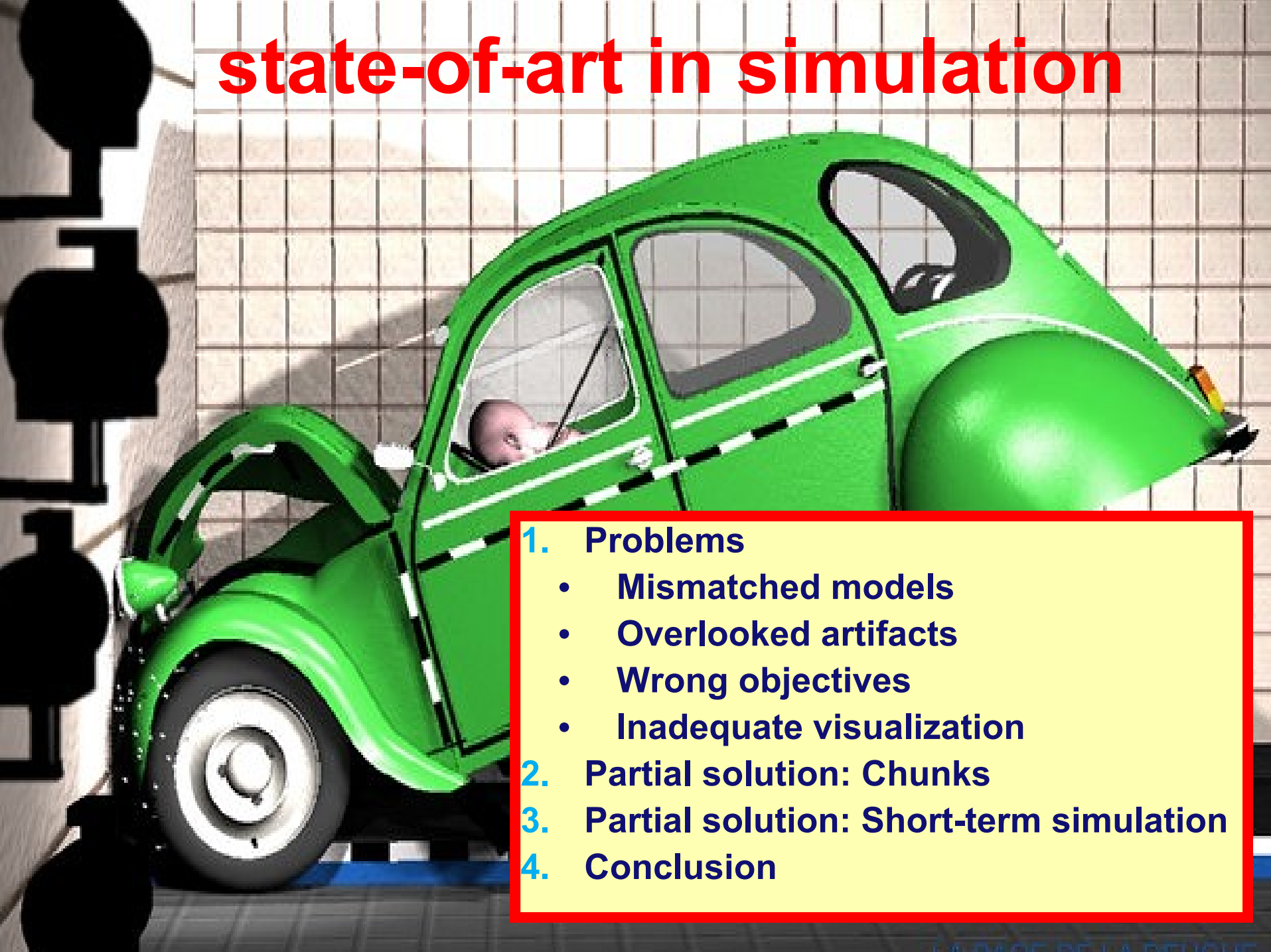
Business Process Simulation Revisited

prof.dr.ir. Wil van der Aalst

TU/e Technische Universiteit Eindhoven University of Technology

Where innovation starts

state-of-art in simulation



1. Problems
   - Mismatched models
   - Overlooked artifacts
   - Wrong objectives
   - Inadequate visualization
2. Partial solution: Chunks
3. Partial solution: Short-term simulation
4. Conclusion

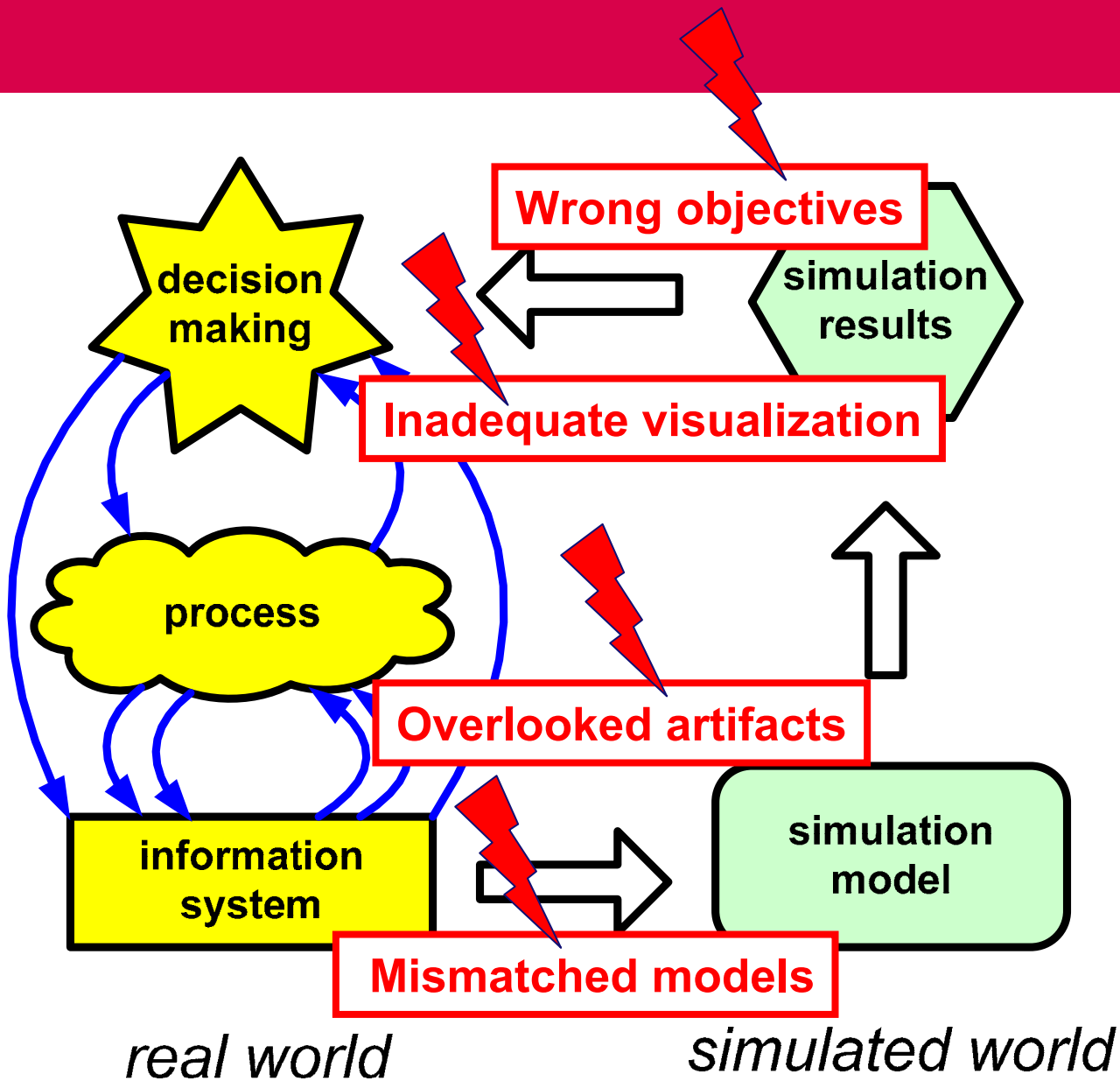
Acknowledgements

- W.M.P. van der Aalst, J. Nakatumba, A. Rozinat and N. Russell. Business Process Simulation: How to Get it Right? *Handbook of Business Process Management,* 2008 (to appear, available as BPM Center Report).
- A. Rozinat, M.T. Wynn, W.M.P. van der Aalst, A.H.M. ter Hofstede, and C. Fidge. Workflow Simulation for Operational Decision Support Using Design, Historic and State Information. *International Conference on Business Process Management (BPM 2008)*, 2008.
- A. Rozinat, R.S. Mans, M. Song, and W.M.P. van der Aalst. Discovering Simulation Models. *Information Systems*, 2008 (to appear, available as BETA report).
- A. Rozinat, R.S. Mans, M. Song, and W.M.P. van der Aalst. Discovering Colored Petri Nets From Event Logs. *International Journal on Software Tools for Technology Transfer*, 10(1):57-74, 2008.
- H.A. Reijers and W.M.P. van der Aalst. Short-Term Simulation: Bridging the Gap between Operational Control and Strategic Decision Making. *Proceedings of the IASTED International Conference on Modelling and Simulation*, 1999.

Problems

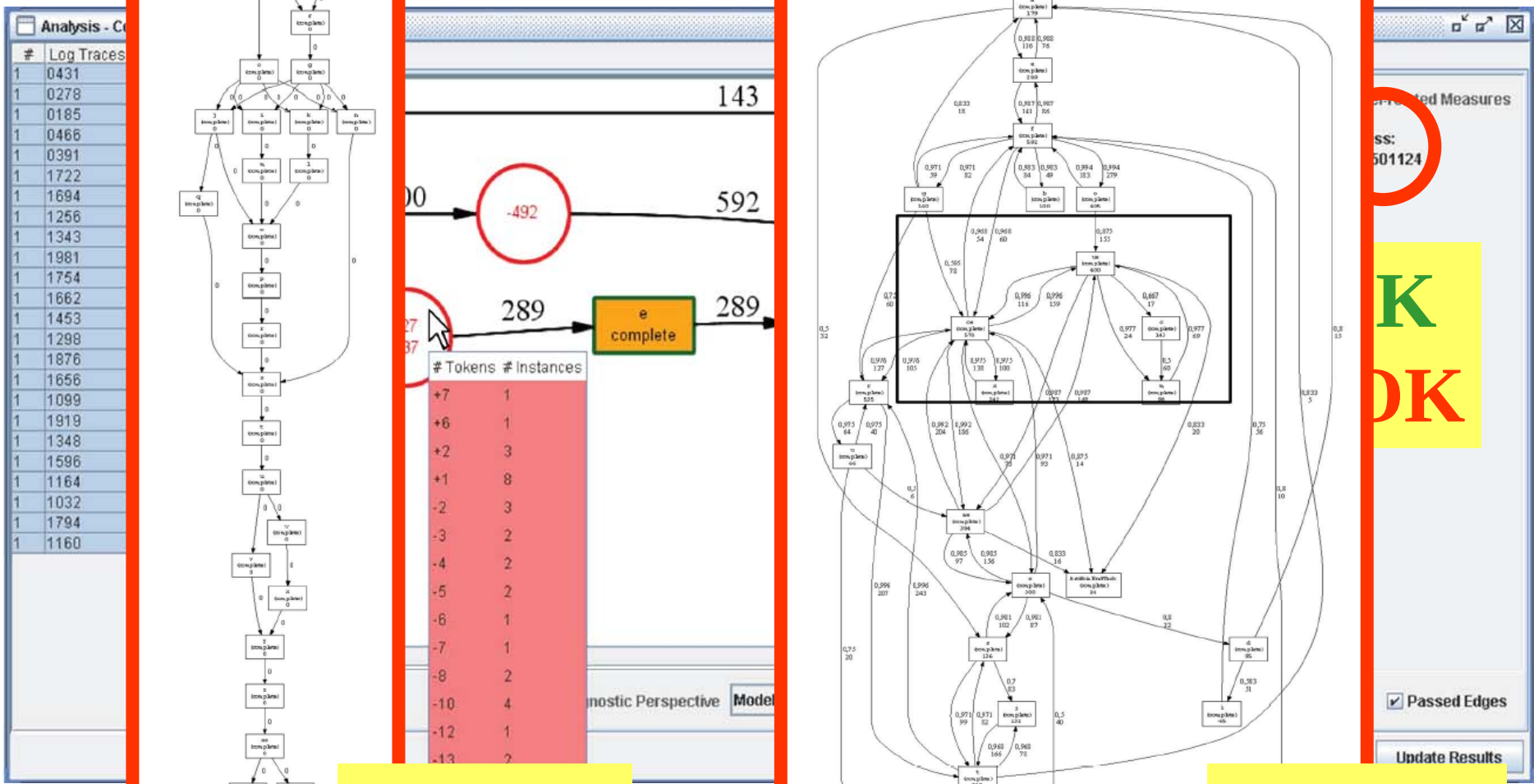
a) Mismatched models
b) Overlooked artifacts
c) Wrong objectives
d) Inadequate visualization

Some of the problems related to BPS



Problem 1: Mismatched models

Ex ASML test p



design

reality

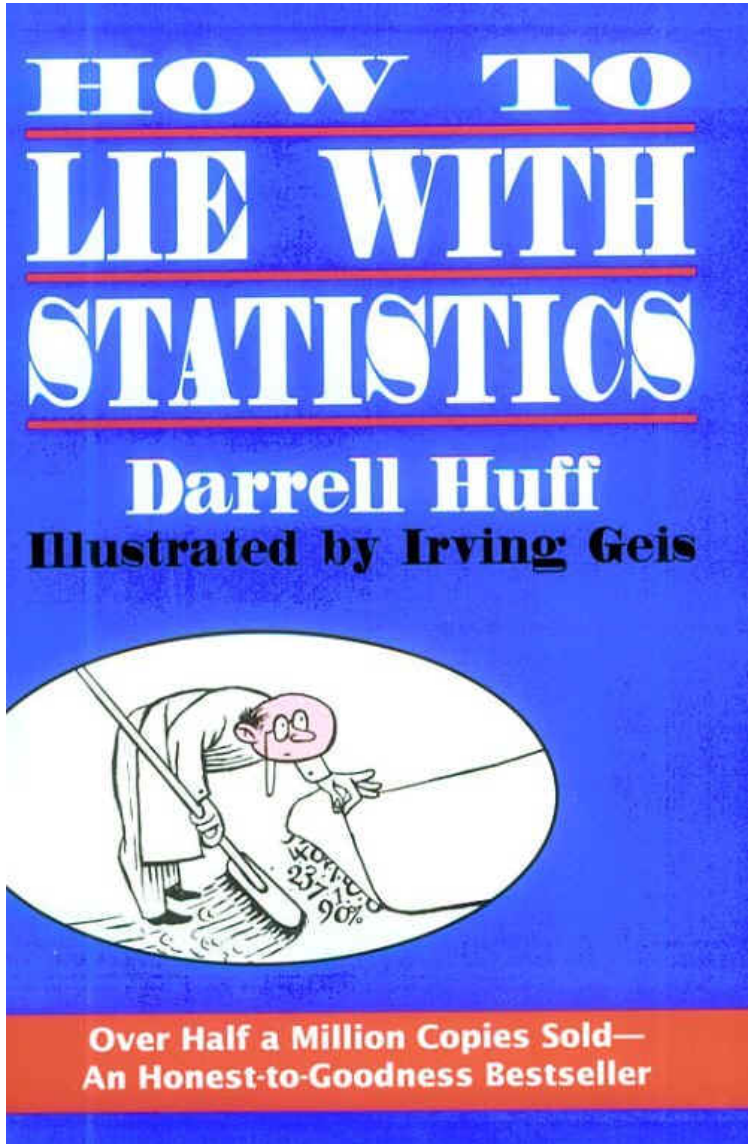
Example: My first real-life simulation

- TASTE (The Advanced Studies of Transport in Europe) project (1997-2003)
- ExSpect simulation model of supply chain of DAF to Spain.
- Interesting problems such as 200.000+ different spare parts.
  - Taking the characteristic ones or just the fast-movers does not work!
  - Consider for example truck loads, warehouse, etc.
- Key Performance Indicators (KPIs) in initial simulations dramatically different from reality.

