UNIVERSITY OF PISA

School of Engineering

Master of Science in Computer Engineering

Performance Evaluation of Computer Systems and Networks

PROJECT DOCUMENTATION

FACULTY BAR

WORKGROUP:

Diego Casu

Iacopo Pacini

ACADEMIC YEAR 2019/2020

INDEX

1 DESCRIPTION AND MODELING

The subject of this study is a faculty bar, an organization that manages the orders of arriving customers, divided into two classes with different service priorities (normal and VIP), and a set of tables, on which it is possible to eat.
An arriving guest can issue an order to a cashier, which has two queues corresponding to each class of customers: both queues are served following a FIFO (First In First Out) policy, where the VIP membership guarantees non-preemptive priority over the normal one, i.e. a VIP user is always served before a normal user, unless the latter is being served by the cashier.
After the completion of her order, the user joins the queue for a table: the queue is shared by both the customer classes, which are served according to a FIFO policy independently from their membership. The bar has a fixed number of tables, with a certain number of seats each, which can contain both normal and VIP customers.
The objective of the study is to analyze, for each class and under a varying workload, the behaviour of the queues, the customer waiting and response times. The customer inter-arrival times, cashier service time and customer eating time will be modeled as constantly or exponentially distributed.
Supposing that a customer:

1. never leaves the area of the cashier before being served, unless she finds the queue full at the arrival;
2. is obliged to issue an order before taking a seat;
3. always queues for a table after an order completion, i.e. she does not eat outside the bar, unless the queue for tables is full;
4. leaves the bar as soon as she finishes eating;

it is possible to model the system as an open queueing network [Figure 1].

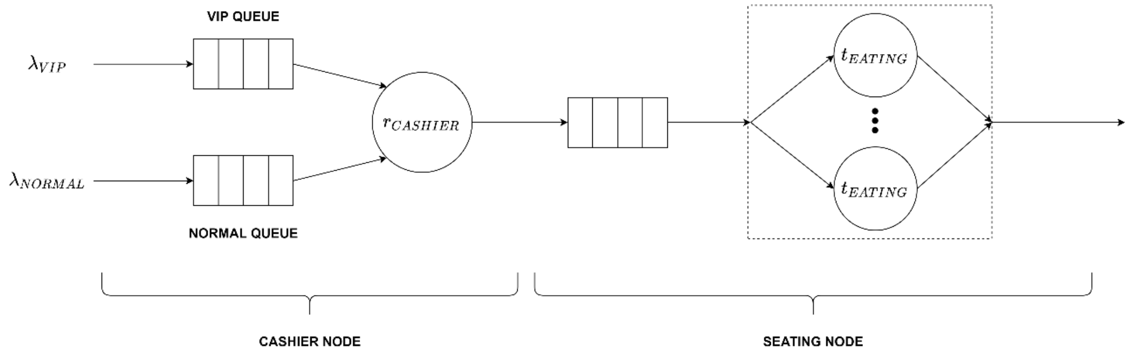

Figure 1 The faculty bar seen as an OQN, with two service centers: cashier node and seating node.

Since the study does not focus on how to optimize the physical space occupied by the tables or on how to group the number of seats according to a certain policy, it is possible to consider only the overall seating capacity and to manage the table area as a flat set of seats.
The parameters related to the customers are expressed as follows:

- λ_{VIP} and λ_{NORMAL} for the arrival rates, T_{VIP} and T_{NORMAL} for the inter-arrival times;
- t_{EATING} and T_{EATING} for the eating rate and eating time of a customer, once she takes a seat;
- W^{VIP} and W^{NORMAL} for the waiting times, R^{VIP} and R^{NORMAL} for the response times;
- N_q^{VIP} and N_q^{NORMAL} for the number of customers in the queues.

