Business Process Management

Paolo Bottoni

Dipartimento di Informatica

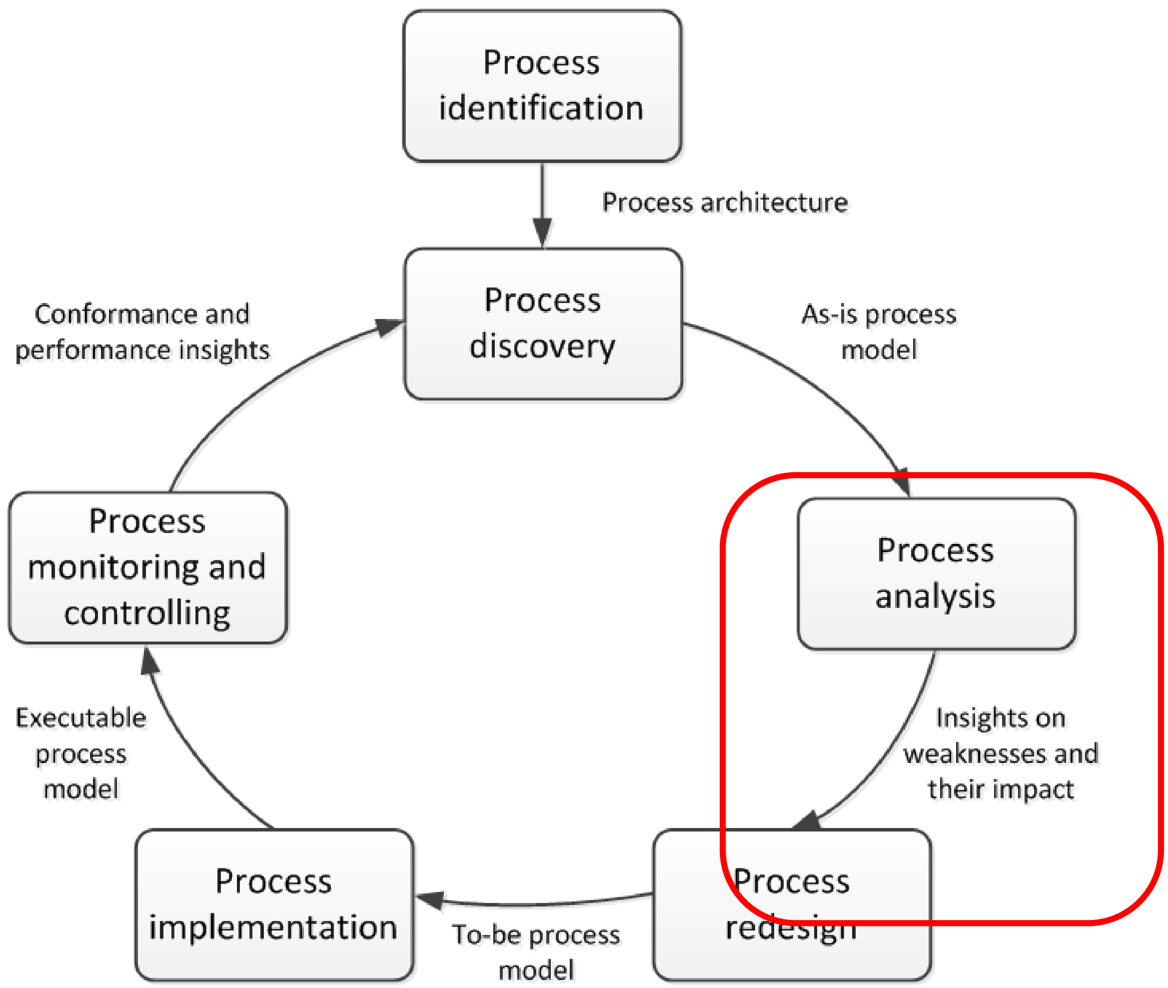
Sapienza Università di Roma

Lecture 9: Quantitative Analysis (Queueing and Simulation)

Adapted from the slides for the book :
Dumas, La Rosa, Mendling & Reijers: Fundamentals of Business Process Management, Springer 2013

http://courses.cs.ut.ee/2013/bpm/uploads/Main/ITlecture6.ppt

Business Process Analysis



Process Analysis Techniques

Qualitative analysis

- Value-Added Analysis
- Root-Cause Analysis
- Pareto Analysis
- Issue Register

Quantitative Analysis

- Quantitative Flow Analysis
- Queuing Theory
- Process Simulation

Why is flow analysis not enough?

Flow analysis does not consider waiting times due to resource contention

Queuing analysis and simulation address these limitations and have a broader applicability

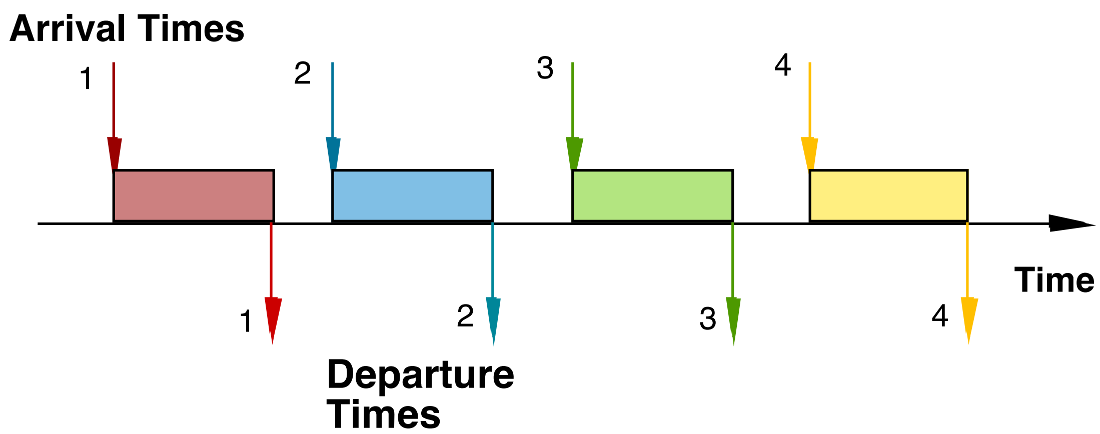
Why is Queuing Analysis Important?

