

# EP2200

# Queueing theory and teletraffic systems

Viktoria Fodor

Laboratory of Communication Networks

School of Electrical Engineering

## **Lecture 1**

*"If you want to model networks  
Or a complex data flow  
A queue's the key to help you see  
All the things you need to know."*

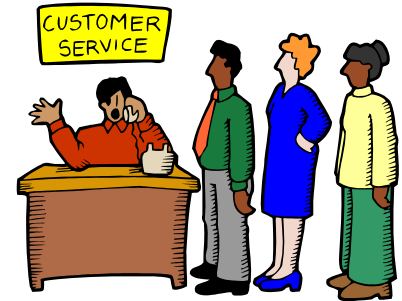
*(Leonard Kleinrock, Ode to a Queue  
from IETF RFC 1121)*

# What is queuing theory?

## What are teletraffic systems?

### Queuing theory

- Mathematical tool to describe resource sharing systems, e.g., telecommunication networks, computer systems
  - Requests arrive **dynamically**
  - Request may form a **queue** to wait for service
- Applied probability theory



### Teletraffic systems

- Systems with telecommunication traffic (data networks, telephone networks)
- Are designed and evaluated using queuing theory



Why do we need a whole theory for that?

# Why do we need teletraffic theory?

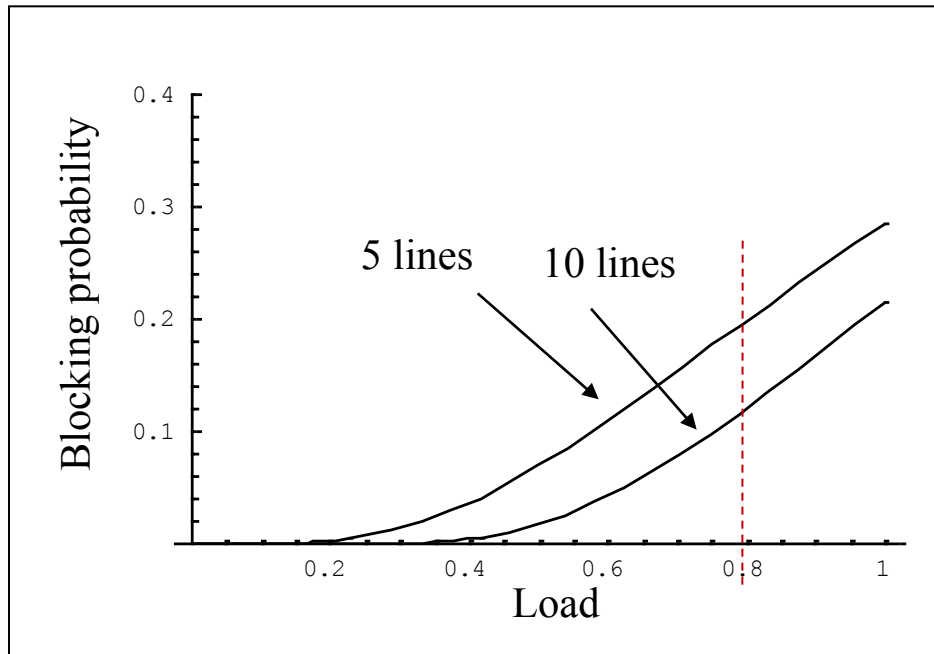
- Waiting delays at output buffers in network routers
- How will the delay change if the number of packets arriving within a second doubles?