If related to a specific node, the latter will be paired with an explanatory subscript.
For what concerns the faculty bar, the related parameters are:

- $r_{CASHIER}$ and $T_{CASHIER}$ for the service rate and service time of the cashier;
- $K_{q,CASHIER}^{VIP}$, $K_{q,CASHIER}^{NORMAL}$ and $K_{q,SEATING}$ for the dimension of the queues;
- N_{SEAT} for the overall number of seats provided by the tables.

2 SIMULATION OF THE SYSTEM

2.1 Omnet++ model

To analyze the system and gather data under different working conditions, an Omnet++ based simulator has been used to reproduce the flow of actions inside the bar. The customers are represented via their issued orders in the form of messages and handled by three main modules:

- OrderProducer, which produces customer orders and assigns a priority to them;
- Cashier, which manages the customer queues, fulfills the orders and routes a customer to the seating area;
- SeatManager, which handles the seating area and the related queue, as well as the final customer leaving.

All the queues are implemented exploiting the std::queue container offered by the Standard Template Library, while the messages are of type OrderMessage, a custom extension of the basic Omnet message type; furthermore, the modules are predisposed such that the arrival/service rates can be constant or extracted from an exponential distribution. The system is obtained instantiating a network of two OrderProducers (one for each membership class), one Cashier and one SeatManager [Figure 2].

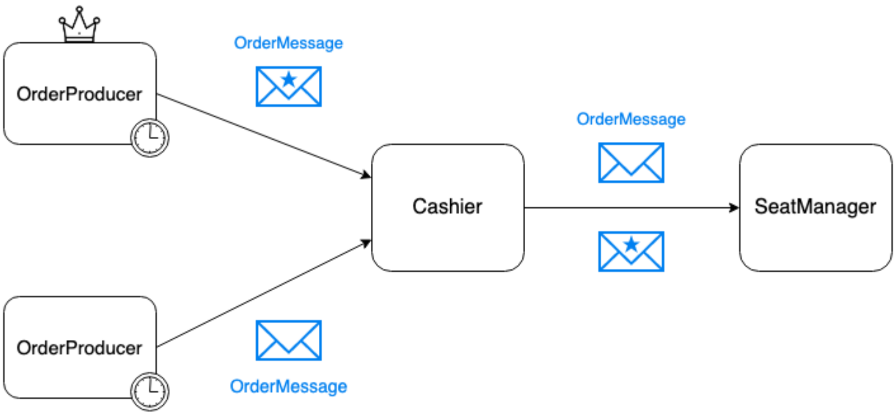

Figure 2 The faculty bar seen as an Omnet++ network.

2.2 Estimation of the warmup and simulation times

The analysis of the system is performed after it reaches the steady state: the warmup time is chosen as the time needed by the throughput to converge to the average arrival rate [Figure 3].
The total simulation time is determined by the sample size needed to reach accurate confidence intervals.