- Capacity problems are very common in industry and one of the main drivers of process redesign
  - Need to balance the cost of increased capacity against the gains of increased productivity and service
- Queuing and waiting time analysis is particularly important in service systems
  - Large costs of waiting and of lost sales due to waiting

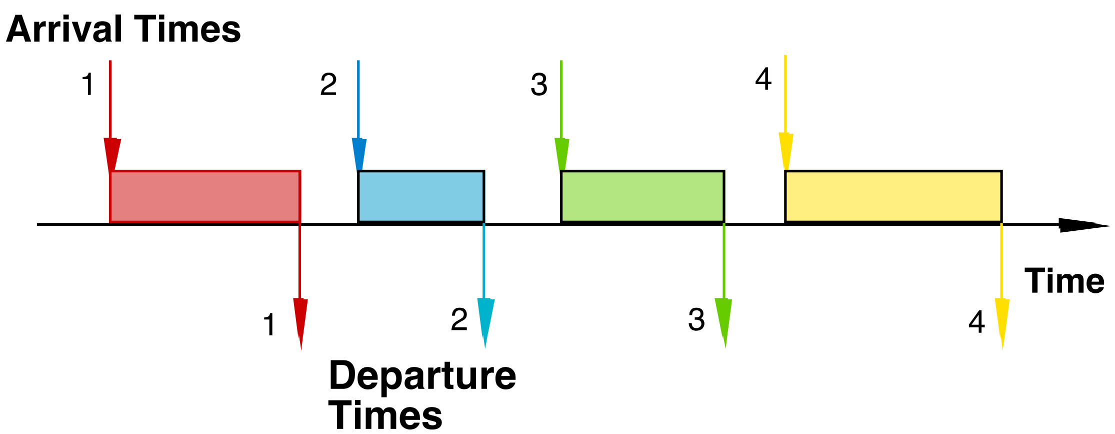
Prototype Example – ER at a Hospital

- Patients arrive by ambulance or by their own accord
- One doctor is always on duty
- More patients seek help ⇒ longer waiting times
- ➢ ***Question***: ***Should another MD position be instated?***

