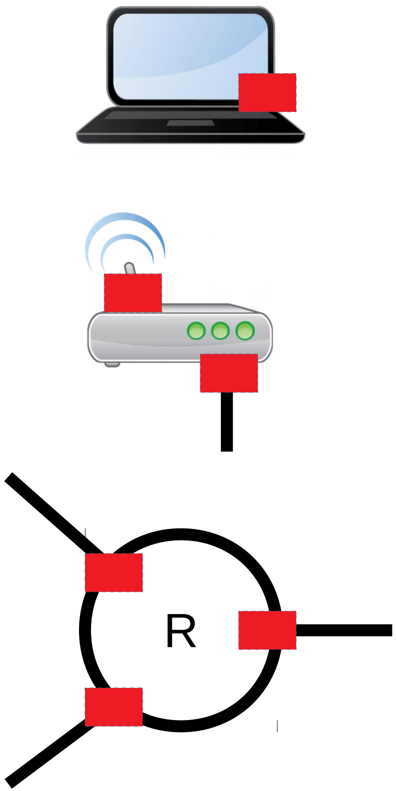


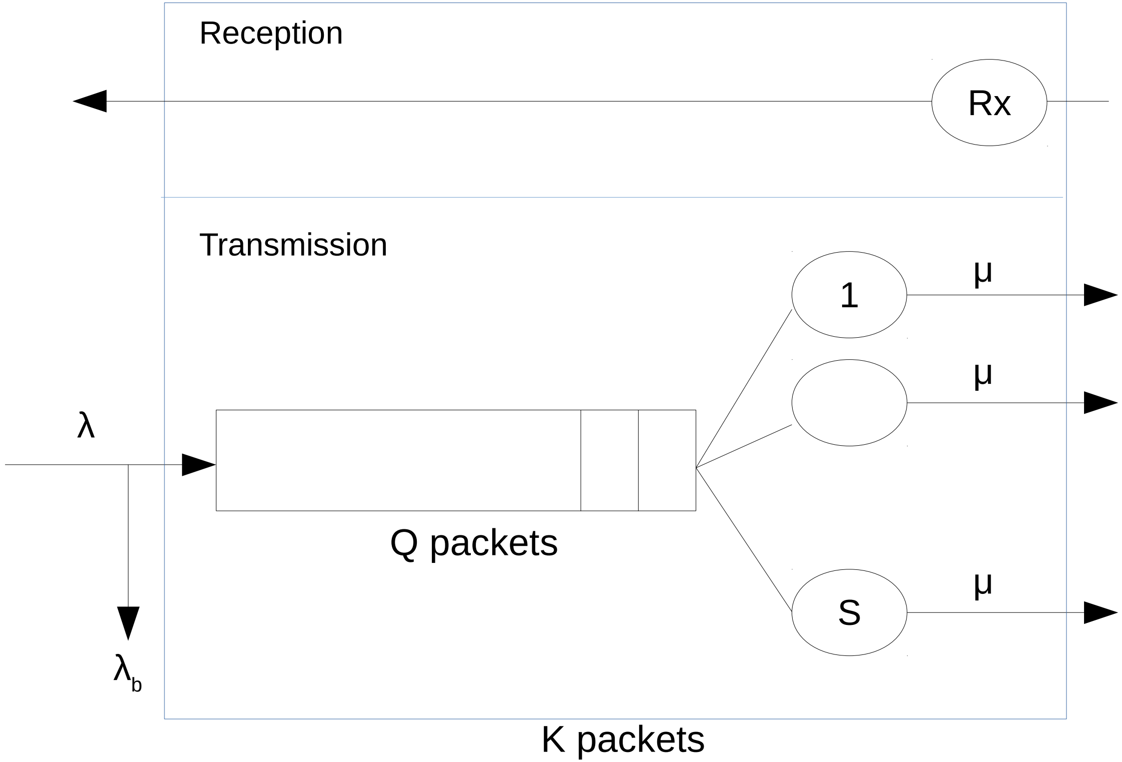


Network Interface



Host, Switch / Access Network, Router





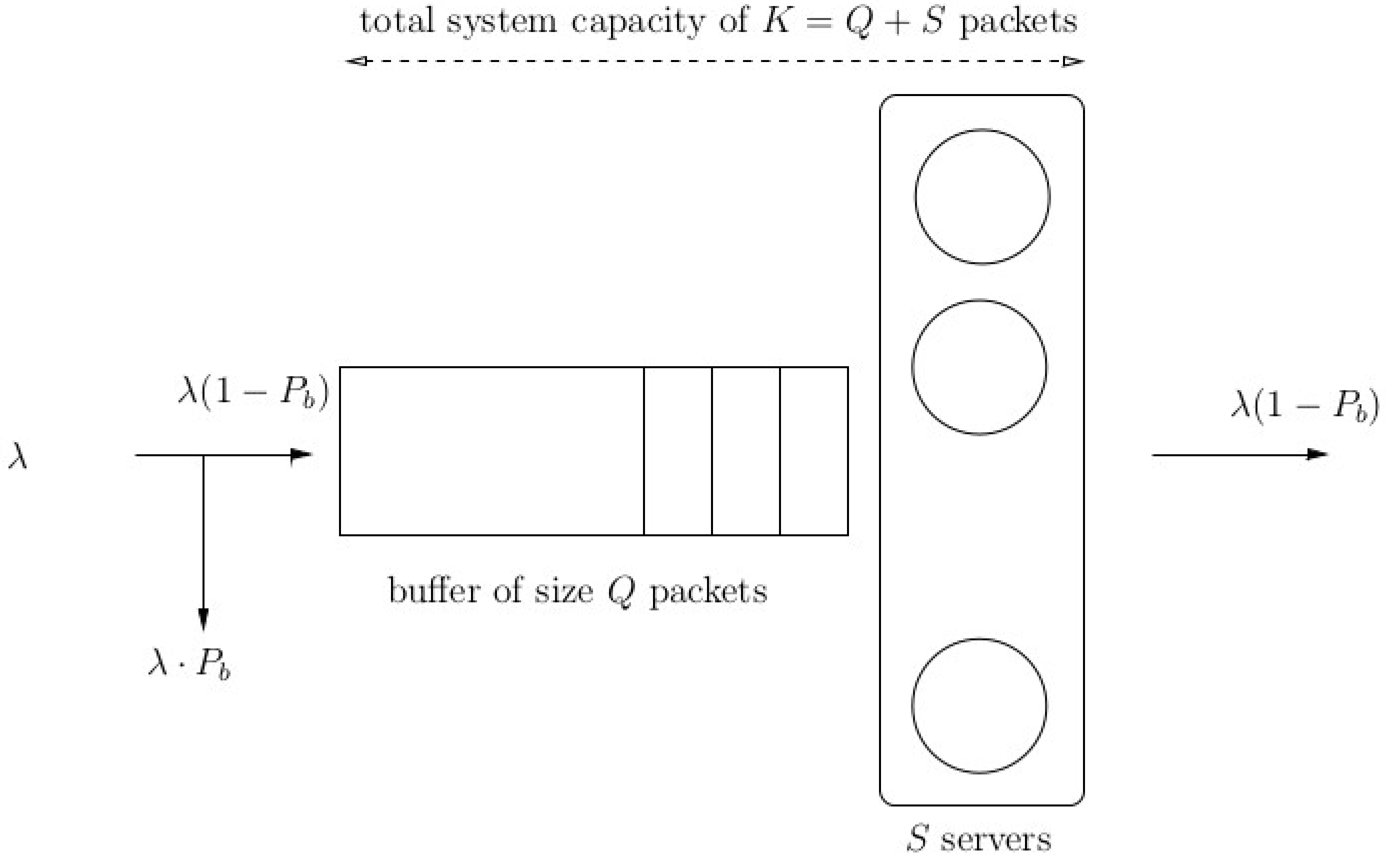
* Unless otherwise stated, all links are bidirectional (half or full duplex)

Network Interface

A Network interface can be abstracted by assuming it contains:

- A single buffer and a single transmitter (usual).
- A single buffer and multiple transmitters (rare). However, it is a more general case containing the previous one.
- Multiple buffers and multiple transmitters, one for each buffer. An example of this case would be an Ethernet switch. However, in this case, we will simply analyze each queue independently of the others, returning to the first case.