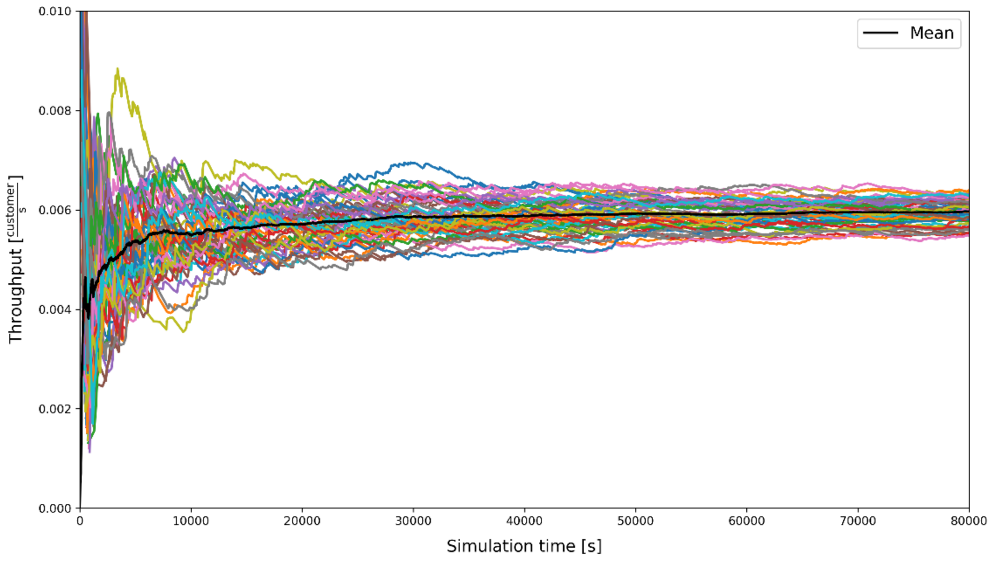

Figure 3 An example of the convergent behaviour of the throughput over time.

3 ANALYSIS: CONSTANT INTER-ARRIVAL AND SERVICE TIMES

3.1 Cashier node

3.1.1 Stability condition and queue occupancy

The cashier is a $D/D/1/NP$ system, where NP denotes a non-preemptive priority queueing, for which the stability condition is:

$$\rho_{CASHIER} = \rho_{CASHIER}^{VIP} + \rho_{CASHIER}^{NORMAL} = \frac{\lambda_{VIP} + \lambda_{NORMAL}}{r_{CASHIER}} \leq 1$$

where $\rho_{CASHIER}$ is the total utilization and $\rho_{CASHIER}^{VIP}, \rho_{CASHIER}^{NORMAL}$ are the per-class utilizations.
If the stability condition is not respected, the number of customers in the system grows indefinitely; in particular:

- if $\lambda_{VIP} > r_{CASHIER}$, the normal queue is never served and goes in starvation, with the exception of the first normal customer if $\lambda_{NORMAL} > \lambda_{VIP}$. Both the numbers of normal and VIP customers grow without bounds, along with their waiting and response times;
- if $\lambda_{VIP} = r_{CASHIER}$, the normal queue is never served and goes in starvation, with the same above exception, while the VIP customers experience constant waiting and response times. If $\lambda_{NORMAL} \leq \lambda_{VIP}$, then $W_{CASHIER}^{VIP} = 0$, $R_{CASHIER}^{VIP} = T_{CASHIER}$; if $\lambda_{NORMAL} > \lambda_{VIP}$, the first normal arrival introduces queueing in the VIP service, resulting in $W_{CASHIER}^{VIP} = T_{NORMAL}$, $R_{CASHIER}^{VIP} = T_{NORMAL} + T_{CASHIER}$.

If the stability condition is respected, the system is always in steady state and the sequence of arrivals given by the combination of T_{VIP} and T_{NORMAL} is periodic.

Moreover, since $r_{CASHIER} \geq \lambda_{VIP} + \lambda_{NORMAL}$, at any time:

- at most one queue out of two is occupied;
- at most one single customer can reside in a queue, regardless of the membership.

Therefore, $K_{q,CASHIER}^{VIP} = K_{q,CASHIER}^{NORMAL} = 1$ is a worst case dimensioning for both queues. From now on, it is supposed that the cashier is a stable system and that no arriving customers are lost.

3.1.2 Waiting and response times

Given the deterministic nature of the node, it is possible to find some useful conditions and upper bounds for the waiting and response times. In this configuration, an arriving customer either finds the cashier available or waits for the current residual service time; hence, the conditions:

1. $0 \leq W_{CASHIER}^{VIP}, W_{CASHIER}^{NORMAL} \leq T_{CASHIER}$;
2. $T_{CASHIER} \leq R_{CASHIER}^{VIP}, R_{CASHIER}^{NORMAL} \leq 2T_{CASHIER}$;

are always verified, where the actual values will depend on the combination of inter-arrival times. Furthermore, within the period P, all the possible relative positions of a VIP arrival with respect to a normal arrival and vice versa are given by $|kT_{VIP}|_{T_{NORMAL}}$ and $|kT_{NORMAL}|_{T_{VIP}}$, with $k \in \mathbb{N}^+$ and $|x|_y = x - \left\lfloor \frac{x}{y} \right\rfloor \cdot y$. Therefore:

