WoPeD project group

Quick Reference for the quantitative analysis with WoPeD

Manual of capacity planning and simulation

Christian Czech
09.09.2007

GENERAL REMARKS

This document is intended to explain the handling of the quantitative analysis functionality within WoPeD.

There are two methods of quantitative analysis: the capacity planning algorithm as you may know it from van der Aalst and van Hee and the (quantitative) simulation.

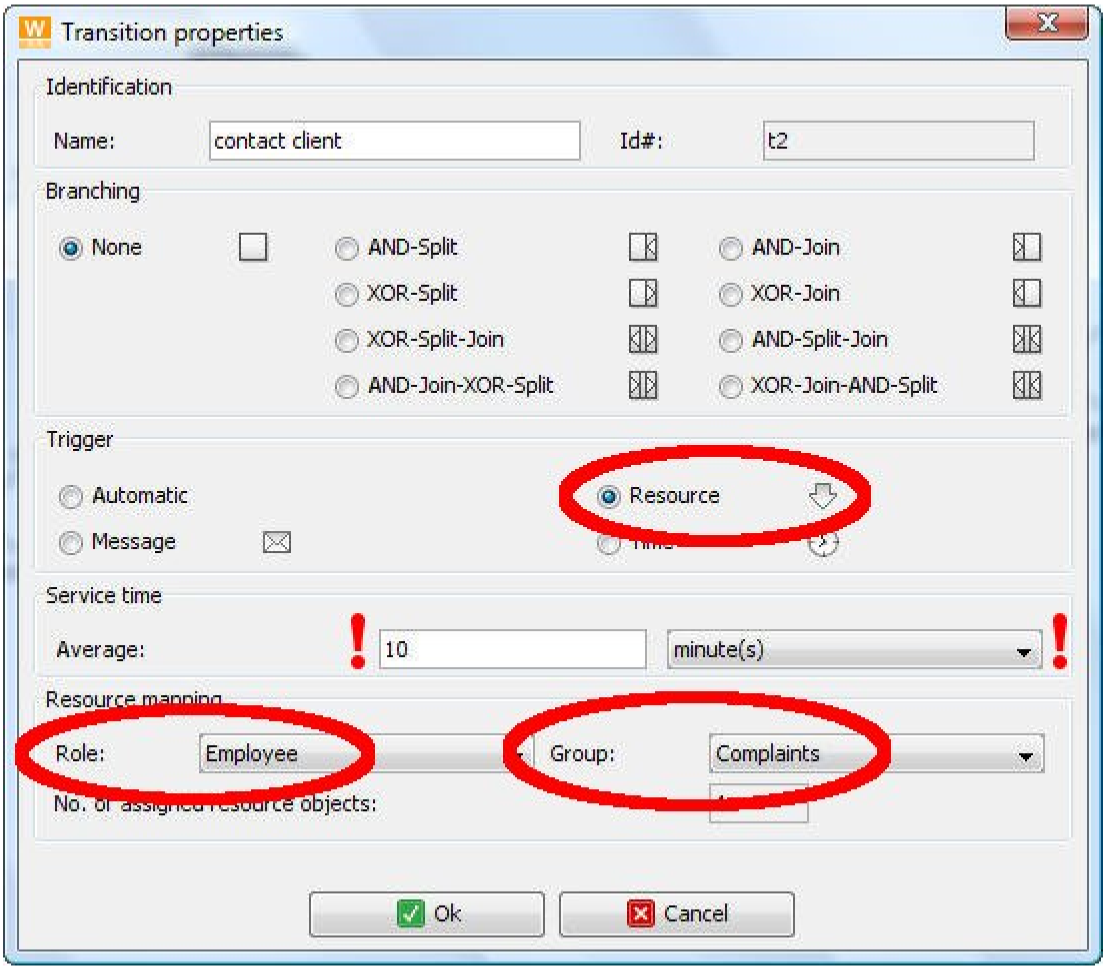
MODEL EXTENSIONS

In order to be able to accomplish quantitative analysis, a Workflow net must be supplemented with statistical data. Those are in detail:

- average service times of transitions
- split probabilities at XOR splits
- assignment of a group and a role for all transitions, which use resources

The necessary data can be given in the properties dialogues of the transitions and/or the arcs and can also be changed later.

TRANSITION PROPERTIES

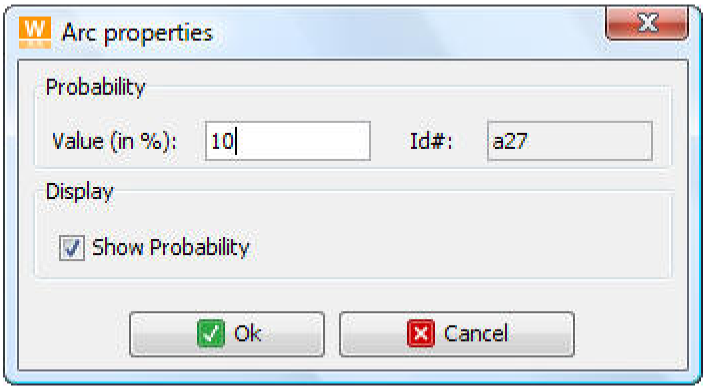


First the resource trigger is to be selected. Then the average service time can be entered in the appropriate text field. In addition the time unit can be selected in the accompanying combination list field. Within the editor the service time appears as green number in the transition. Below the current value a symbol for the time unit is indicated:

s	seconds	m	minutes	h	hours	d	days
w	weeks	M	months	y	years		

If a resource classification is given a group and a role can be chosen.