# Why do we need teletraffic theory?

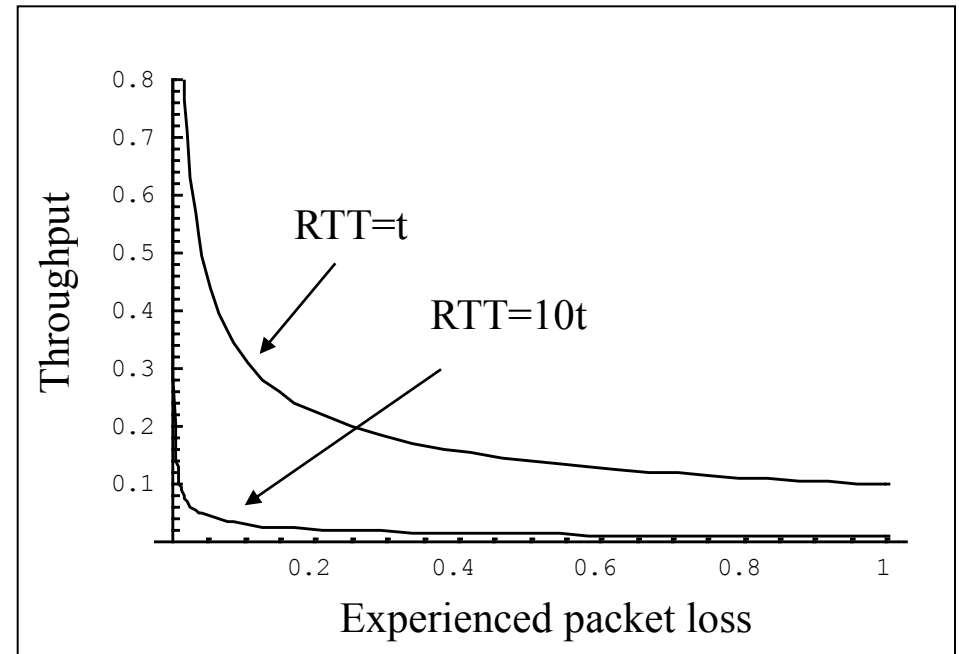
- Throughput (useful transmissions) in a wireless network with random access
- If transmissions may collide, how would the throughput change if the number of packets to be sent doubles? What is the effect of packet collisions?

# Why do we need teletraffic theory?

Call blocking probability  
in a telephone network vs. load



TCP throughput vs. packet loss



- Teletraffic systems are non-linear, and mathematical tools are needed to predict their performance

# Course objectives

- Basic theory
  - understand the theoretical background of queuing systems, apply the theory for systems not considered in class
- Applications
  - find appropriate queuing models of simple problems, derive performance metrics
- Basis for modeling more complex problems
  - advanced courses on performance evaluation
  - master thesis project
  - industry (telecommunication engineer)
- Prerequisites
  - mathematics, statistics, probability theory, stochastic systems
  - communication networks, computer systems

# Course organization

- Course responsible
  - Viktoria Fodor <[vfodor@s3.kth.se](mailto:vfodor@s3.kth.se)>
- Lectures
  - Viktoria Fodor
- Recitations
  - Ioannis (John) Glarapoulos
  - Liping Wang
- Course web page
  - KTH Social EP2200 (<https://www.kth.se/social/course/EP2200/>)
  - Home assignments, project, messages, updated schedule and course information
    - Your responsibility to stay up to date!
  - Useful resources: Erlang and Engset calculators, Java Aplets
  - Useful links: on-line books
  - [Links to probability theory basics](#)

# Course material

- Course binder
  - Lecture notes by Jorma Virtamo, HUT, and Philippe Nain, INRIA
    - Used with their permission
  - Excerpts from L. Kleinrock, *Queueing Systems*
  - Problem set with outlines of solutions
  - Old exam problems with outlines of solutions
  - Erlang tables (get more from course web, if needed)
  - Formula sheet, Laplace transforms

For sale at STEX, Q2 building. Costs 100 SEK.
- No text book needed!
  - If you would like a book, then you can get one on your own
    - Ng Chee Hock, *Queueing Modeling Fundamentals*, Wiley, 1998. (simple)
    - L. Kleinrock, *Queueing Systems, Volume 1: Theory*, Wiley, 1975 (well known, engineers)
    - D. Gross, C. M. Harris, *Fundamentals of Queueing Theory*, Wiley, 1998 (difficult)
  - Beware, the notations might differ