- if T_{NORMAL} is an integer multiple of T_{VIP}, a normal customer arrives simultaneously with a VIP one, while the latter can arrive alone. If $T_{CASHIER} \leq \frac{1}{2} T_{VIP}$, it holds $W_{CASHIER}^{VIP} = 0$, $R_{CASHIER}^{VIP} = T_{CASHIER}$ and $W_{CASHIER}^{NORMAL} = T_{CASHIER}$, $R_{CASHIER}^{NORMAL} = 2T_{CASHIER}$. In fact, if both arrive at the same time, the cashier must prepare the VIP and normal orders for $t = 2T_{CASHIER}$; then, if another VIP arrives during this slice of time, she finds the cashier busy and goes in queue. Under the previous condition, the cashier is able to serve both customers before a new VIP arrival [Figure 4];

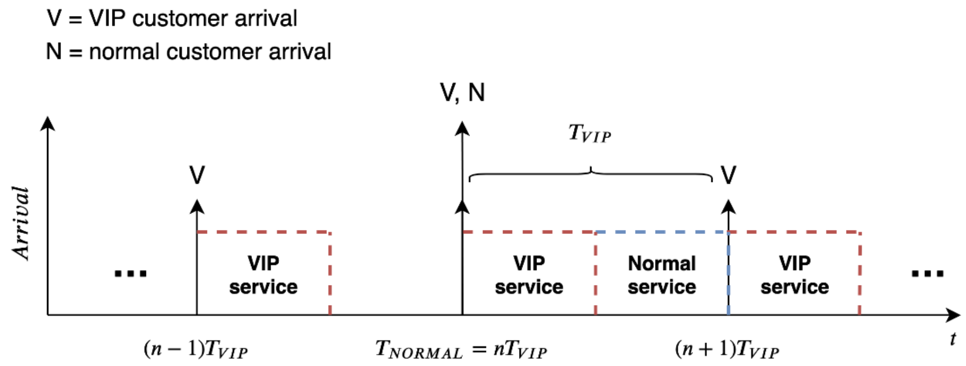

Figure 4 Interplay between normal and VIP customers when T_{NORMAL} is an integer multiple of T_{VIP}.

- if T_{VIP} is an integer multiple of T_{NORMAL}, such that $T_{VIP} = nT_{NORMAL}$, the VIP customer arrives simultaneously with a normal customer, while the latter can arrive alone. Then, if $T_{CASHIER} \leq \frac{n}{n+1} T_{NORMAL}$, it is $W_{CASHIER}^{VIP} = 0$, $R_{CASHIER}^{VIP} = T_{CASHIER}$. The condition is more understandable if written as $(n+1)T_{CASHIER} \leq nT_{NORMAL}$: observing that the period in this case is $P = nT_{NORMAL}$, it means that the cashier is able to serve n normal customers and one VIP customer before the next VIP arrival. Moreover, few computations show that the condition is equivalent to $\frac{n+1}{n} \lambda_{NORMAL} \leq r_{CASHIER}$, which is also the stability condition, meaning that a stable system with T_{VIP} integer multiple of T_{NORMAL} verifies $W_{CASHIER}^{VIP} = 0$ and $R_{CASHIER}^{VIP} = T_{CASHIER}$ [Figure 5];