Delay is Caused by Job Interference



Deterministic traffic

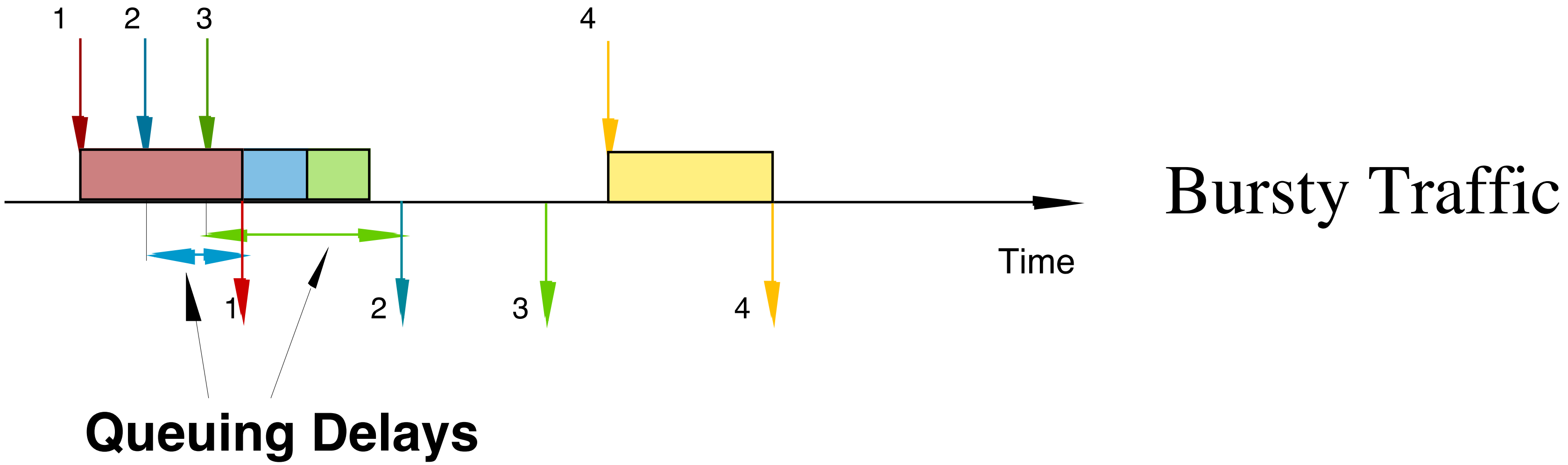


Variable but spaced apart traffic

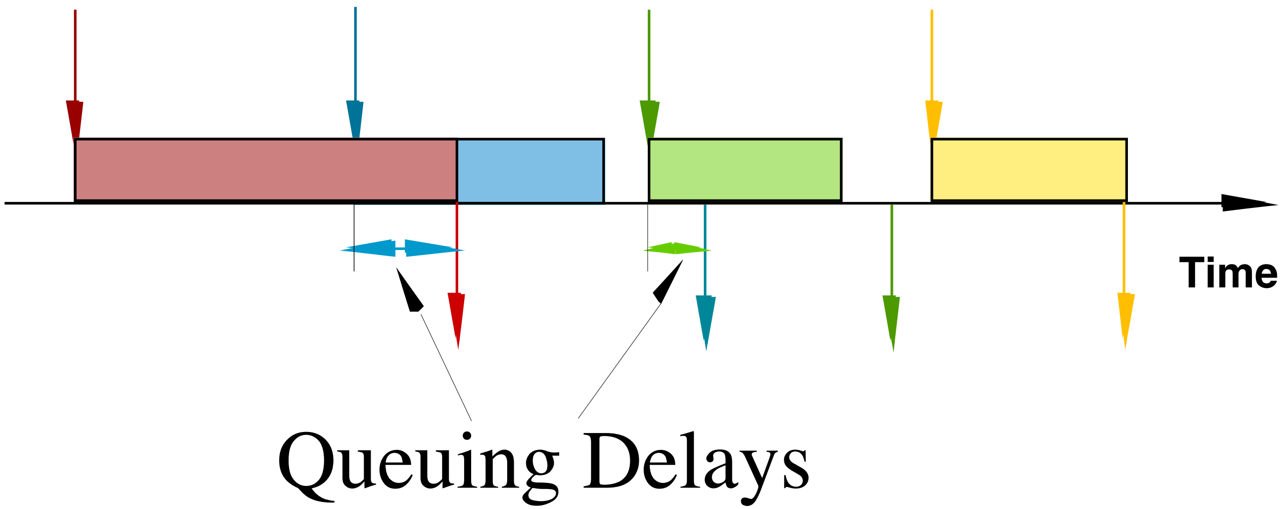
- If arrivals are regular or sufficiently spaced apart, no queuing delay occurs