# Course organization

- 12 lectures – cover the theoretical part
- 12 recitations – applications of queuing models
- Home assignments and project (1.5 ECTS, compulsory, pass/fail)
- Home assignment
  - problems and (later) solutions on the web
  - individual submission, only handwritten version
  - you need 75% satisfactory solution to pass this moment
  - submission on Nov. 20, submit at the STEX office
- Small project
  - computer exercise
  - details later
  - submission deadline: Jan 4
  - +5 points for outstanding projects

# Exam

- There is a written exam to pass the course, 5 hours
  - Consists of five problems of 10 points each
    - Passing grade usually  $20 \pm 3$  points
  - Allowed aid is the Beta mathematical handbook (or similar) and simple calculator. **Probability theory and queuing theory books are not allowed!**
  - The sheet of queuing theory formulas will be provided, also Erlang tables and Laplace transforms, if needed (same as in the course binder and on the web)
- Possibility to complementary oral exam if you miss E by 2-3 points (Fx)
  - Complement to E
- Registration is mandatory for all the exams
  - At least two weeks prior to the exam
- Students from previous years: contact STEX (stex@ee.kth.se) if you are not sure what to do

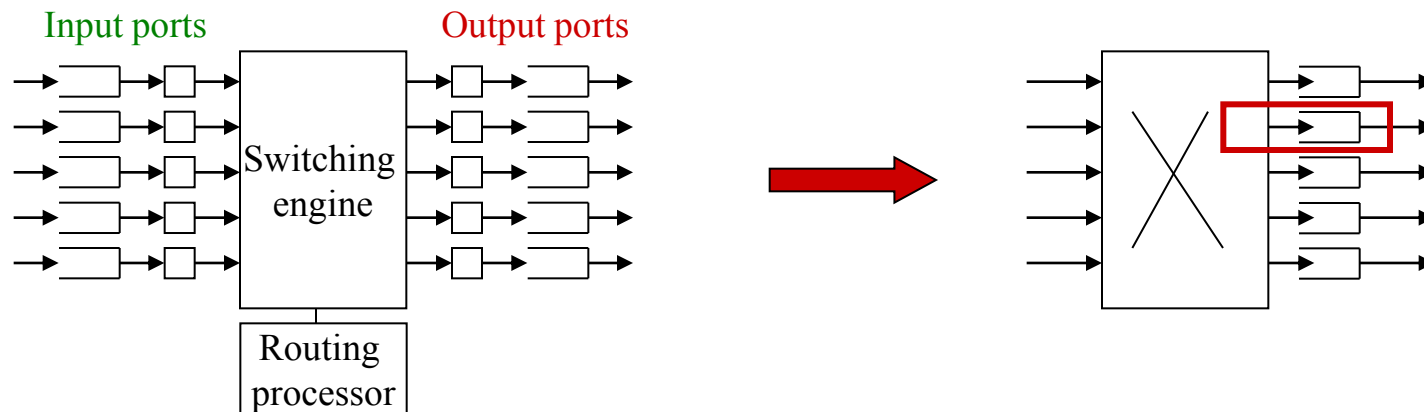
# Lecture 1

## Queuing systems - introduction

- Teletraffic examples and the performance triangle
- The queuing model
  - System parameters
  - Performance measures
- Stochastic processes recall

# Example

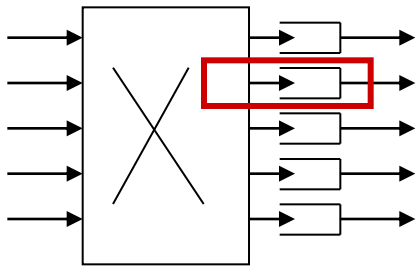
- Packet transmission at a large IP router



- We simplify modeling
  - typically the switching engine is very fast
  - the transmission at the output buffers limits the packet forwarding performance
  - we do not model the switching engine, only the output buffers