Network Interface



Parameters Network Interface

- The number S transmitters or servers.
- Maximum number K of packets in the system (buffer + servers). A real-life buffer is limited in size, although it might be large. This is modeled by the maximum number of packets it can hold.
- Maximum number $Q = K - S$ of packets in the buffer.
- Transmission rate of each transmitter: R bps. We assume the same transmission rate for all transmitters, i.e., the same transmission technology is used in the device.
- Aggregate packet arrival rate: λ packets/second. This is the number of packets that arrives at the buffer per time unit (typically seconds), averaged over a

Parameters Network Interface

long time scale. Bursts in the packet arrival process, i.e., periods in which more than average packets arrive in a short time, are modeled by the variance of the packet interarrival time distribution, which is in the following typically assumed to be exponential.

- Packet Length: L bits. This is a random variable describing the distribution of the number of bits over the arriving packets.
- Service/Transmission time: $E[D_s] = \frac{E[L]}{R}$ seconds. Since we assume R to be constant, the variation of the transmission time therefore depends on the variation of the packet length distribution.
- Packet Departure rate: $\lambda(1 - P_b)$ packets/second.

Parameters Network Interface

- Packet Departure rate: $\lambda(1 - P_b)$ packets/second.
- Maximum Packet Departure rate (assuming that all the transmitters are always busy): $S\mu = \frac{S}{E[D_s]}$ packets / second.
- Offered traffic or traffic intensity: $a = \frac{\lambda}{\mu}$ [Erlangs] (note that this value is independent of the number of servers).
- System utilization: $\rho = 1 - P_e$ (assuming that the system is working always when it is not empty).

Definitions

- Normalized Traffic load: $a = \lambda/\mu$ [Erlangs]
- Normalized Throughput: $\rho = \lambda(1\text{-}P_b)/(S\mu)$ [Erlangs]
- Utilization: $\rho = \lambda(1\text{-}P_b)/(S\mu)$
  - Fraction of the time the servers are busy (i.e., the transmitters are active) in average.
  - Using the equilibrium distribution, we know that the probability the servers are busy is $1\text{-}\pi_0$.
  - Then, it must be that: $\rho = \lambda(1\text{-}P_b)/(S\mu) = 1\text{-}\pi_0$; and $\lambda(1\text{-}P_b) = S\mu(1\text{-}\pi_0) \rightarrow \lambda(1\text{-}\pi_K) = S\mu(1\text{-}\pi_0)$

Erlang Notation

- A/B/C/D:SP/Population
  - A: distribution of the arrival process (deterministic (D), Poisson (M), uniform (U), etc.)
  - B: distribution of the service process (deterministic (D), exponential (M), uniform (U), etc.)
  - C: number of servers (S)
  - D: maximum number of entities of the system (K=Q+S)
  - SP: Service policy (FIFO, Non-preemptive priority, preemptive priority)
  - Population: finite or infinite

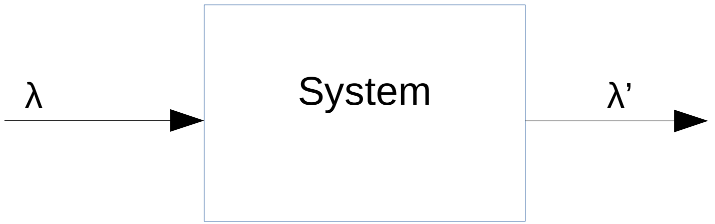
Erlang Notation

- **M/M/1/K**: Poisson arrival process, exponentially distributed service times, 1 server and a total system capacity of K packets, with $Q = K - 1$. A FIFO policy is considered.
- **M/M/3/K**: Poisson arrival process, exponentially distributed service times, 3 servers and a total system capacity of K packets, with $Q = K - 3$. A FIFO policy is considered.
- **M/M/1**: Poisson arrival process, exponentially distributed service times, 1 server and a total system capacity of ∞ packets. A FIFO policy is considered.