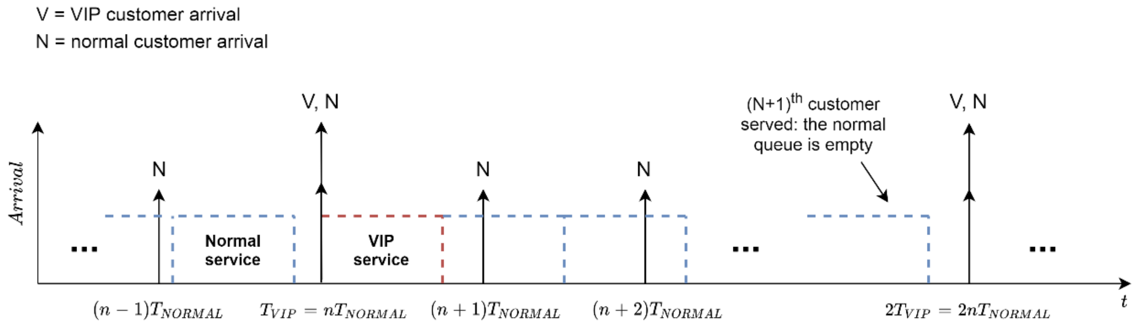

Figure 5 Interplay between normal and VIP customers when T_{VIP} is an integer multiple of T_{NORMAL}.

- if T_{NORMAL} and T_{VIP} are not integer multiples, $W_{CASHIER}^{VIP} = 0$, $R_{CASHIER}^{VIP} = T_{CASHIER}$ is verified if $T_{CASHIER} \leq \min_{k\in\mathbb{N}^+} |kT_{VIP}|_{T_{NORMAL}} : |kT_{VIP}|_{T_{NORMAL}} \neq 0$. This means that the cashier is able to serve a normal customer before a VIP arrival in the worst case of minimum time distance between class arrivals. The non-null condition excludes the cases where a simultaneous arrival occurs, in which the VIP has always the priority [Figure 6];

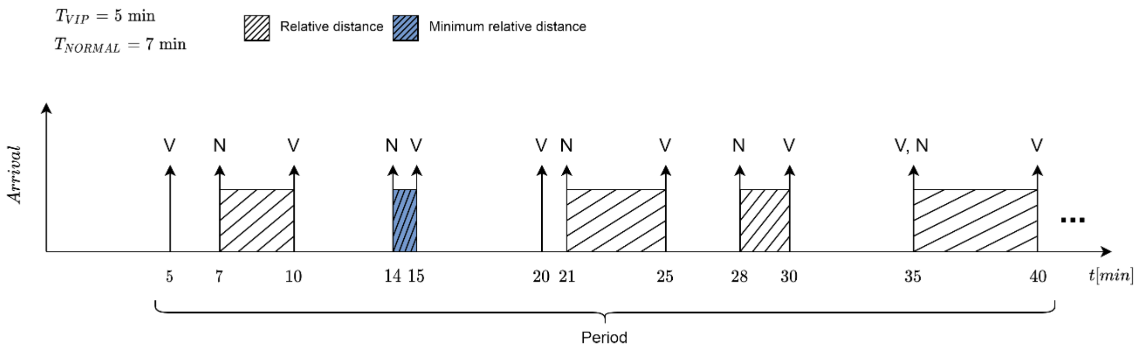

Figure 6 Example of $T_{CASHIER}$ condition for $W_{CASHIER}^{VIP} = 0$ when T_{VIP} and T_{NORMAL} are not integer multiples.

- if T_{NORMAL} and T_{VIP} are not integer multiples and reasoning as above, it is $W_{CASHIER}^{NORMAL} = 0$, $R_{CASHIER}^{NORMAL} = T_{CASHIER}$ if $T_{CASHIER} \leq \min_{k\in\mathbb{N}^+} |kT_{NORMAL}|_{T_{VIP}} : |kT_{NORMAL}|_{T_{VIP}} \neq 0$ and it does not exist a $k \in \mathbb{N}^+$ such that $|kT_{NORMAL}|_{T_{VIP}} = 0$. If the latter exists, a simultaneous arrival occurs, in which the VIP has always the priority and $W_{CASHIER}^{NORMAL}$ cannot be null.