# Example

- Packet transmission at the output link of a large IP router – packets arrive randomly and wait for free output link

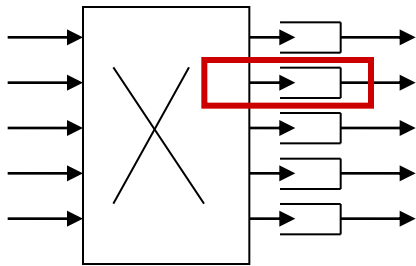


- Performance:

- Depends on:

# Example

- Packet transmission at the output link of a large IP router - packets wait for free output link

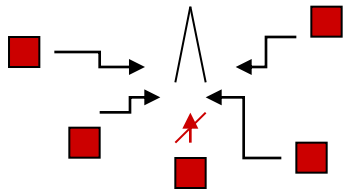


- Performance:

- Depends on:

# Example

- Voice calls in a GSM cell – calls arrive randomly and occupy a “channel”. Call blocked if all channels busy.

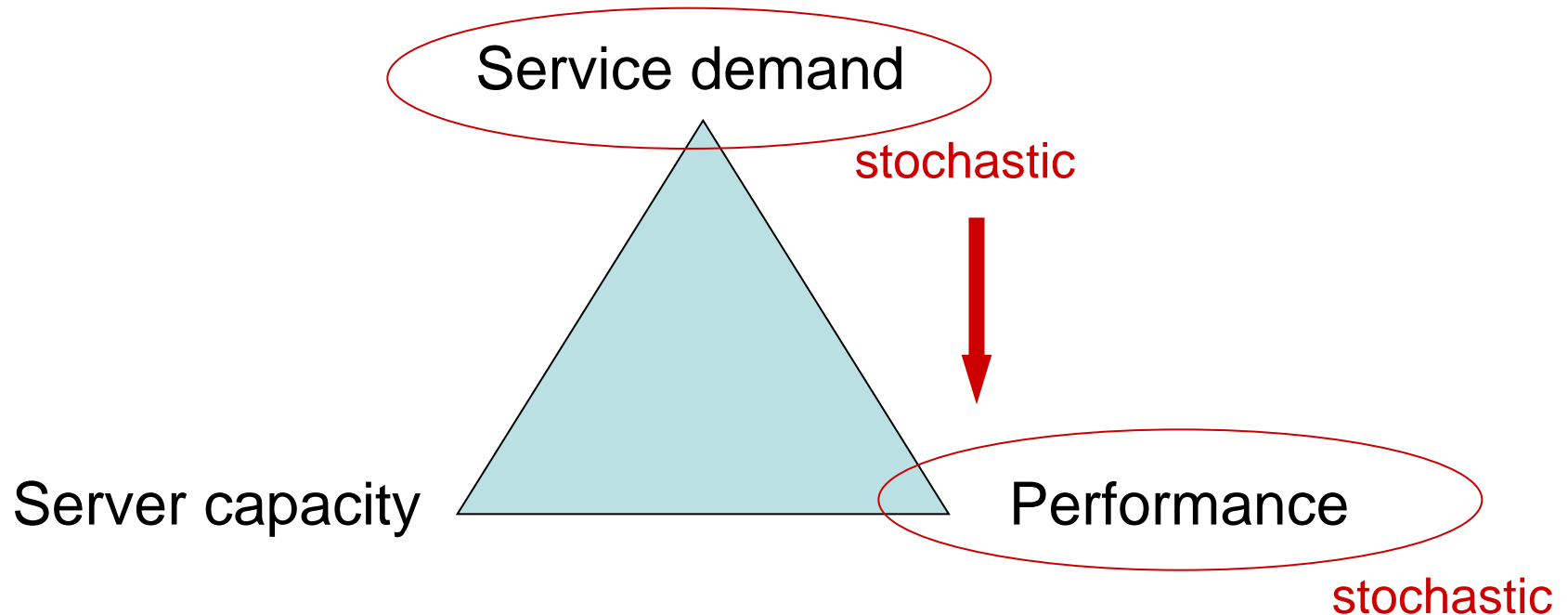


- Performance

- Depends on:

# Performance of queuing systems

- The triangular relationship in queuing



- Works in 3 directions
  - Given service demand and server capacity → achievable performance
  - Given server capacity and required performance → acceptable demand
  - Given demand and required performance → required server capacity



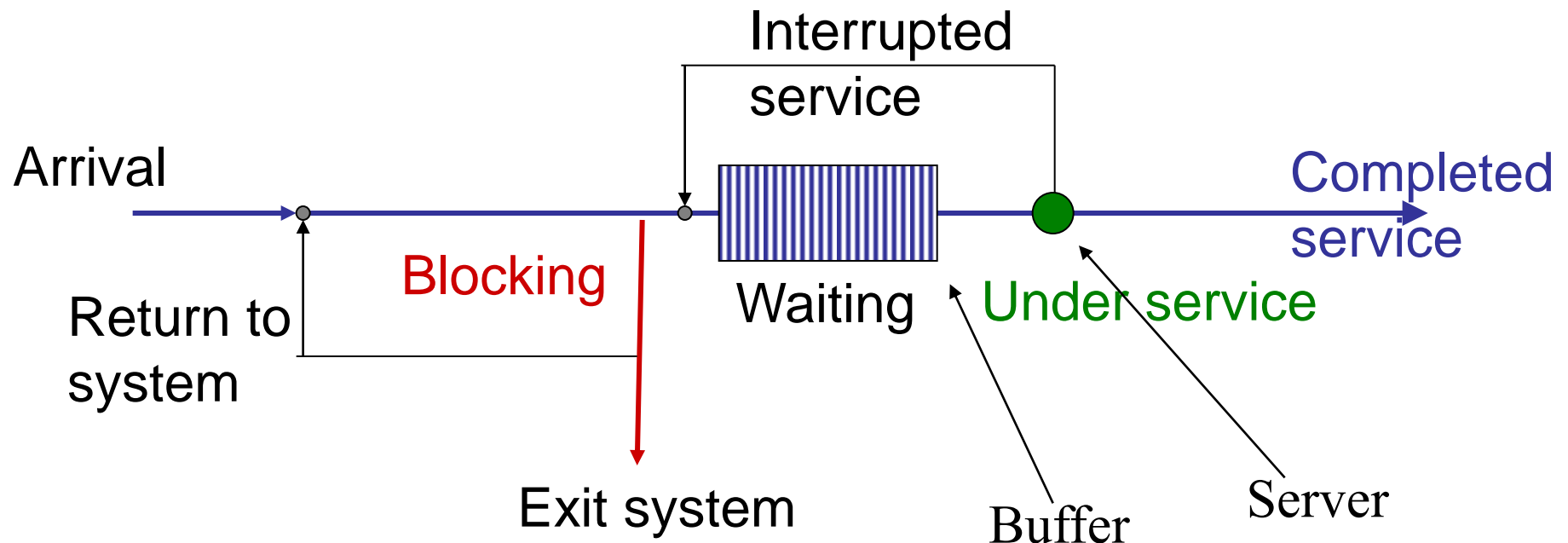
# Lecture 1

## Queuing systems - introduction

- Teletraffic examples and the performance triangle
- The queuing model
  - System parameters
  - Performance measures
- Stochastic processes recall

