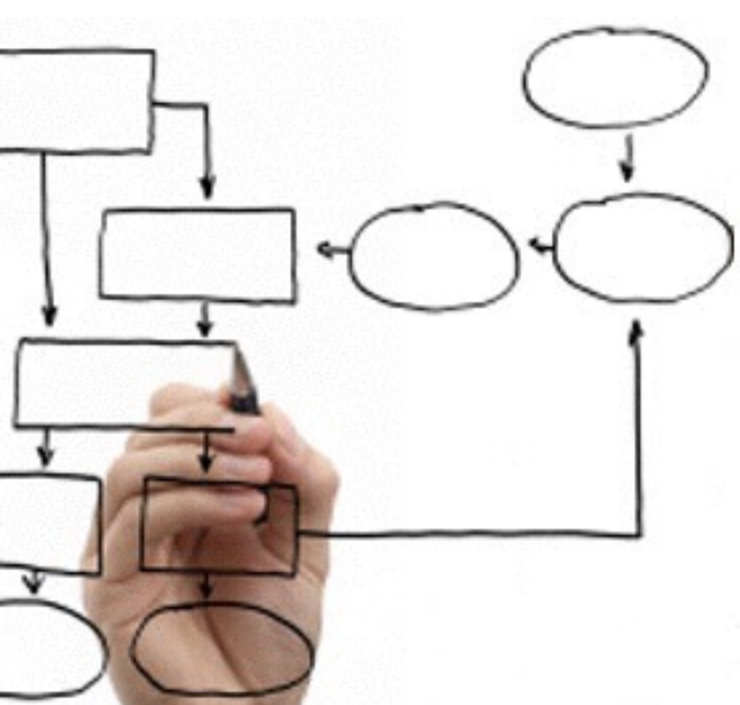
Methods for the specification and verification of business processes

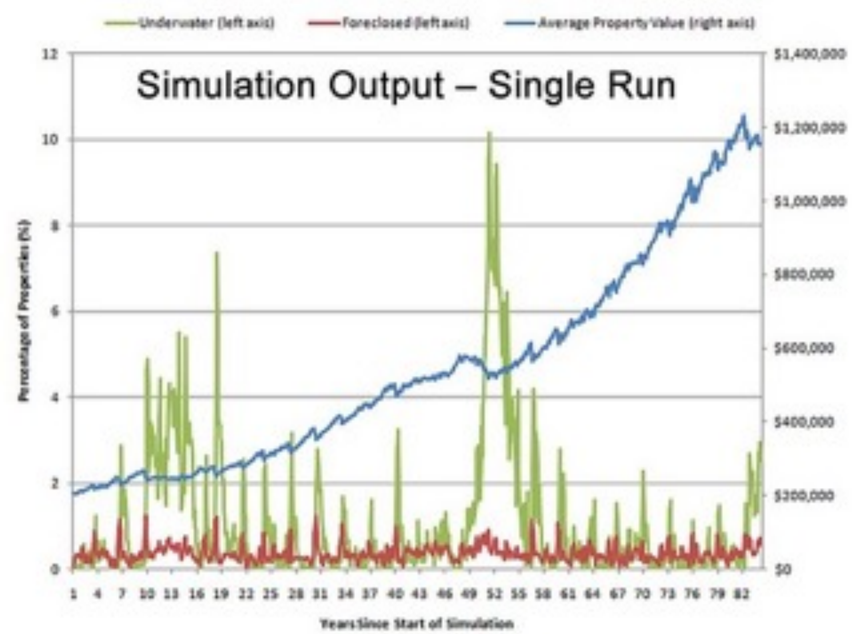
MPB (6 cfu, 295AA)

Roberto Bruni
http://www.di.unipi.it/~bruni



23 - Simulation

Object



We overview some principles for the quantitative simulation of business processes

Ch.7 of Fundamental of Business Process Management. M. Dumas et al.
(inspired by slides available at https://courses.cs.ut.ee/2014/bpm/)

Simulation

Process simulation is the most popular and widely supported technique for quantitative analysis of process models.

It is a very flexible analysis technique.
It is applicable to almost any workflow.
It is accessible to people without mathematical background.

It boils down to **computer aided repeated execution** of paths in the reachability graph:
A large number of hypothetical instances of the process are generated and executed step-by-step;
the produced output can include logs as well as statistics about (average) cycle/waiting times and resource utilization

System: parameters

For each **task**:
the probability distribution for the processing time
other performance attributes (e.g., cost, value)
the set of resources able to perform the task

For each **(XOR) split**:
branching probability of every outgoing arc

For each **resource class/group**:
the number of resources in the pool,
other performance attributes (e.g., hourly cost)

Task duration: fixed

Fixed:
the processing time is relatively constant

rare when humans are involved,
common for automated tasks
(e.g., automatic report generation)

Task duration

(Negative) Exponential distribution:
applicable when the processing time is most often around a **mean value**, but sometimes considerably longer
(e.g., assess insurance claim)