Example: Numerous master projects

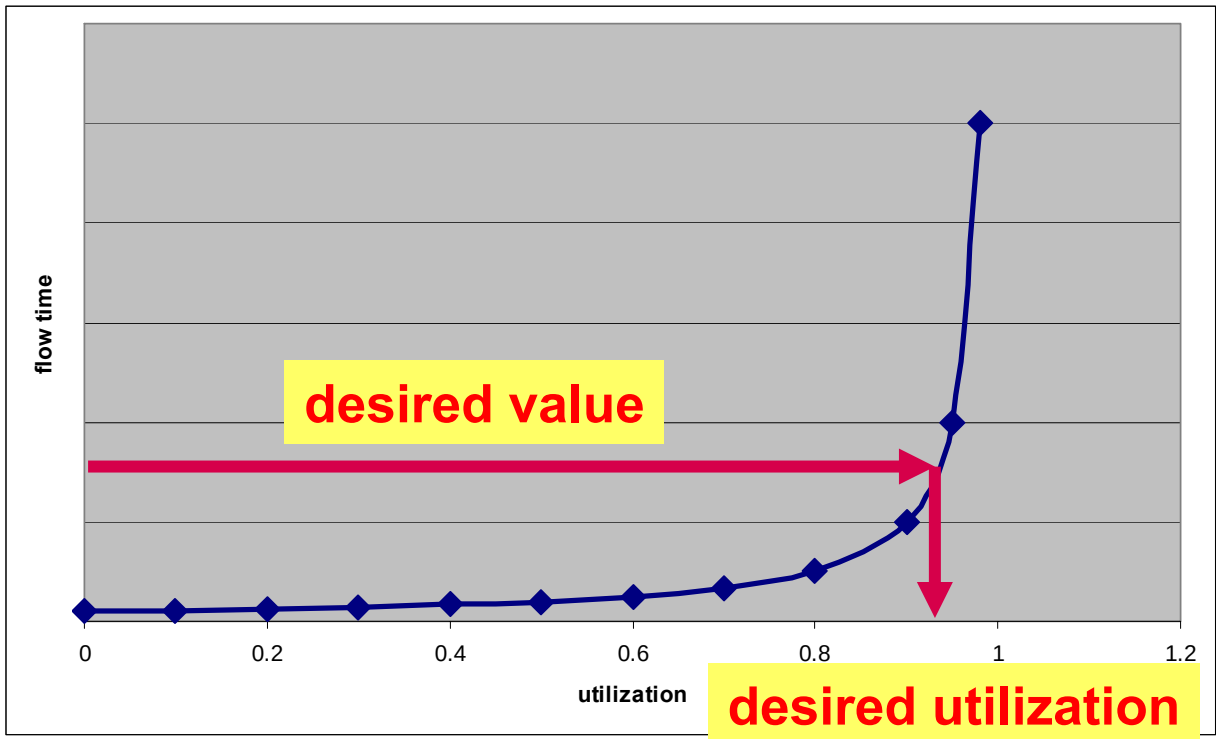
- After measuring times and routing probabilities, a faithful simulation model is constructed.
- However, KPIs in initial simulations dramatically different from reality.
  - Simulated flow times of hours correspond to real-life flow times of weeks.
  - Simulated utilizations of 30 percent in processes where employees complain about workload (burnout, boreout, or simout?).
- Simulation model is "massaged" until reality and simulation match.
- Observation: processes adapt based on context (when busy; skip checks, work longer, etc.).

How to Lie With Statistics?



How to Lie With Simulation?

- M/M/1 queue: arrival rate λ, service rate μ, utilization ρ = λ/μ. T1
- Flow time = 1/(μ-λ), # in system = ρ/(1- ρ)

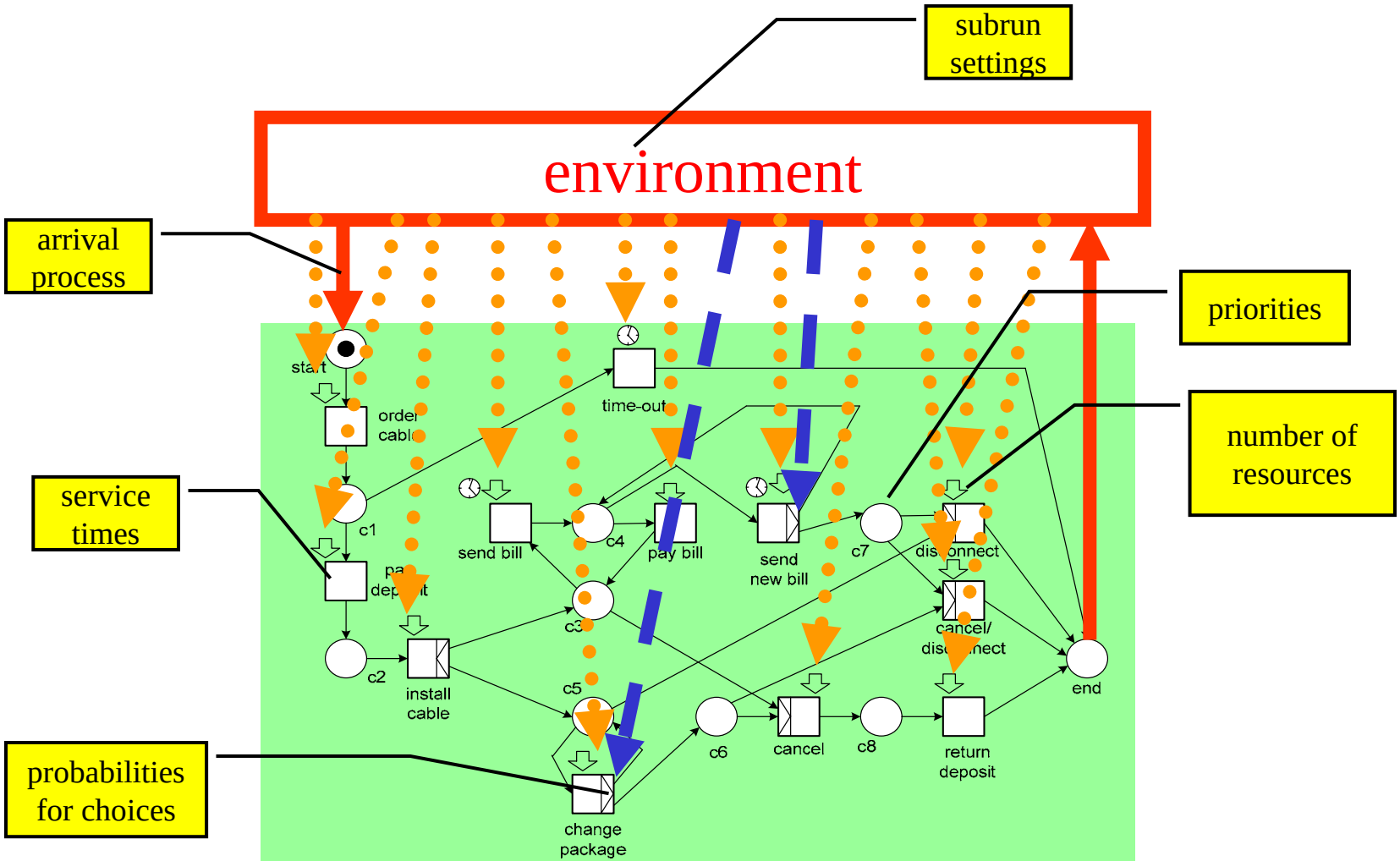


T1 TM, 10/4/2008

Problem 2: Overlooked artifacts

- **Simulation models are typically built from scratch ignoring a wealth of information:**
  - **process models (e.g., workflow models implicit or explicit)**
  - **historic data (event logs, information about arrival rates, service rates, etc.)**
  - **current state**

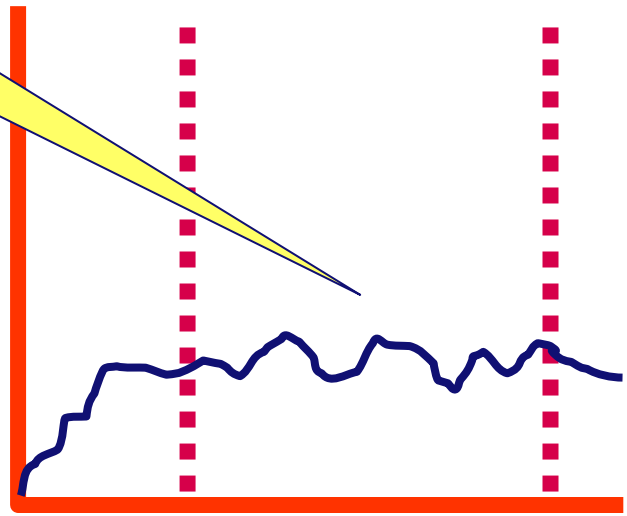
Example: How to turn a workflow model into a simulation model?



Use historic data!

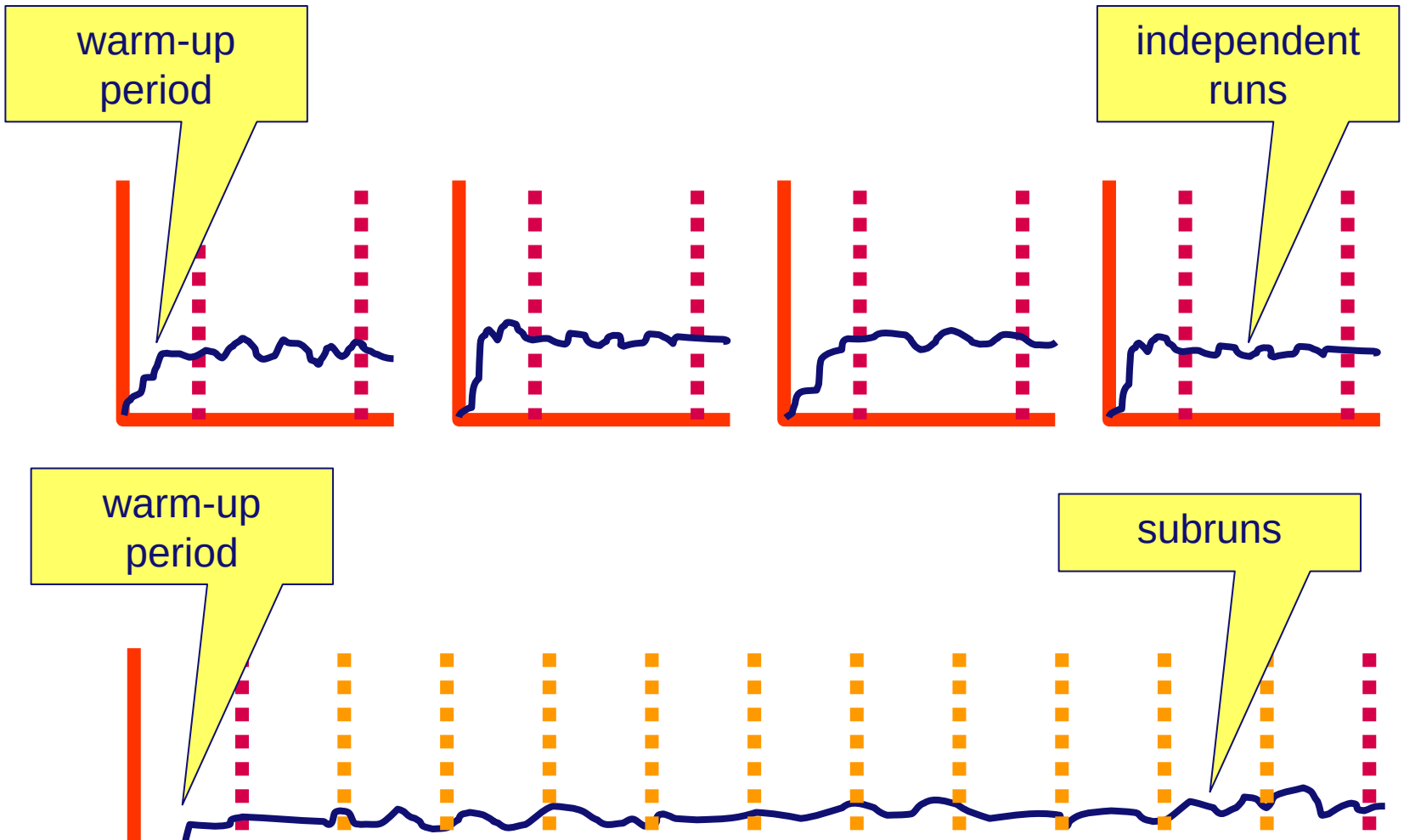
Problem 3: Wrong objectives

traditional steady state analysis focusing on long-term average behavior



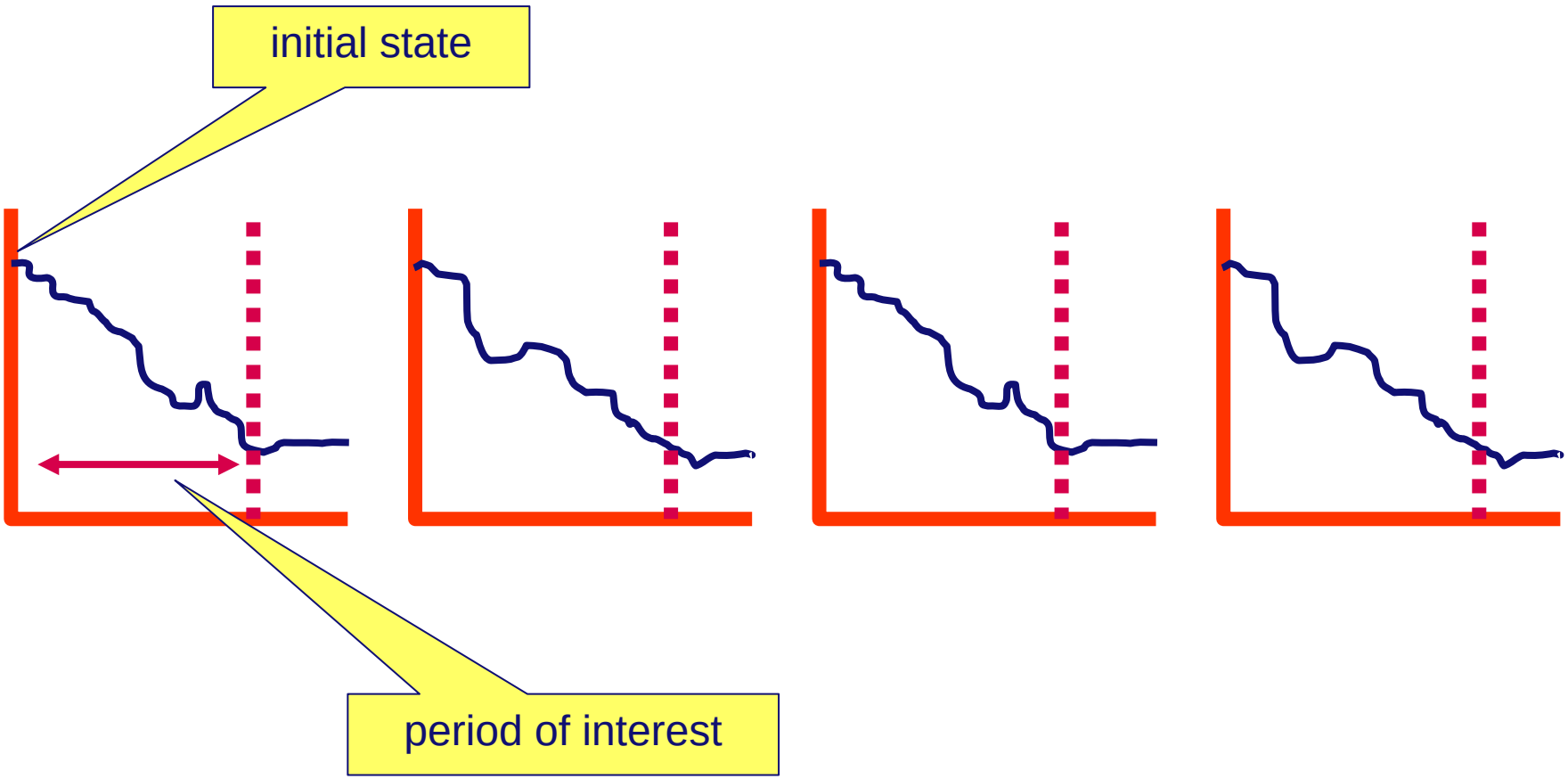
classical focus of simulation (tools)