# Block diagram of a queuing system

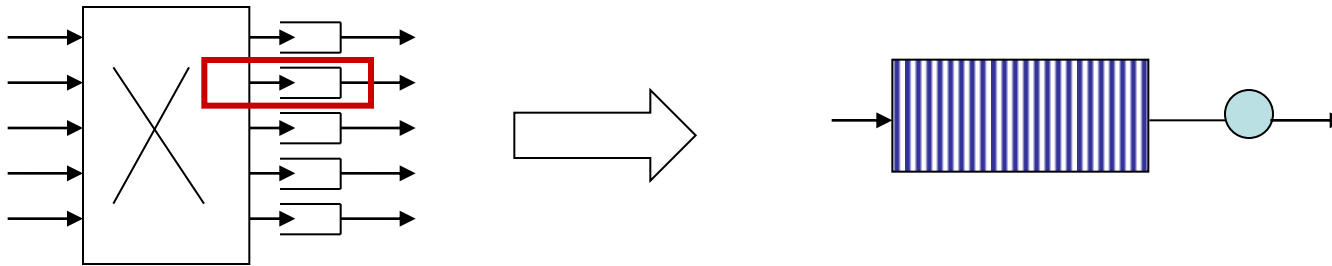
- Queuing system: abstract model of a resource sharing system
  - buffer and server(s)
- **Customers** arrive, wait, get served and leave the queuing system
  - customers can get blocked, service can be interrupted





# Examples in details

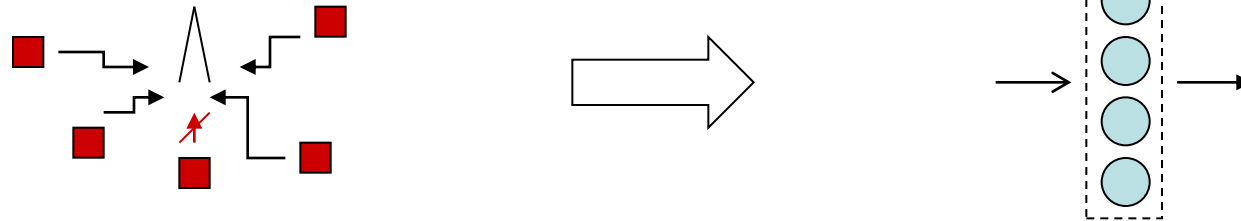
- Packet transmission at the output link of a large IP router



- Number of servers: 1
- Buffer capacity: max. number of IP packets
- Order of service: FIFO
- Arrivals: IP packet multiplexed at the output buffer
- Services: transmission of one IP packet  
(service time = transmission time = packet length / link transmission rate)

# Examples in details

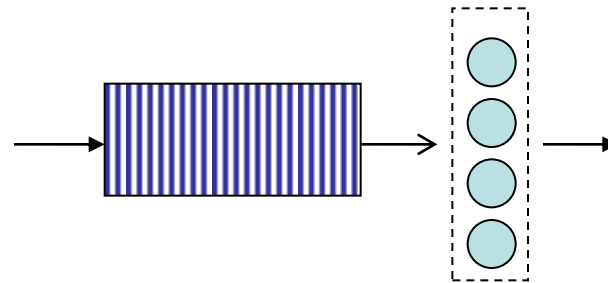
- Voice calls in a GSM cell
  - channels for parallel calls, each call occupies a channel
  - if all channels are busy the call is blocked



- Number of servers: number of parallel channels
- Buffer capacity: no buffer
- Order of service: does not apply
- Arrivals: call attempts in the GSM cell
- Service: the phone call (service time = length of the phone call)

# Group work

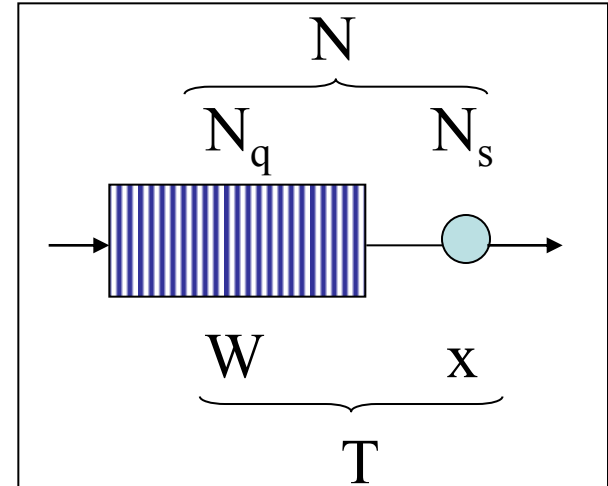
- Service at a bank, with “queue numbers” and several clerks
- Draw the block diagram of the queuing systems



