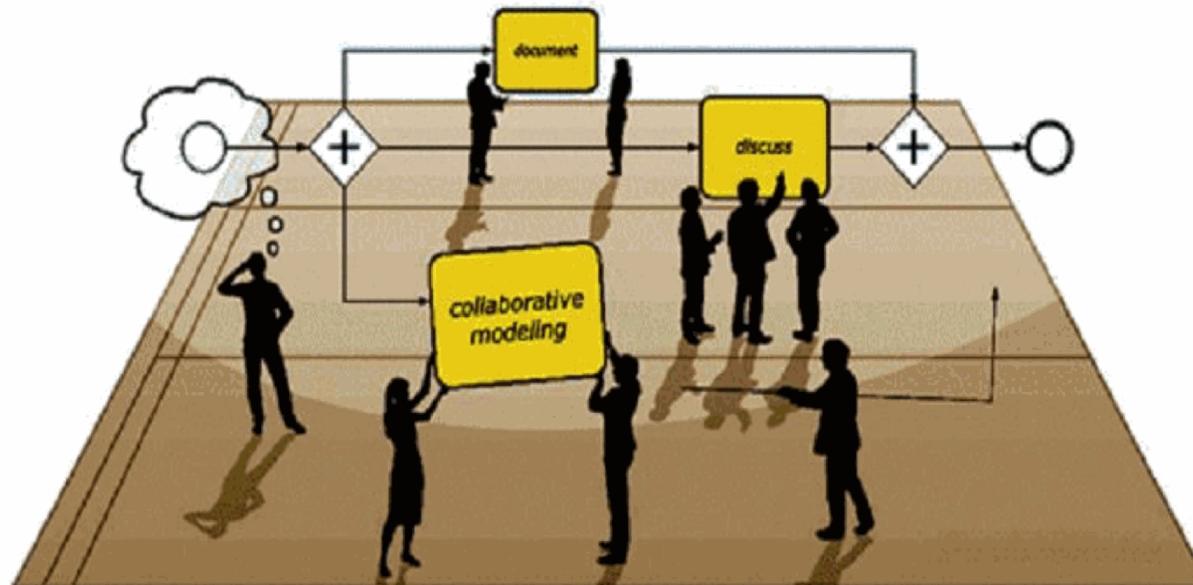


University of Pisa
MSc in Computer Engineering
Supply Chain Operation Management

"Large and complex organizations are a tangible manifestation of advanced technology, more than machinery itself." (J.K. Galbraith)

BPMN Modeling and Simulation



Lectures

Mario G. Cimino, Department of Information Engineering, Center for Logistics Systems
Pisa, March-May 2016, Monday 14.30-17.30, Room: ADInform2

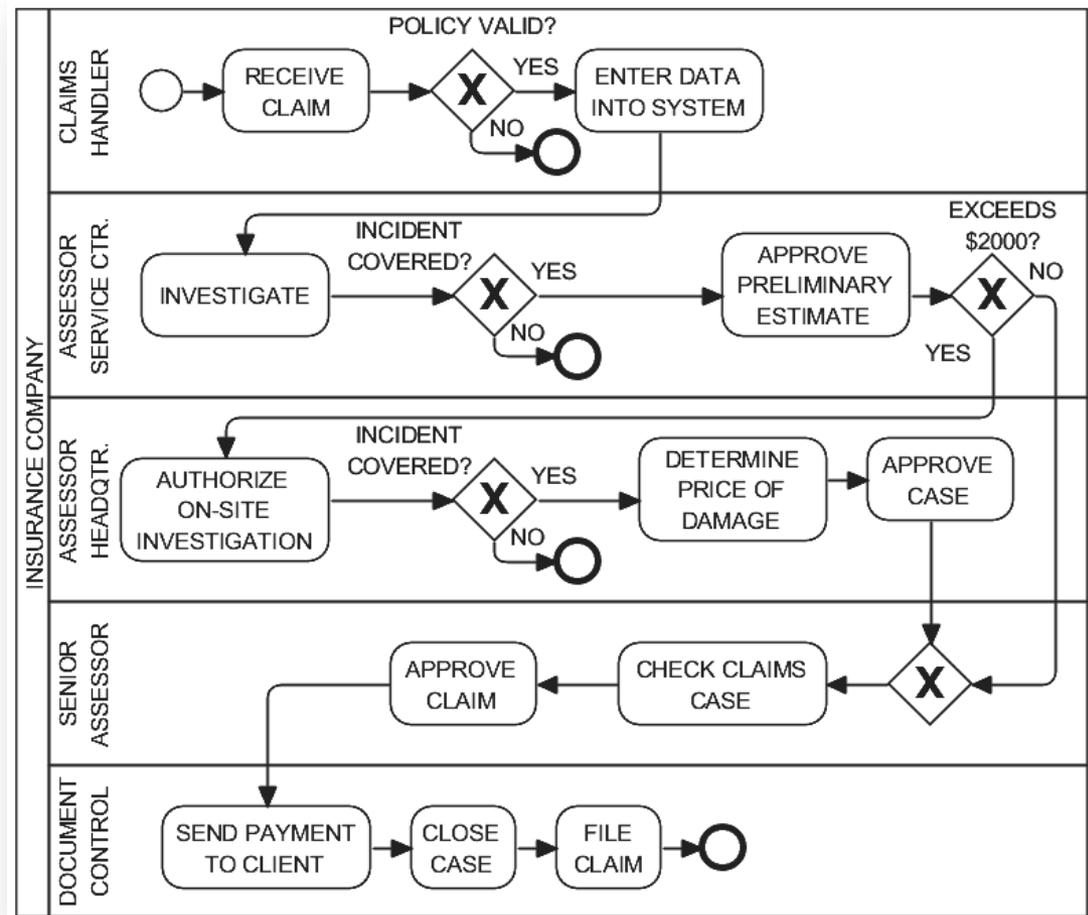
Insurance company: selecting the best staffing level for a claims process

- The problem is to find the most efficient staffing levels for each of the five resource types. Each resource type has a maximum limit of 20 people, and the overall headcount in the process cannot exceed 70. For 100 claims in a peak scenario, the total duration should be lower than 200 minutes.

- Perform a what-if analysis of possible solutions to this problem.

- To decide which configuration best aligns with service levels and process goals, analyze the trade-offs between headcount and total duration.

- Finally, consider also **sensitivity**: which type of resource produces a lower difference in total duration when reduced by a unit.



- Scenarios

SCENARIO	N. OF CASES
1) POL-NOT-VAL	5
2) POL-VAL & INC-NOT-COV	5
3) POL-VAL & INC-COV & >2000 & INC-NOT-COV	1
4) POL-VAL & INC-COV & >2000 & INC-COV	31
5) POL-VAL & INC-COV & <2000	58
TOTAL CASES	100