parameter: rate λ

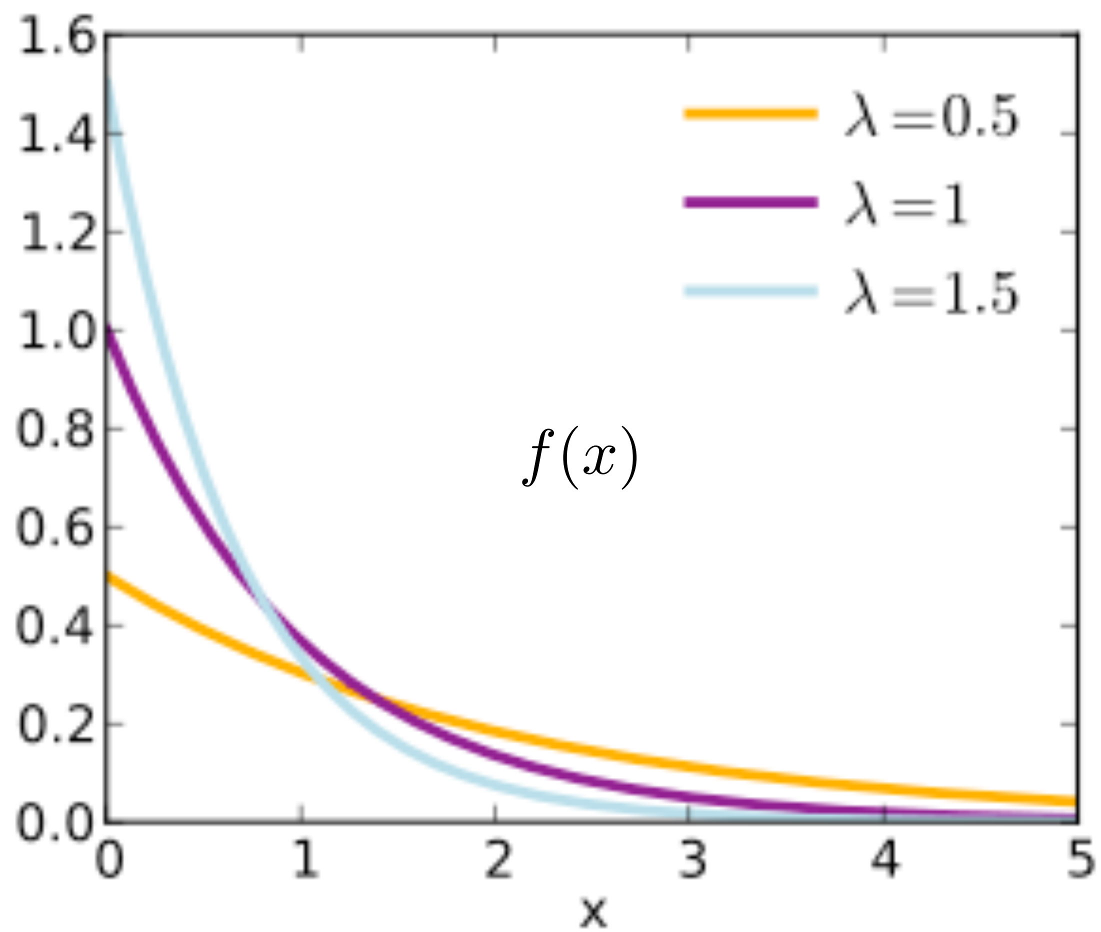
probability density function: $f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$

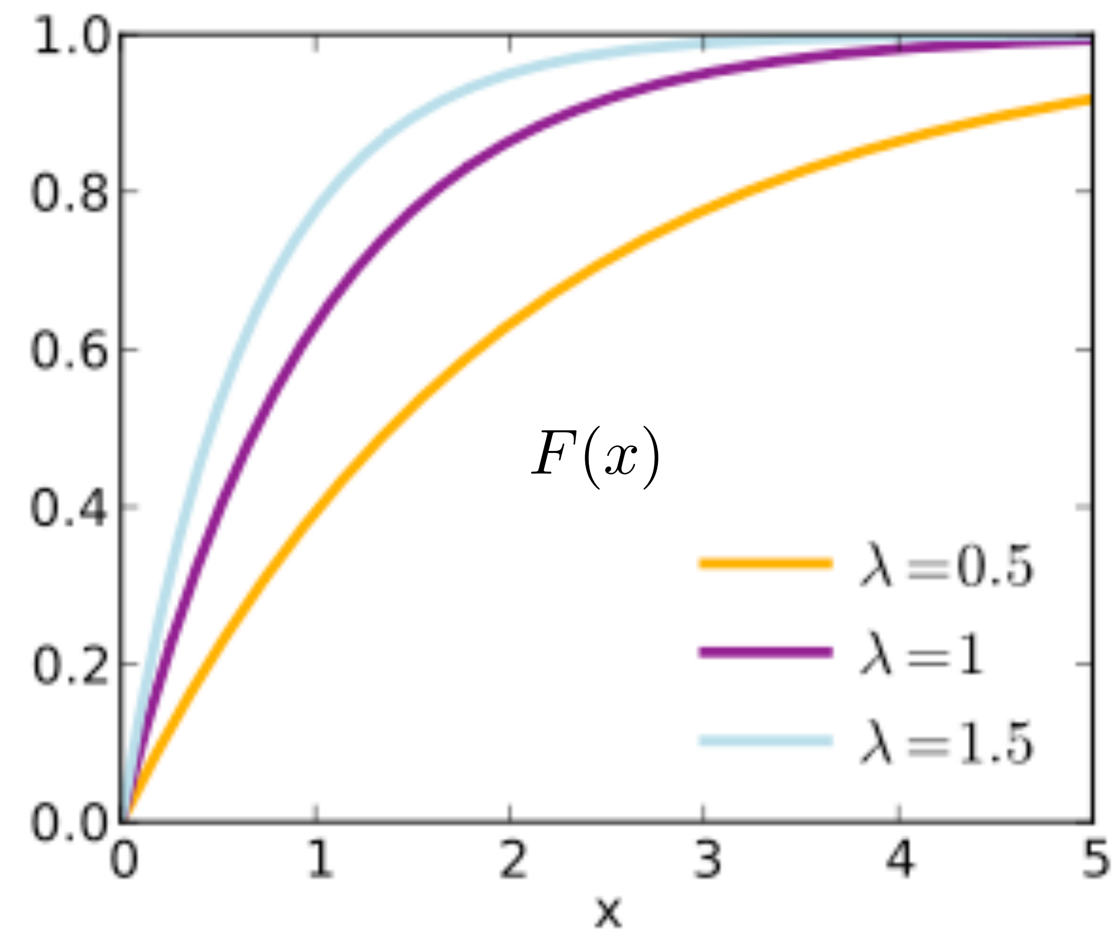
cumulative distribution function: $F(x) = \begin{cases} 1 - e^{\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$

mean value: $\frac{1}{\lambda}$

Task duration

Exponential distribution:





Task duration

Normal distribution:
applicable when the processing time is around a given **average** and the **deviation** around this value is symmetric (e.g., paper-form check)

parameters: mean value $\boldsymbol{\mu}$, standard deviation $\boldsymbol{\sigma}$

probability density function: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

cumulative distribution function: $\Phi(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} \mathrm{d}t$

Task duration

Normal distribution:

applicable when the processing time is around a given **average** and the **deviation** around this value is symmetric (e.g., paper-form check)

parameters: mean value $\boldsymbol{\mu}$=0, standard deviation $\boldsymbol{\sigma}$=1