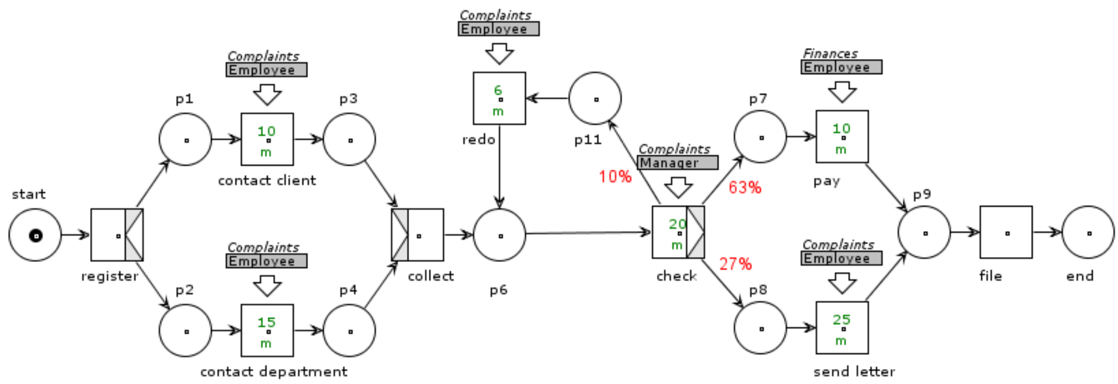
ARC PROPERTIES



With XOR split transitions the split probabilities can be entered in the arc properties dialog. Those are integers between 0 and 100. It is to be made certain that the sum of all probabilities of a transition is equal to 100%.

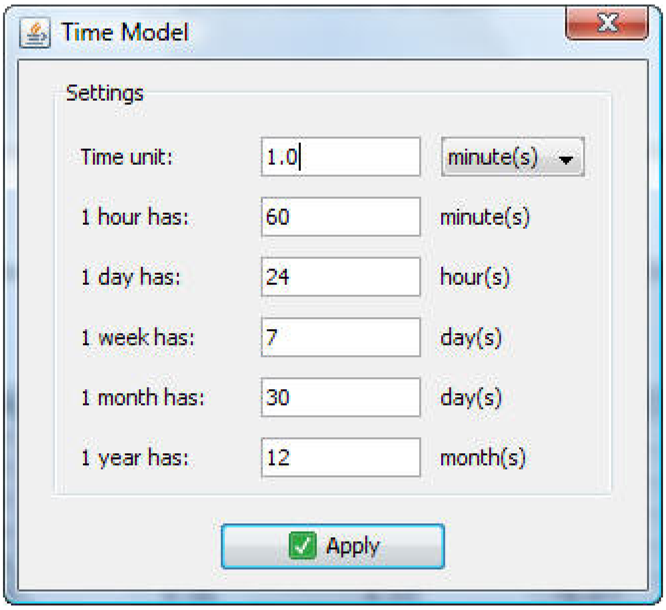
Furthermore it can be chosen whether the value shall be displayed in the editor.

An example workflow net looks as follows:



THE TIME MODEL

In the workflow model different time units can be used. For the computations all numerical data must be converted therefore into one time unit. Therefore the standard time unit can be changed in a dialog.



The default time unit is 1 minute. You have also the possibility to choose rather unusual time units like 1 week or 30 seconds etc. The results of the quantitative analysis are then to be understood as multiple of these units.

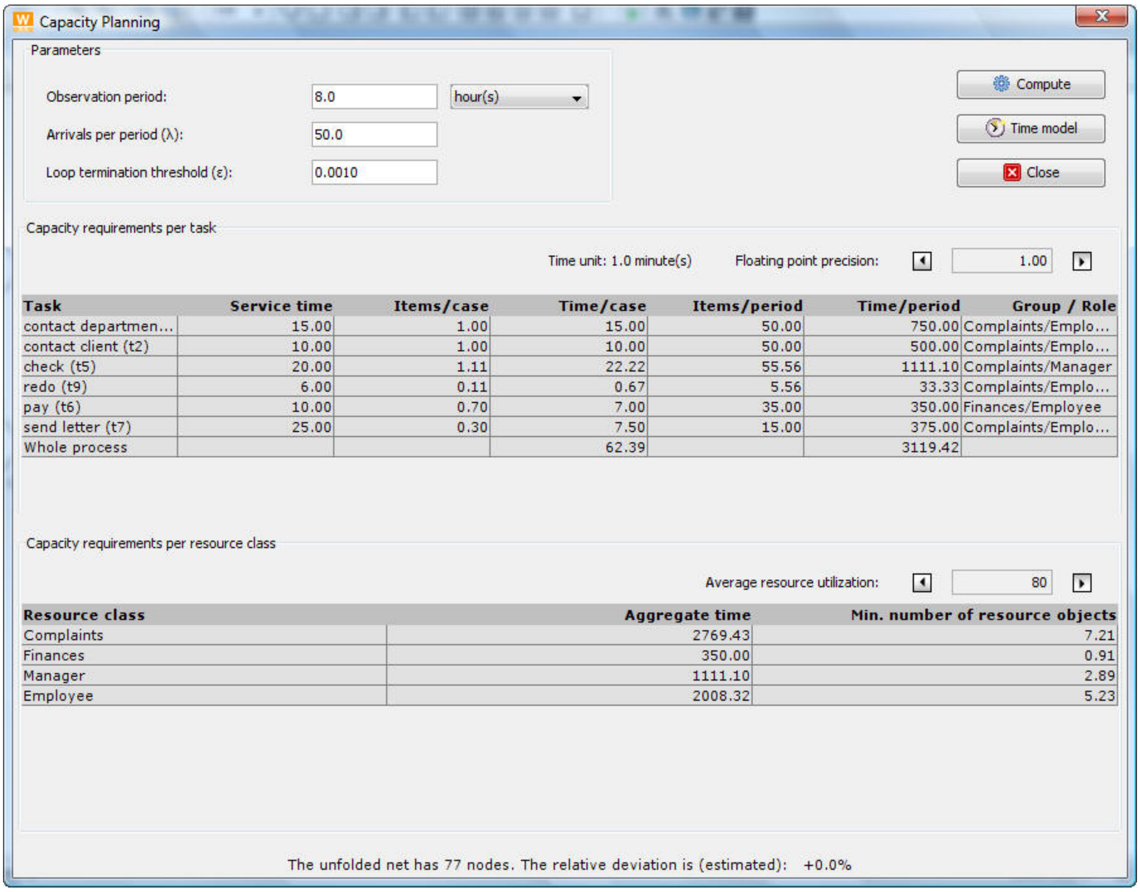
In addition more realistic conversion factors can be specified. Then you would work for example with 1 working day of 8 hours and 5 working days per week.