- Tasks duration

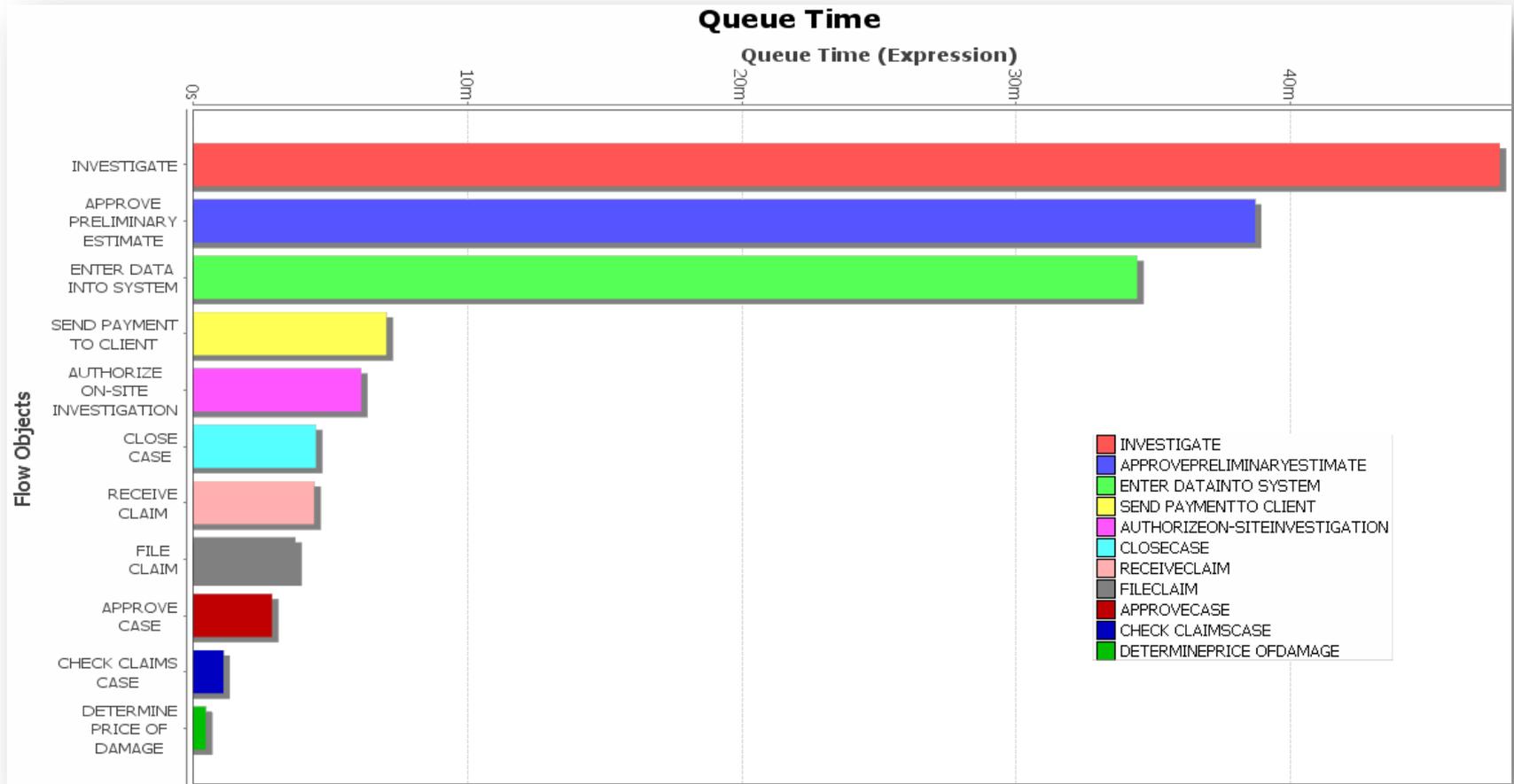
ACTIVITY	AVG. DURATION
RECEIVE CLAIM	2m 12s
ENTER DATA INTO SYSTEM	10m 32s
INVESTIGATE	19m 28s
APPROVE PRELIMINARY ESTIMATE	3m 54s
AUTHORIZE ON-SITE INVESTIGATION	2m 46s
DETERMINE PRICE OF DAMAGE	37m 16s
APPROVE CASE	2m 10s
CHECK CLAIMS CASE	3m 38s
APPROVE CLAIM	1m 18s
SEND PAYMENT TO CLIENT	7m 11s
CLOSE CASE	1m 46s
FILE CLAIM	3m 22s

- Optimization problem: minimize the headcount, i.e., the total number of lane instances, under the following constraints: (i) each lane type has a maximum limit of 20; (ii) the total number of lane instances in the process cannot exceed 70; (iii) the total duration should be lower than 200 minutes.
- An experiment with 20 instances available for all lanes (20, 20, 20, 20, 20) produces a total duration of **3h 26m 5s** (206 minutes).
- The longest queues are located in the second lane (Fig.1, *investigate, approve preliminary estimate*), i.e., “Assessor Service Ctr”. Moreover, the second lane exploits all 20 instances for a considerable amount of time (Fig.2, green line). Thus, a reduction of the available instances of the second lane may produce a substantial increase in the total duration. **Second lane => 20 instances.**
- Queues in the second lane may also depend on the inter-arrival time, which is determined by resources in the first lane. There could be a positive effect in the reduction of some resource in the first lane:

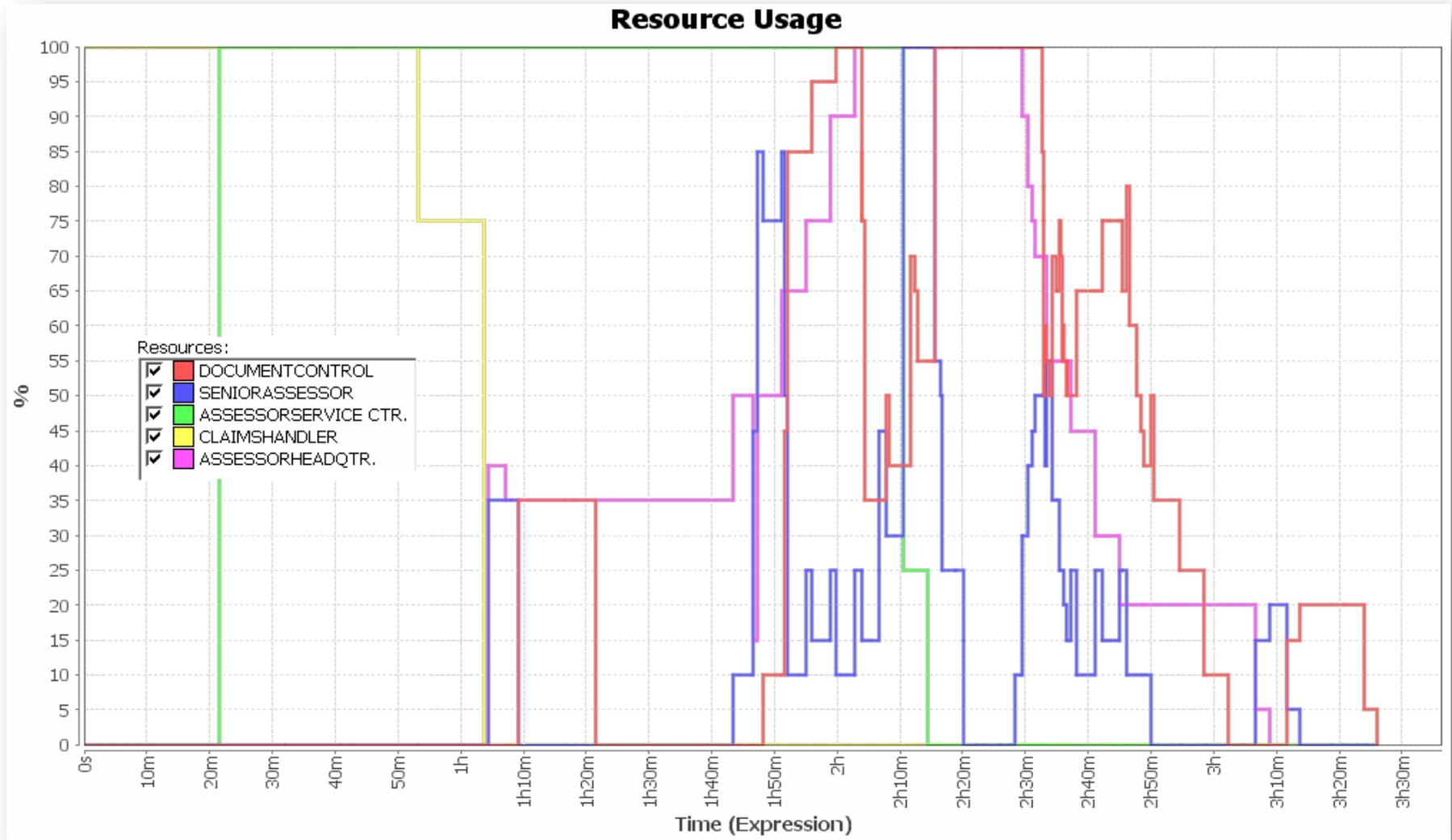
Inst.	19	18,17,15	16	14	13,12	11	10	9	8
Dur.	3:18:55	3:11:05 (191m)	3:12:13	3:13:17	3:15:29	3:14:55	3:12:05	3:11:59 (192m)	3:24:01

First lane => 9 instances (9, 20, 20, 20, 20). New results are shown in Fig. 4, 5, 6.