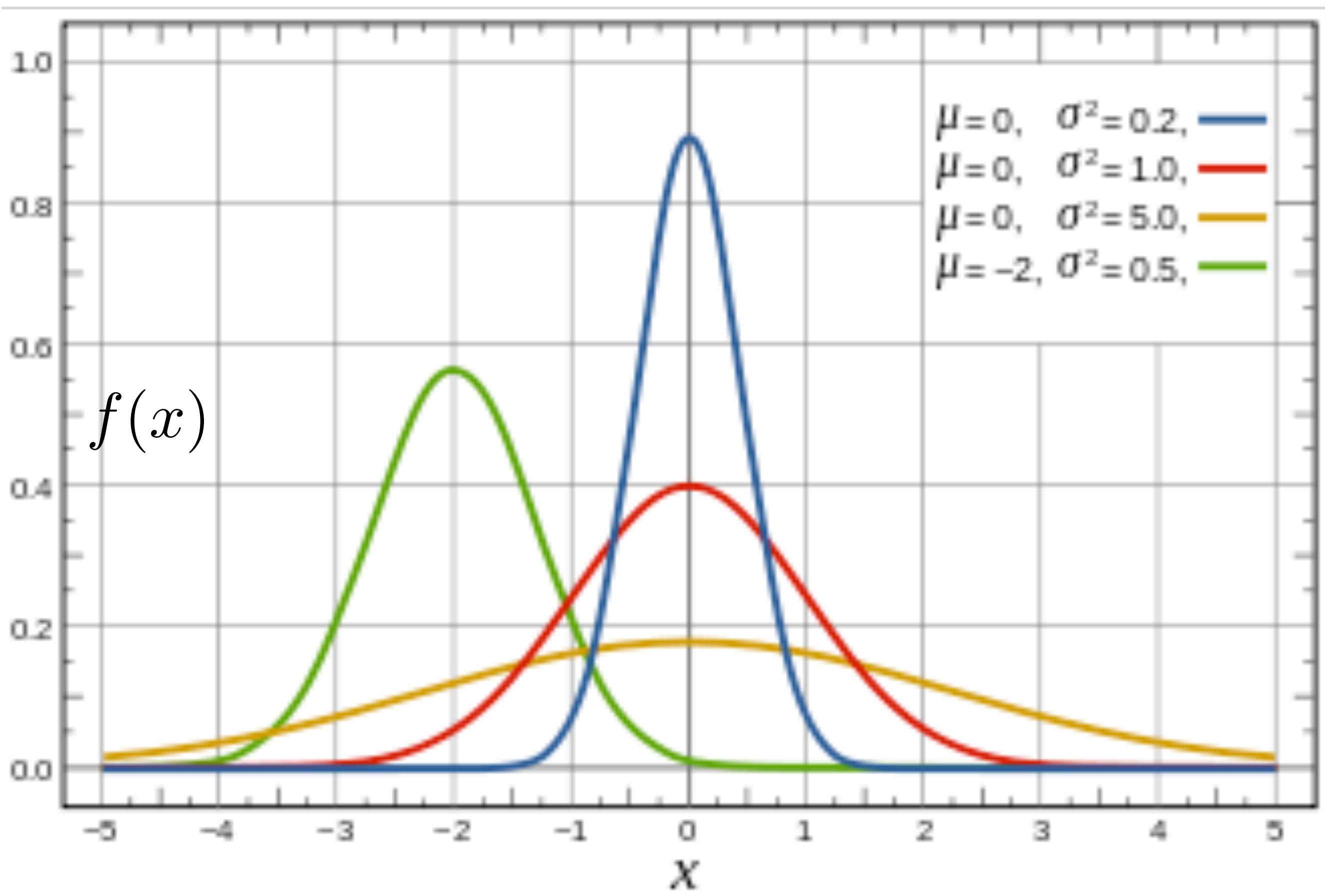
probability density function: $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

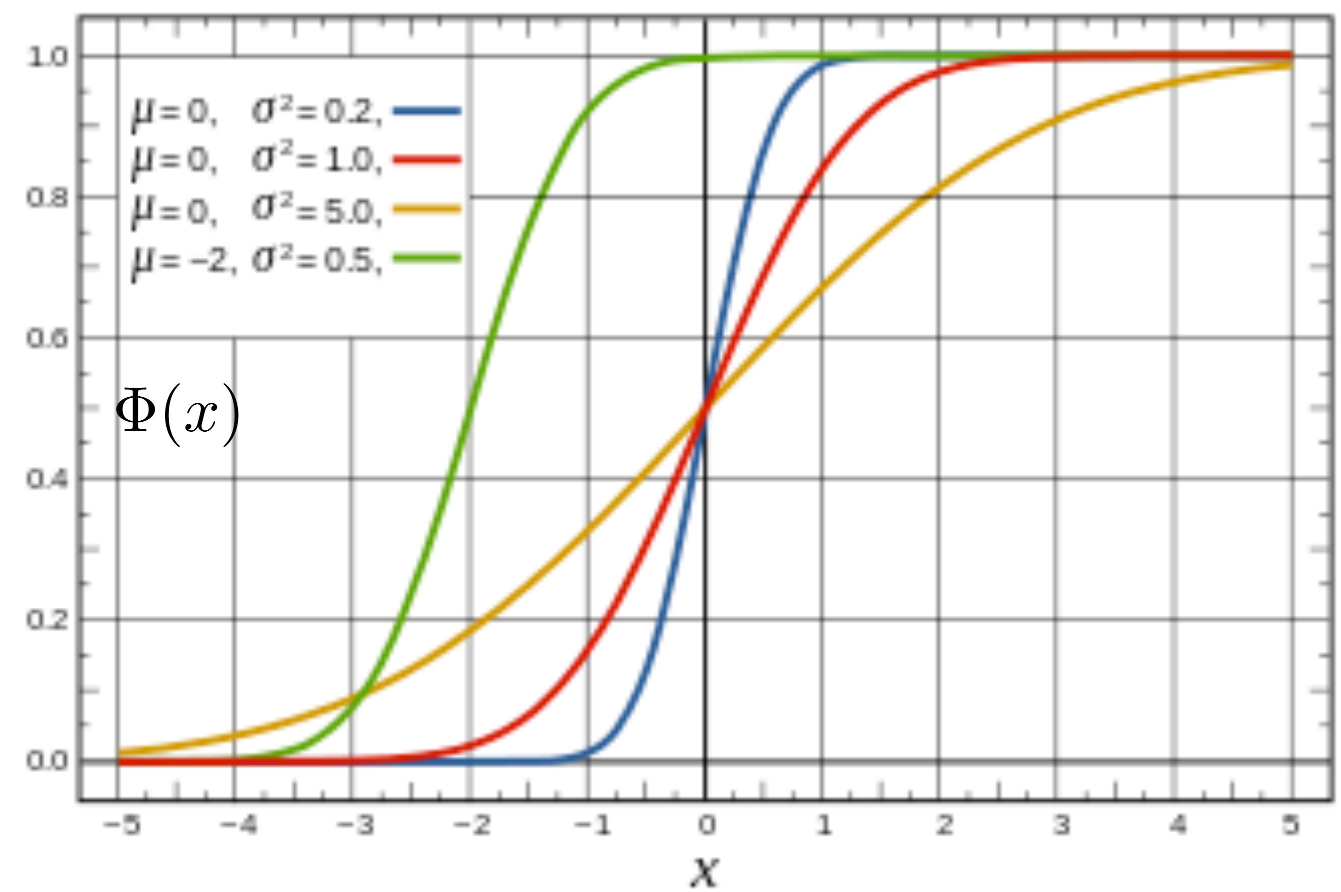
cumulative distribution function: $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} \mathrm{d}t$

(standard normal distribution $\boldsymbol{\mu}$=0, $\boldsymbol{\sigma}$=1)

Task duration

Normal distribution:





Task duration: estimate

How to estimate the parameters?