Steady state analysis



(re-)design-time analysis, i.e., not for operational decision making

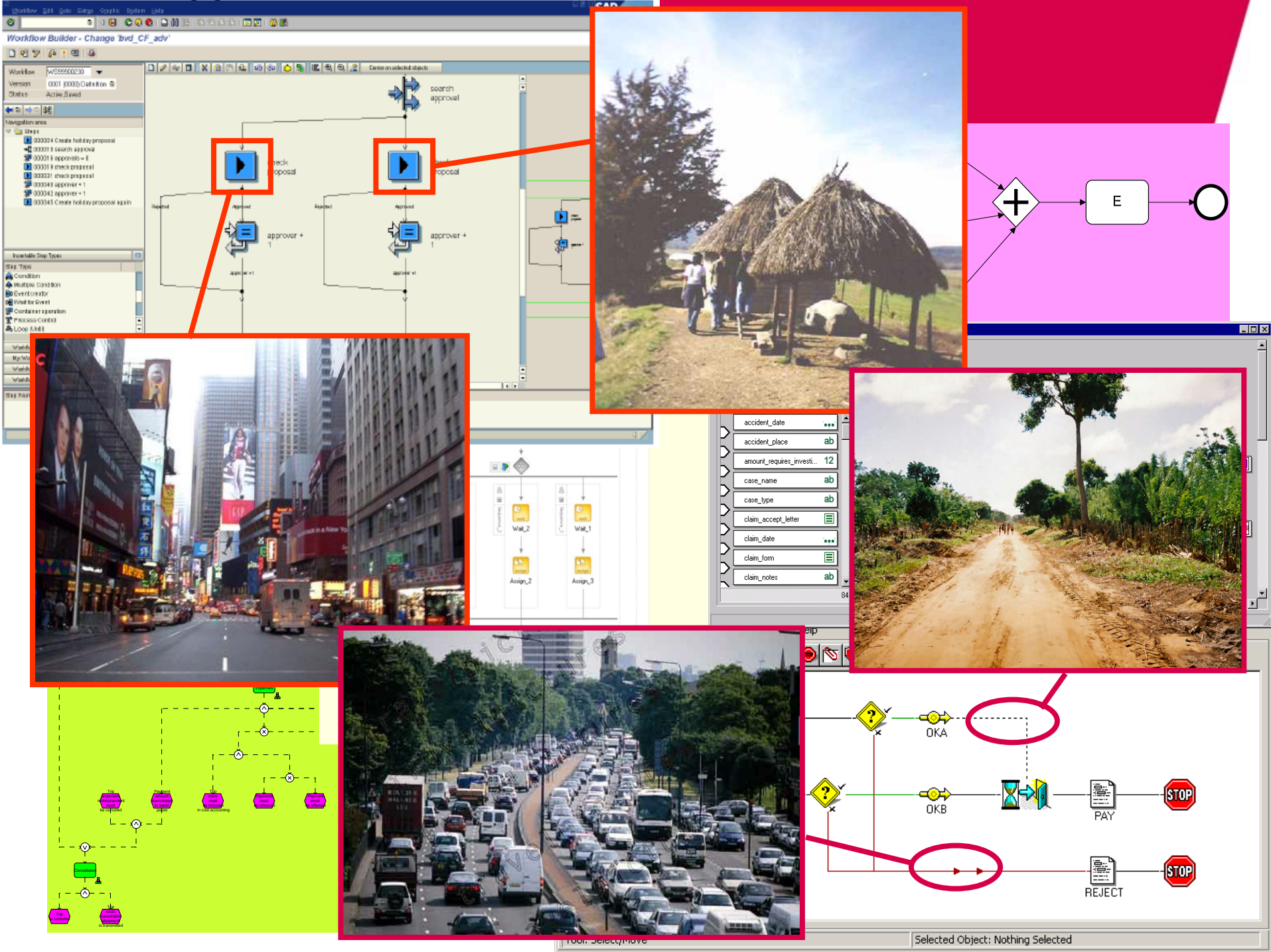
Transient analysis

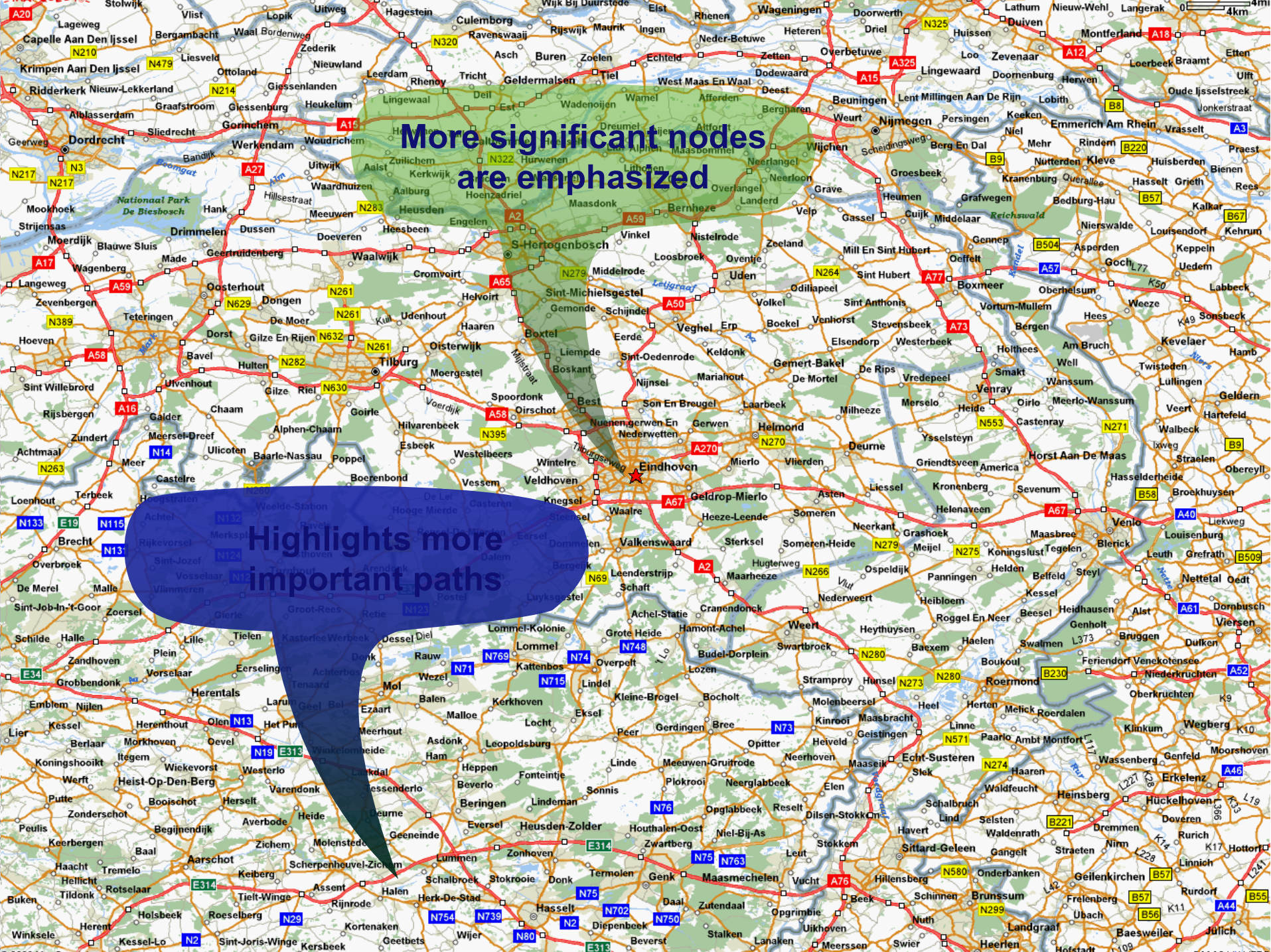


Steady-state may not exist and may not be relevant!

Problem 4: Inadequate visualization

- Simulation tools show "tables" and "graphs", and, if they are advanced also "token game animations".
- Management dashboards show "tables" and "graphs", and, if they are advanced also "speedometers", etc.
- Problems:
  - Observing the "simulated world" and the "real world" should be unified to allow for a better interpretation of the results.
  - Management dashboards are not looking "inside the process".
  - Process visualization is rather primitive and tries to show design artifacts rather than the process itself.


Workflow Builder - Change 'bvd_CF_adv'
search approval
approver + 1
accident_date
accident_place
amount_requires_investi...
case_name
case_type
claim_accept_letter
claim_date
claim_form
claim_notes
Wait_2
Wait_1
Assign_2
Assign_3
E
OKA
OKB
PAY
REJECT
STOP
STOP
Selected Object: Nothing Selected


More significant nodes are emphasized
Highlights more important paths

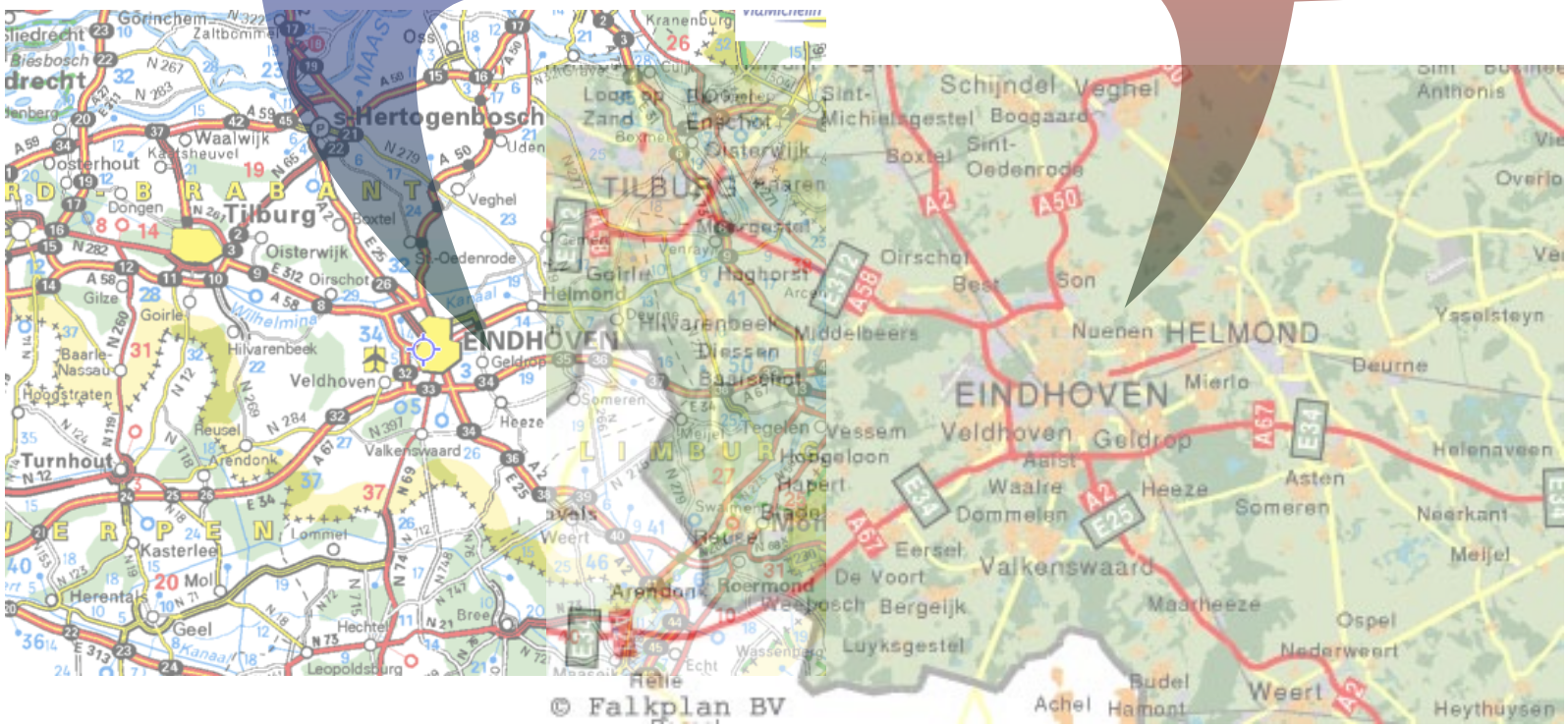
More to learn from maps...

Aggregation

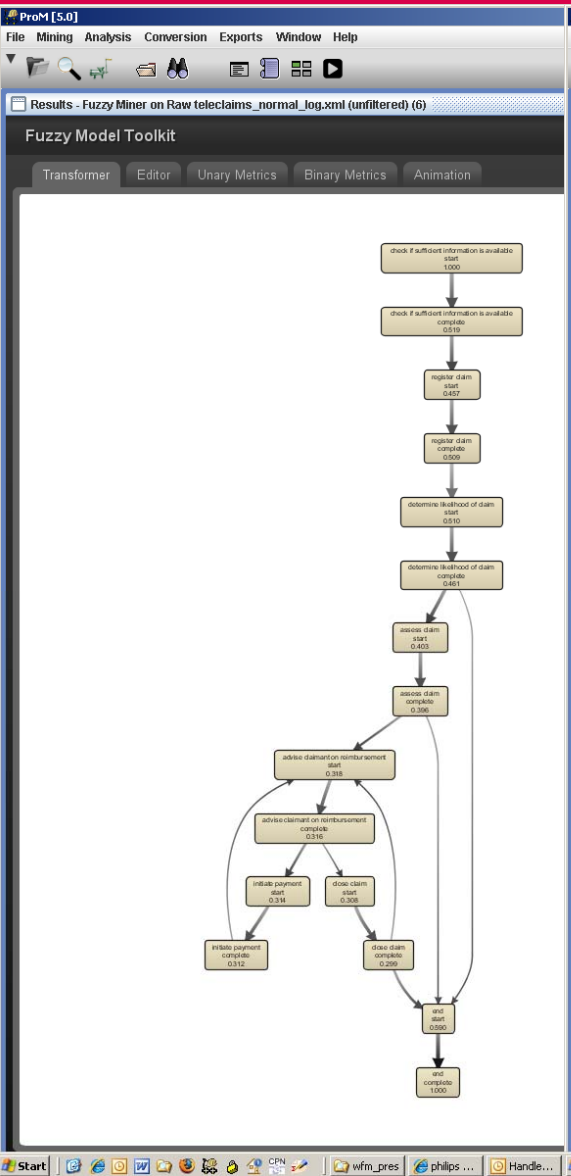
Clustering of coherent, less significant structures

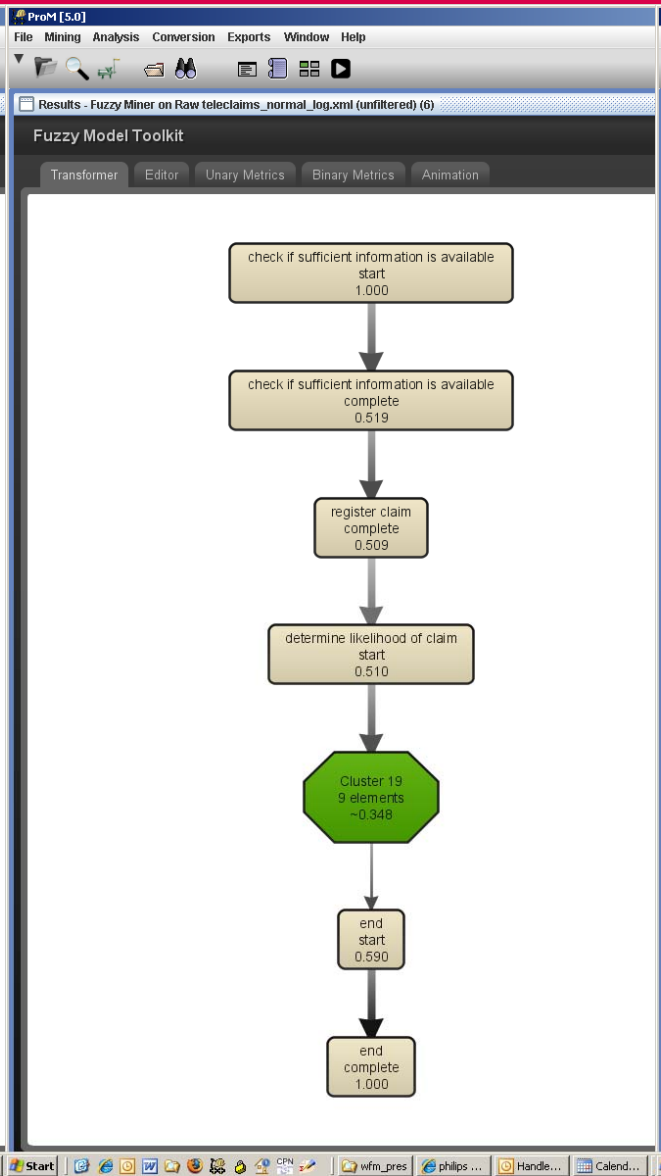
Abstraction

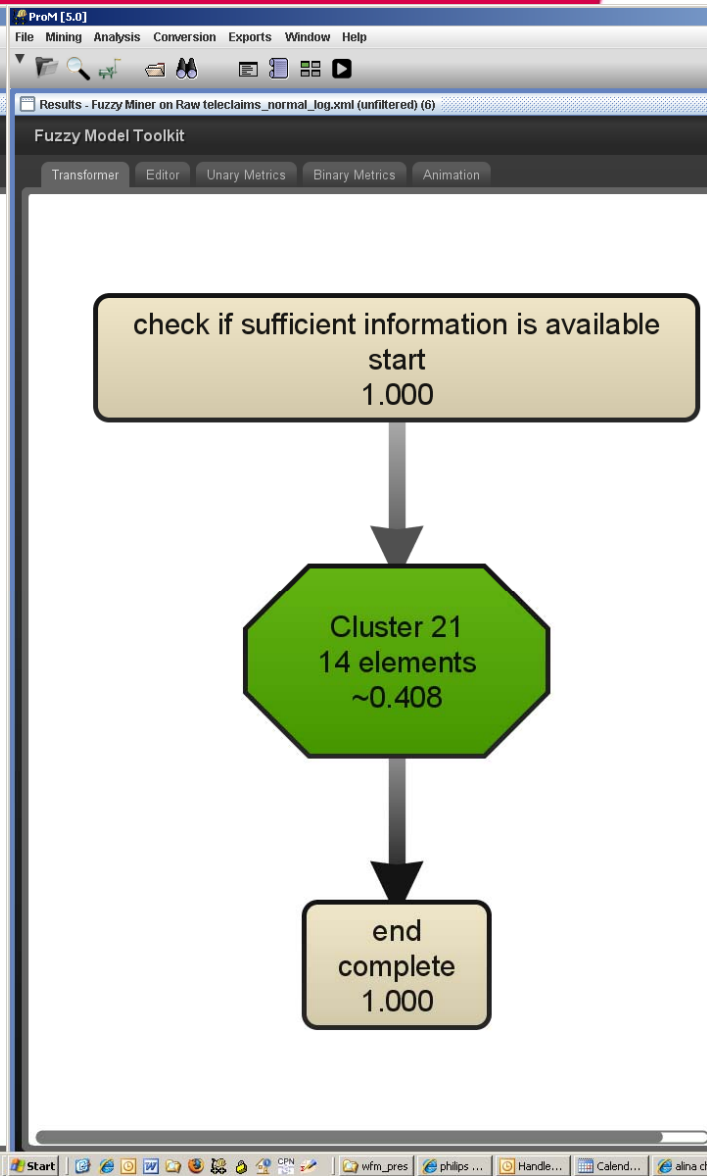
Removing isolated, less significant structures