Erlang Notation

- **M/D/2**: Poisson arrival process, deterministically distributed service times, 2 servers and a total system capacity of ∞ packets. A FIFO policy is considered.
- **D/D/2**: Deterministic arrival process, deterministically distributed service times, 2 servers and a total system capacity of ∞ packets. A FIFO policy is considered.
- **M/G/1**: Poisson arrival process, generally distributed service times, 1 server and a total system capacity of ∞ packets. A FIFO policy is considered. The general distribution is usually characterized by both the expected value and the coefficient of variation.

Stability

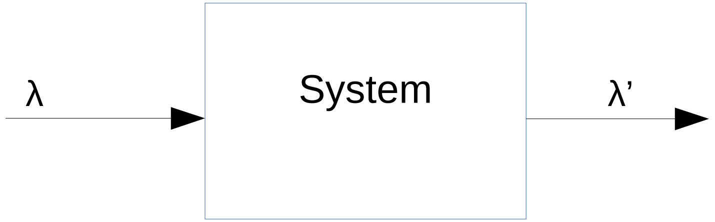


If the buffer is finite $\rightarrow$ the system is always stable: $\lambda' = \lambda\,(1\text{-}Pb)$

If the buffer is infinite $\rightarrow$ the system is stable only if $S\mu > \lambda$

$$\lambda(1 - P_\mathrm{b}) < S\mu \quad \rightarrow \quad \frac{\lambda(1 - P_\mathrm{b})}{S\mu} < 1 \quad \rightarrow \quad \frac{a(1 - P_\mathrm{b})}{S} < 1$$

Little’s formula



Little’s law, applied to our use-case says that the long-term average number of packets in a system is the average arrival rate of packets to that system multiplied by the average time it spends in the system. This can be applied to just the number of waiting packets (if just the buffer is interpreted as the system), just the number of packets in the server (with just the transmitter being the system), or the complete system. Formally, this means

$$\mathrm{E}[D_q] = \frac{\mathrm{E}[N_q]}{\lambda(1-P_{\mathrm{b}})}, \qquad \mathrm{E}[D_s] = \frac{\mathrm{E}[N_s]}{\lambda(1-P_{\mathrm{b}})} \tag{5.23}$$

Performance Metrics

- Probability that the system is empty at any arbitrary time: $P_e = \pi_0$.
- Probability that there are i packets in the system at any arbitrary time: π_i.
- Blocking or packet loss probability: $P_b = \pi_K$.
- System utilization: the fraction of time that the system is active (i.e., transmitting packets), given by: $\rho = 1 - \pi_0$.
- Expected number of packets in the queue: $E[Q] = \sum_{i=S+1}^{K} (i - S) \cdot \pi_i = E[N] - E[N_s]$.
- Expected number of packets in service: $E[N_s] = \sum_{i=0}^{K} \min(i, S) \cdot \pi_i$.
- Expected number of packets in the system: $E[N] = \sum_{i=0}^{K} i \cdot \pi_i$.

Performance Metrics

- Expected delay of a packet in the buffer:

$$E[D_q] = \frac{E[N_q]}{\lambda(1 - P_b)}.$$

- Expected delay of a packet in the server:

$$E[D_s] = \frac{E[N_s]}{\lambda(1 - P_b)}.$$

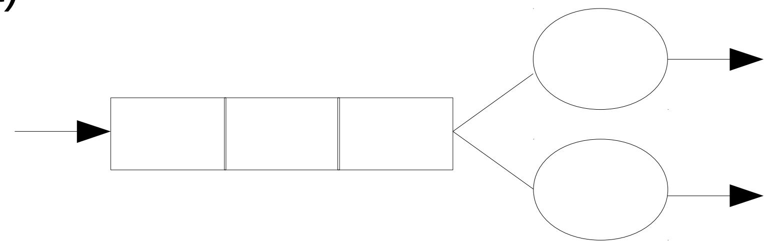
- Expected delay of a packet in the system:

$$E[D] = \frac{E[N]}{\lambda(1 - P_b)}.$$

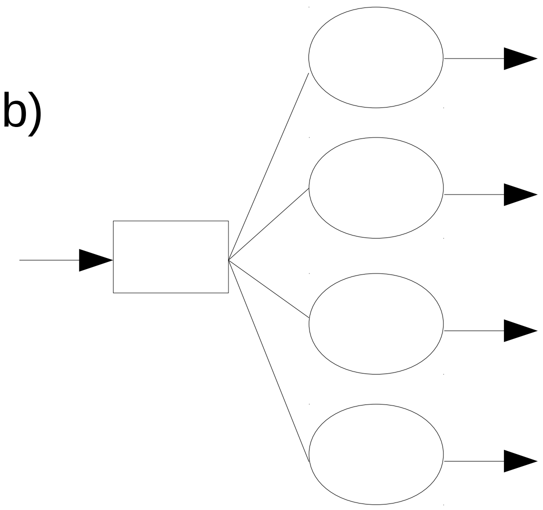
Exercise

- Q: Which system is empty during a larger fraction of time:
  - a) M/M/2/5 vs b) M/M/4/5,
  - with Ra = 1 Mbps, Rb = 0.5 Mbps,
  - For both a) and b) : Traffic Load = 1.8 Mbps, L = 10000 bits

a)



b)



M/M/S/K system

The M/M/S/K system has been widely used in the past to plan telephone networks, where the number of servers S models the number of calls that can be active simultaneously in a link or a cell (i.e., lines or channels).

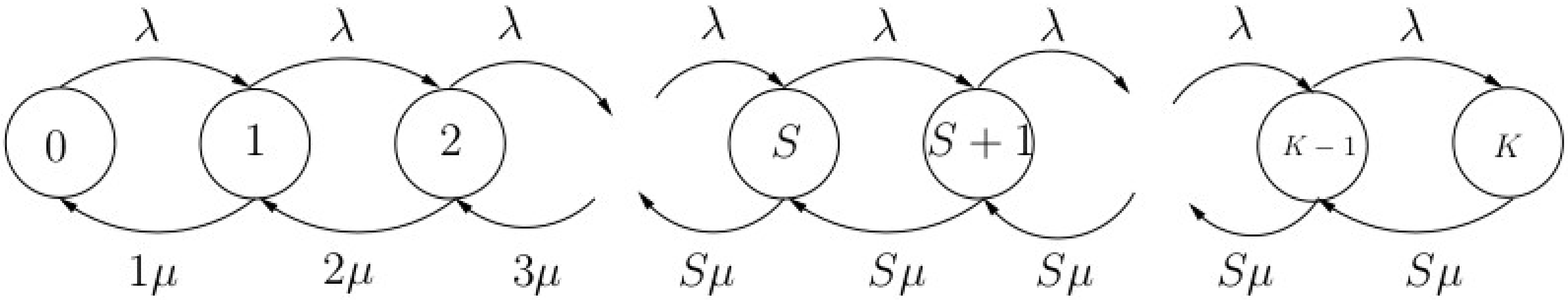


Figure 5.3: Markov Chain for the M/M/S/K queue

The M/M/S/K queueing system implicitly assumes Poisson arrivals with rate λ and exponentially distributed service times with average $E[D_s] = 1/\mu$.

M/M/S/K system

$$\pi_0 \lambda = \pi_1 \mu$$

$$\pi_1 \lambda = \pi_2 2\mu$$

$$\cdots$$

$$\pi_{i-1} \lambda = \pi_i i \mu, \qquad i \leq S$$

$$\cdots$$

$$\pi_{i-1} \lambda = \pi_i S \mu, \qquad i \geq S$$

$$\cdots$$

$$\pi_{K-1} \lambda = \pi_K S \mu$$

This, together with the normalization condition:

$$\sum_{i=0}^{K} \pi_i = 1,$$

M/M/S/K system

$$\pi_0 = \frac{1}{\sum_{j=0}^{S} \frac{a^j}{j!} + \sum_{j=S+1}^{K} \frac{1}{S^{j-S}} \frac{a^j}{S!}}.$$

$$\pi_i = \frac{\lambda}{i\mu} \pi_{i-1} = \frac{a^i}{i!} \pi_0, \qquad i \leq S \qquad (5.28)$$

$$\cdots$$

$$\pi_i = \frac{\lambda}{S\mu} \pi_{i-1} = \frac{a^{i-S}}{S^{i-S}} \pi_S = \frac{a^{i-S}}{S^{i-S}} \frac{a^S}{S!} \pi_0 = \frac{a^i}{S^{i-S} S!} \pi_0, \qquad i > S \qquad (5.29)$$