Informed guess: interviews with stakeholders

sampling: collect sample data from real executions

logs import: allowed by some tools
(this functionality is called **simulation input analysis**)

Simulation: other inputs

Inter-arrival times and mean arrival rate
(e.g., exponential distribution, normal distribution)

Starting date and time of the simulation

Either:
end date and time of the simulation; or
real-time duration of the simulation; or
number of process instances to be simulated.

Simulation: execution

During a process simulation, the tasks in the process are not actually executed

When a task is ready to be executed a so-called work item is created and the simulator first tries to find a resource to which it can assign this work item

If no resource is available the work item is put in waiting mode until a resource is freed up

Once a resource is assigned to a work item, the simulator determines the duration of the activity by drawing a random number according to a probability distribution

Simulation: execution

Once the duration has been determined, the work item is put in sleeping mode for that duration
(to simulate the execution of the task)

Once the time interval has passed (according to the simulation's clock), the work is completed and the resource that was assigned to it is available again

Simulators exploit smart algorithms to complete the simulation as fast as possible: they do not effectively wait for tasks to come back from their sleeping mode and do not wait for the whole activity duration either (thousands of process instance can be simulated in matter of seconds)

Simulation: execution

For each work item created during a simulation, the simulator records the identifier of the resource that was assigned to this instance as well as three time-stamps:

the time when the task was **ready** to be executed

the time when the it was (assigned to a resource and) **started**

the time when the task was **completed**

Thanks to the collected data, the simulator can compute the **average waiting time** for each task
(important to identify bottlenecks in the process)

It can also compute the **total amount of time** during which a resource is **busy** handling work items and its **resource utilization** (the average percentage of time that it is busy)

BIMP

http://bimp.cs.ut.ee

BIMP

BIMP is a free, simple online simulator of BPMN models.

1. Upload BPMN2.0 models created with BPMN-compliant tools

2. Create a simulation scenario including parameters such as:
the number of process instances to be simulated;
their arrival rates;
the number, types and timetables of resources;
branching probabilities;
the duration and fixed cost of each task (uniform, normal and exponential distributions for durations are supported).
The (parameters of the) simulation scenario can be saved.

3. Run the simulation

BIMP: output

A dashboard is eventually displayed that includes:

Cost information: total cost of the scenario, min, average and max costs of individual process instance and a diagram of the process instance cost distribution.

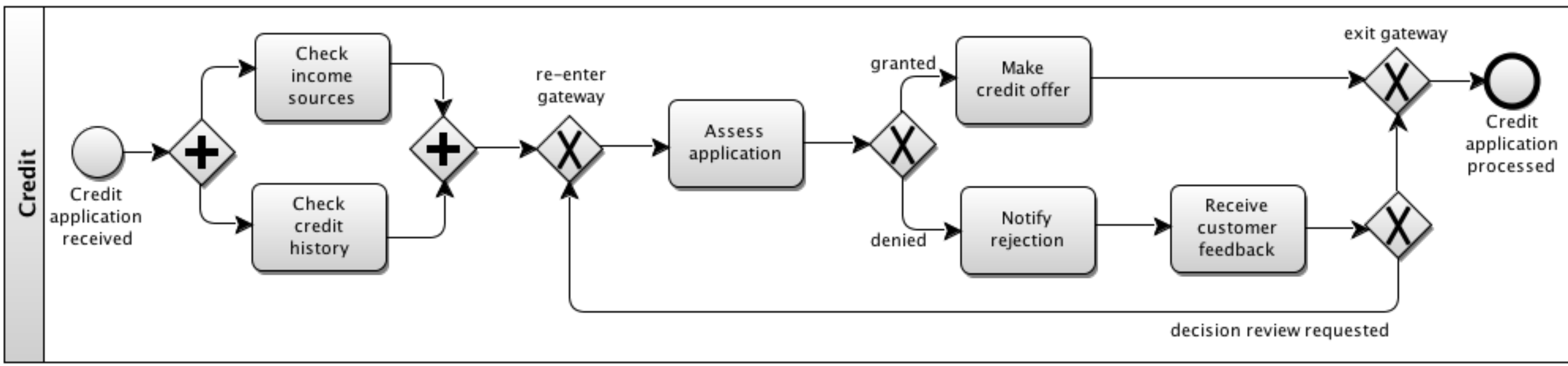
Bottlenecks: If resources are not sufficient to handle the scenario, then tasks will queue, causing high waiting times and cycle times.
A diagram shows the distribution of waiting times.

Resource utilization: Average utilization percentage of each resource.

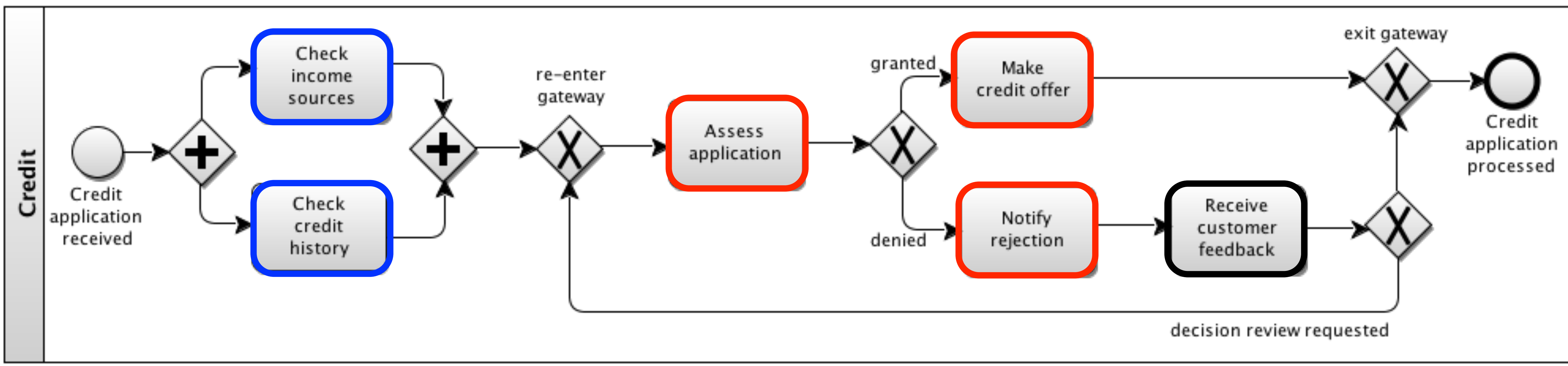
Cycle times (**process duration**): Total cycle time of the scenario, and diagrams about duration and cycle time distribution.