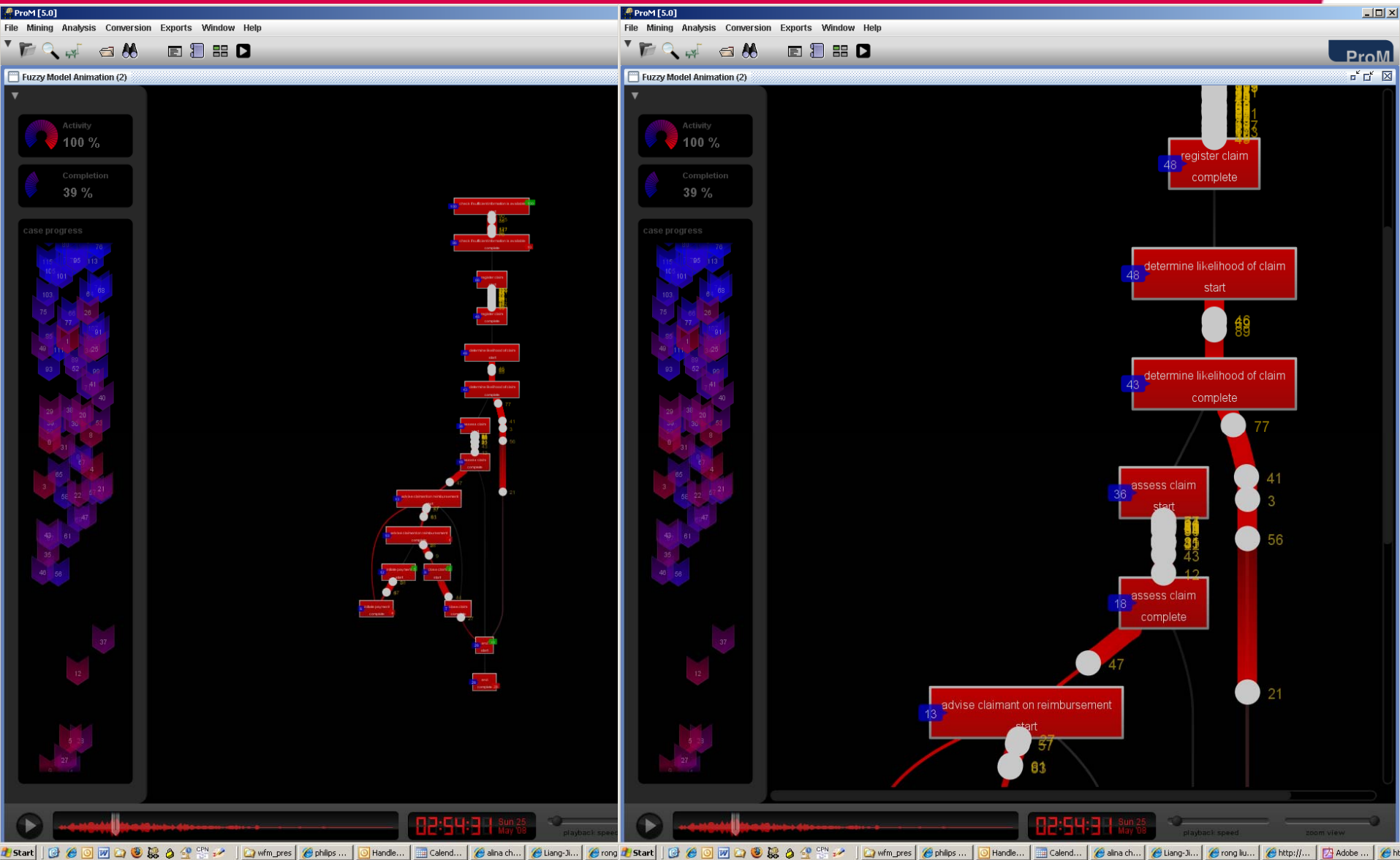
Fuzzy miner (Christian Güther)



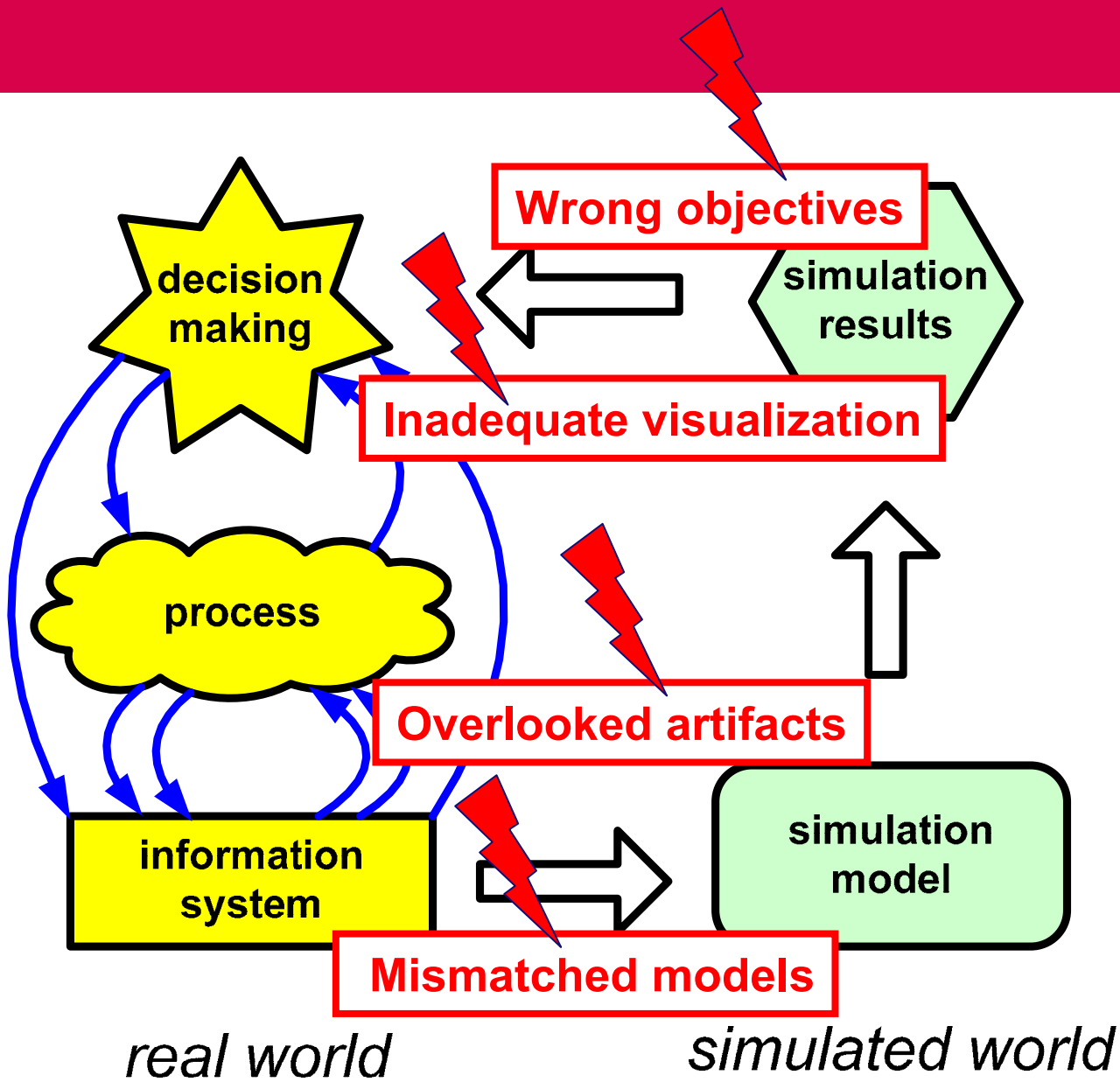




Showing reality (Christian Güther)



Overview problems



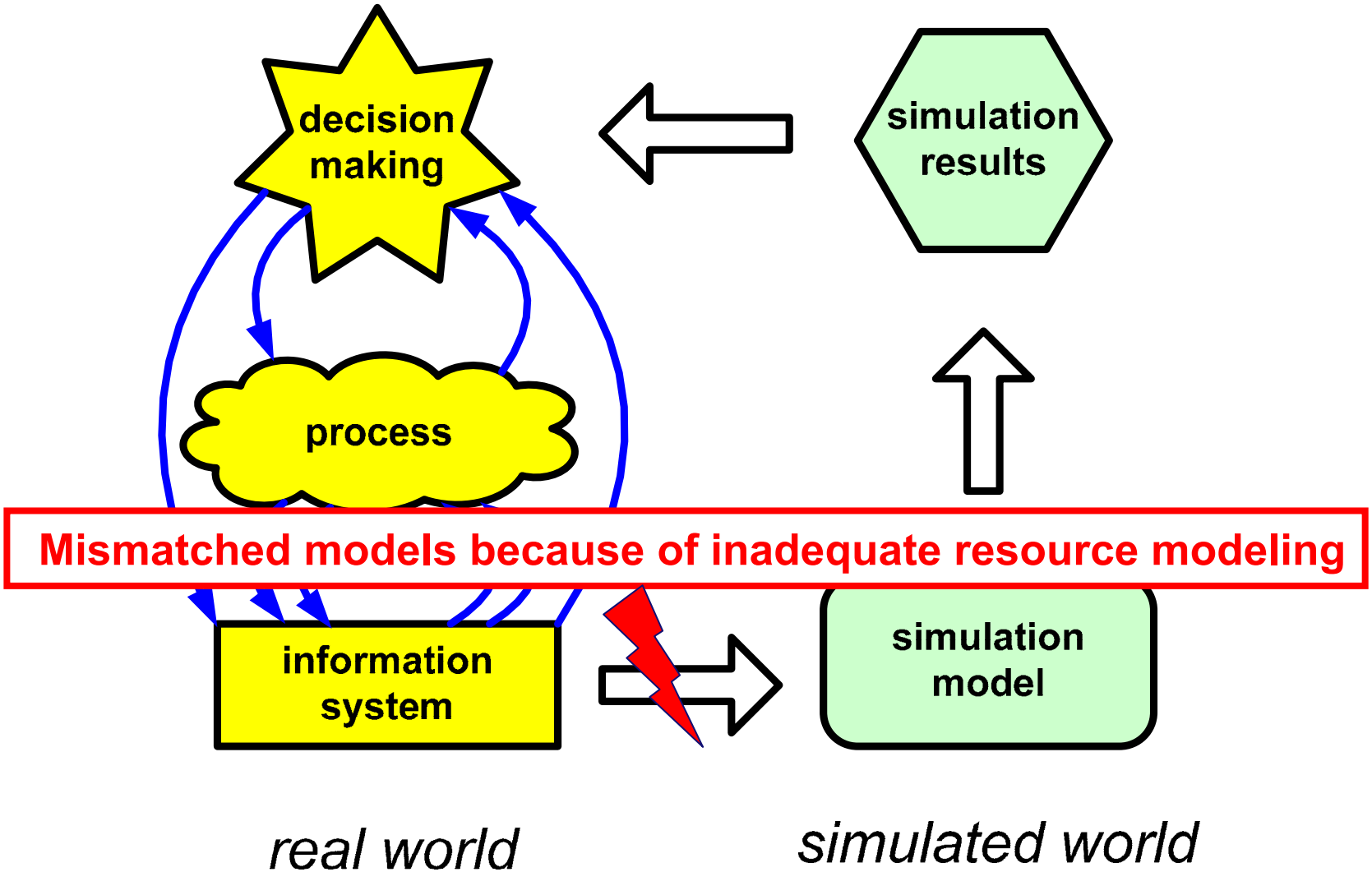
Partial solution:
Chunks

W.M.P. van der Aalst, J. Nakatumba, A. Rozinat and N. Russell. Business Process Simulation: How to Get it Right? *Handbook of Business Process Management,* 2008 (to appear, available as BPM Center Report).

TU/e Technische Universiteit Eindhoven University of Technology

Where innovation starts

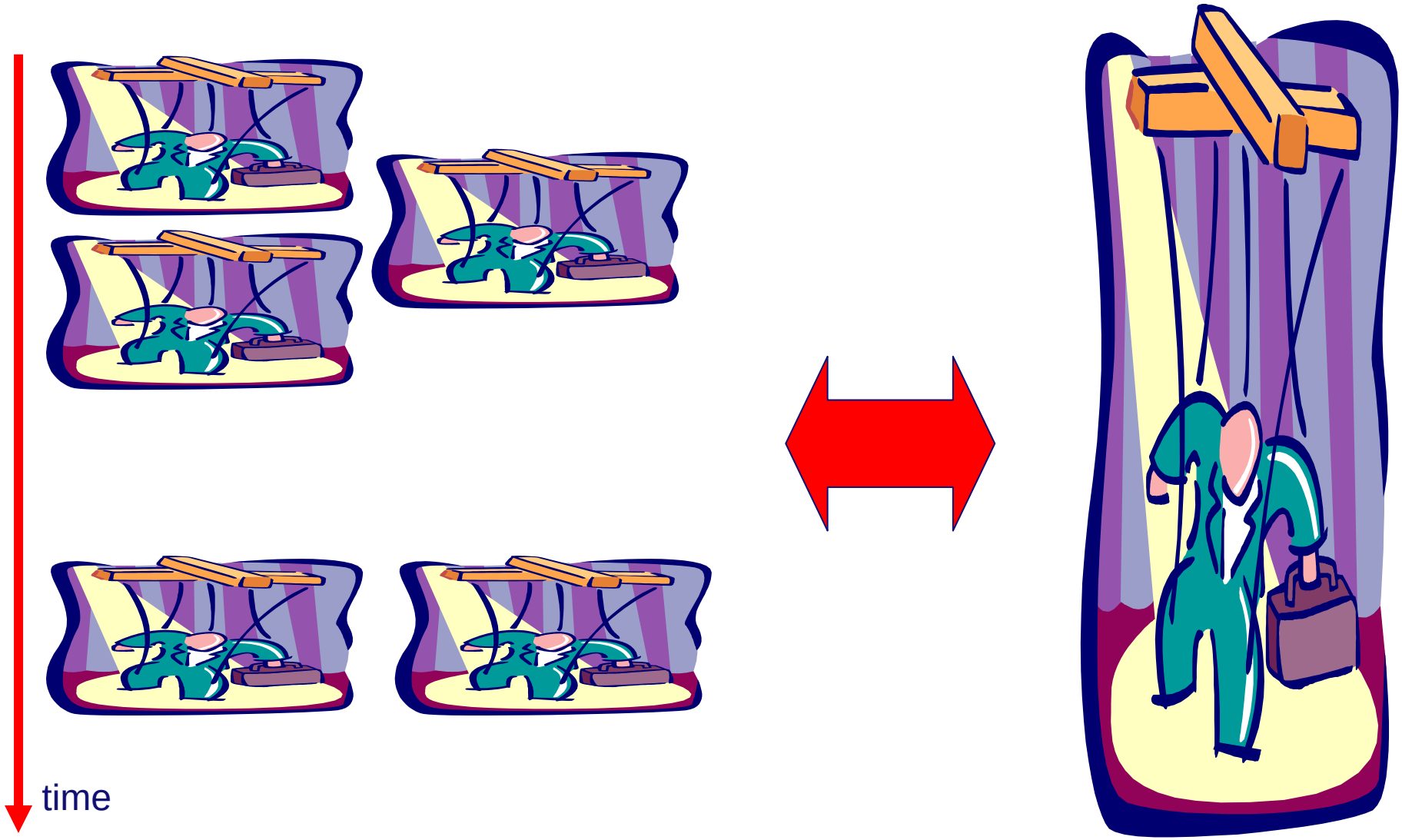
Focus


decision making
simulation results
process
Mismatched models because of inadequate resource modeling
information system
simulation model
real world
simulated world

Chunks?

- Remember: Simulated flow times of hours while real-life flow times are weeks!
- Hypothesis: Primary cause is inadequate resource modeling.
- People:
  - are working part-time, have breaks, holidays, sick leaves, etc.
  - are involved in multiple processes and need to assign priorities dynamically,
  - do not work at a constant speed,
  - etc.

5*0.2 ≠ 1

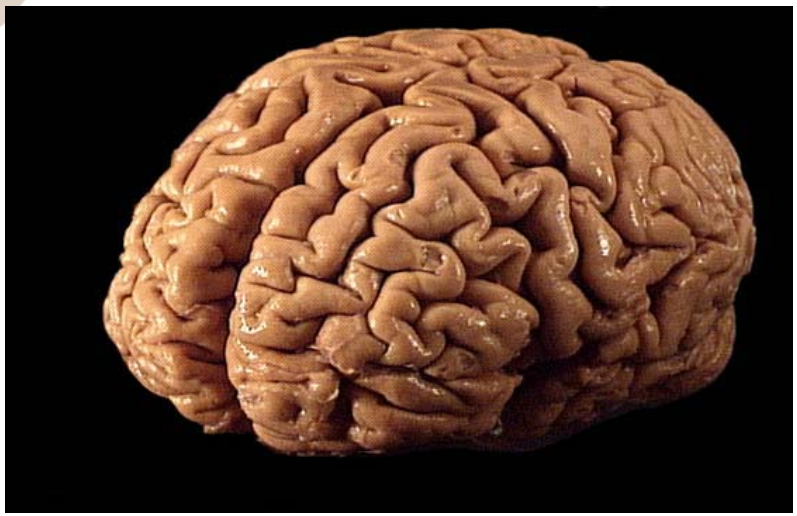


Classical simulation assumptions

- A resource is:
  - eager to start working,
  - dedicated to a single process,
  - works at a constant speed,
  - does not work in batches,
  - does not have coffee breaks,
  - etc.