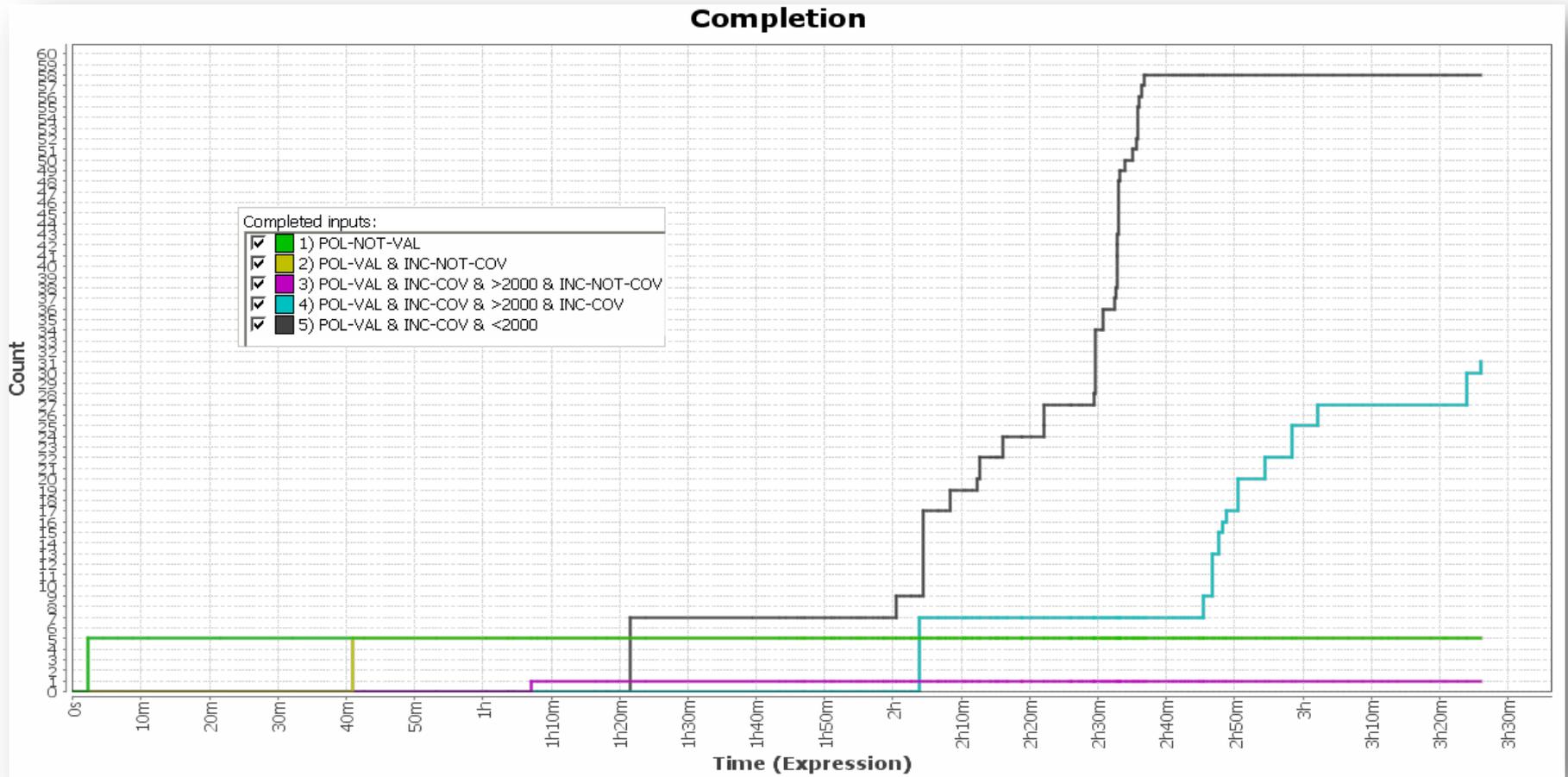
- Queue Time with (20, 20, 20, 20, 20) instances