3.2 Seating node

If the cashier node is stable, in steady state and without customer losses, the average arrival rate $\lambda_{VIP} + \lambda_{NORMAL}$ is also the average departure rate of the node, where the departure process is deterministic. Therefore, the seating node is a $D/D/N_{SEAT}/FCFS$ system with an average arrival rate equal to $\lambda_{VIP} + \lambda_{NORMAL}$. The system is always in steady state and has no queueing as long as the stability condition is respected, which is:

$$\rho_{SEATING} = \frac{\lambda_{VIP} + \lambda_{NORMAL}}{N_{SEAT} \cdot t_{EATING}} \leq 1$$

Hence, it is possible to set $N_{SEAT} \geq \frac{\lambda_{VIP}+\lambda_{NORMAL}}{t_{EATING}}$ and to dimension the queue with $K_{q,SEATING} = 0$, obtaining $W_{SEATING}^{VIP} = W_{SEATING}^{NORMAL} = 0$, $R_{SEATING}^{VIP} = R_{SEATING}^{NORMAL} = T_{EATING}$.

4 ANALYSIS: EXPONENTIAL INTER-ARRIVAL AND SERVICE TIMES

4.1 Cashier node

4.1.1 Stability condition

The cashier is an $M/M/1/NP$ system, for which the stability condition is:

$$\rho_{CASHIER} = \rho_{CASHIER}^{VIP} + \rho_{CASHIER}^{NORMAL} = \frac{\lambda_{VIP} + \lambda_{NORMAL}}{r_{CASHIER}} < 1$$

If the stability condition is not respected, the number of customers in the system grows indefinitely. From now on, it is supposed that the cashier is a stable system in steady state, with infinite queues: considerations about finite queues are discussed in the queue occupancy section.

4.1.2 Waiting and response times

In absence of a direct theoretical method, the waiting and response times can be studied exploiting the $2^k r$ factorial technique, involving the performance indexes $E[R_{CASHIER}^{VIP}]$, $E[R_{CASHIER}^{NORMAL}]$, $E[W_{CASHIER}^{VIP}]$ and $E[W_{CASHIER}^{NORMAL}]$, with factor intervals chosen as $T_{VIP},\ T_{NORMAL} \in [5.5min, 10min]$, $T_{CASHIER} \in [1min, 2.5min]$. Applying a logarithmic transformation to meet the hypotheses of the method, the resulting models have negligible errors ($\cong 1\%$) and show that:

- 95.77% of the variation of $E[R_{CASHIER}^{VIP}]$ and 91.45% of the variation of $E[W_{CASHIER}^{VIP}]$ are due to $T_{CASHIER}$, with negligible impact due to other factors or interactions between them (at most $\cong 5\%$);
- 80.67% of the variation of $E[R_{CASHIER}^{NORMAL}]$ and 83.11% of the variation of $E[W_{CASHIER}^{NORMAL}]$ are due to $T_{CASHIER}$, with negligible impact due to other factors or interactions between them (at most $\cong 8\%$ due to T_{VIP}).

Thus, the waiting and response times can be studied fixing $T_{VIP} = T_{NORMAL} = 5.5min$ and varying $T_{CASHIER}$ in the set $\{1min, 1.5min, 2min, 2.5min\}$ [Figure 7-8].