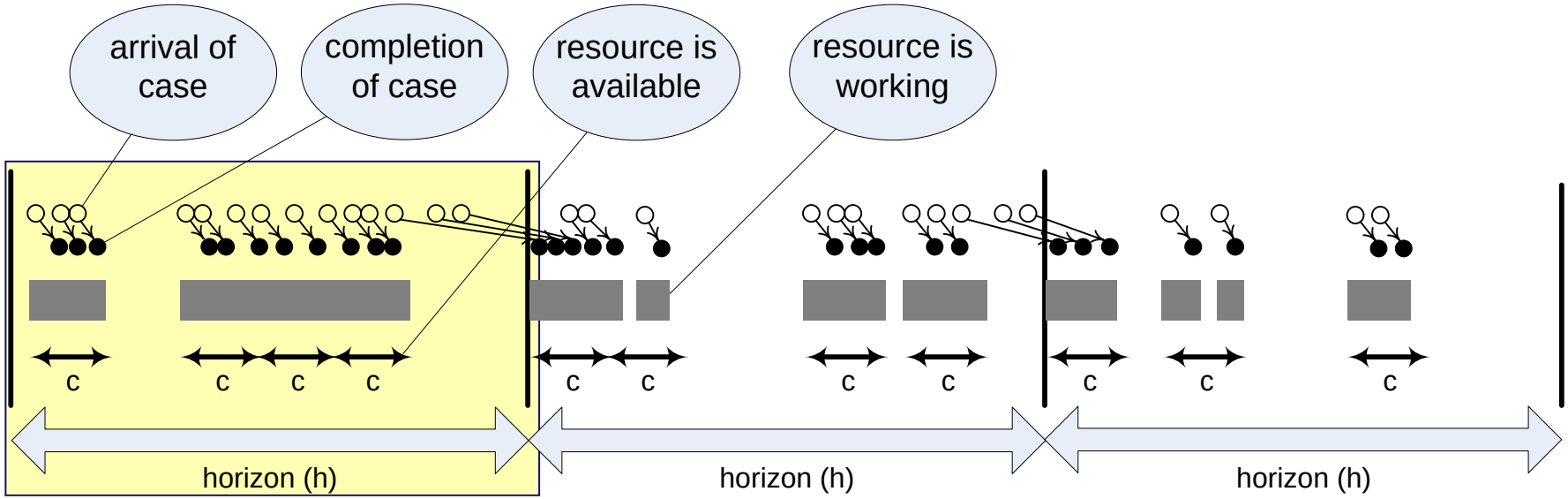
- Do you know this person?
- Chunks: towards a more accurate modeling of resource availability

Avoid modeling the world in a detailed manner



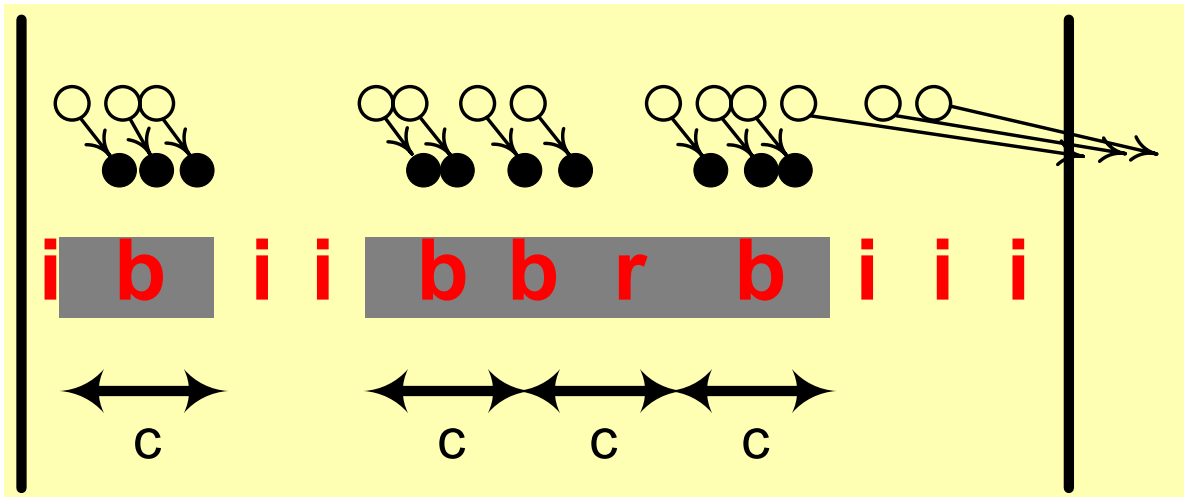
Goal: Characterize resource availability with just a few parameters

Chunks: Basic Idea

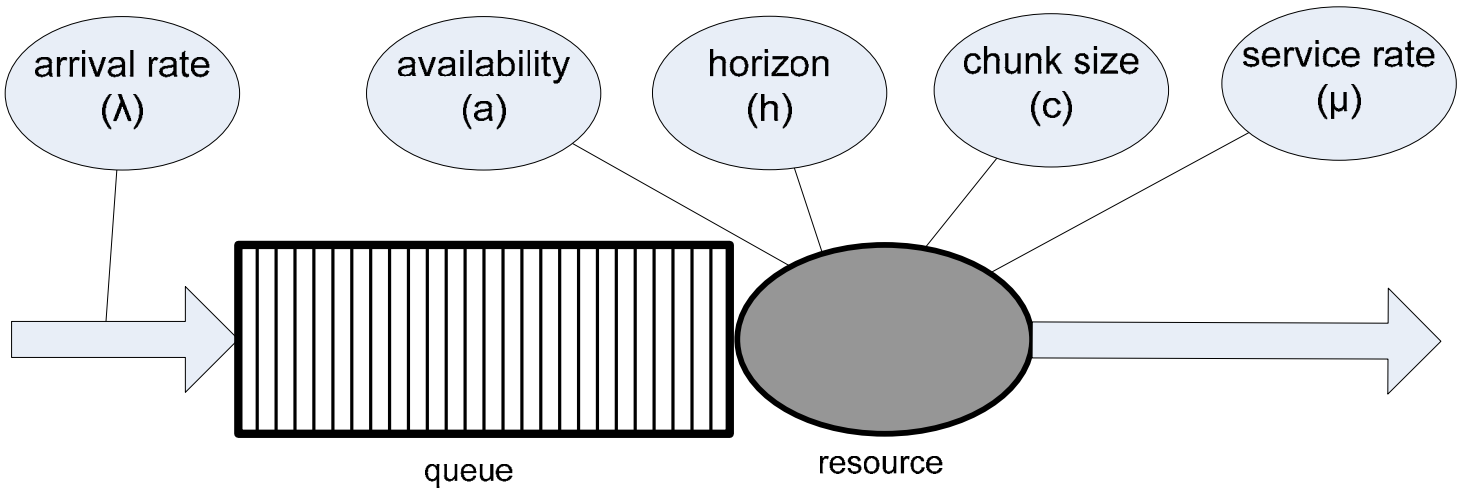


resource:

- **inactive**
- **ready**
- **busy**

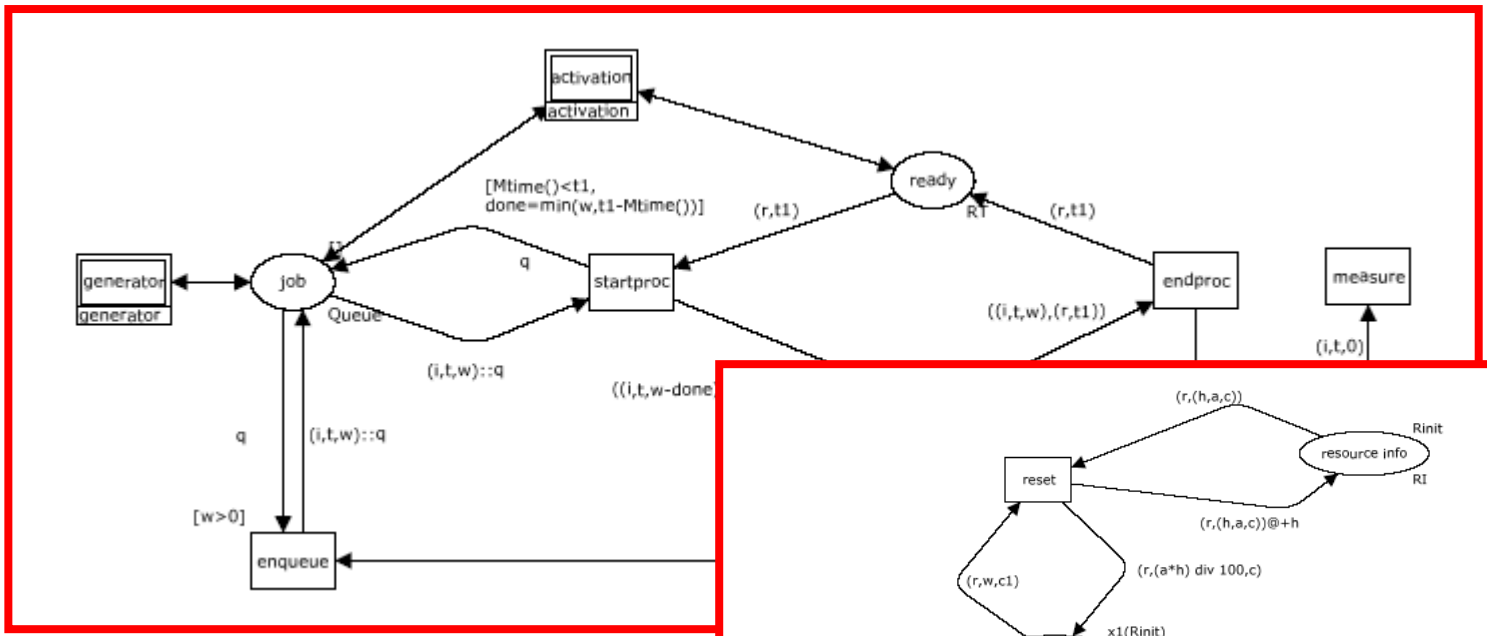


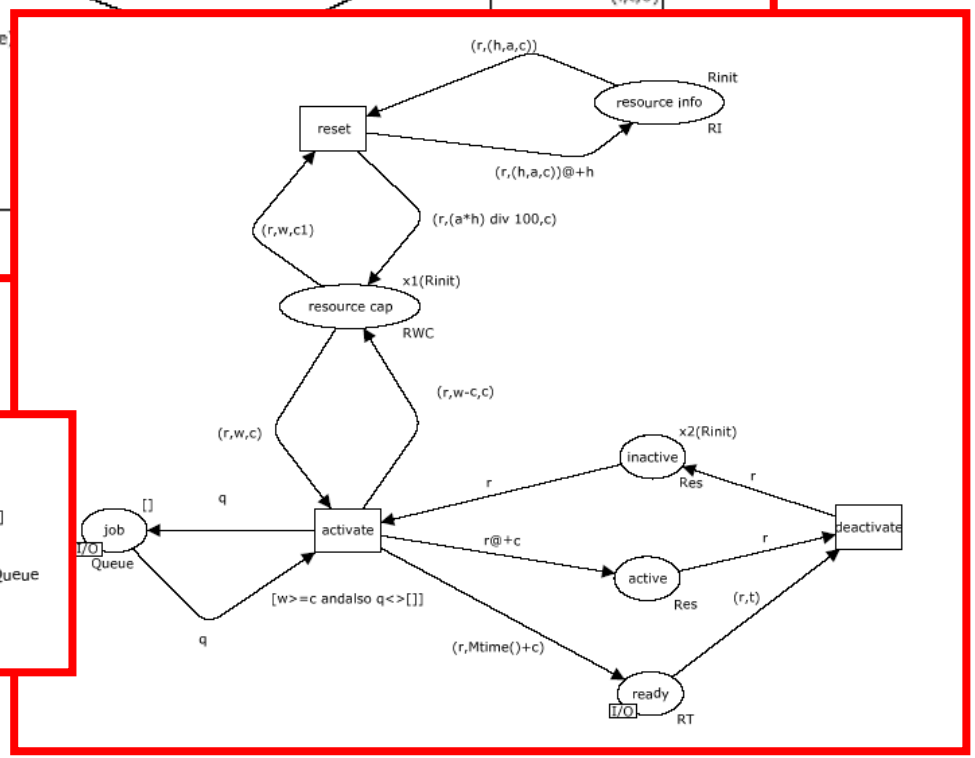
Parameters

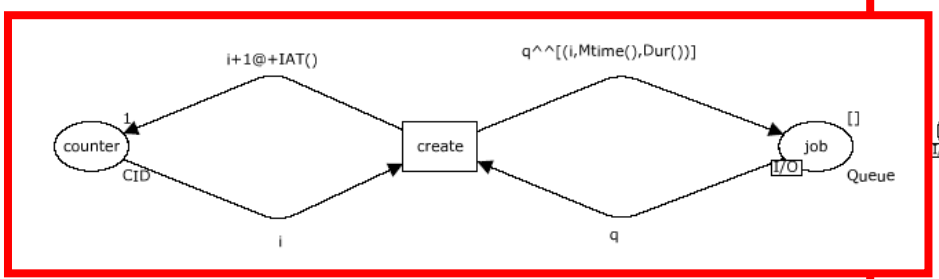


- ρ = λ/μ ≤ a, i.e., utilization is less than availability
- c ≤ h, i.e., chunk size cannot be larger than the horizon
- (a*h) mod c = 0 in experiments to avoid unusable availability

CPN model


activation
activation
[Mtime()<t1,
done=min(w,t1-Mtime())]
(r,t1)
ready
RT
(r,t1)
generator
generator
job
Queue
q
startproc
endproc
measure
((i,t,w),(r,t1))
(i,t,0)
(i,t,w)::q
((i,t,w-done)
q
(i,t,w)::q
[w>0]
enqueue


(r,(h,a,c))
Rinit
resource info
RI
reset
(r,(h,a,c))@+h
(r,w,c1)
(r,(a*h) div 100,c)
x1(Rinit)
resource cap
RWC
(r,w-c,c)
(r,w,c)
x2(Rinit)
inactive
Res
r
r
[]
q
job
I/O
Queue
activate
r@+c
active
Res
r
deactivate
[w>=c andalso q<>[]]
q
(r,t)
(r,Mtime()+c)
ready
I/O
RT


i+1@+IAT()
q^^[(i,Mtime(),Dur())]
1
counter
CID
create
[]
job
I/O
Queue
i
q

Effect of availability (a)

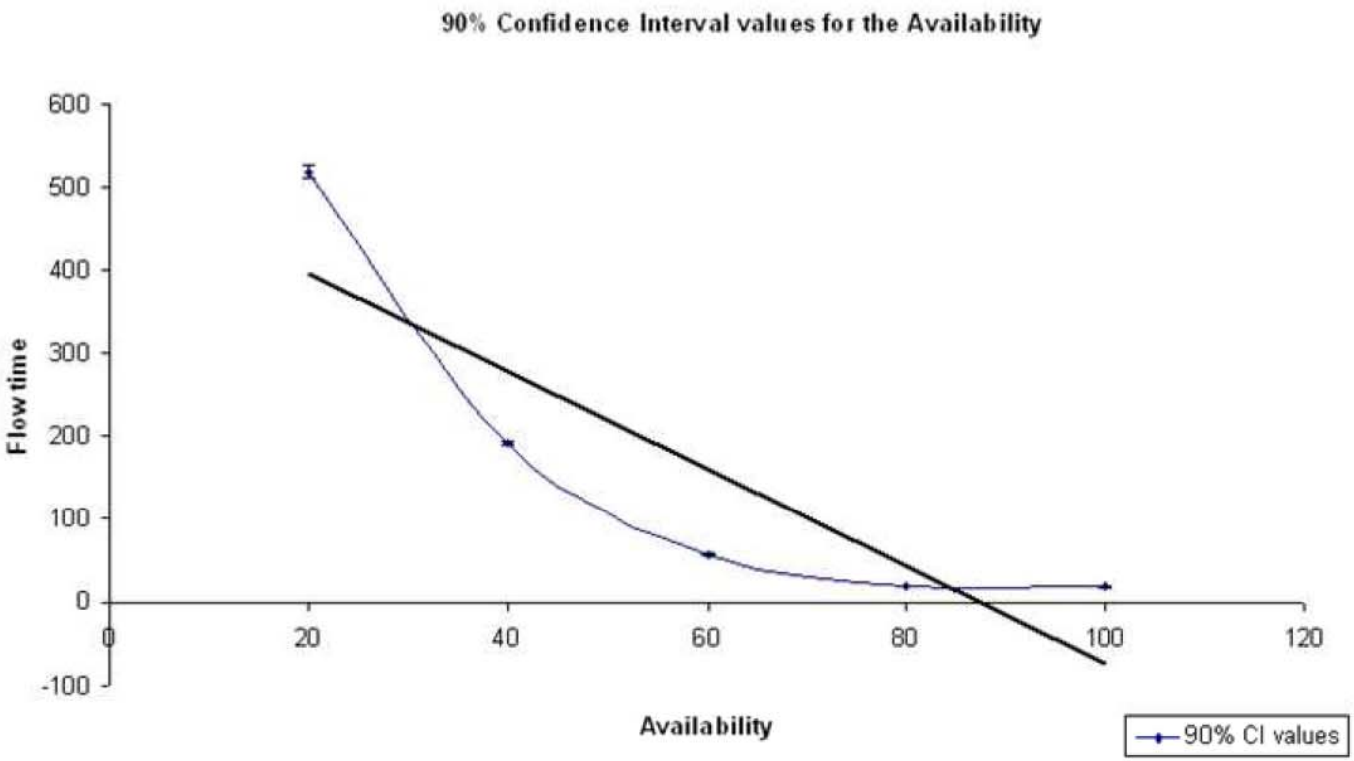


Fig. 9. Graph showing availability against flow time ($\lambda = \frac{1}{100}$, $\mu = \frac{1}{15}$, $\rho = 0.15$, $c = 200$, and $h = 1000$). The flow time reduces as the availability increases. (The straight line shows the trend using linear regression.)