After clicking the button [Apply] the data are stored and inside the program all current values are converted.

WHAT IS MISSING

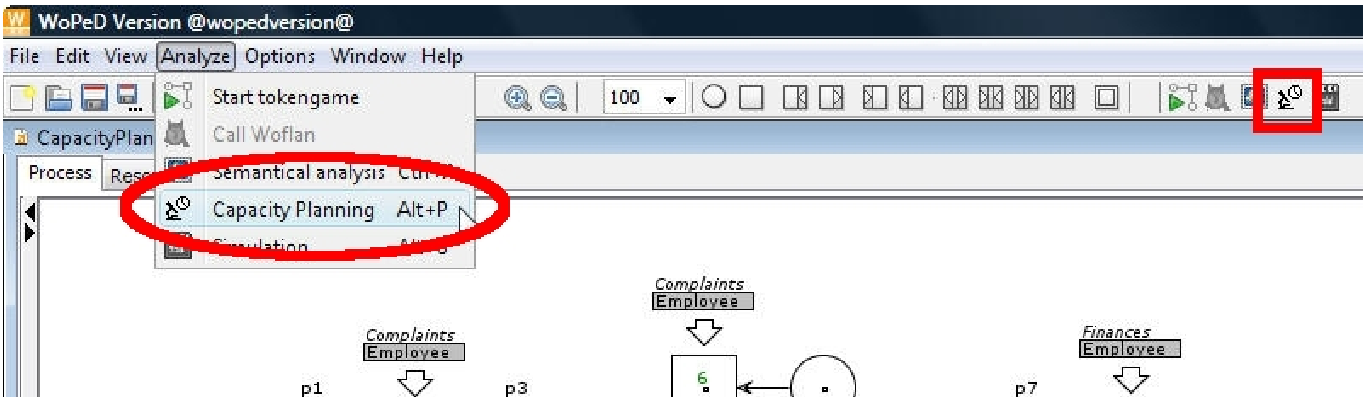
At present subprocesses are not supported. It will be treated like a usual transition and they must be assigned an average service time explicitly.

At first the operation of the dialogue for capacity planning is described.



CALL THE DIALOG

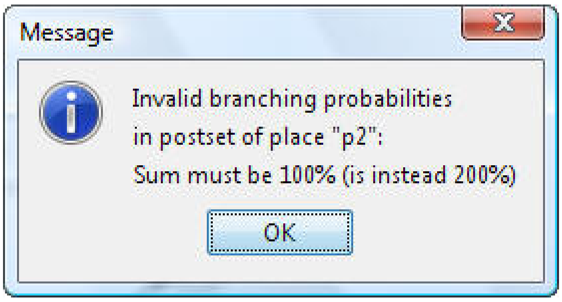
The capacity planning dialog can be started by clicking in the toolbar, the menu item Analyze ▹ Capacity Planning or with ALT+**P**.



The dialog for the capacity planning can only be opened, if there is a semantically and syntactically correct workflow net in the editor. For that purpose an internal examination is accomplished, which the user usually does not recognize. However two typical errors can occur:

1. The net is not sound. Then the soundness examination supplies a negative result and an error message is indicated:

2. All XOR splits have by default a split probability of 100%. If they are not changed or if their sum is not exactly 100%, a further error message follows:

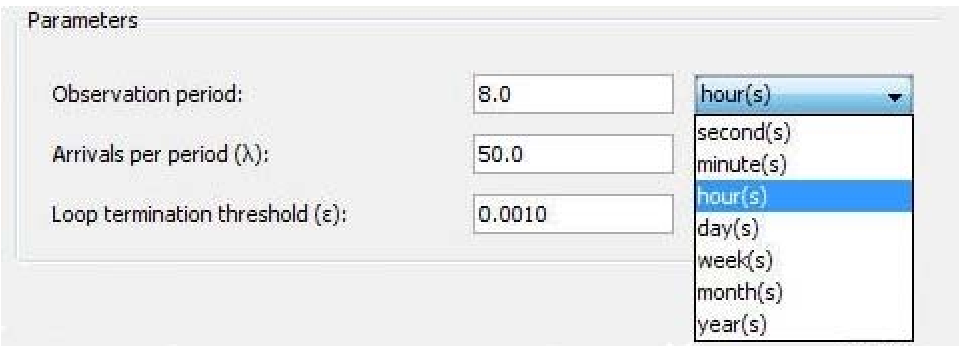


When there are no errors the dialog shows up. It is a modal dialog, i.e. underlying windows cannot be achieved until the capacity planning is closed. That is to prevent that during the analysis the model in the editor is changed or another net is opened.