- Resource Usage with (20, 20, 20, 20, 20) instances

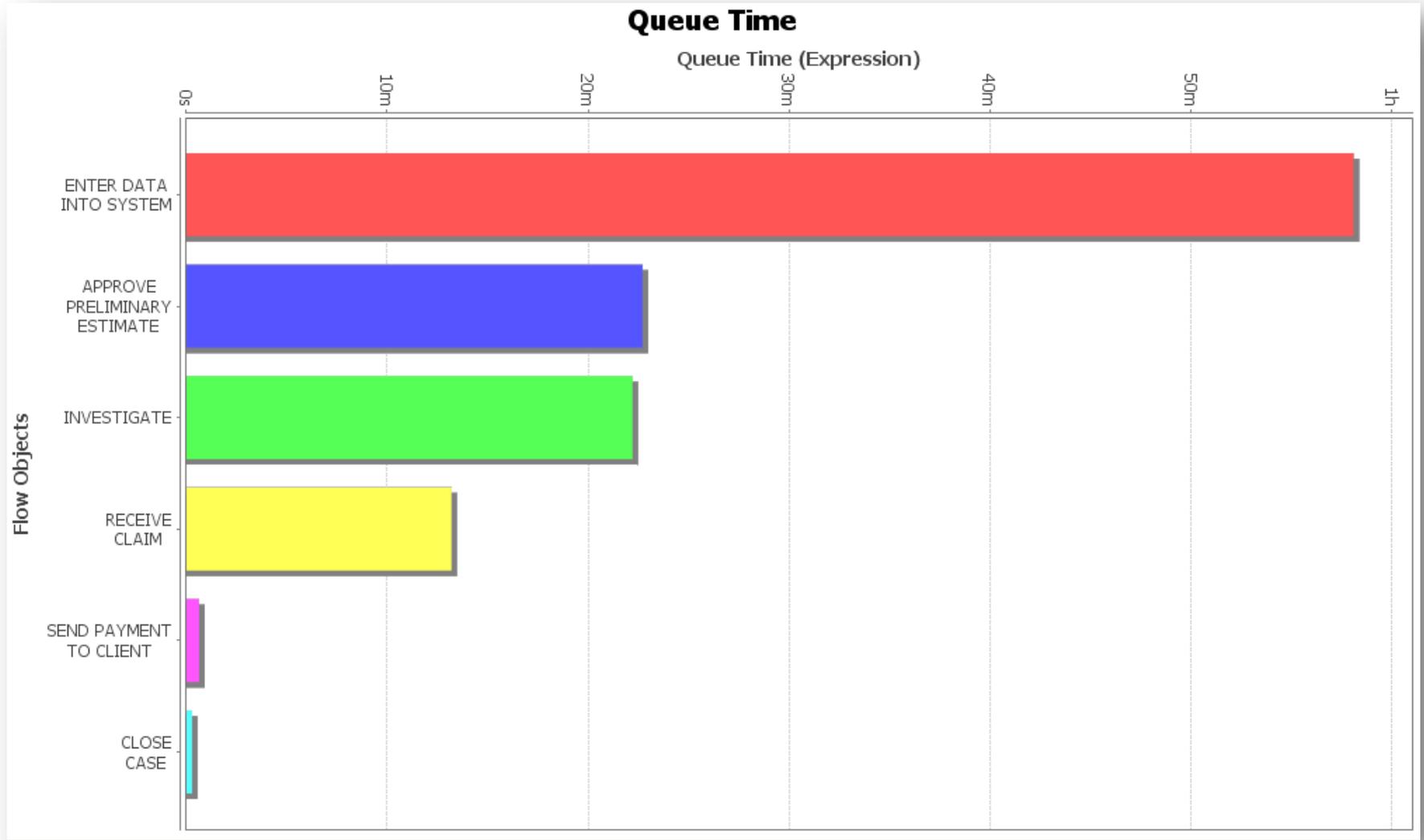


- Completion Time with (20, 20, 20, 20, 20) instances

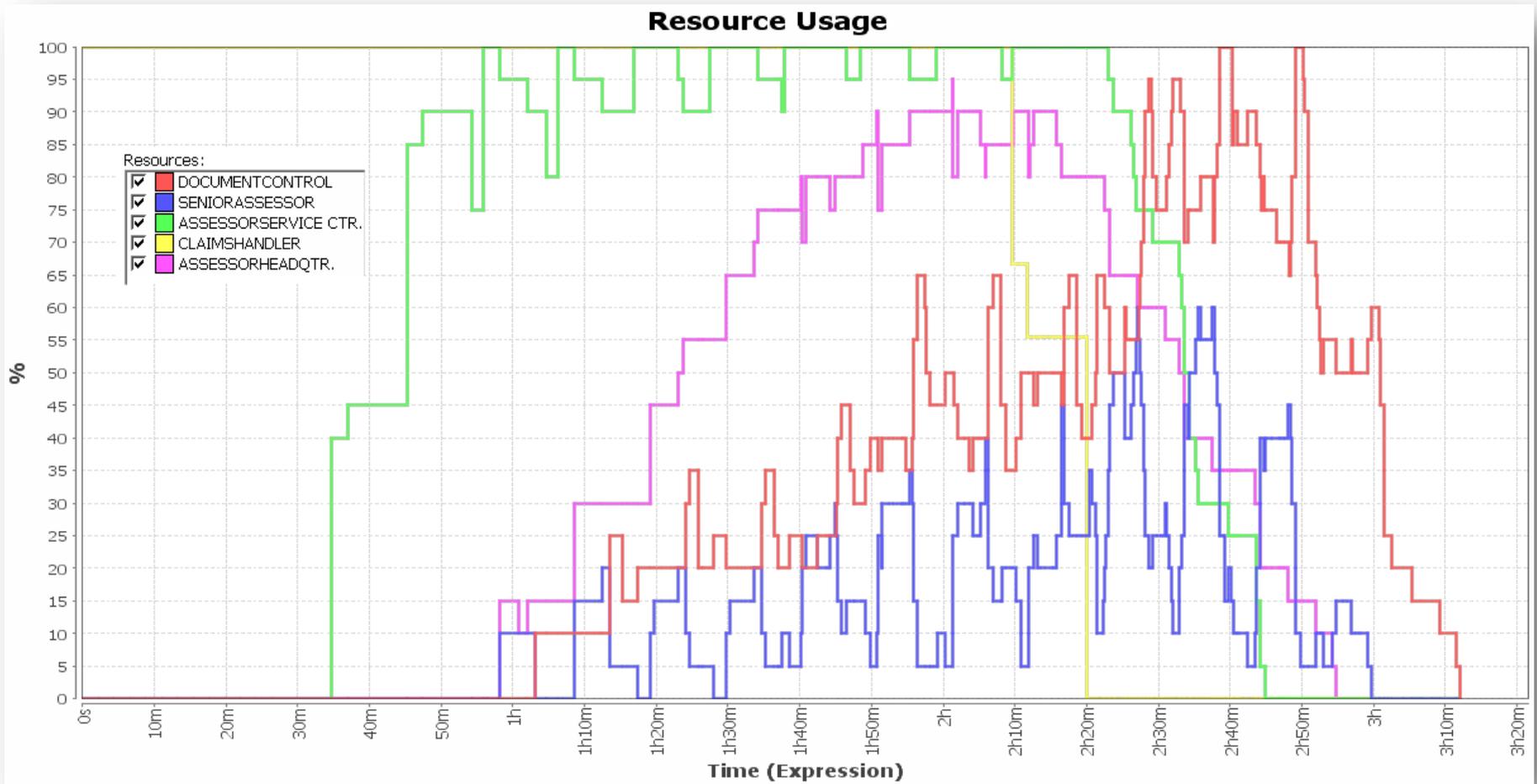


- Note that scenarios 4 and 5 (the most relevant in terms of tokens) cannot be carried out in parallel: when there is an increase in the completed instances of a scenario, the other scenario is stationary, and vice versa.

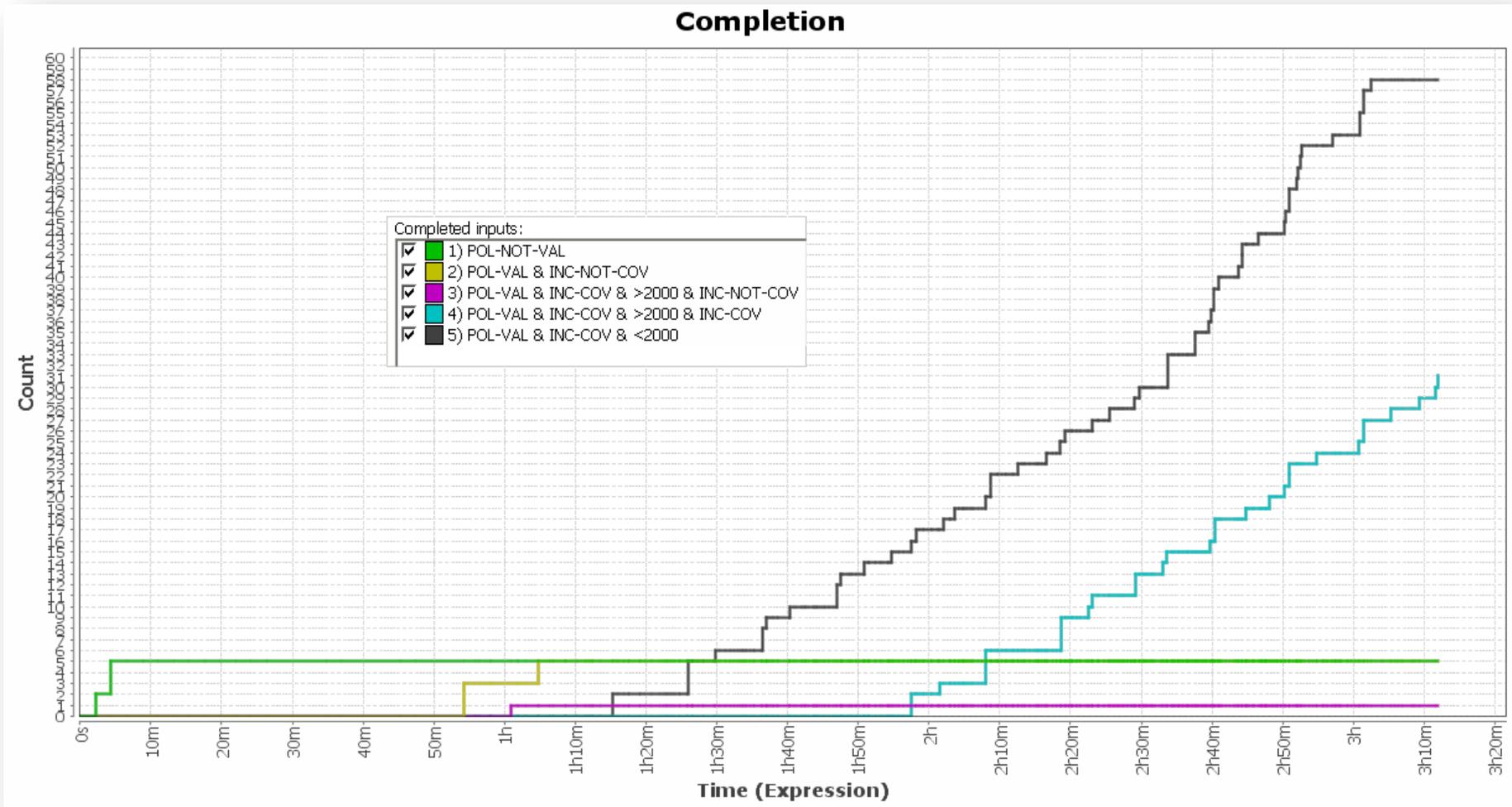
- Queue Time with (9, 20, 20, 20, 20) instances



- Resource usage with (9, 20, 20, 20, 20) instances



- Completion Time with (9, 20, 20, 20, 20) instances



- Note that scenarios 4 and 5 (the most relevant in terms of tokens) can be now carried out in parallel: both scenarios can increase the completed instances.