Simulation logs can be exported in MXML format and then imported in the ProM toolset for more detailed analysis.

Example: Credit application



Example: Credit application



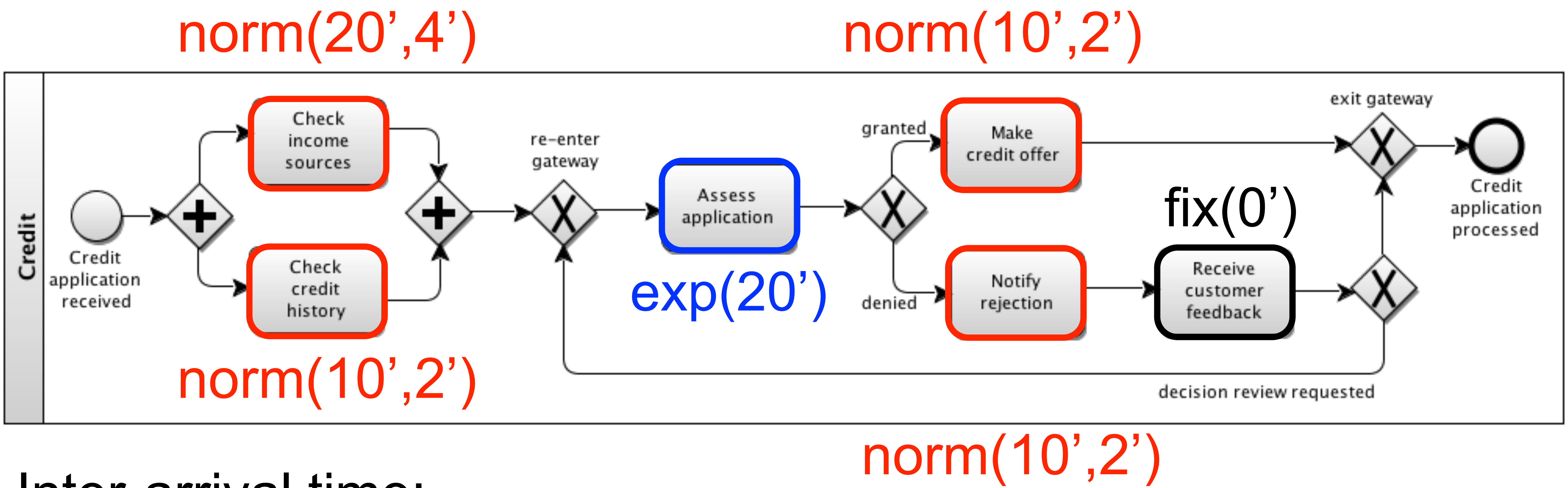
#3 Clerks (cost 25€/h)

#3 Credit officers (cost 25€/h)

System

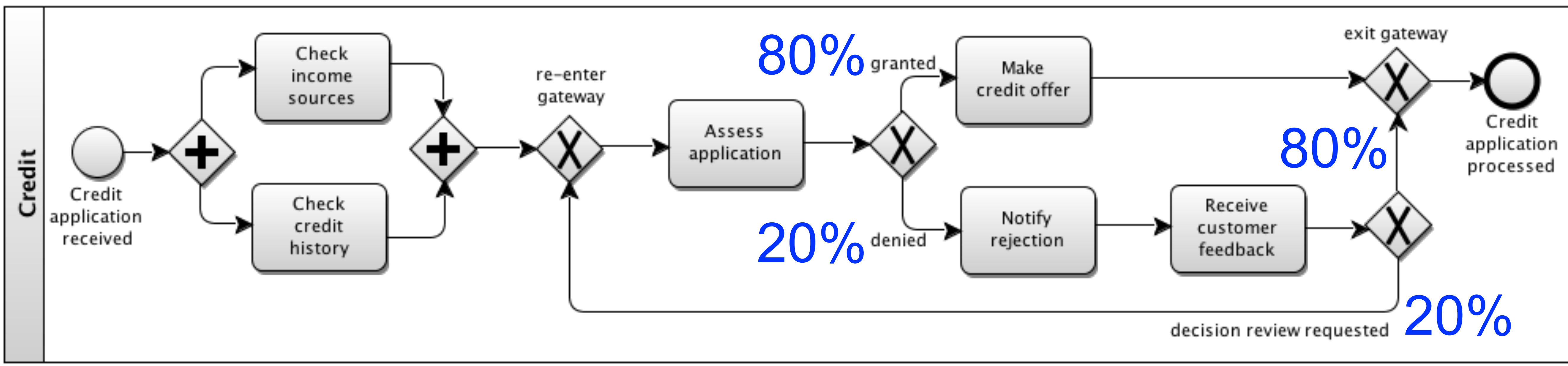
Working hours: monday/friday, 9:00/17:00