The GUI of the dialog consists of three parts: the input part, the result part and an information border.

THE INPUT PART

Within the left range the input part permits entering the parameters:



The first parameter is the observation period, e.g. one working day or one month etc. Beneath it the average arrival rate λ, related to the period indicated before, is given.

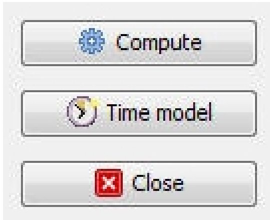
The third parameter is worth a special remark. Here a ε smaller than one is indicated. This parameter does not have to do anything with the actual logic of the capacity planning. For the computation the Petri net is unfolded internally. That is, it will go through the original net beginning with the starting place by means of breadth first search and for each node and each edge a copy is provided. This generates an unfolded net, which contains no cycles.

By traversing the original net a value φ_K is assigned to each node, which becomes ever smaller for each repeated reaching of the same node. Cycles are in principle infinitely often gone through thereby. Therefore the indication of an abort condition is necessary. For any node of the Petri net with the value φ_K and one of its output edges with probability p the path ends when: $\boldsymbol{\varphi_K * p < \varepsilon}$.

When this condition is fulfilled - which happens sooner or later definitely - the algorithm breaks off, if it does not do it before. For the case $\lambda=1$ ε can be regarded as accuracy measure for the results.

The results are the more exact, the smaller ε is. However the computations are becoming thereby more complex and more storage is used. Perhaps the program breaks off then with a heap overflow.

In the right range of the input part there are the buttons, in order to start the computation, to change the time model or to close the dialog.



THE RESULT PART

In the result part the results are displayed in two tables. In the upper table the execution frequencies and time demands of the transitions are shown.

Capacity requirements per task

Time unit: 1.0 minute(s) Floating point precision: 1.00

Task	Service time	Items/case	Time/case	Items/period	Time/period	Group / Role
contact departmen...	15.00	1.00	15.00	50.00	750.00	Complaints/Emplo...
contact client (t2)	10.00	1.00	10.00	50.00	500.00	Complaints/Emplo...
check (t5)	20.00	1.11	22.22	55.56	1111.10	Complaints/Manager
redo (t9)	6.00	0.11	0.67	5.56	33.33	Complaints/Emplo...
pay (t6)	10.00	0.70	7.00	35.00	350.00	Finances/Employee
send letter (t7)	25.00	0.30	7.50	15.00	375.00	Complaints/Emplo...
Whole process			62.39		3119.42	

Additionally the used time unit is indicated. With the buttons ◄ or ► the number of decimal places of the results in both tables can be adjusted between 0 and 6.

The table beneath indicates the time demands per resource class and the computed number of resources needed.

Capacity requirements per resource class

Average resource utilization: 80

Resource class	Aggregate time	Min. number of resource objects
Complaints	2769.43	7.21
Finances	350.00	0.91
Manager	1111.10	2.89
Employee	2008.32	5.23

Within the computation an utilization of 80% of all resources of a resource class is assumed. Again another value can be set between 5% and 100% in increments of 5 by clicking ◄ or ►.

THE INFORMATION BORDER

In the information border the number of nodes in the unfolded net and an approximate estimation of the relative error are specified.

The unfolded net has 77 nodes. The relative deviation is (estimated): +0.0%

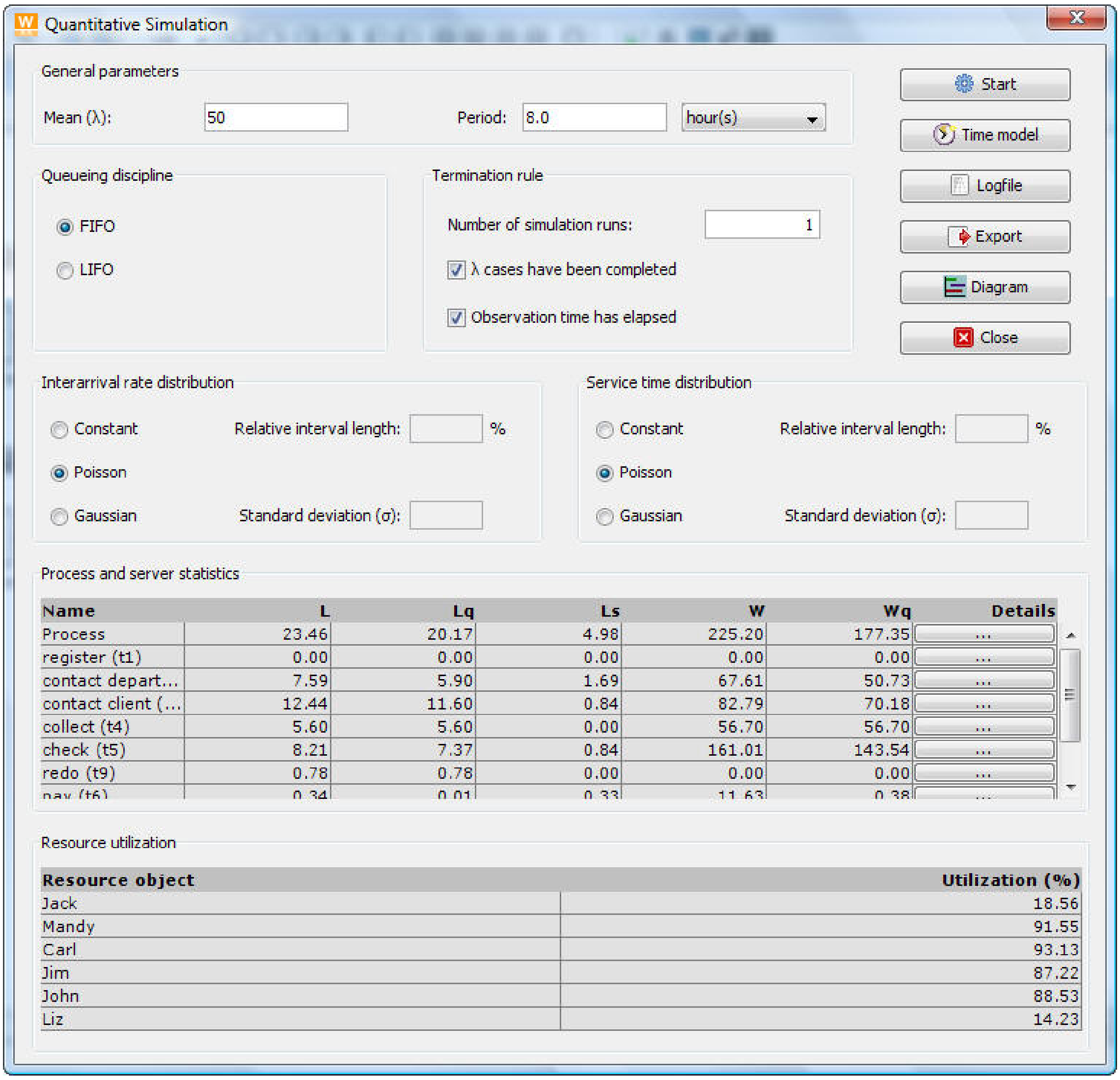
This information can be used for a purposeful change of the threshold value ε. Thus, for nets of similar size as the example, but a more complex loop structure, the unfolding can consist of several thousand nodes. A decrease of ε by a decimal place can lead then to several seconds computation duration or to the program crash because of memory lacking.