- The activities included in the IV lane (Senior Assessor) do not produce significant queues and the maximum resource usage is 55-60% (Fig.5). Thus, we can reduce the available instances without effects on the total duration:

Instances	19,...,6	5
Duration	3:11:59	3:15:53

- Fourth lane => 6 instances (9, 20, 20, 6, 20).
- Similarly, the maximum resource usage of the III lane (Assessor Headqtr) is about 90% (Fig.5). Thus, we can reduce the available instances without effects on the total duration:

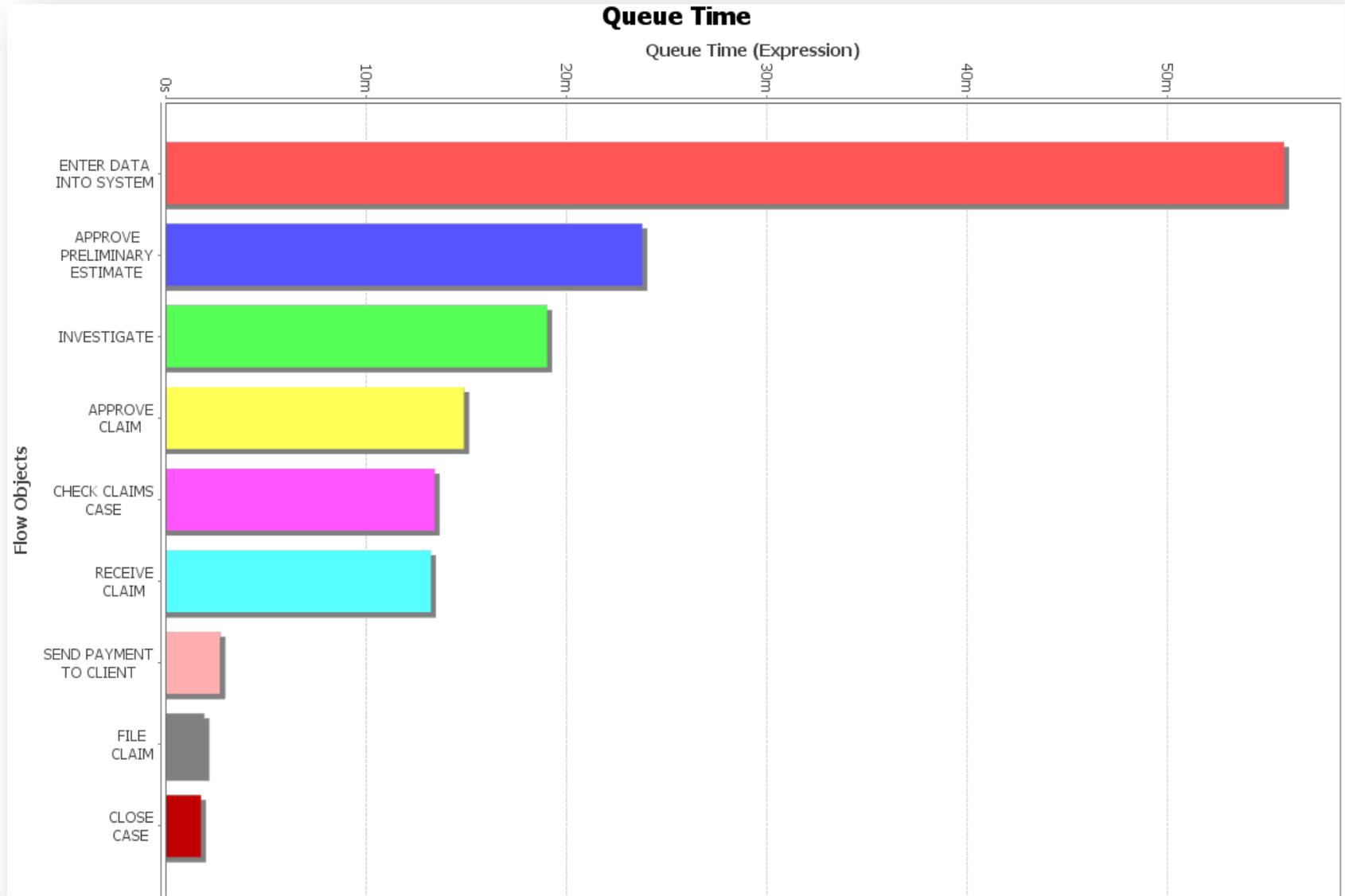
Instances	19,18	17
Duration	3:11:59	3:13:05

- Third lane => 18 instances (9, 20, 18, 6, 20).
- Again, the maximum resource usage of the V lane (Document Control) is about 85% (Fig.5). Thus, we can reduce the available instances without effects on the total duration:

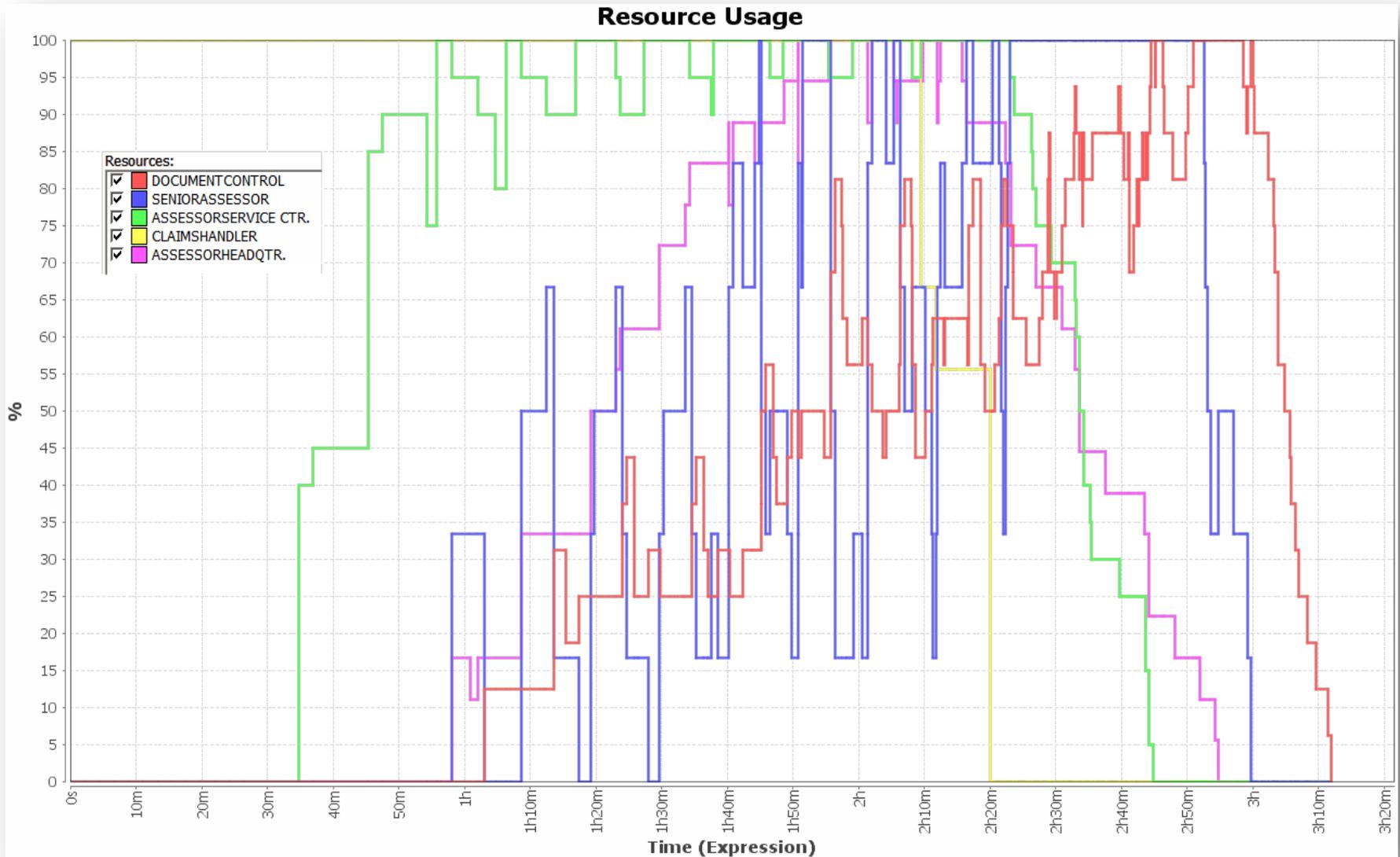
Instances	19,..16	15
Duration	3:11:59	3:12:36