Example: Credit application



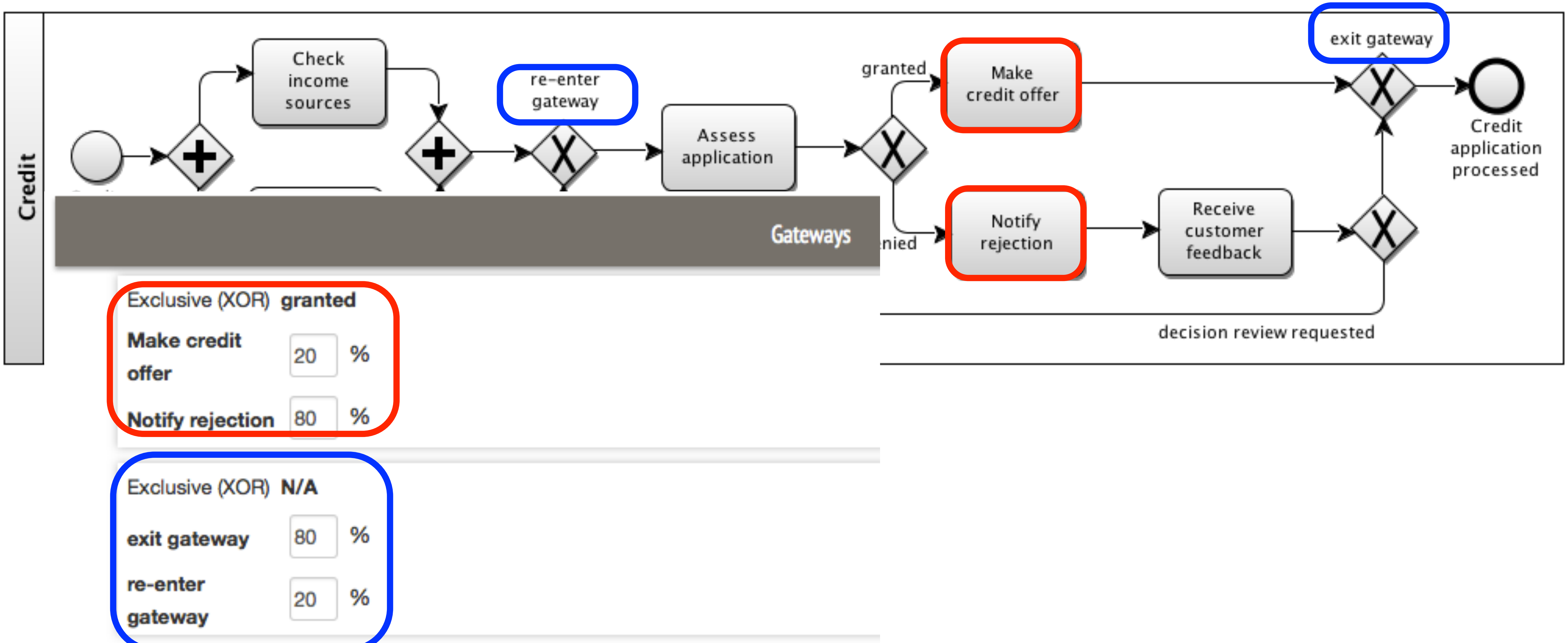
Inter-arrival time:
exp(30’)

Example: Credit application



Example: Credit application

Always assign names to items that follow any decision gateway: BIMP use those names to indicate branching probabilities

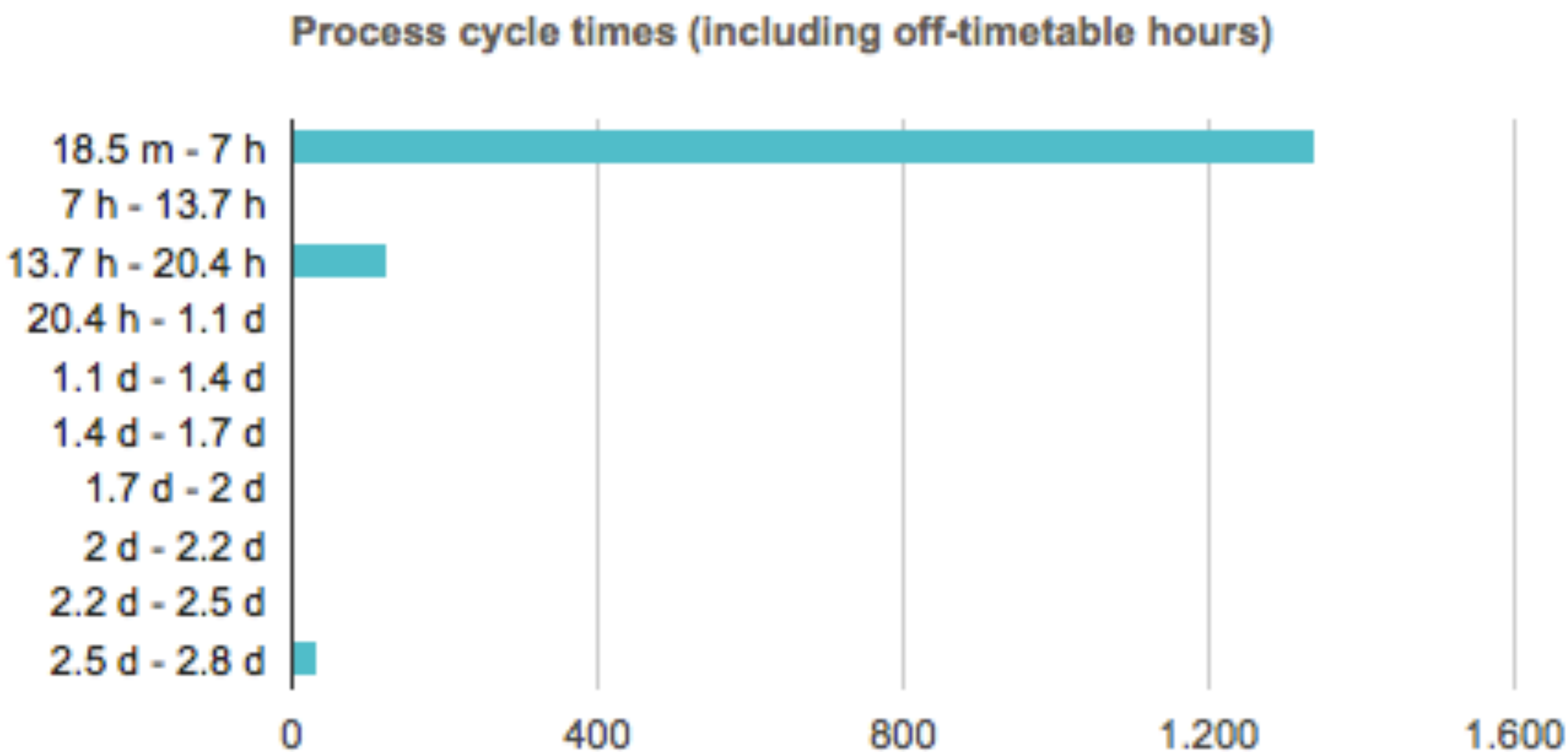


Simulation Results

General information

Completed process instances 1500
Total cost 62174.5 EUR
Total simulation time 19 weeks