In the example the original net possesses 20 nodes. The net unfolding with $\lambda = 50$ and $\varepsilon = 0.001$ consists of 77 nodes. The relative deviation amounts to 0.0%, rounded on a decimal place.

ANNOTATION:
Since the example net contains a cycle, for some nodes a smaller execution frequency is computed than analytically correct result would be - because the unfolding was broken off. The 0.0% error are therefore only an estimation and does not stand for a 100% correct result. Indeed the maximum deviation is 0.1%.

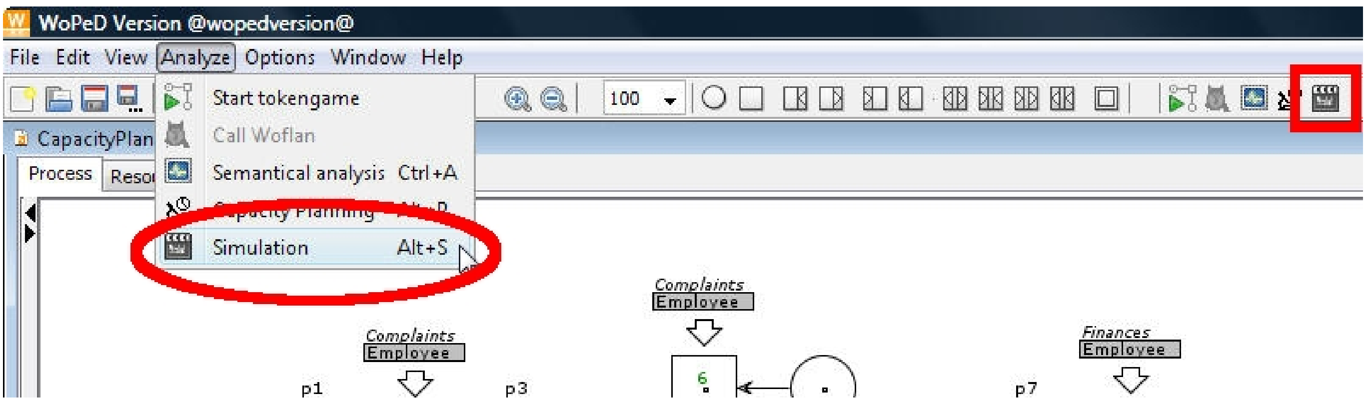
SIMULATION

The simulation dialog is just as compactly arranged as for the capacity planning.



CALL THE DIALOG

Calling is similar to capacity planning. It is done by clicking on in the toolbar or selecting the menu item Analyze ▹ Simulation or pressing ALT + S.



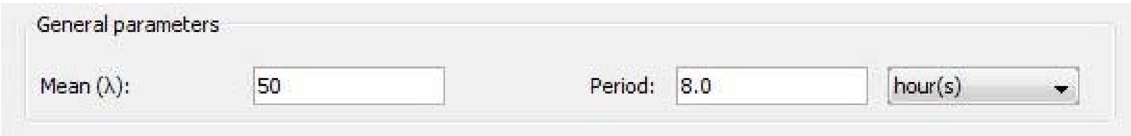
The soundness check of the workflow net also takes place here and if necessary displays the same error messages.

The dialog essentially consists of two parts, the configuration part above and the result part underneath it. It is clearly more complex than the dialog for the capacity planning.

CONFIGURATION PART

Here a variety of adjustments can be done. Actions are invoked by clicking on the buttons. The various possibilities to take control are introduced in the following paragraphs.