- Fifth lane => 16 instances (9, 20, 18, 6, 16). Total duration **3:11:59**

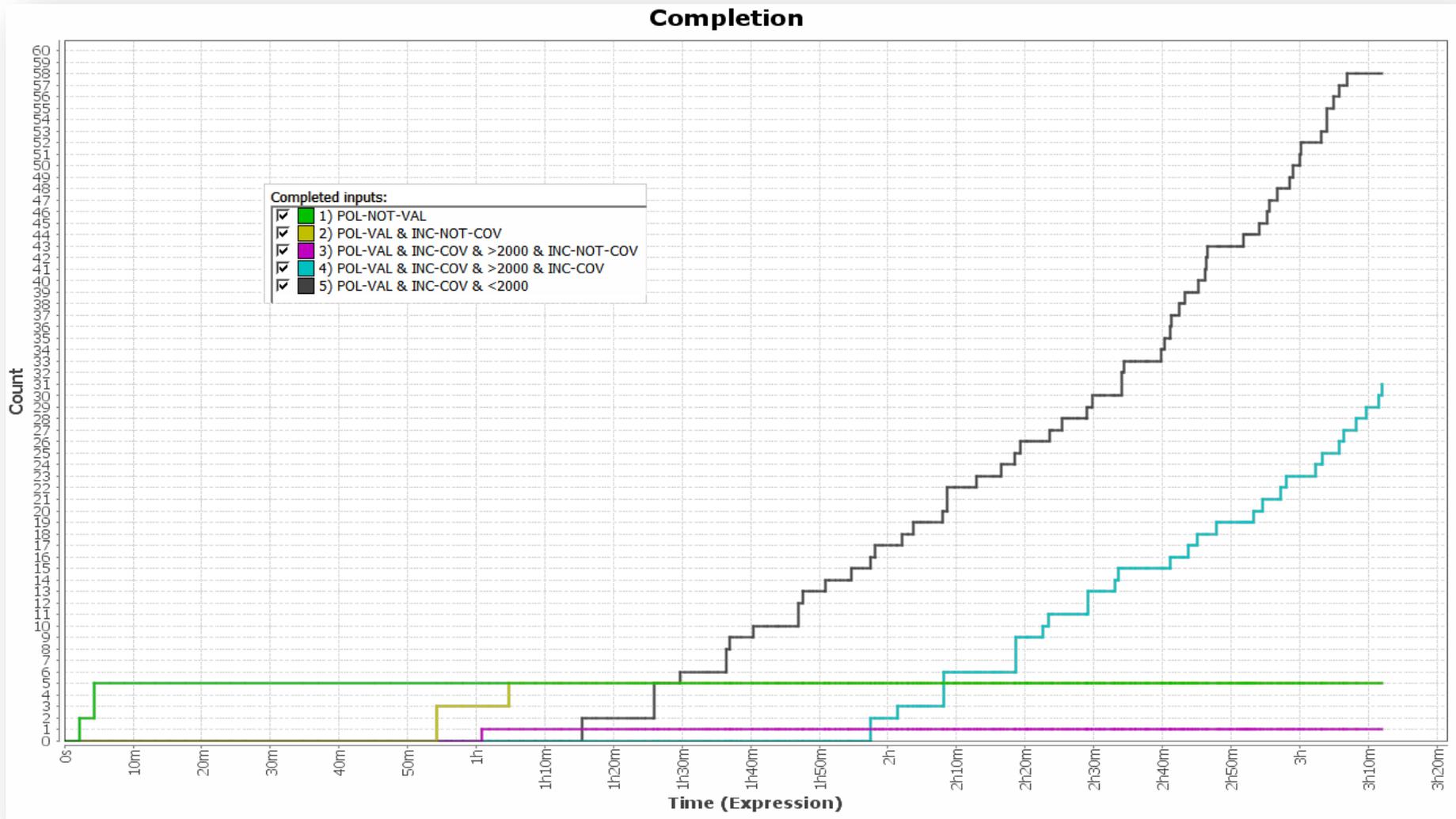
- Queue Time with (9, 20, 18, 6, 16) = 69 instances



- Resource usage with (9, 20, 18, 6, 16) = 69 instances



- Completion Time with (9, 20, 18, 6, 16) = 69 instances

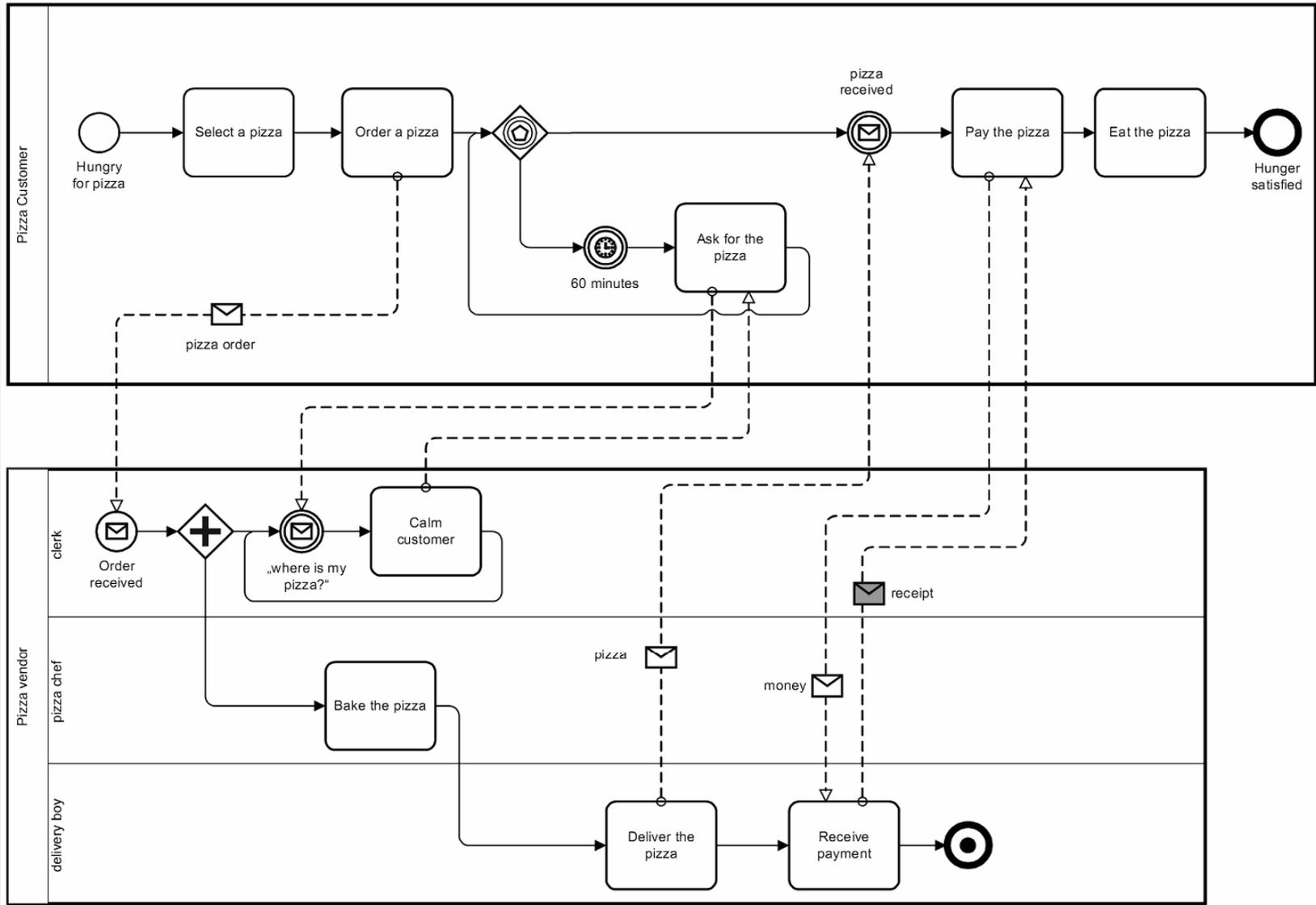


- Compare the **parameters sensitivity**, i.e., the outcome variation with respect to the unit variation of the parameter. Starting from (9, 20, 18, 6, 16), decrease a resource by one unit and assess the total duration.