Example: Credit application



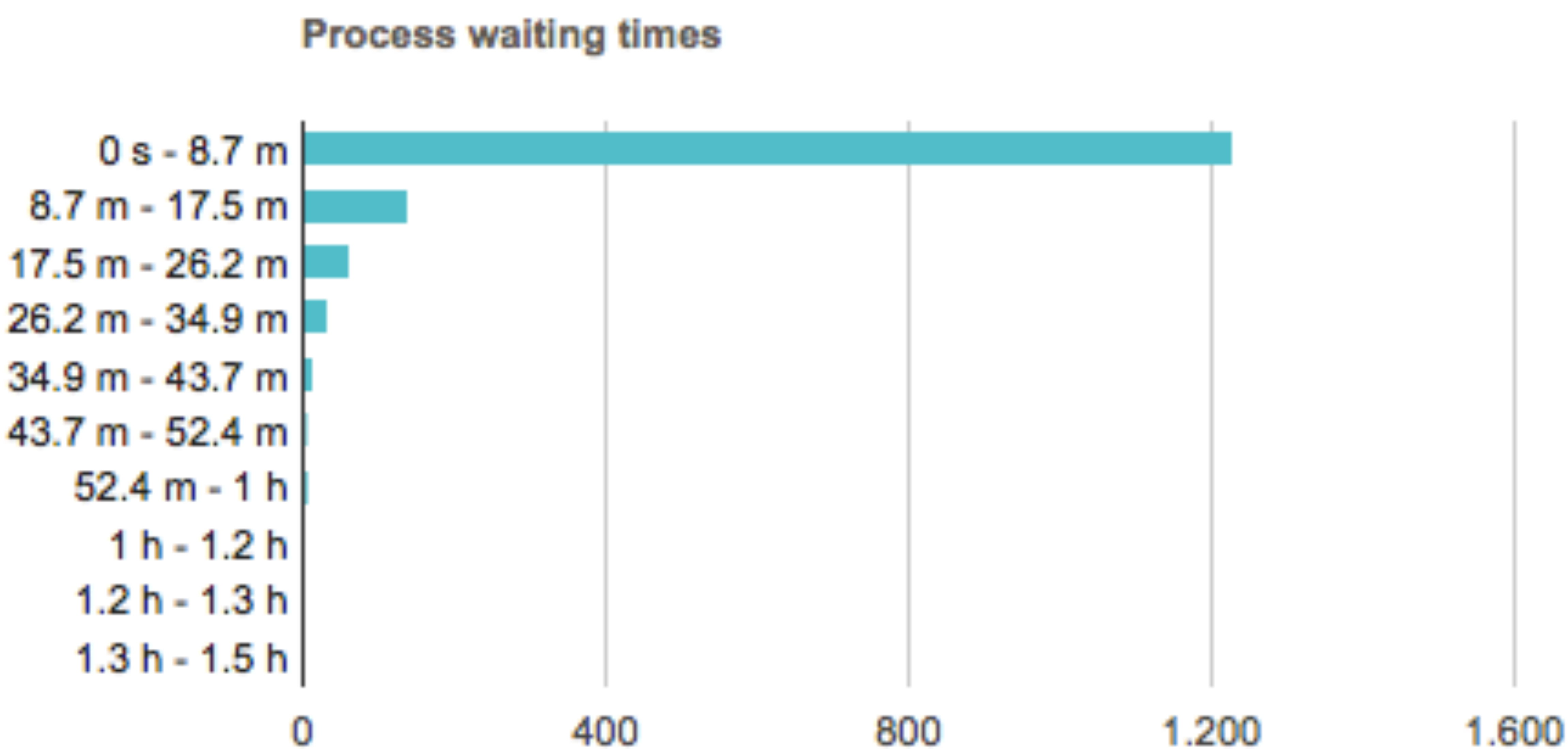
#instances

Example: Credit application

Simulation Results

General information

Completed process instances 1500
Total cost 62174.5 EUR
Total simulation time 19 weeks



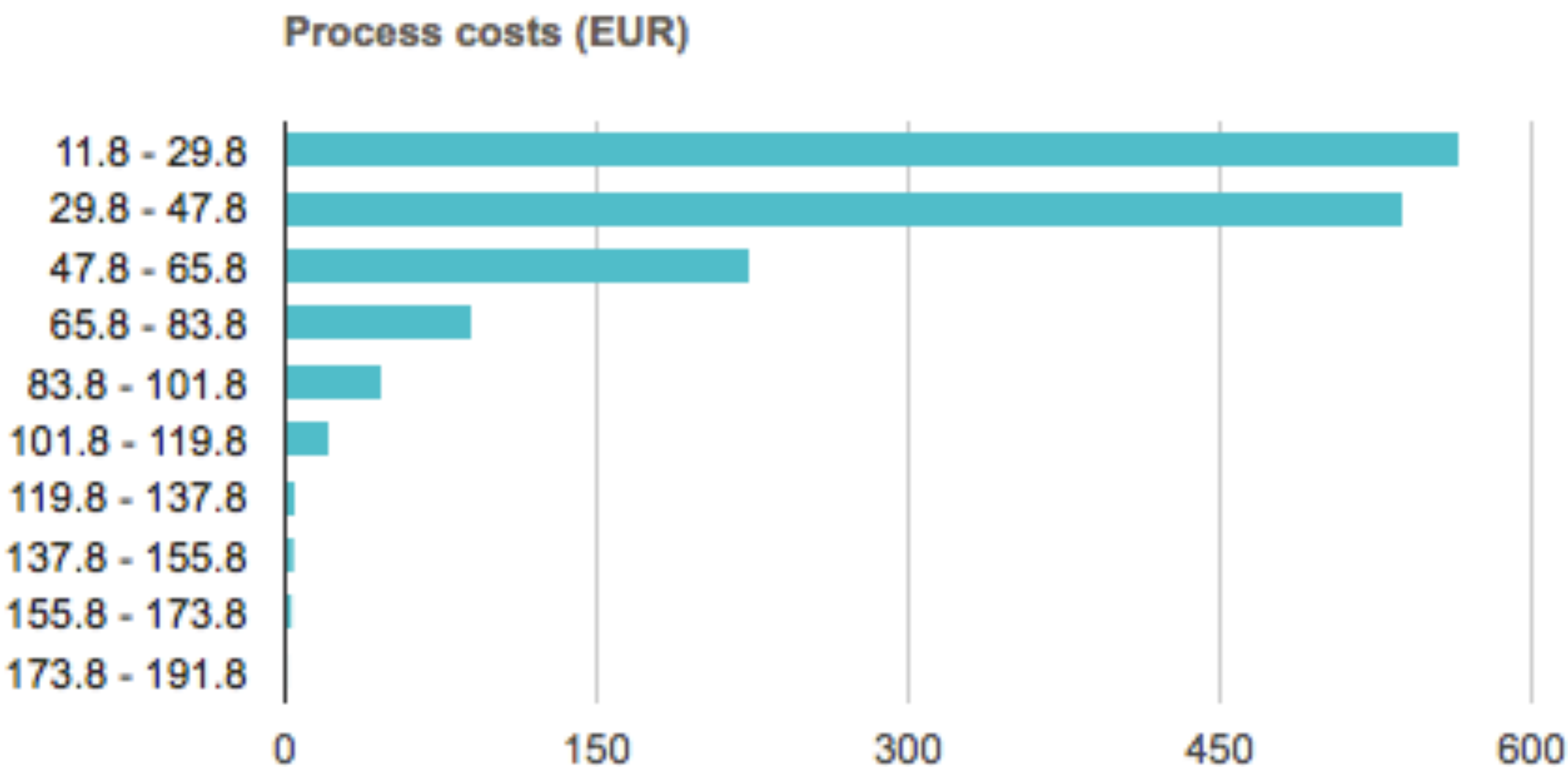
#instances

Simulation Results

Example: Credit application

General information

Completed process instances 1500
Total cost 62174.5 EUR
Total simulation time 19 weeks