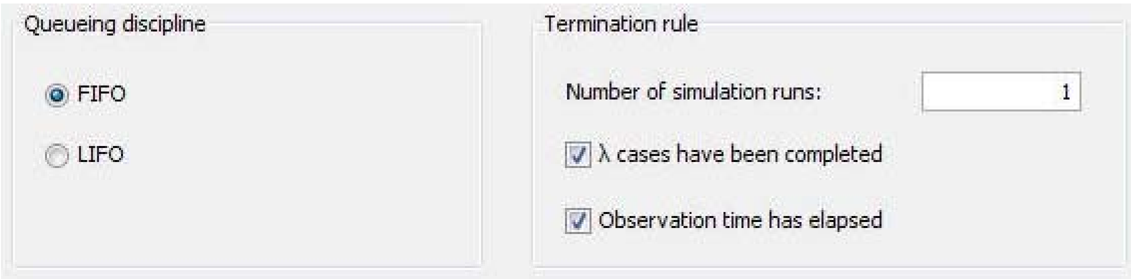
GENERAL PARAMETERS



As is the case for the capacity planning the average arrival rate λ measured in cases per period and the length of the observation period are indicated. By increasing the length of the period the significance of the simulation results can be improved.

The longer the period is, the better are the results. However the cost of computation rises enormously and it can last some time, until the simulation is terminated.

CONTROL FLOW PARAMETERS



The control flow of the simulation is steered on the one hand by the queue discipline and on the other hand by the termination rule.

The queuing discipline can be set by choosing between first-in-first-out (FIFO) and last-in-first-out (LIFO). Thus the order, in which waiting cases are taken out of the queue, is specified. The default is FIFO.

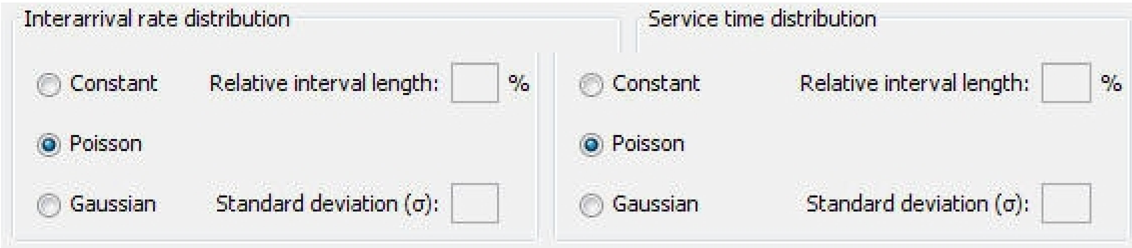
The termination of the simulation takes place, when the event list is empty. It is not predictable, when this is the case. However, in order to be able to control the break off, an termination rule can be selected. There is the choice between

- terminate, if λ or more cases have been finished or
- terminate, if the simulation time is larger or equal to the period or
- terminate as soon as the first of these two cases occurs

In addition the number of simulation runs can be indicated. The results of the individual simulation runs are stored. At termination of the simulation from all runs the arithmetic means are formed then.

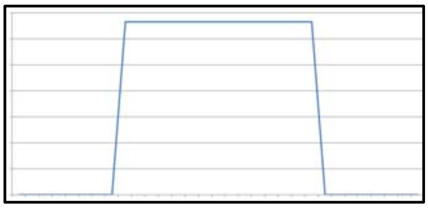
It applies here too, that the significance of the results is the better, the more simulation runs are performed (Law of large numbers). Just like with the period it is to be noted, however, that the complexity of the computation rises extremely.

PROBABILITY DISTRIBUTIONS

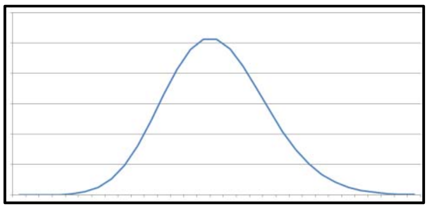


The distribution of the arrival process and the kind of distribution of the service times of the servers (transitions) can be specified. The proceeding is alike for both distributions. There is the choice between

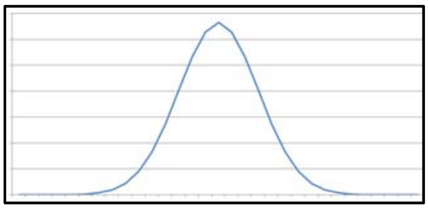
- **Constant distribution within an interval**
  The parameter is the relative interval length ς, an integer between {0, 1,..., 100}. It is related to the average arrival rate λ and means, that the expected value λ is surrounded by a symmetric interval of length 2ςλ. Within this interval all values have the same probability $P(X = \lambda^*) = \frac{1}{2\varsigma\lambda^*}$.



- **Poisson distribution**
  The arrival process has a Poisson distribution with parameter λπ, where π is the period. Then the interarrival times of successive cases are exponentially distributed with parameter $\zeta = \frac{\pi}{\lambda}$ and the density function $f(t) = \zeta e^{-\zeta t}$.



- **Gaussian distribution**
  The arrivals are normal distributed with expected value λ and standard deviation σ.