Effect of chunk size (c)

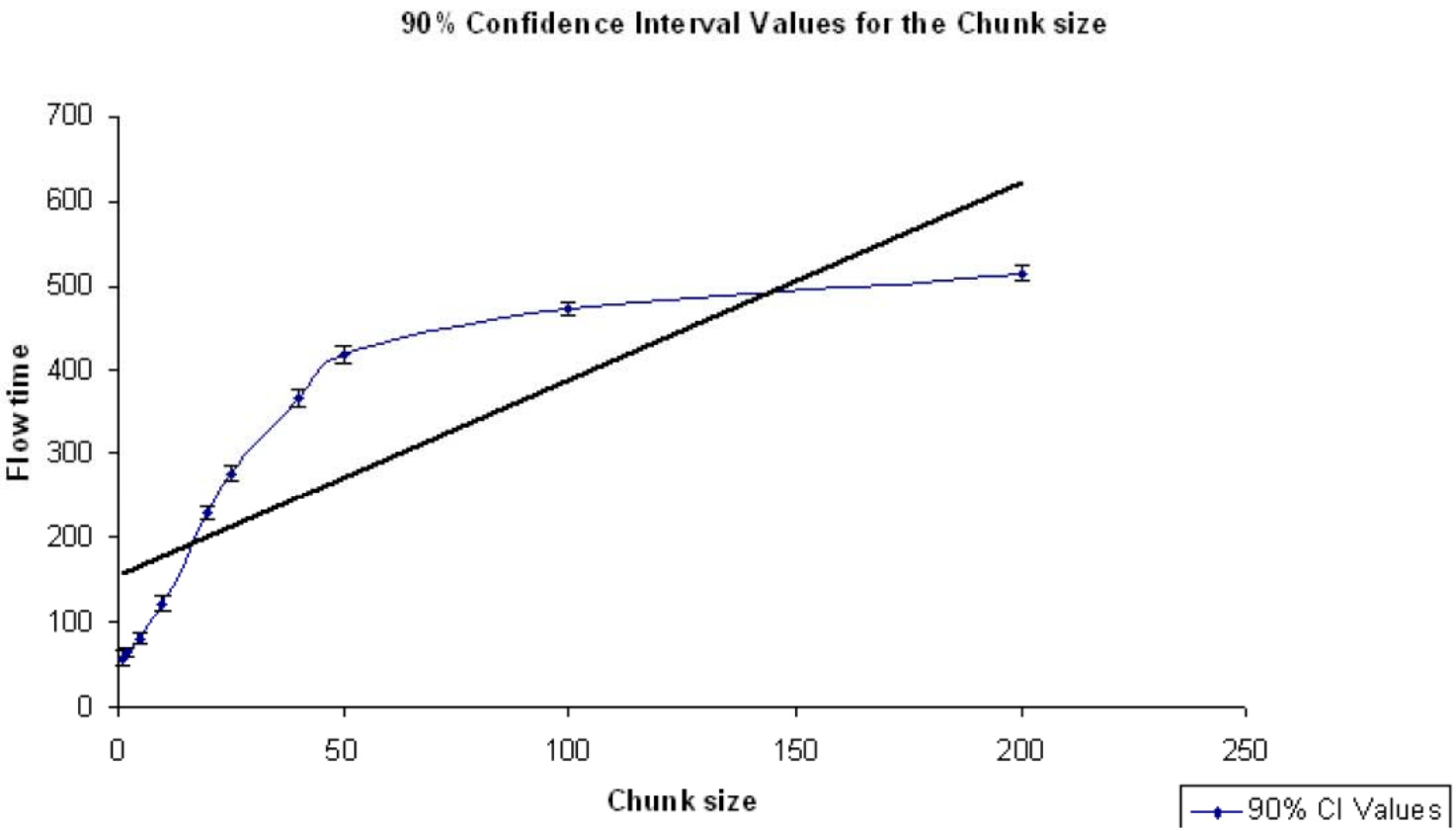


Fig. 10. Graph showing chunk size against flow time ($\lambda = \frac{1}{100}$, $\mu = \frac{1}{15}$, $\rho = 0.15$, $a = 0.2$, and $h = 1000$). The flow time increases as the chunk size increases.

Effect of horizon (h)

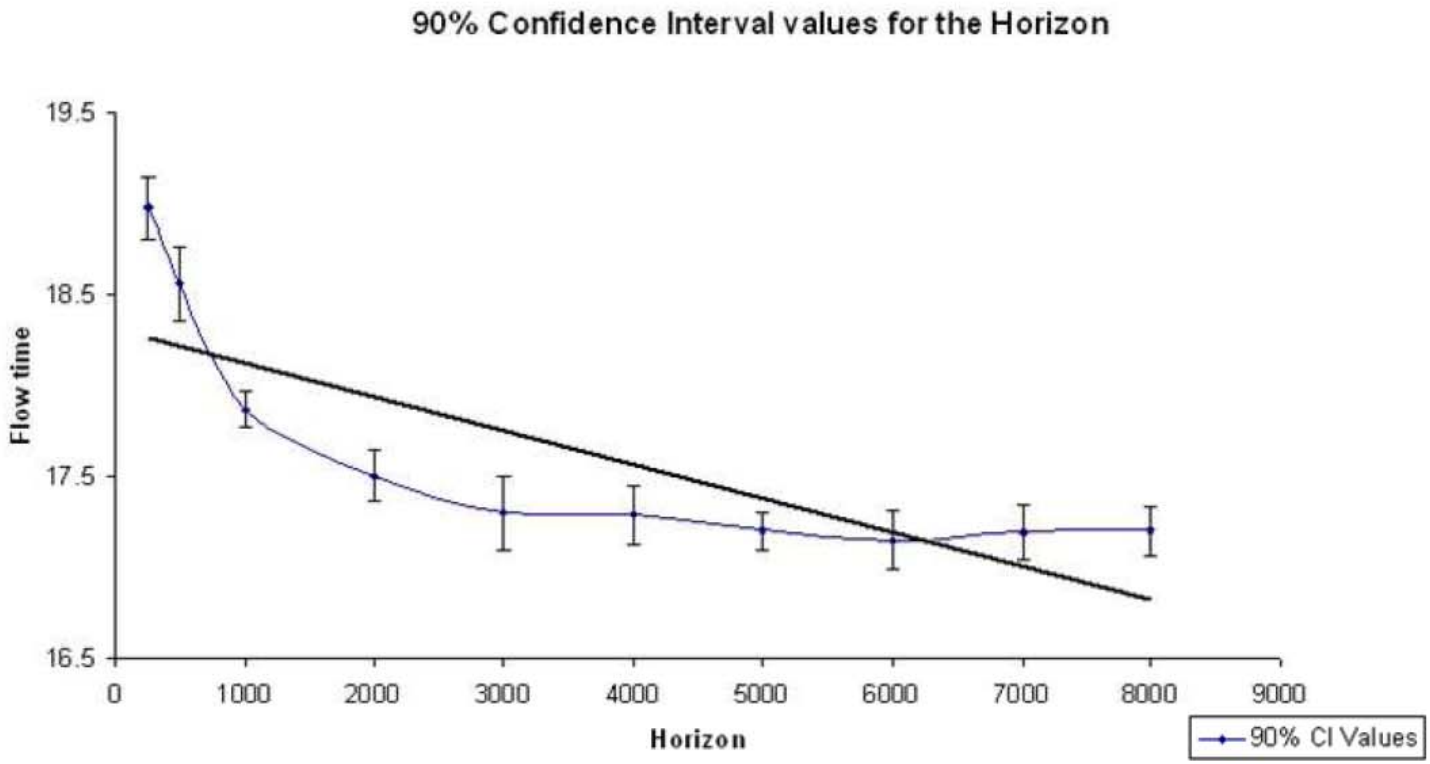


Fig. 11. Graph showing the horizon against the flow times ($\lambda = \frac{1}{100}$, $\mu = \frac{1}{15}$, $\rho = 0.15$, $c = 200$, and $a = 0.8$). The flow time decreases as the horizon increases.

Effect of utilization (ρ)

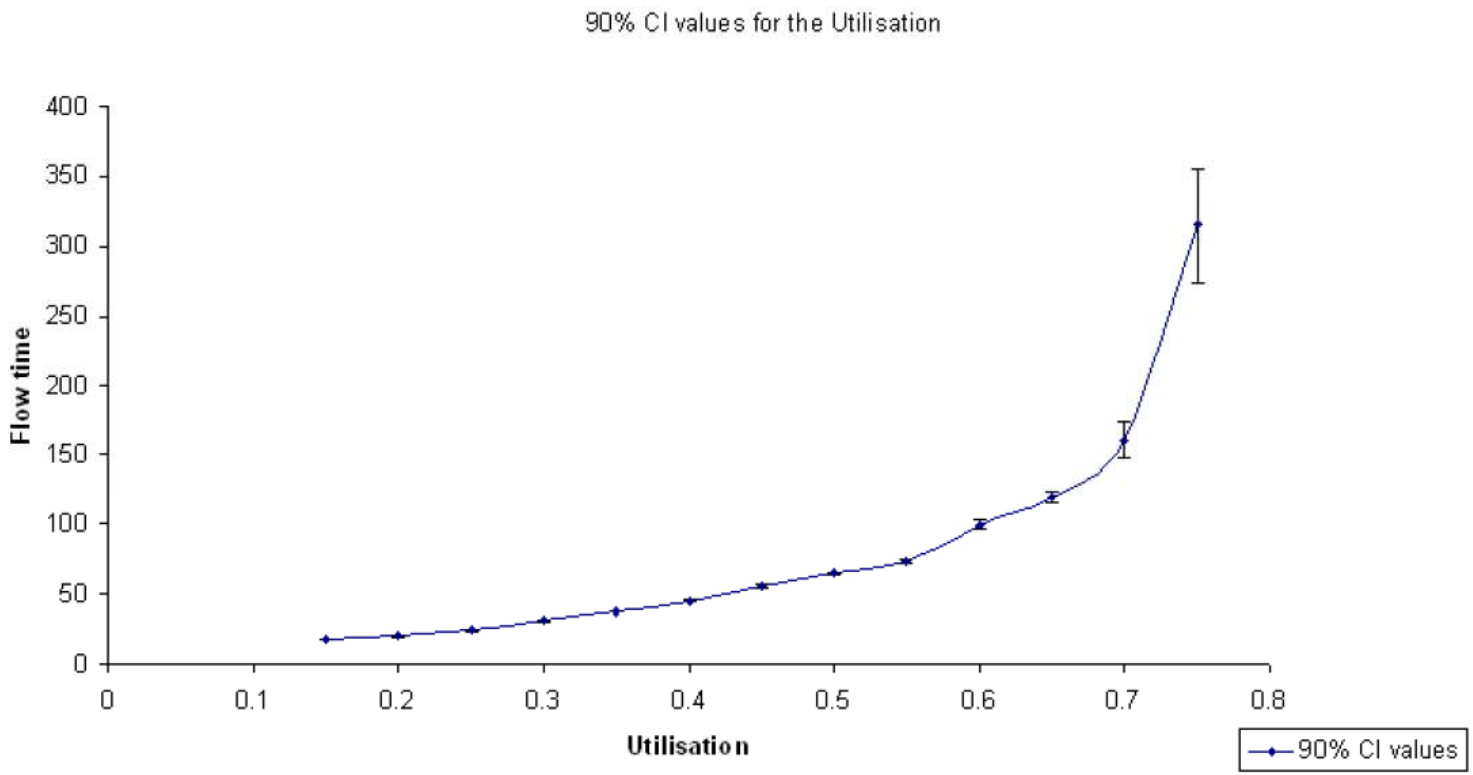
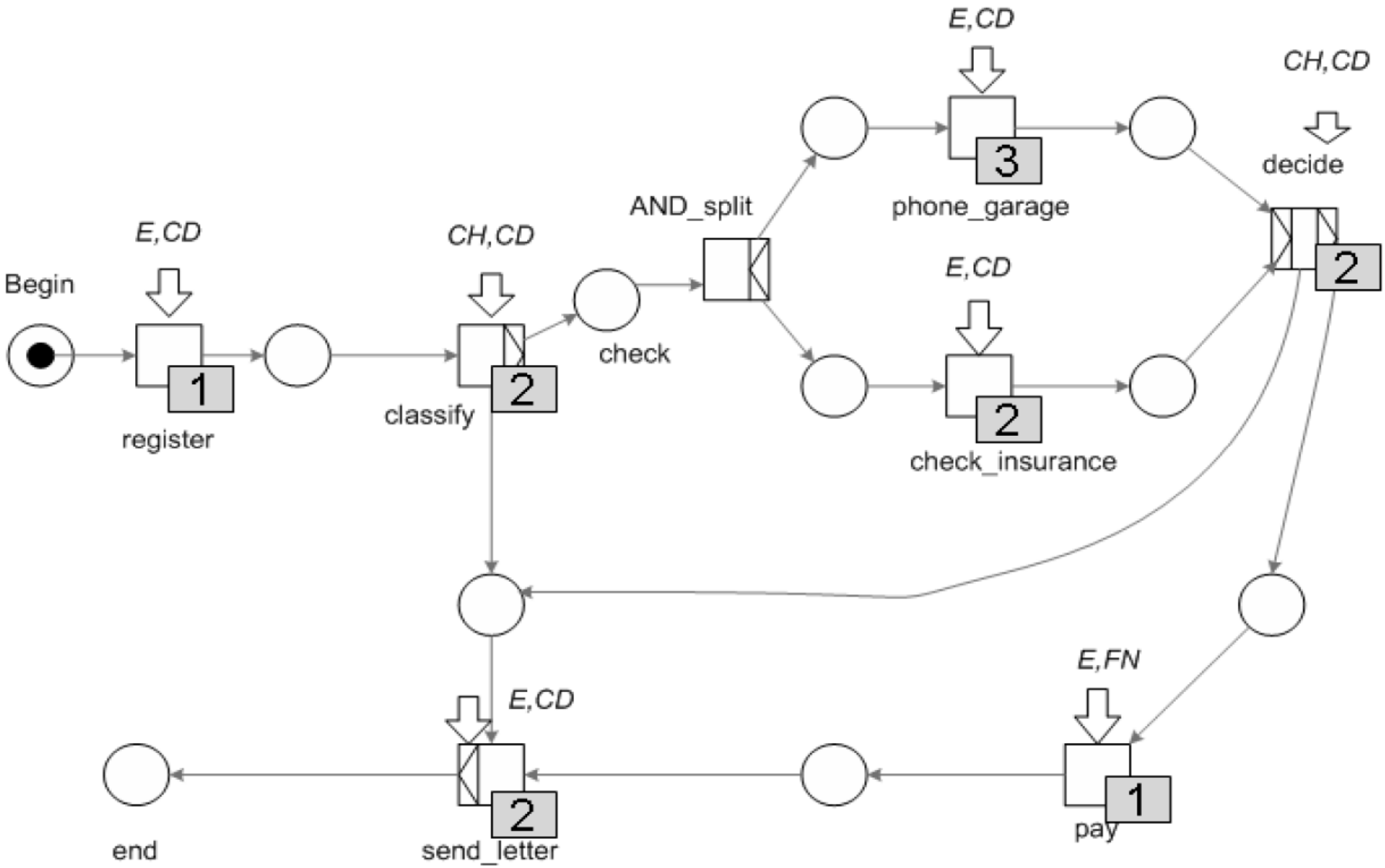
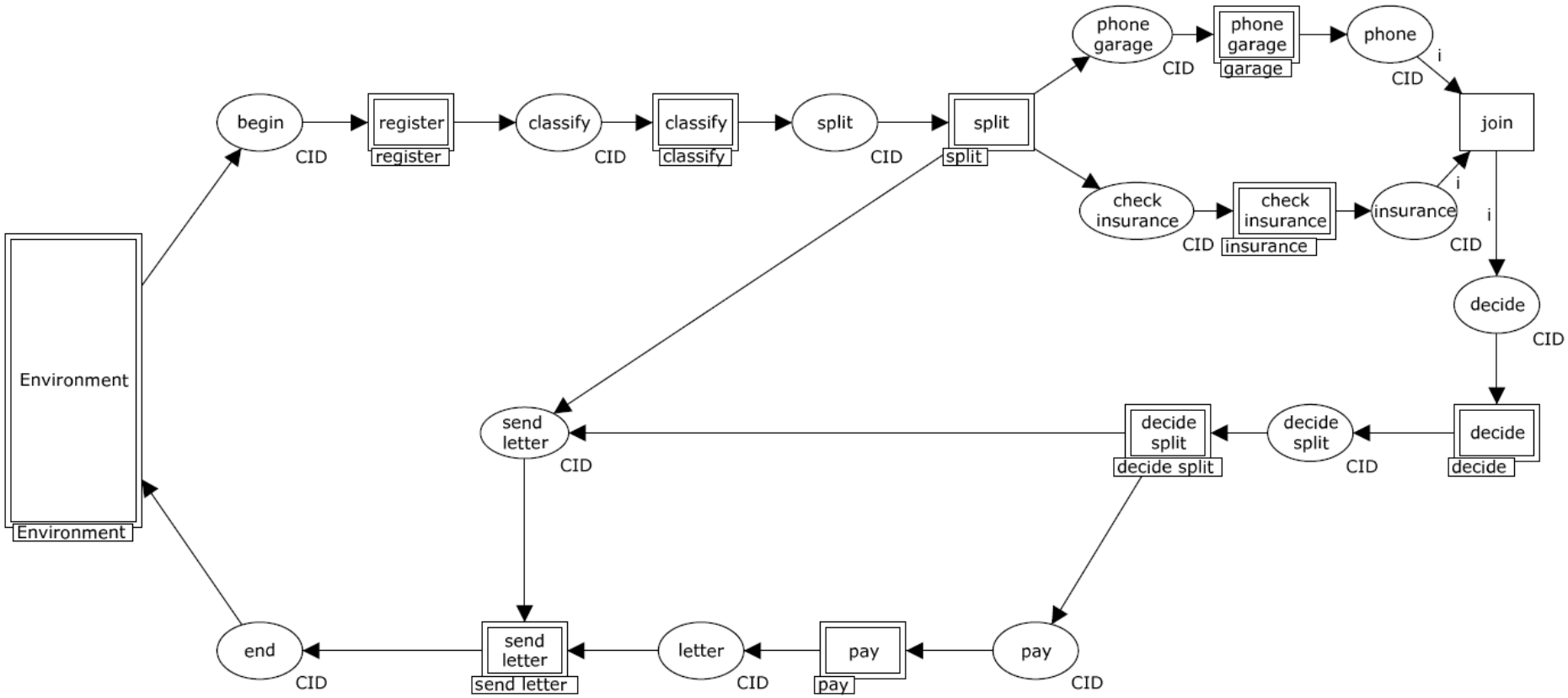


Fig. 12. Graph showing utilization against flow time ($\mu = \frac{1}{15}$, $c = 200$, $a = 0.8$, and $h = 1000$). The flow time increases as utilization increases.