Burstiness Causes Interference

*

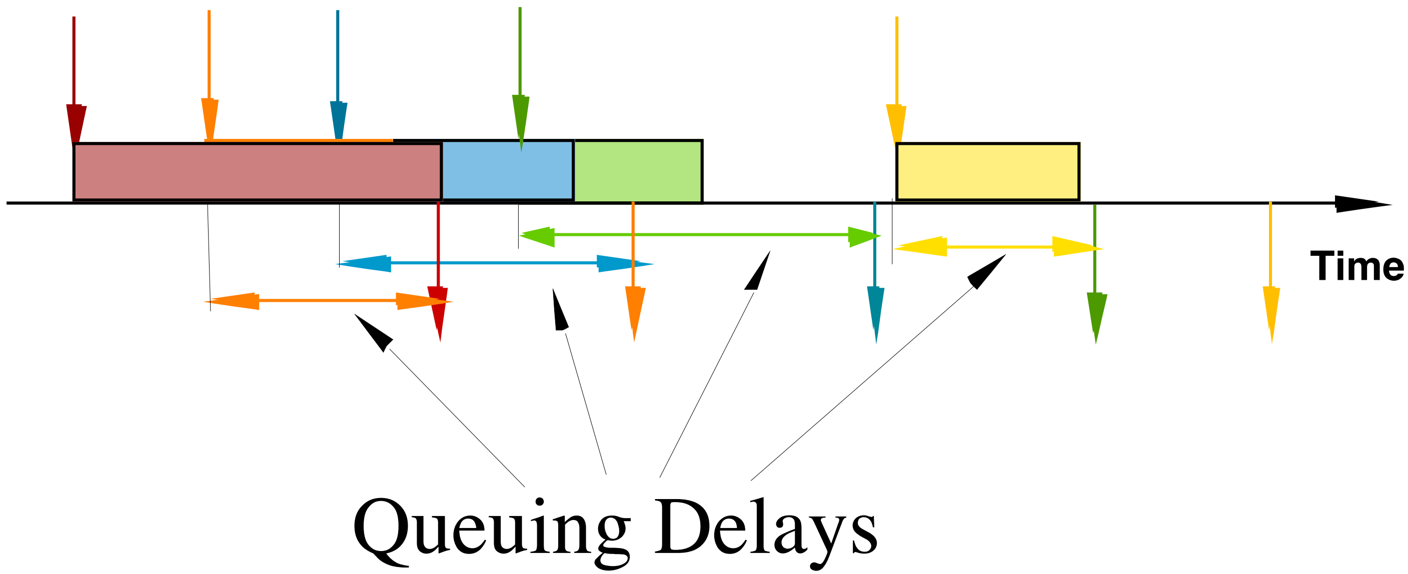


Job Size Variation Causes Interference



- Deterministic arrivals, variable job sizes

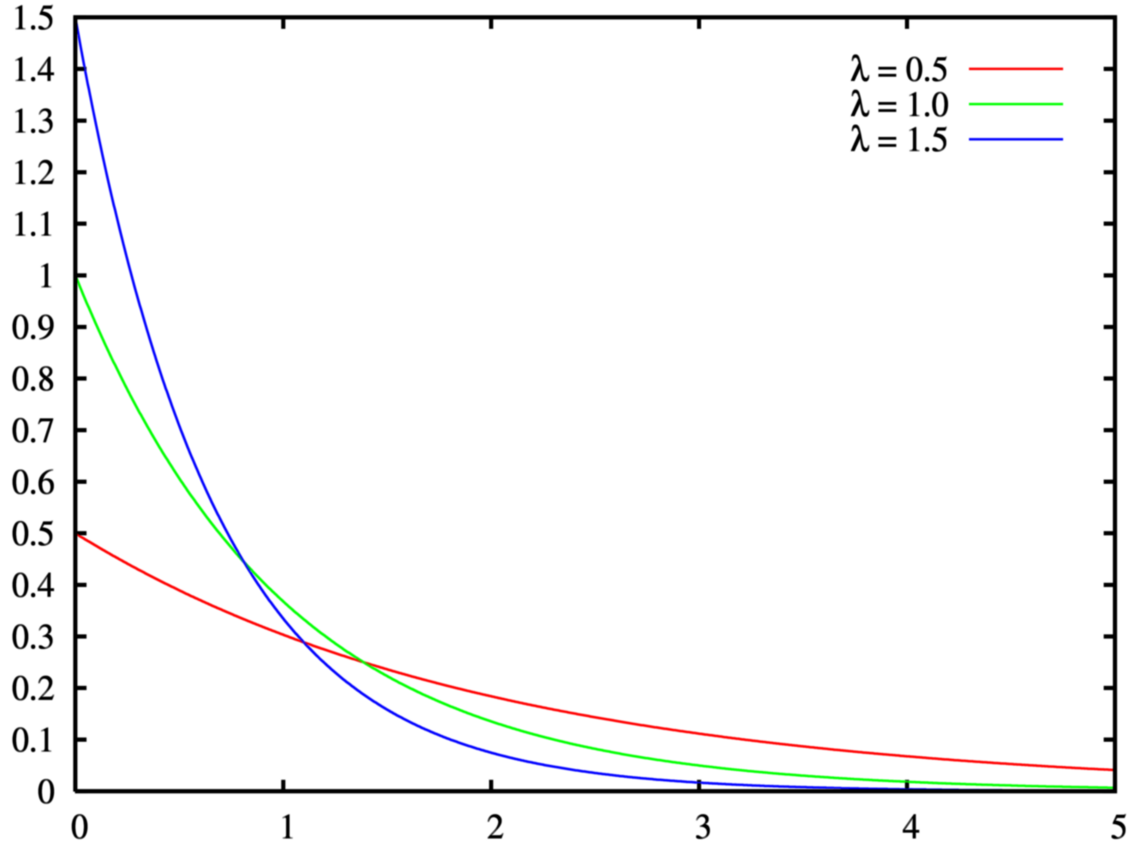
High Utilization Exacerbates Interference



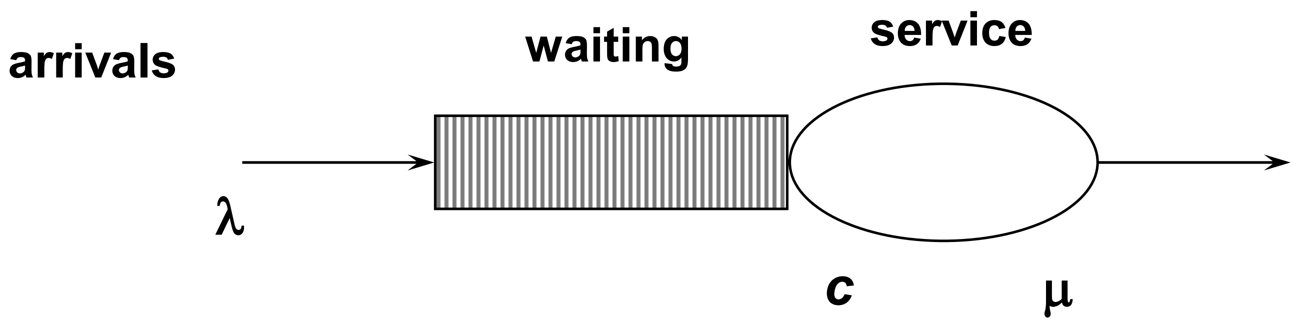
The Poisson Process

- Common arrival assumption in many queuing and simulation models
- The times between arrivals are independent, identically distributed and exponential
  - P (arrival < t) = 1 – $e^{-\lambda t}$
- Key property: The fact that a certain event has not happened tells us nothing about how long it will take before it happens
  - e.g., P(X > 40 | X >= 30) = P (X > 10)

Negative Exponential Distribution



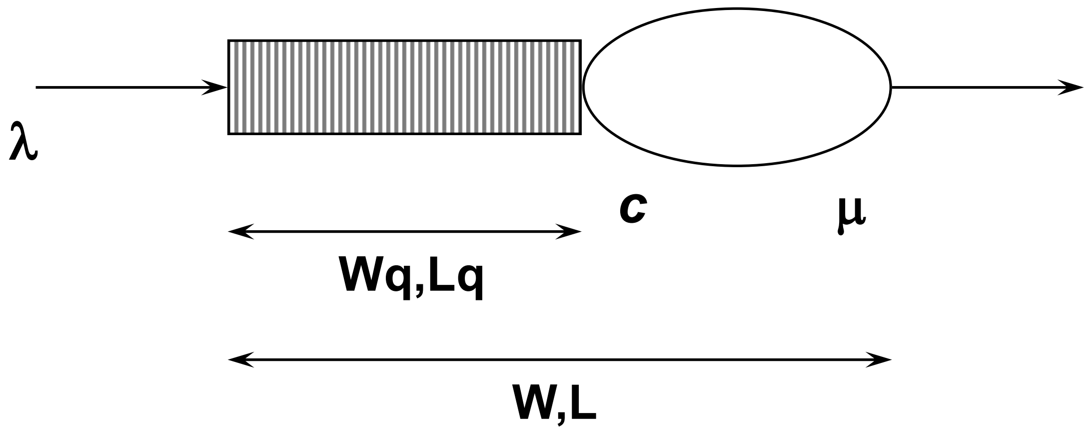
Queuing theory: basic concepts



Basic characteristics:

- λ (mean arrival rate) = average number of arrivals per time unit
- μ (mean service rate) = average number of jobs that can be handled by one server per time unit:
- c = number of servers

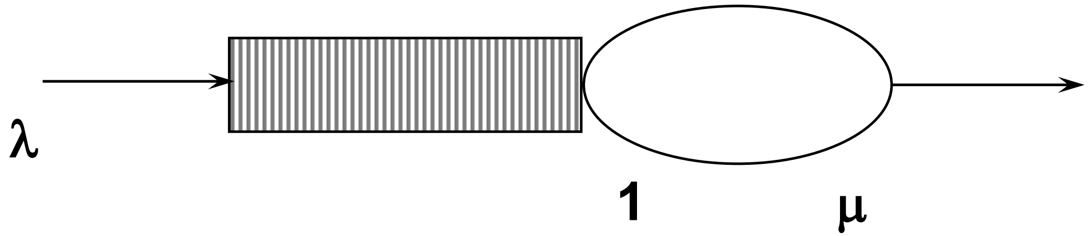
Queuing theory concepts (cont.)