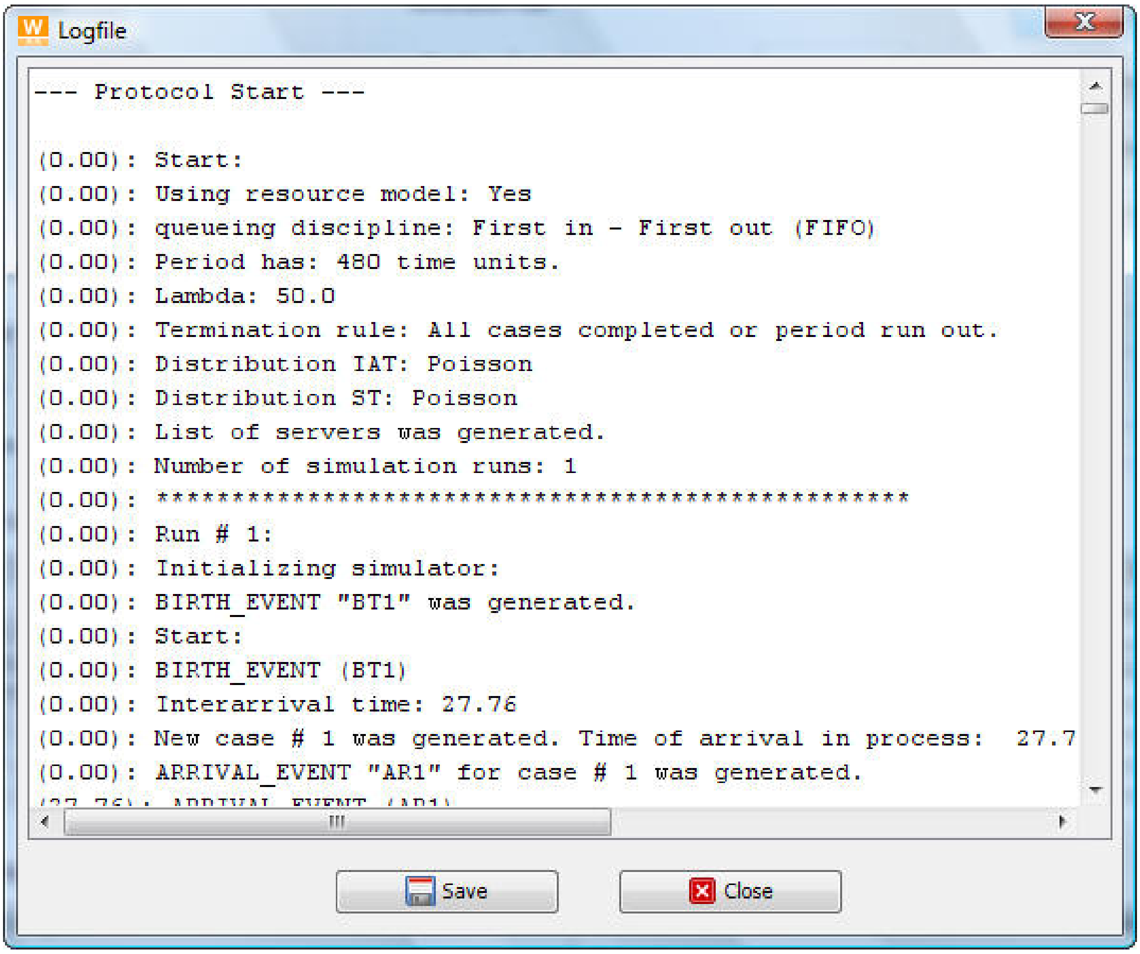


The buttons are located in the right area.

When all adjustments are made, the simulation can be started by clicking on [Start]. Then a waiting dialog appears which requests the user to wait until the computations are accomplished.

Clicking on [Time model] allows the user to change the time units, as known from the capacity planning.

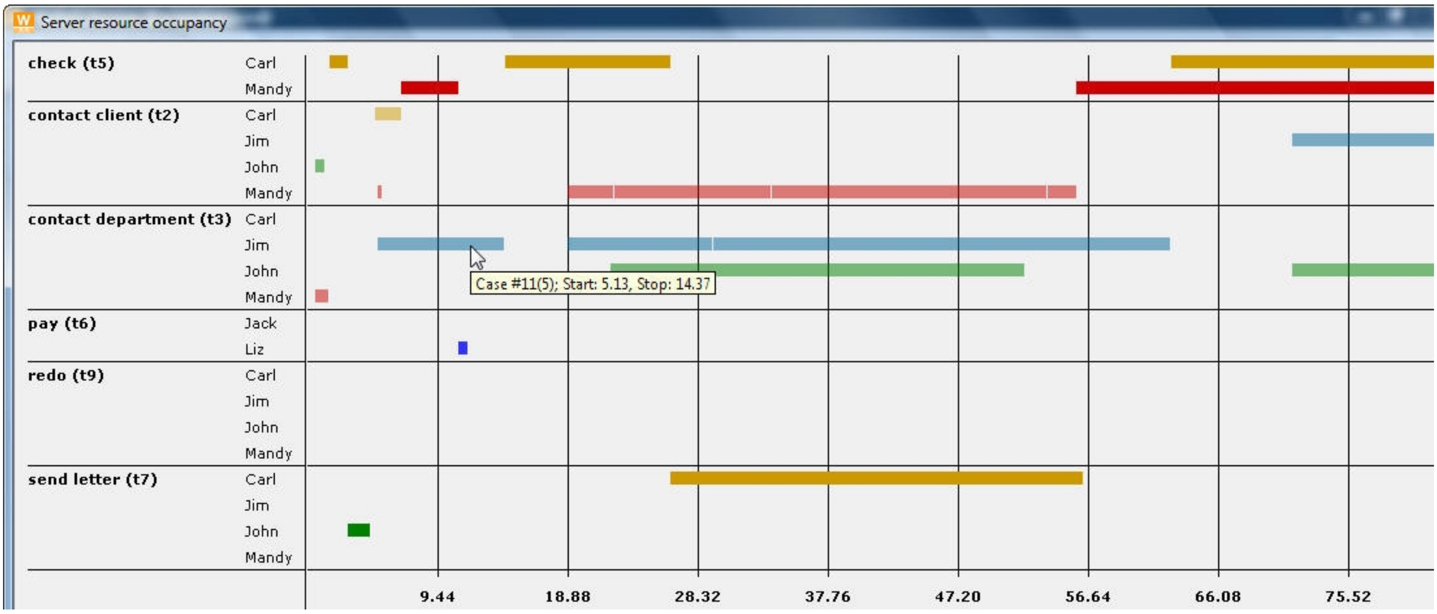
[Protocol] opens up a dialog, in which the protocol is displayed as plain text. It can be stored as ASCII text by giving of the file extension „txt“ or in the XML format through appendices of „xml“ at the file names.