Experiment: Note multiple resources and potential accumulation of effects



CPN model


Environment
Environment
begin
CID
register
register
classify
CID
classify
classify
split
CID
split
split
phone garage
CID
phone garage
garage
phone
CID
i
join
check insurance
CID
check insurance
insurance
insurance
CID
i
i
decide
CID
send letter
CID
decide split
decide split
decide split
CID
decide
decide
end
CID
send letter
send letter
letter
CID
pay
pay
pay
CID

Some findings

		Parameters	Flow Time
a)		Base Case Scenario ($c = 5$, $h = 2000$, $\lambda = \frac{1}{50}$ and $a = 0.4$, see Appendix B for all other parameters)	757.6 $\pm$65.0
b)	i)	Divide the horizon by 20 ($h = 100$)	1218.9 $\pm$72.3
	ii)	Divide the horizon by 40 ($h = 50$)	1247.8 $\pm$51.8
c)	i)	Multiply the chunk size by 5 ($c = 25$)	1158.7 $\pm$47.2
	ii)	Multiply the chunk size by 20 ($c = 100$)	1698 $\pm$139
	iii)	Multiply the chunk size by 80 ($c = 400$)	1950 $\pm$83.7
	iv)	Multiply the chunk size by 160 ($c = 800$)	2025 $\pm$99
d)	i)	Decrease availability and arrival rate by 2 ($a = 0.2$, $\lambda = \frac{1}{100}$)	1634 $\pm$105
	ii)	Decrease availability and arrival rate by 4 ($a = 0.1$, $\lambda = \frac{1}{200}$)	3420.32 $\pm$252

"Chunks Conclusion"

- It is important not to assume that people are always available and eager to work when cases arrive.
- The assumptions heavily impacts flow time, e.g., the bigger the chunk size, the longer the flow times of cases.
- The "chunk model" is rather simple, however, the typical assumptions made in today's simulation tools (i.e. $a = 1$, $c = 0$, and h=inf), may result in flow times of minutes or hours while with more realistic settings for a, c, and h the flow time may go up to weeks or months and actually coincide with the actual flow times observed.

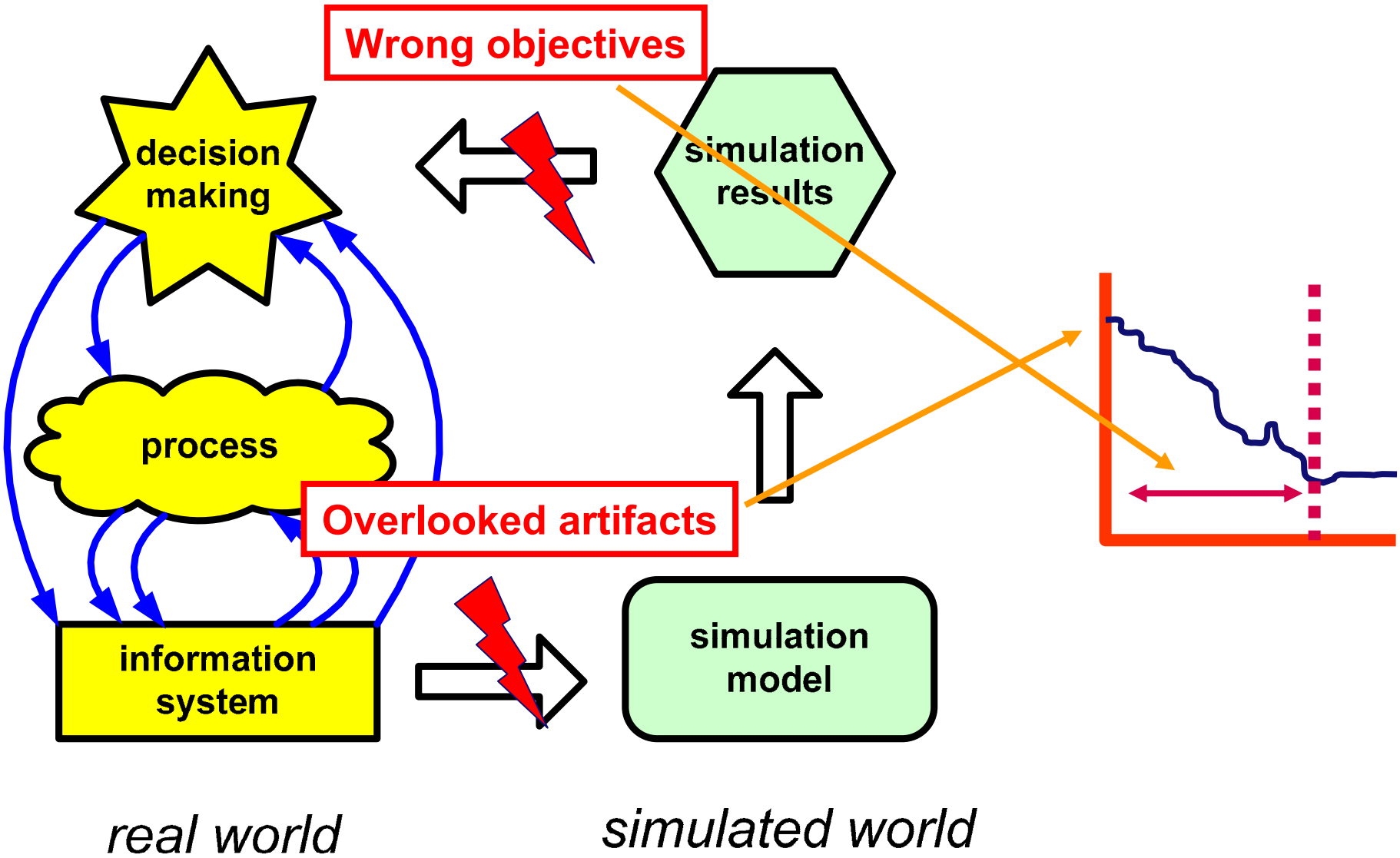
Partial solution:
Short-Term Simulation

A. Rozinat, M.T. Wynn, W.M.P. van der Aalst, A.H.M. ter Hofstede, and C. Fidge. Workflow Simulation for Operational Decision Support Using Design, Historic and State Information. *International Conference on Business Process Management (BPM 2008)*, 2008.

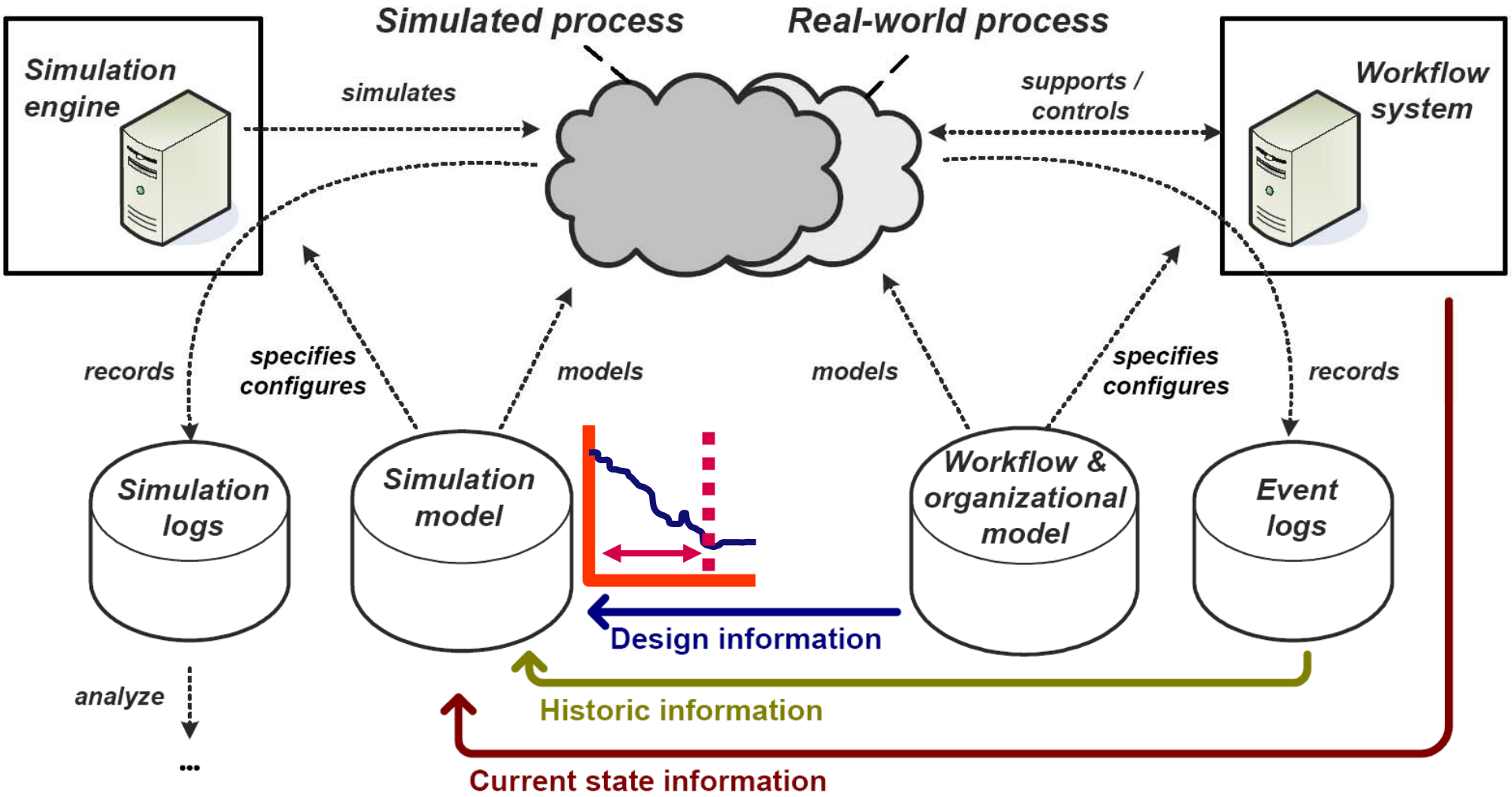
TU/e Technische Universiteit Eindhoven University of Technology

Where innovation starts

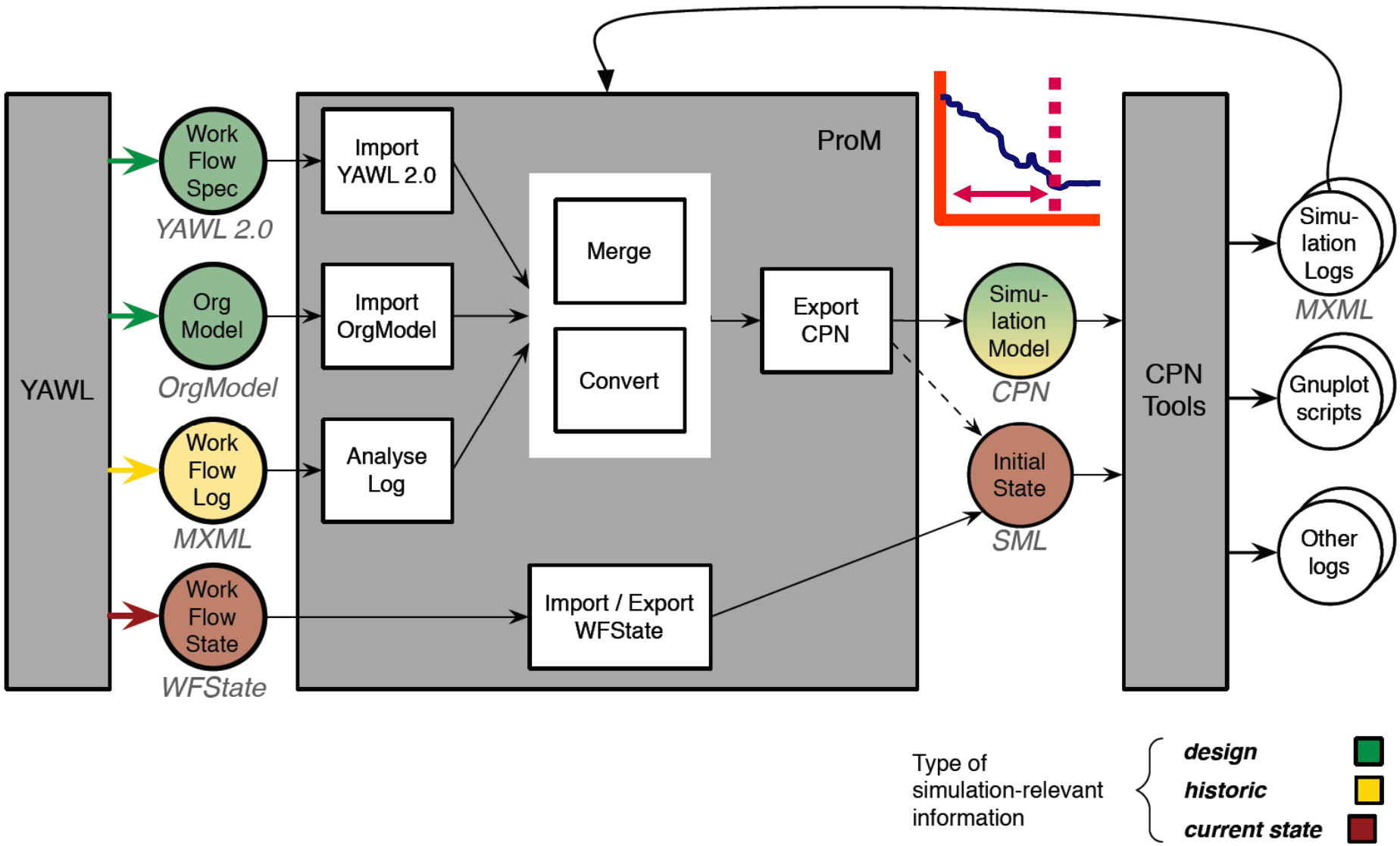
Focus: transient analysis using design, historic, and current state information.