- Arrivals:
- Service:
- Number of servers:
- Buffer capacity:
- Order of service:

# Performance measures

- Number of customers in the system ( $N$ )
  - Number of customers in the queue ( $N_q$ )
  - Number of customers in the server ( $N_s$ )
- System time ( $T$ )
  - Waiting time of a customer ( $W$ )
  - Service time of a customer ( $x$ )
- Probability of blocking (blocked customers / all arrivals)
- Utilization of the server (time server occupied / all considered time)
- **Transient measures**
  - how will the system state change in the near future?
- **Stationary measures**
  - how does the system behave on the long run?
  - average measures
  - often considered in this course



# Lecture 1

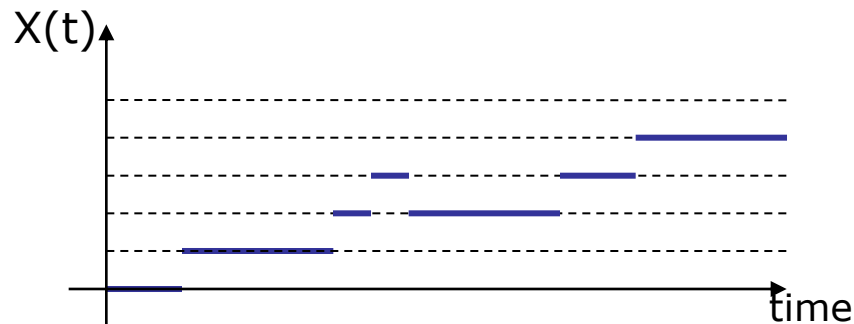
## Queuing systems - introduction

- Teletraffic examples and the performance triangle
- The queuing model
  - System parameters
  - Performance measures
- Stochastic processes recall

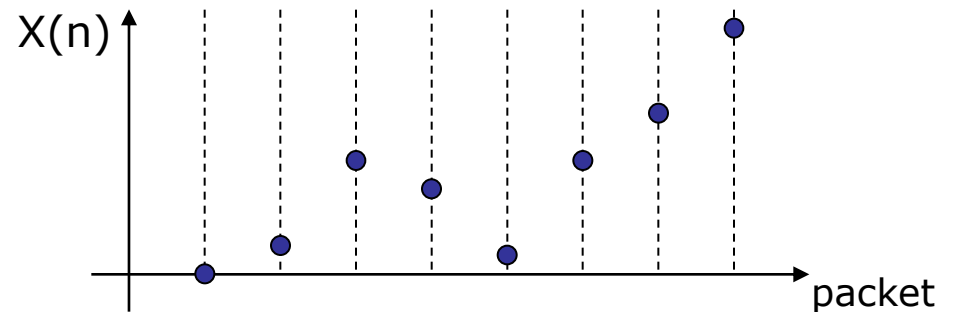


# Stochastic process

- Stochastic process
  - A system that evolves – changes its state - in time in a random way
  - Family of random variables
  - Variables indexed by a time parameter
    - Continuous time:  $X(t)$ , a random variable for each value of  $t$
    - Discrete time:  $X(n)$ , a random variable for each step  $n=0,1,\dots$
  - State space: the set of possible values of r.v.  $X(t)$  (or  $X(n)$ )
    - Continuous or discrete state



- Number of packets waiting:
  - Discrete space
  - Continuous time



- Waiting time of consecutive packets:
  - Discrete time
  - Continuous space

# Stochastic process - statistics

- We are interested in quantities, like:
  - time dependent (transient) state probabilities (statistics over many realizations, an *ensemble* of realizations, *ensemble average* ):