Given λ , μ and c, we can calculate :

- **occupation rate: ρ**
- **Wq = average time spent in queue**
- **W = average time spent in system (i.e. cycle time)**
- **Lq = average number of jobs in queue (i.e. length of queue)**
- **L = average number of jobs in system (i.e. Work-in-Progress)**

M/M/1 queue



Assumptions:

- **time between arrivals and service time follow a negative exponential distribution**
- **1 server (*c* = 1)**
- **FIFO**

$$\rho = \frac{\text{Capacity Demand}}{\text{Available Capacity}} = \frac{\lambda}{\mu}$$

$L=\rho/(1-\rho)$

$W=L/\lambda=1/(\mu-\lambda)$

$L_q=\rho^2/(1-\rho) = L-\rho$

$W_q=L_q/\lambda=\lambda/(\mu(\mu-\lambda))$

M/M/c queue

- **Now there are c servers in parallel, so the expected capacity per time unit is then $c^*\mu$**

$$\rho = \frac{\text{Capacity Demand}}{\text{Available Capacity}} = \frac{\lambda}{c * \mu}$$

Little's Formula $\Rightarrow$ $W_q = L_q / \lambda$

$W = W_q + (1/\mu)$

Little's Formula $\Rightarrow$ $L = \lambda W$

Tool Support

- For M/M/c systems, the exact computation of L_q is rather complex…

$$L_q = \sum_{n=c}^{\infty}(n-c)P_n = \ldots = \frac{(\lambda/\mu)^c \rho}{c!(1-\rho)^2} P_0$$

$$P_0 = \left(\sum_{n=0}^{c-1} \frac{(\lambda/\mu)^n}{n!} + \frac{(\lambda/\mu)^c}{c!} \cdot \frac{1}{1-(\lambda/(c\mu))} \right)^{-1}$$

- Consider using a tool, e.g.
  - http://apps.business.ualberta.ca/aingolfsson/qtp/
  - http://www.stat.auckland.ac.nz/~stats255/qsim/qsim.html

Example – ER at County Hospital

- Situation
  - Patients arrive according to a Poisson process with intensity λ ($\Leftrightarrow$ the time between arrivals is exp(λ) distributed.
  - The service time (the doctor's examination and treatment time of a patient) follows an exponential distribution with mean $1/\mu$ (=exp(μ) distributed)
  - $\Rightarrow$ *The ER can be modeled as an M/M/c system where c=the number of doctors*
- **Data gathering**
  - $\Rightarrow$ **λ = 2 patients per hour**
  - $\Rightarrow$ **μ = 3 patients per hour**
- **Question**
  - **Should the capacity be increased from 1 to 2 doctors?**

Queuing Analysis – Hospital Scenario

- Interpretation
  - To be in the queue = to be in the waiting room
  - To be in the system = to be in the ER (waiting or under treatment)

Characteristic	One doctor (c=1)	Two Doctors (c=2)
ρ	2/3	1/3
L_q	4/3 patients	1/12 patients
L	2 patients	3/4 patients
W_q	2/3 h = 40 minutes	1/24 h = 2.5 minutes
W	1 h	3/8 h = 22.5 minutes

- Is it warranted to hire a second doctor ?

Process Simulation

- Drawbacks of queuing theory:
  - Generally not applicable when system includes parallel activities
  - Requires case-by-case mathematical analysis
  - Assumes "steady-state" (valid only for "long-term" analysis)
- Process simulation is more versatile (also more popular)
- Process simulation = run a large number of process instances, gather data (cost, duration, resource usage) and calculate statistics from the output

Process Simulation

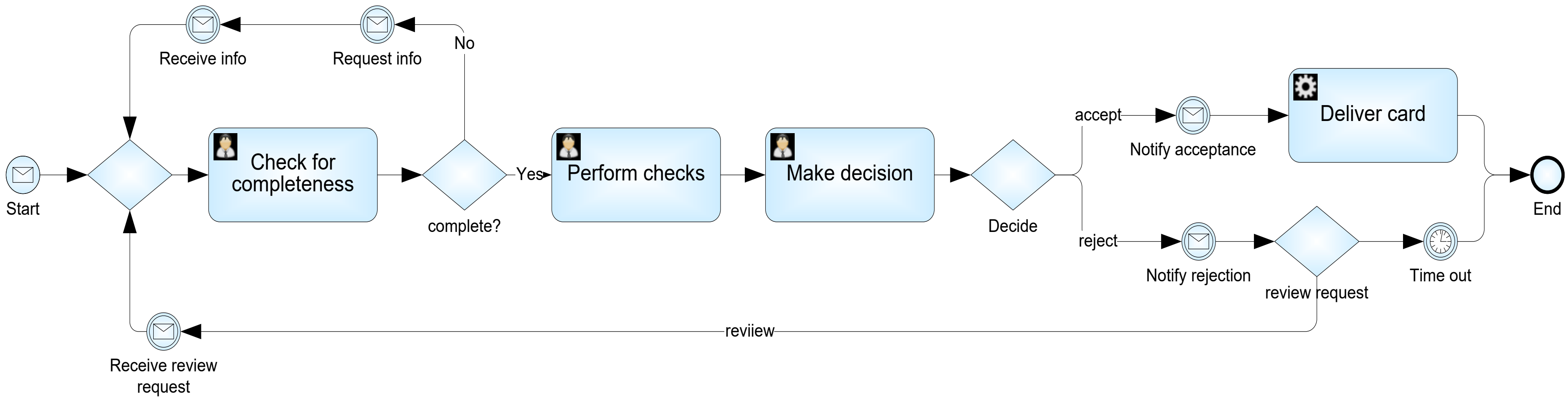
Steps in evaluating a process with simulation