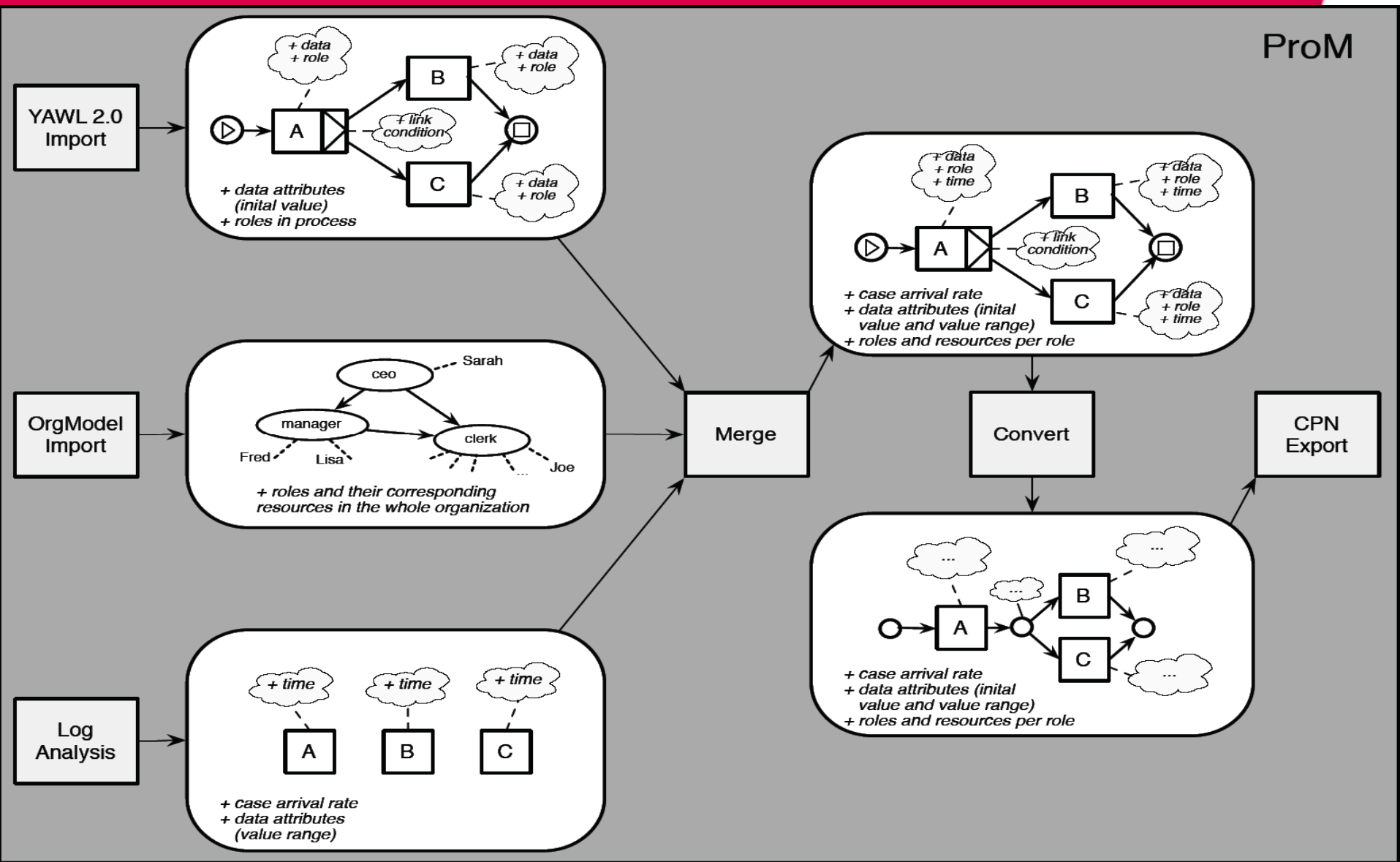
Overview: Short-Term Simulation



Implementation using YAWL, ProM, and CPN Tools

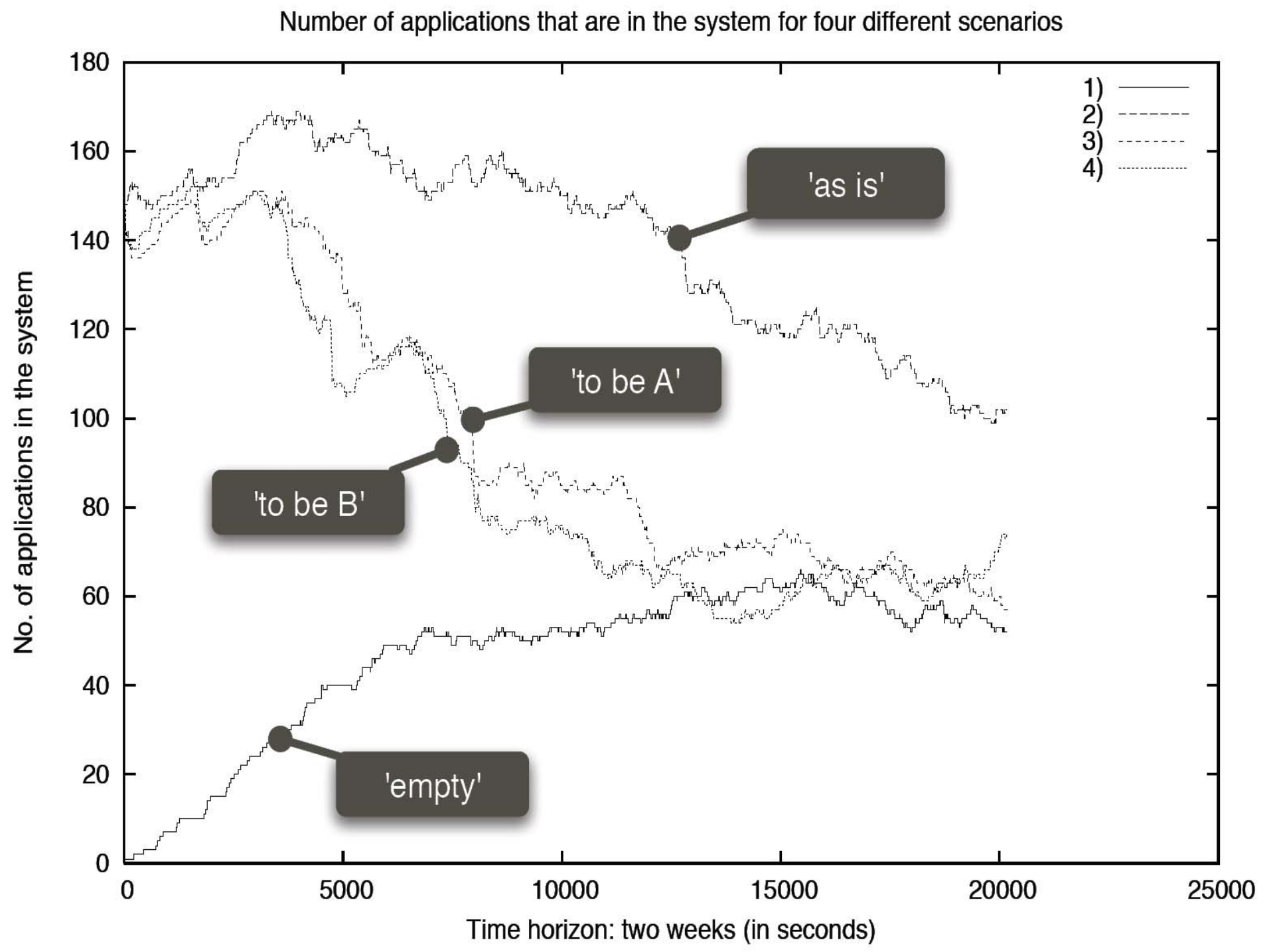


ProM: Merging and converting models covering different aspects

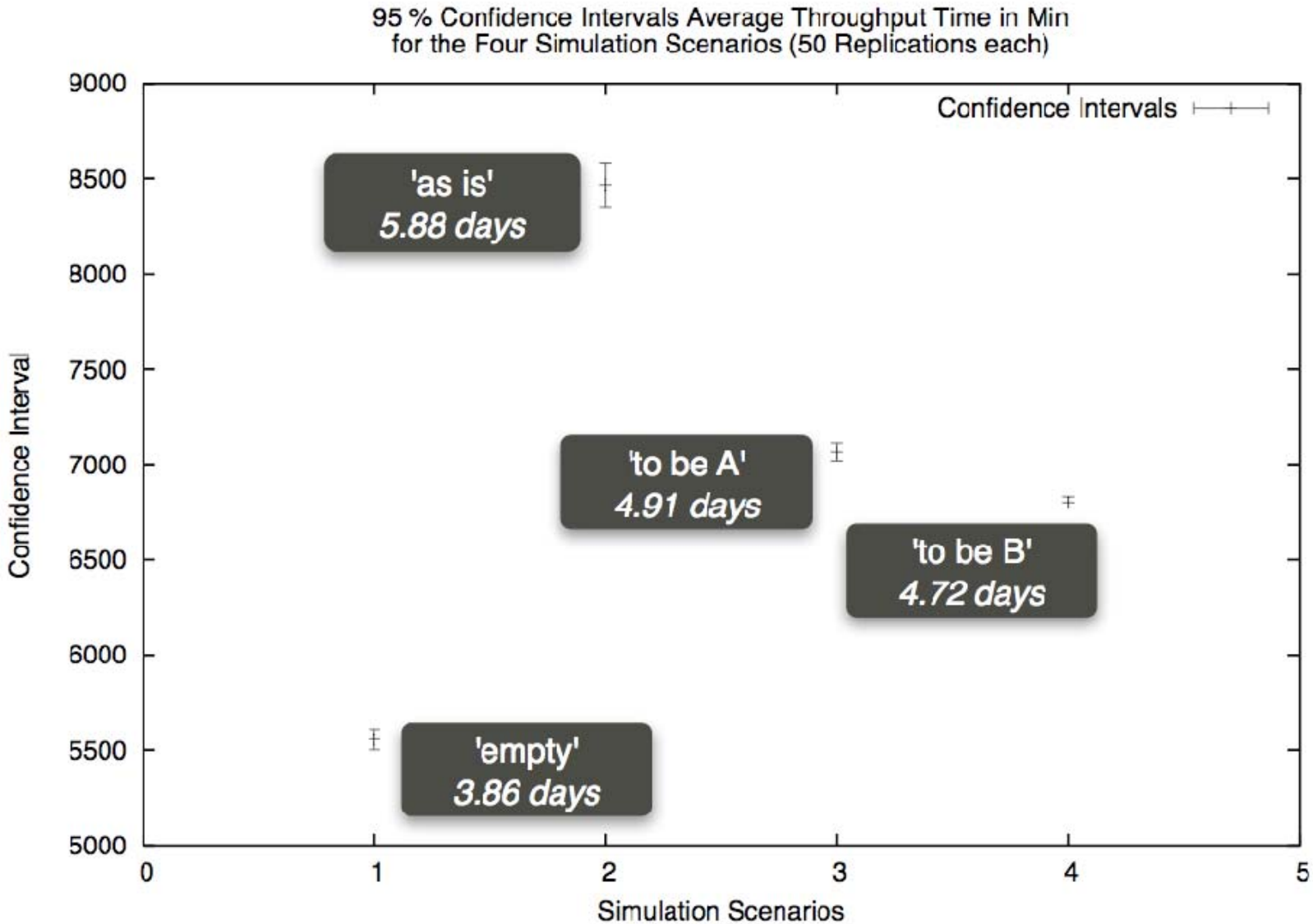


Example: 4 different simulation scenarios

1. An empty initial state (‘empty’)
2. After loading the current state file with the 150 applications currently in the system (‘as is’)
3. After loading the current state file but adding four extra resources (‘to be A’)
4. After loading the current state file and adding eight extra resources (‘to be B’)


Number of applications that are in the system for four different scenarios
No. of applications in the system
Time horizon: two weeks (in seconds)
0
20
40
60
80
100
120
140
160
180
0
5000
10000
15000
20000
25000
1)
2)
3)
4)
'as is'
'to be A'
'to be B'
'empty'

Confidence intervals

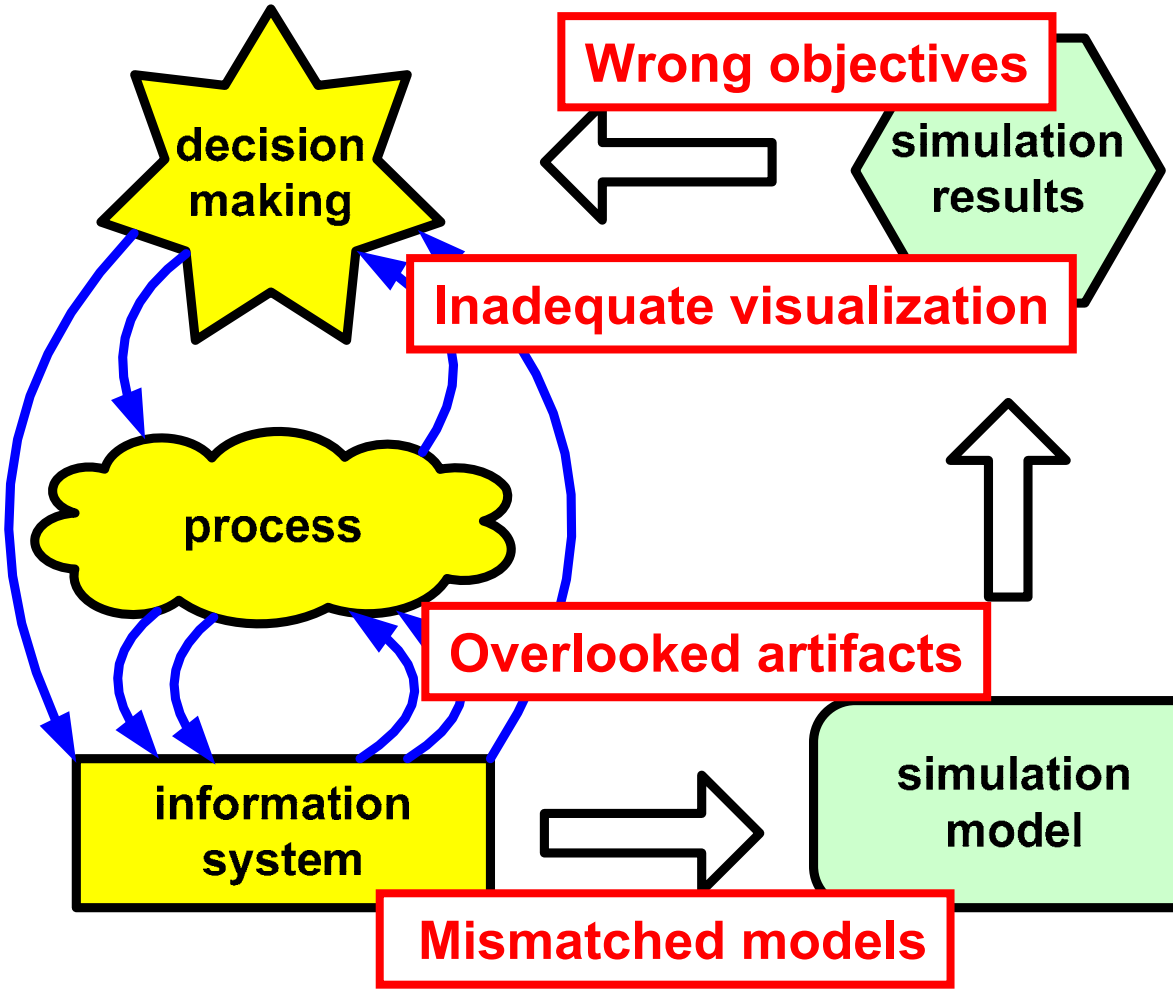


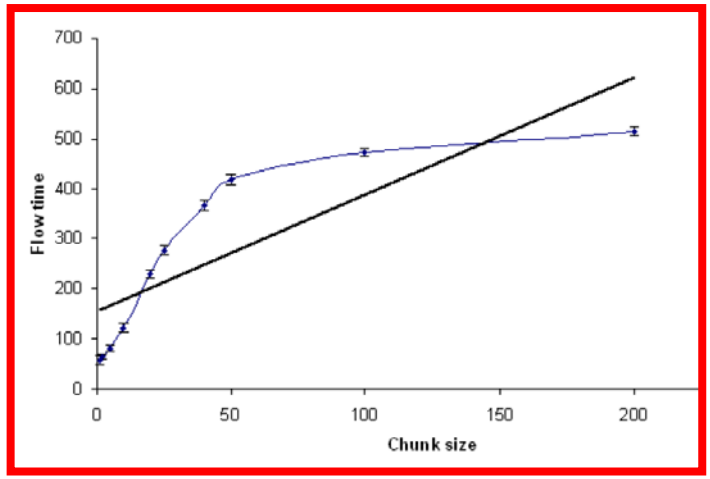
Conclusion Short-Term Simulation

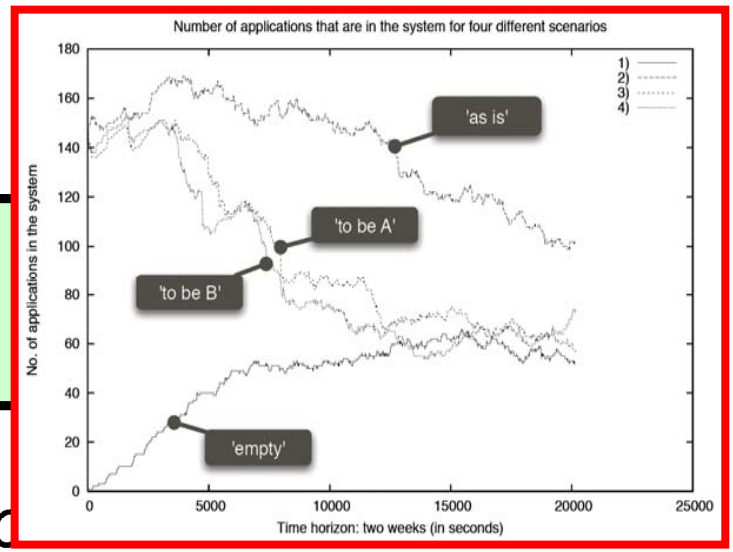
- Transient analysis is essential for operational decision making!
- The initial state matters!
- Artifacts (design, historic, and current state information) from a workflow management systems like YAWL can be used!
- Interesting side effect of the YAWL, ProM, CPN Tools integration: the real and simulated process can be viewed in a unified manner using process mining!

Conclusion

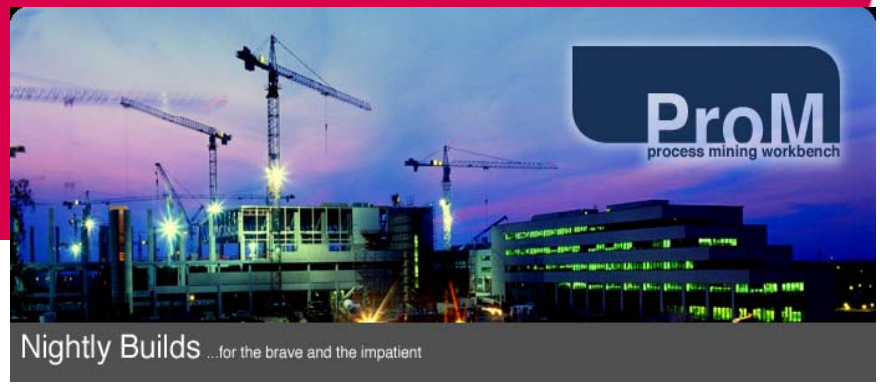
Problems and some solutions ...







Relevant WWW sites



- http://www.processmining.org
- **http:// promimport.sourceforge.net**
- **http://prom.sourceforge.net**
- **http://www.workflowpatterns.com**
- **http://www.workflowcourse.com**
- **http://www.vdaalst.com**

SUPER

LaQuSo
Laboratory for Quality Software

SIKS

EIT informatica onderwijs

NWO

stw

Beta