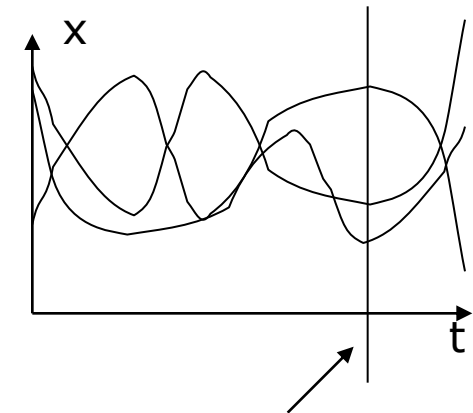
$$f_x(t) = P(X(t) = x), \quad F_x(t) = P(X(t) \leq x)$$

- $n^{\text{th}}$  order statistics – joint distribution over  $n$  samples

$$F_{x_1, \dots, x_n}(t_1, \dots, t_n) = P(X(t_1) \leq x_1, \dots, X(t_n) \leq x_n)$$

- limiting (or stationary) state probabilities (if exist) :

$$f_x = \lim_{t \rightarrow \infty} P(X(t) = x), \quad F_x = \lim_{t \rightarrow \infty} P\{X(t) \leq x\}$$



ensemble average

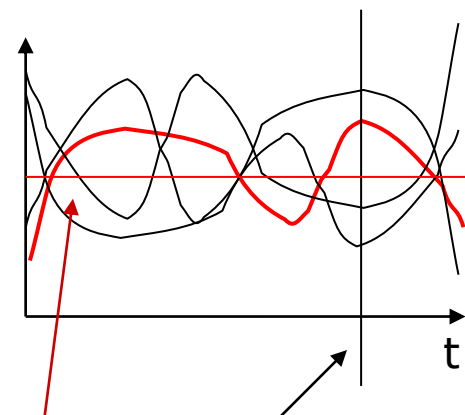
# Stochastic process - terminology

- The stochastic process is:
  - **stationary**, if all  $n^{\text{th}}$  order statistics are unchanged by a shift in time:

$$F_x(t + \tau) = F_x(t), \quad \forall t$$

$$F_{x_1, \dots, x_n}(t_1 + \tau, \dots, t_n + \tau) = F_{x_1, \dots, x_n}(t_1, \dots, t_n), \quad \forall n, \quad \forall t_1, \dots, t_n$$

- **ergodic**, if the ensemble average is equal to the time average of a single realization
- consequence: if a process ergodic, then the statistics of the process can be determined from a single (infinitely long) realization and vice versa

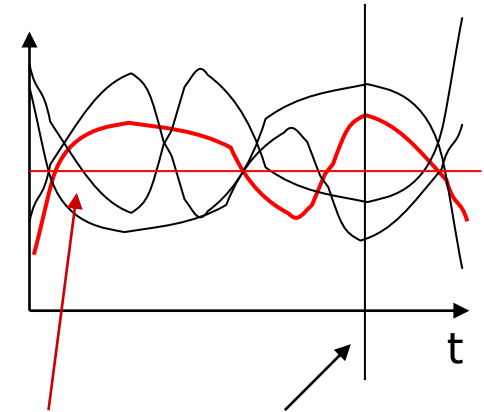


time average

ensemble average

# Stochastic process

- Example on stationary versus ergodic
- Consider a source, that generates the following sequences with the same probability:
  - ABABABAB...
  - BABABABA...
  - EEEEEEEEE...
- Is this source stationary?
- Is this source ergodic?



time average

ensemble average

# Summary

Today:

- Queuing systems - definition and parameters
- Stochastic processes

Next lecture:

- Poisson processes and Markov-chains, the theoretical background to analyze queuing systems

Recitation:

- Probability theory and transforms
- **Prepare for the recitation: read Virtamo 1-3 in the course binder or download from the course web**
  - Definition of probability of events
  - Conditional probability, law of total probability, Bayes formula, independent events
  - Random variables, distribution functions (discrete and continuous)
  - Z and Laplace transforms