1. Model the process (e.g. BPMN)
2. Enhance the process model with simulation info ➔ simulation model
   - Based on assumptions or better based on data (logs)
3. Run the simulation
4. Analyze the simulation outputs
   1. Process duration and cost stats and histograms
   2. Waiting times (per activity)
   3. Resource utilization (per resource)
5. Repeat for alternative scenarios

Elements of a simulation model

- The process model including:
  - Events, activities, control-flow relations (flows, gateways)
  - Resource classes (i.e. lanes)
- *Resource assignment*
  - Mapping from activities to resource classes
- Processing times
  - Per activity or per activity-resource pair
- Costs
  - Per activity and/or per activity-resource pair
- Arrival rate of process instances
- Conditional branching probabilities (XOR gateways)

Simulation Example – BPMN model



Resource Pools (Roles)

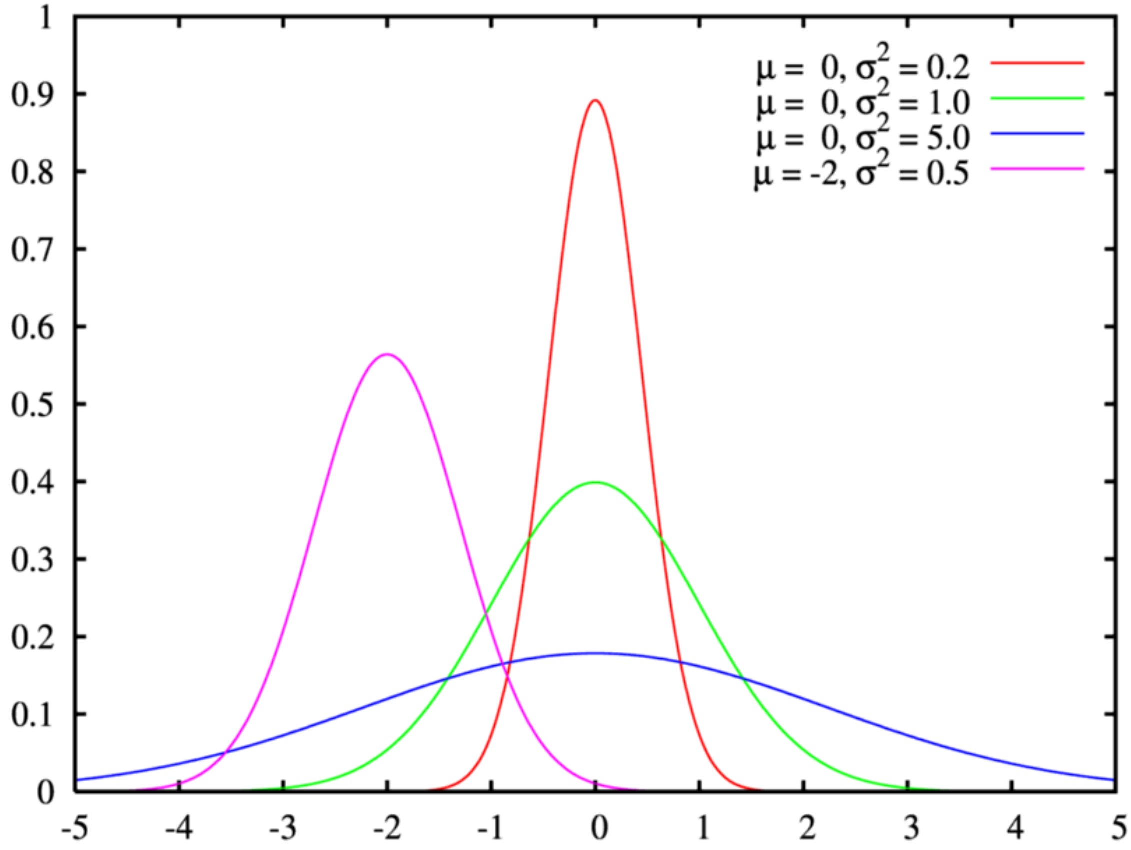
- Two options to define resource pools
  - Define individual resources of type clerk
  - Or assign a number of “anonymous” resources all with the same cost
- E.g.
  - 3 anonymous clerks with cost of € 10 per hour, 8 hours per day
  - 2 individually named clerks
    - Jim: € 12.4 per hour
    - Mike: € 14.8 per hour
  - 1 manager John at € 20 per hour, 8 hours per day

Resource pools and execution times

Task	Role	Execution Time Normal distribution: mean and std deviation	
Receive application	system	0	0
Check completeness	Clerk	30 mins	10 mins
Perform checks	Clerk	2 hours	1 hour
Request info	system	1 min	0
Receive info (Event)	system	48 hours	24 hours
Make decision	Manager	1 hour	30 mins
Notify rejection	system	1 min	0
Time out (Time)	system	72 hours	0
Receive review request (Event)	system	48 hours	12 hours
Notify acceptance	system	1 min	0
Deliver Credit card	system	1 hour	0