Resources	Total Duration	Difference w.r.t 3:11:59 (192m)
(8, 20, 18, 6, 16)	3:24:01	+12:02 → NO (more than 200m)
(9, 19, 18, 6, 16)	3:22:05	+10:06 → NO "
(9, 20, 17, 6, 16)	3:13:05	+01:06 → YES (less than 200m)
(9, 20, 18, 5, 16)	3:16:12	+04:13 → YES "
(9, 20, 18, 6, 15)	3:12:36	+00:37 → YES "

Ordering and delivering pizza

- The process represents the interaction between a pizza customer and the vendor.



Ordering and delivering pizza

□ Describe in semi-formal natural language the above BPMN process.

1. The Customer is hungry for pizza
2. The Customer selects a pizza
3. The Customer orders a pizza from the Vendor
4. An order of the Customer has been received by the Clerk
 - 4.a.1. The chef bakes the pizza
 - 4.a.2. The boy delivers the pizza to the Customer
 - 4.a.3. The pizza has been received by the Customer
 - 4.a.4. The Customer pays the pizza to the boy
 - 4.a.5. The boy receives the payment from the Customer
 - 4.a.6. The boy delivers the receipt to the Customer
 - 4.a.7. The Pizza vendor terminates his activities on 4.a and 4.b (Vendor End)
 - 4.a.8. The Customer receives the receipt from the Customer
 - 4.a.9. The Customer eats the pizza
 - 4.a.10. The Customer has been satisfied (Customer End)
 - 4.b.1. When 60 minutes are elapsed:
 - 4.b.1.1. The Customer asks the clerk for the pizza
 - 4.b.1.2. The Clerk has received a request by the Customer for the pizza
 - 4.b.1.3. The Clerk calms the Customer → 4.b.1

- Given 100 starting tokens, determine the number of ending tokens for each scenario (path), via the following statistics: every 60', the 5% of the orders that were not baked 60' earlier are still under baking, and the 90% of the orders that were not completed 60' earlier are completed.

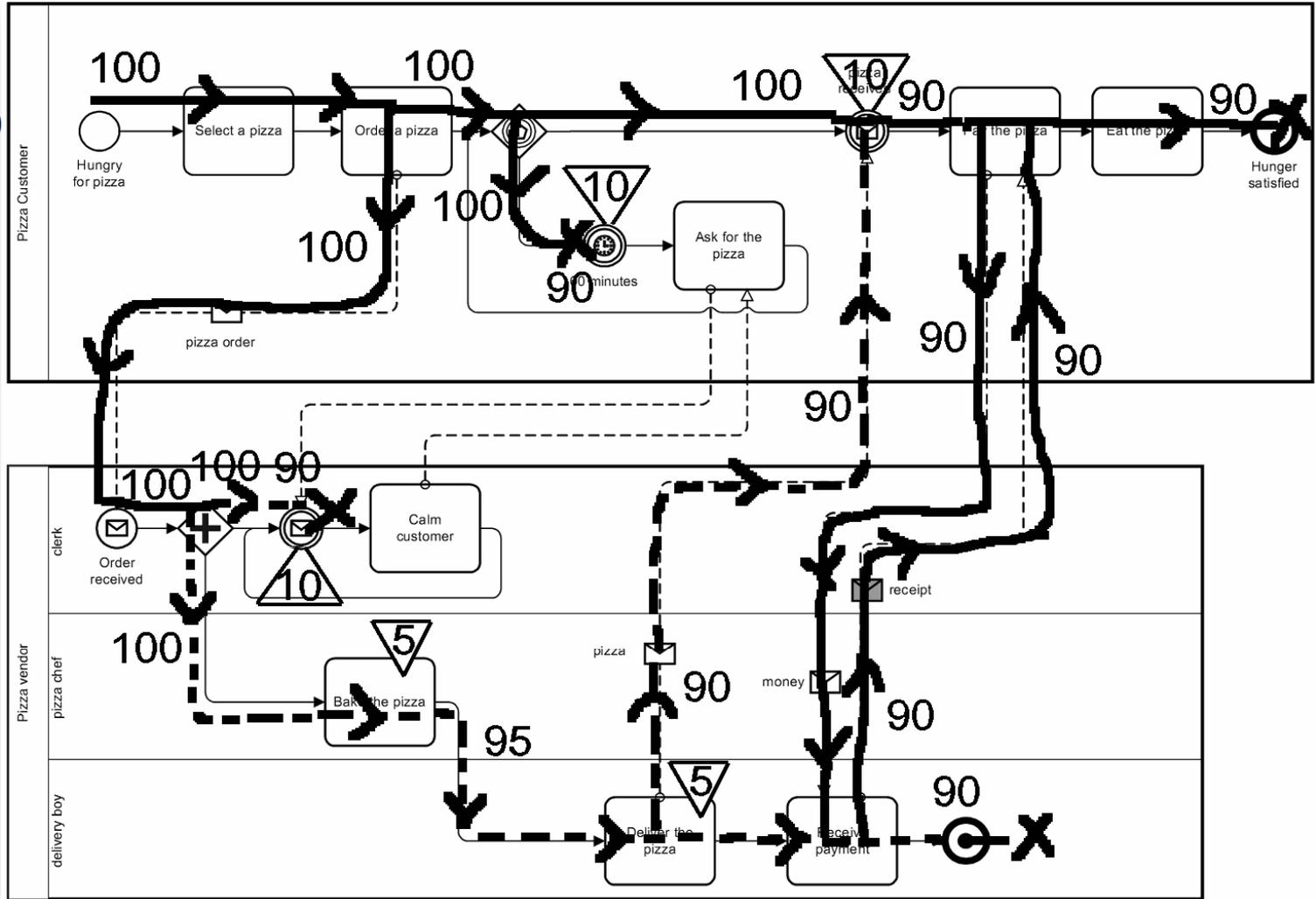
Time: 0'
 Not baked: 100
 Not completed: 100