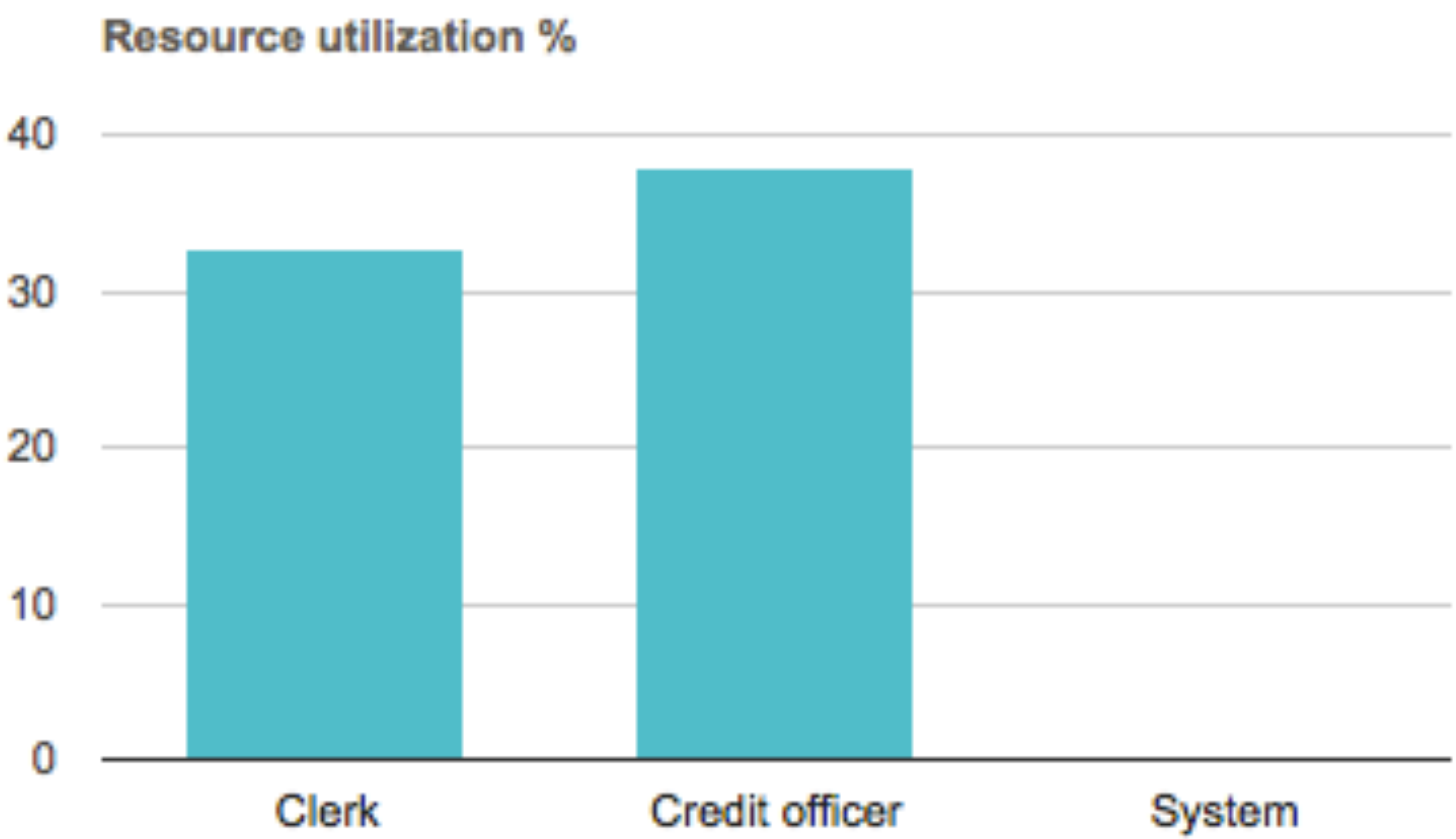
#instances

Example: Credit application

Simulation Results

General information

Completed process instances 1500
Total cost 62174.5 EUR
Total simulation time 19 weeks



Example: Credit application

Simulation Results

General information

Completed process instances 1500
Total cost 62174.5 EUR
Total simulation time 19 weeks

Process instance costs and cycle times (incl. off-timetable hours)

Minimum process cost 11.8 EUR **Maximum process cost** 183.9 EUR **Average cost** 41.4 EUR

Minimum cycle time 18.5 minutes **Maximum cycle time** 2.8 days **Average cycle time** 3.8 hours

Example: Credit application

Simulation Results

General information

Completed process instances 1500
Total cost 62174.5 EUR
Total simulation time 19 weeks

Task costs and waiting times

Task name	Average cost	Average waiting time
Assess application	16.4 EUR	28.2 seconds
Check credit history	4.2 EUR	59.6 seconds
Check income sources	8.3 EUR	2.6 minutes
Make credit offer	8.4 EUR	21.9 seconds
Notify rejection	8.3 EUR	30.1 seconds

Advices

It is recommended to run the simulation multiple times and then take the average of results

Quantitative analysis in general and simulation in particular are based on simplified models: their reliability depends very much on the quality of inputs (check the sensitivity of the analysis w.r.t. small changes)

Process participants are humans, not robot: they are not all the same, they get distracted, get ill, change the way to handle cases, change job, their performance may vary,…

Cross-check simulation results against reality

Exercise

1. Run the BIMP simulation of the example by yourselves
2. Change the inter-arrival time to 15’ and re-run the simulation: observe the changes in the results
3. Change the inter-arrival time to 10’ and re-run the simulation: observe the changes in the results
4. Change the number of Clerks and Credit officers to 5 and re-run the simulation: observe the changes in the results
5. Change the branching probabilities to 50% and re-run the simulation: observe the changes in the results