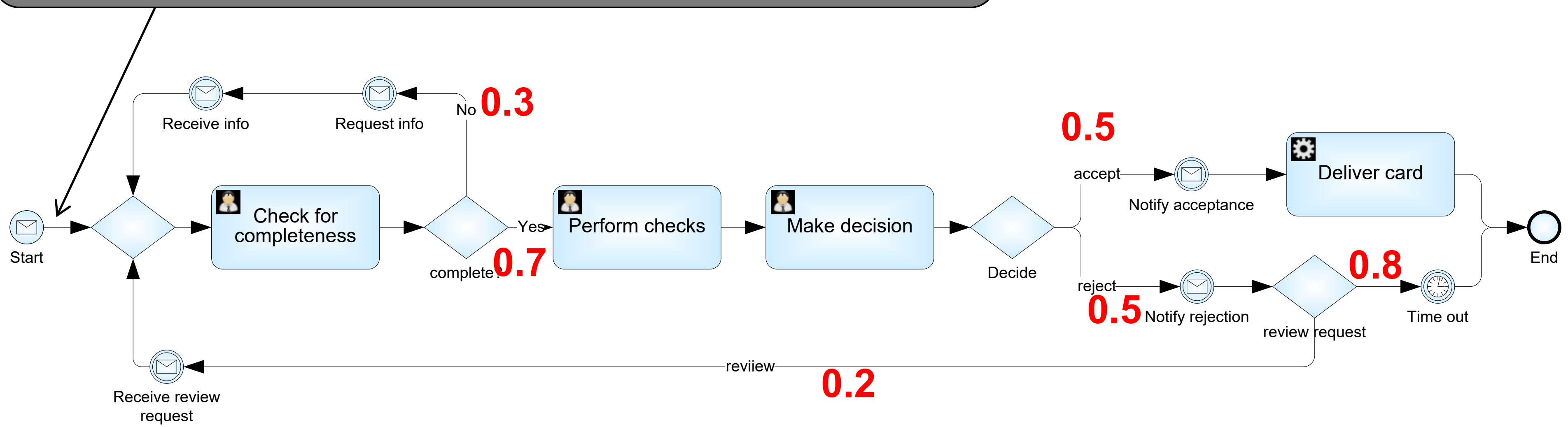
Alternative: assign execution times to the tasks only (like in cycle time analysis)

Reminder: Normal Distribution



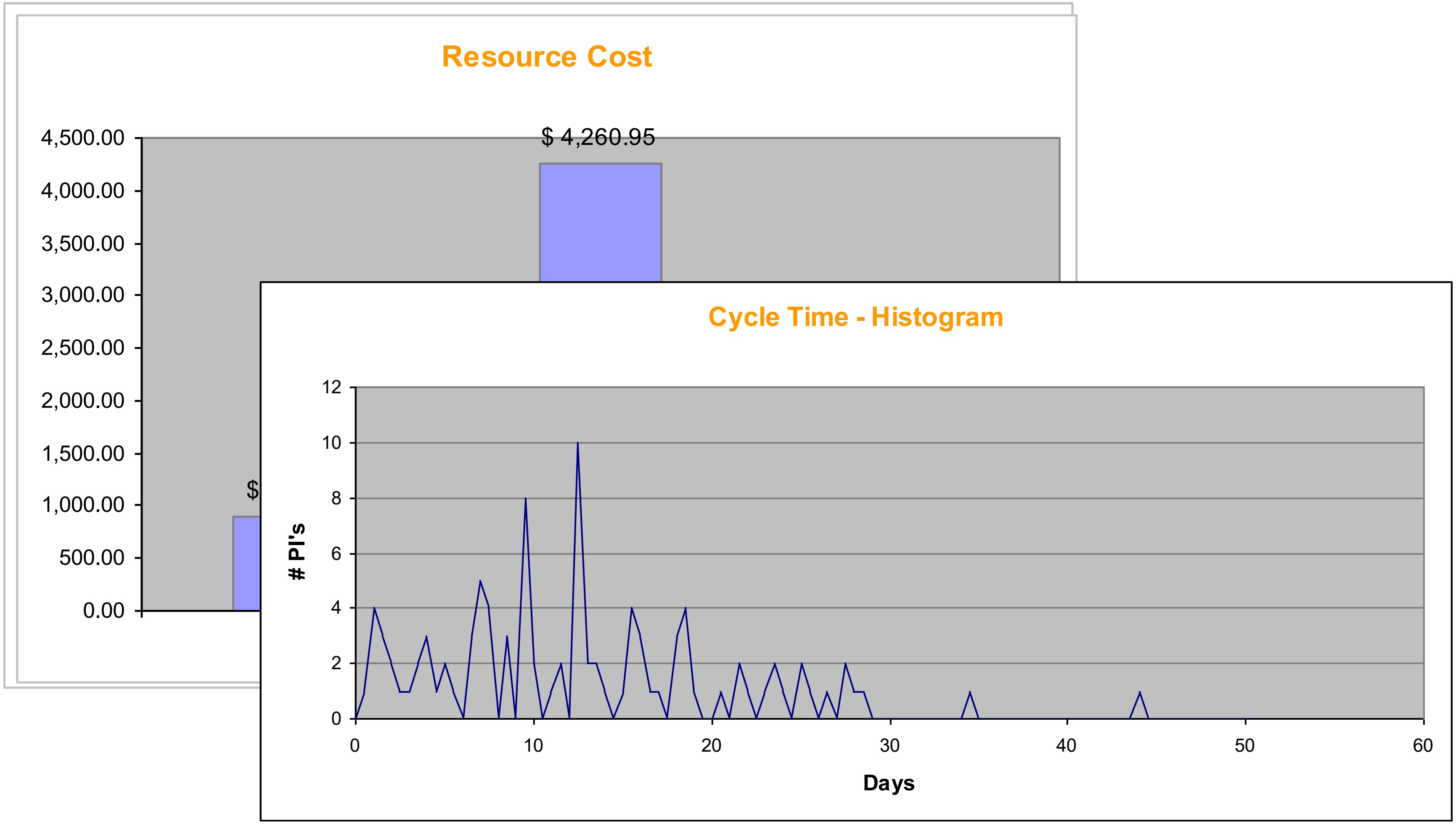
Arrival rate and branching probabilities