S1)
Time: 60'
 • NOT-BAKED:
 $100 \times .05 = 5$

• COMPLETED
 (BAKED &
 DELIVERED):
 $100 \times .9 = 90$

• BAKED &
 NON-DELIV.:
 $100 - 90 - 5 = 5$

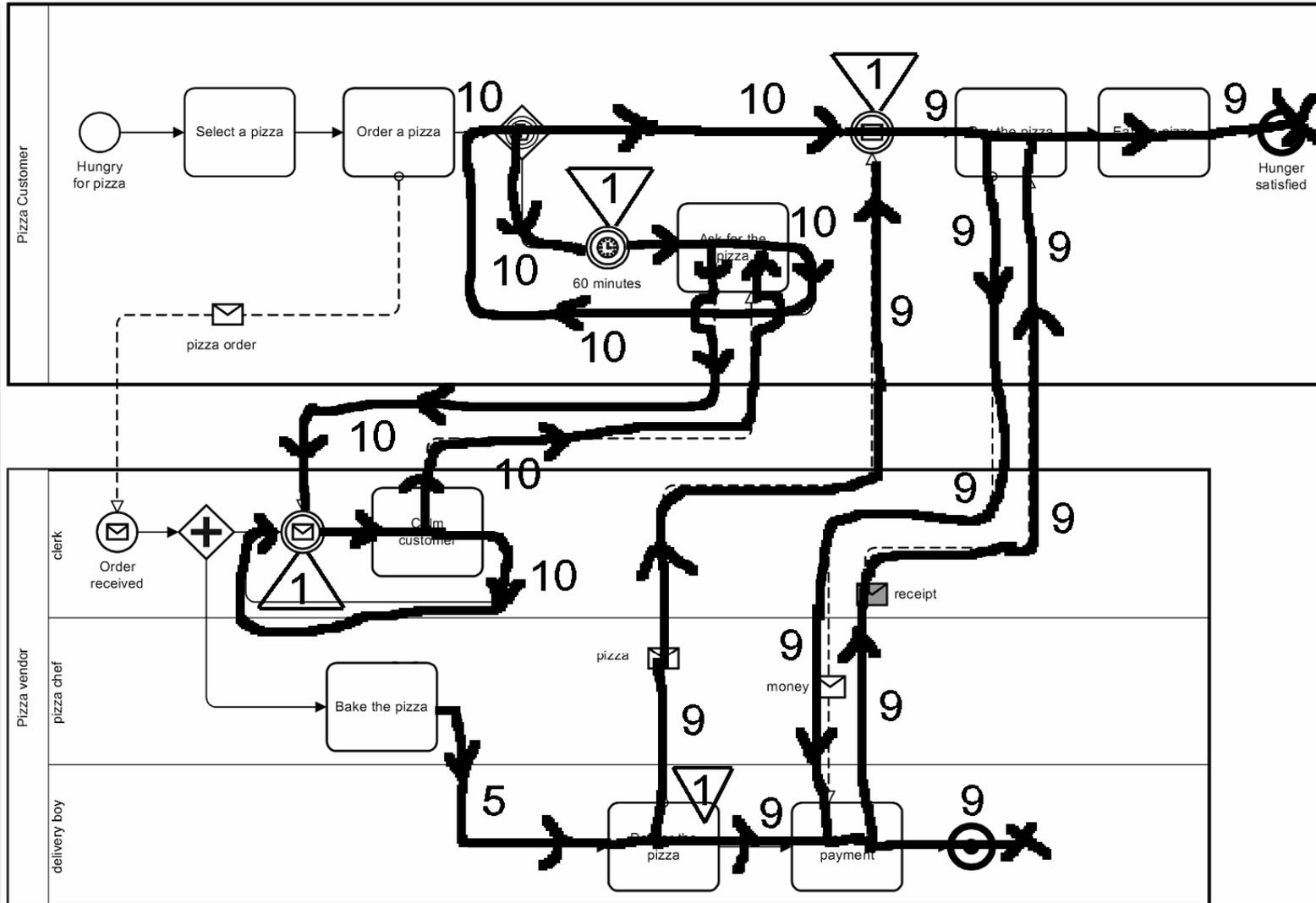
Not baked: 5
 Not-completed: 10



S2) Time: 120'

- NOT-BAKED: $5 \times .05 = 0$
- COMPLETED (BAKED & DELIVERED): $10 \times .9 = 9$
- BAKED & NOT DELIVERED: $10 - 9 - 0 = 1$

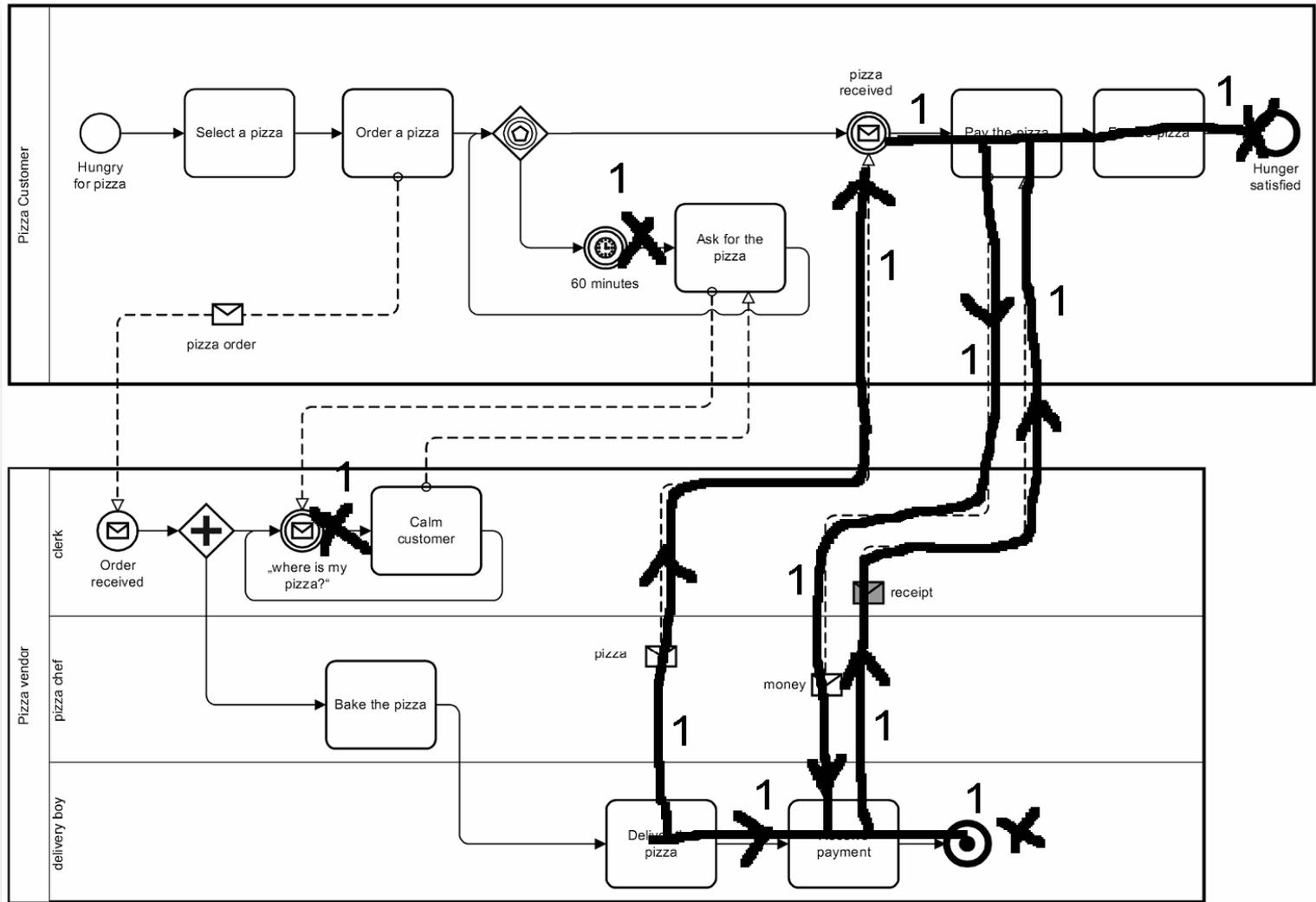
Not baked:0
Not completed:1



S3) Time: 180'

- NOT-BAKED: $0 \times .05 = 0$
- COMPLETED (BAKED & DELIVERED): $1 \times .9 = 1$
- BAKED & NOT DELIVERED: $1 - 1 - 0 = 0$

Not baked: 0
Not completed: 0



- ❑ Motivate the use of a termination event on the Pizza Vendor Pool (Why not an end event?).

The terminate event, like the end event, deletes a token when it arrives. However, it not only deletes this single token, but it also terminate the entire process, i.e. any other tokens in the entire process are also removed at the same time.

Without the terminate event, the tokens arrived at the “Where is my pizza?” event would be, stuck, trapped in the loop “Calm customer”, and the process could not be finished.

This problem is avoided by using terminate event.