With [export] the results, as they are shown in the result part in the tables and the detail dialogs, are exported into a CSV file *<name>.csv*. This file format can be read by usual spread-sheet programs.

In addition the values of the statistic counter variables of the individual simulation runs are stored in further files with the names *<name_i>.csv*, whereby i stands for the number of the simulation run. Thereby you have the possibility to analyze the simulation more exactly.

The button [Diagram] opens another dialog, which gives a graphical view of the allocation of the servers with resources over the time of the simulation.



If the mouse is held over one of the bars, a tooltip appears, in which the ID of the serviced case and the start and stop times are indicated.

A case ID with a number in parentheses behind it means the following: After AND splits several instances of a case are being serviced independently of each other. Therefore copies of the original case are processed. These copies receive successive ID's as the "normal" cases. In order to be able to assign the case to an original case, its ID is indicated in parentheses behind the ID of the copy.

The button [Close] closes the simulation dialog.

RESULT PART

The results of the simulation are arithmetic averages of the values of the counter variables of the individual simulation runs. They are shown once from process and/or server view in the first table, and once from resource view in the second table.

PROCESS VIEW

Process and server statistics

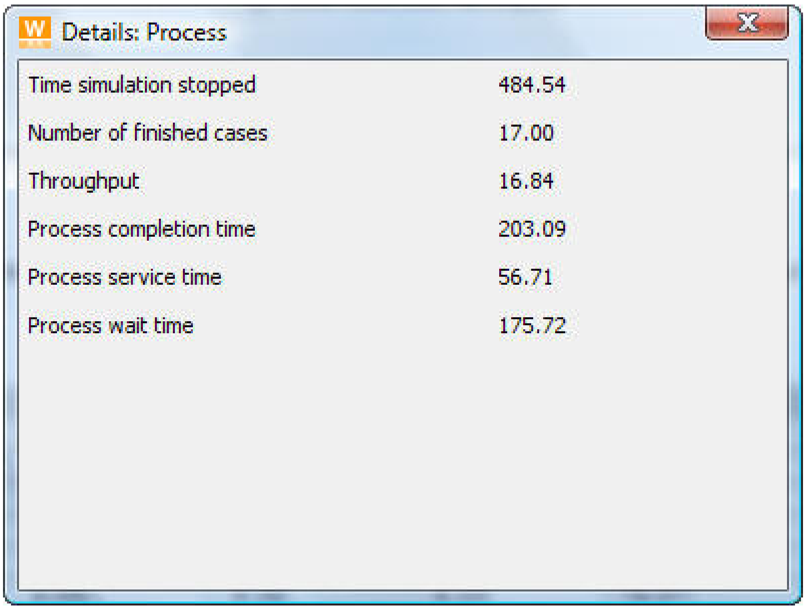
Name	L	Lq	Ls	W	Wq	Details
Process	8.85	7.01	4.22	84.94	44.46	...
register (t1)	0.00	0.00	0.00	0.00	0.00	...
contact depart...	3.86	2.51	1.35	41.83	27.61	...
contact client (...	4.06	3.10	0.96	44.40	33.31	...
collect (t4)	3.64	3.64	0.00	39.27	39.27	...
check (t5)	5.61	4.68	0.93	56.72	39.27	...
redo (t9)	0.27	0.27	0.01	1.49	0.00	...
pay (t6)	0.32	0.00	0.32	8.78	0.00	...

Process and server view are summarized in the upper table. The process data are located in the first line. Clicking on the last column opens a detail dialog with further process specific information.

The meaning of the column headings are geared to the designations usually used in queuing theory:

L	**average number of cases in system/server**	$L = L_q + L_s$, $W = W_q + \frac{1}{\mu}$
L_q	**average number of cases in queue**	μ: service time
L_s	**average number of cases in service**	Little's law
W	**average completion time**	$L = \lambda W$, $L_q = \lambda W_q$, $L_s = \lambda\mu$
W_q	**average wait time**	

The details dialog appears as follows:



SERVER VIEW

The server view is similar to the process view. Further details displayed in the details dialogs are:

(1) *Zero Delays* := number of cases that started service without waiting
(2) *Calls* := number of arrived cases (had to wait if necessary)
(3) *Cases service began* := number of cases that started service
(4) *Finished cases* := number of cases that finished service = (2) – [(8) + (9)]
(5) *Max. length of queue* := highest number of waiting cases
(6) *Max. number parallel cases* := highest number of cases serviced simultaneously
(7) *Max. wait time* := longest time a cases had to wait
(8) *Number served cases when simulation stopped*: number of cases being served when simulation was terminated = (3) – (4)
(9) *Number waiting cases when simulation stopped*: number of cases in the queue when the simulation run was finished = (2) – (3)

Details: contact department (t3)

Number of zero delays	0.00
Number of calls	45.00
Number of cases service began	43.00
Number of cases service finished	41.00
Max. length of queue	7.00
Max. number of parallel cases	3.00
Max. wait time of cases	103.96
Number served cases when simulation stopped	2.00
Number waiting cases when simulation stopped	2.00

RESOURCE VIEW

Resource utilization

Resource object	Utilization (%)
Jack	13.08
Mandy	89.90
Carl	76.35
Jim	71.63
John	75.70
Liz	12.32

In the table below all resources are shown and their computed rates of utilization in per cent are indicated.