10 applications per hour (one at a time)
Poisson arrival process (negative exponential)



Alternative: instead of branching probabilities one can assign "conditional expressions" to the branches based on input data

Simulation output: KPIs



Simulation output: detailed logs

Process Instance	# Activities	Start	End	Cycle Time	Cycle Time (s)	Total Time
6	5	4/06/2007 13:00	4/06/2007 16:26	03:26:44	12403.586	03:26:44
11	5	4/06/2007 18:00	5/06/2007 12:14	18:14:56	65695.612	18:14:56
13	5	4/06/2007 20:00	5/06/2007 13:14	17:14:56	62095.612	17:14:56
16	5	4/06/2007 23:00	5/06/2007 15:06	16:06:29	57989.23	16:06:29
22	5	5/06/2007 5:00	6/06/2007 10:01	29:01:39	104498.797	29:01:39
27	8	5/06/2007 10:00	6/06/2007 12:33	26:33:21	95600.649	26:33:21

Process Instance	Activity ID	Activity Name	Activity Type	Resource	Start	End
	6aed54717-f044-4da1-b543-82a660809ecb	Check for completeness	Task	Manager	4/06/2007 13:00	4/06/2007 13:53
	6a270f5c6-7e16-42c1-bfc4-dd10ce8dc835	Perform checks	Task	Clerk	4/06/2007 13:53	4/06/2007 15:25
	677511d7c-1eda-40ea-ac7d-886fa03de15b	Make decision	Task	Manager	4/06/2007 15:25	4/06/2007 15:26
	6099a64eb-1865-4888-86e6-e7de36d248c2	Notify acceptance	IntermediateEvent	(none)	4/06/2007 15:26	4/06/2007 15:26
	60a72cf69-5425-4f31-8c7e-6d093429ab04	Deliver card	Task	System	4/06/2007 15:26	4/06/2007 16:26
	7aed54717-f044-4da1-b543-82a660809ecb	Check for completeness	Task	Manager	4/06/2007 14:00	4/06/2007 14:31
	7a270f5c6-7e16-42c1-bfc4-dd10ce8dc835	Perform checks	Task	Clerk	4/06/2007 14:31	5/06/2007 8:30

Tools for Process Simulation

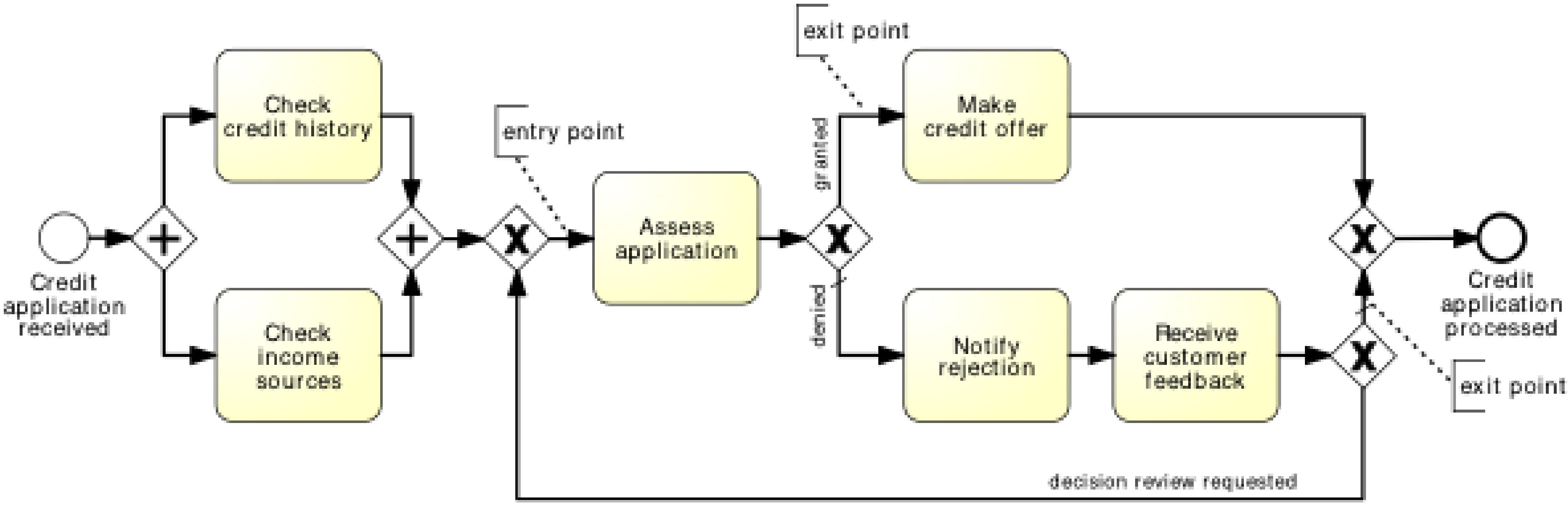
Listed in no specific order:

- ITP Commerce Process Modeler for Visio
  - Models presented earlier are made with ITP Commerce
- Progress Savvion Process Modeler
- IBM Websphere Business Modeler
- Oracle BPA
- ARIS
- ProSim

Simple Online Simulator

- BIMP: http://bimp.cs.ut.ee/
- Accepts standard BPMN 2.0 as input
- Link from Signavio Academic Edition to BIMP
  - Open a model in Signavio and push it to BIMP using the flask icon

BIMP Demo



Credit application received
Check credit history
Check income sources
entry point
Assess application
exit point
granted
denied
Make credit offer
Notify rejection
Receive customer feedback
exit point
decision review requested
Credit application processed