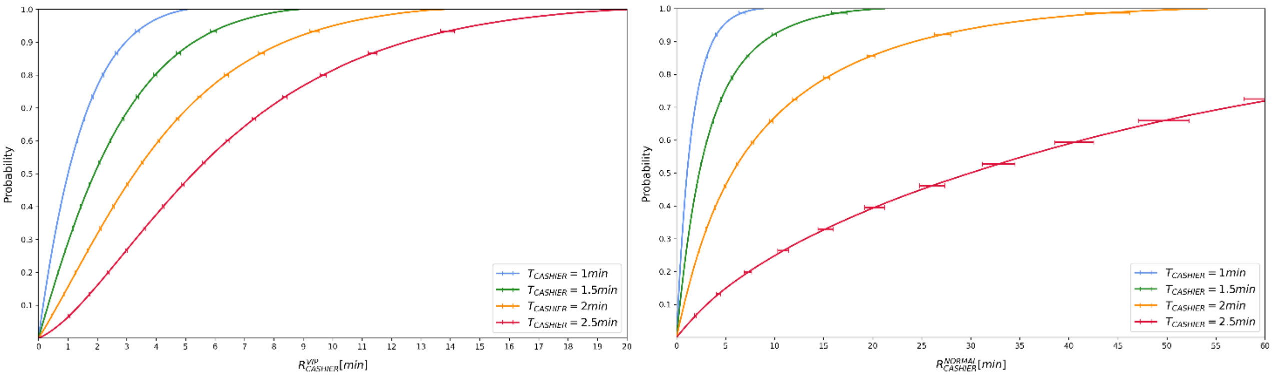

Figure 7 Comparison between the ECDFs of $R_{CASHIER}^{VIP}$ (left) and $R_{CASHIER}^{NORMAL}$ (right). The confidence level is 99%.

When $T_{CASHIER}$ increases, i.e. when the utilization $\rho_{CASHIER}$ increases, both the customer classes experience a substantial growth in the response time. The latter is asymmetric, because of the different priorities: while the VIP customers experience a graceful rise, the normal ones suffer a heavy and quick degradation of the service duration, that reaches peaks of over $60min$ when the system is close to saturation. This behaviour is attributable to the fact that, whatever the load, the response time of a VIP is linked to the VIP queue occupancy only (in the worst case, a VIP sees only one normal guest ahead of her), while the one of a normal customer depends of the crowding of both queues. A VIP membership grants a better treatment in every utilization scenario: the same discrepancy characterizes the waiting times.

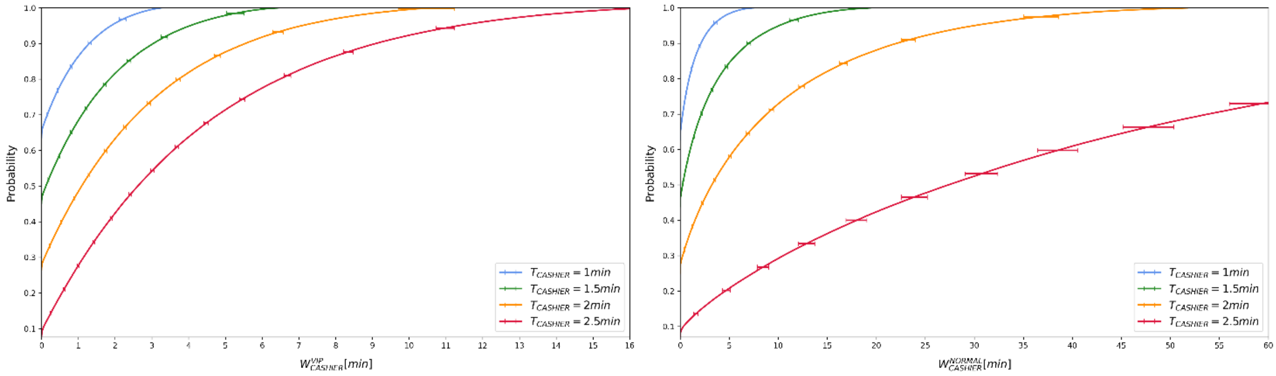

Figure 8 Comparison between the ECDFs of $W^{VIP}_{CASHIER}$ (left) and $W^{NORMAL}_{CASHIER}$ (right). The confidence level is 99%.

Although the presence of a special membership introduces clear advantages, it is also source of great unfairness when the service is facing a high workload. Since the composition of the classed population is hardly manageable (at least for what concerns the normal guests), the best solution to increase the fairness is to reduce $T_{CASHIER}$, eventually with the introduction of multiple employees: this grants low response and waiting times for all the customers, preserves the VIP advantages and guarantees a finer control over the service experience. The benefit given by the reduction of $T_{CASHIER}$ is also noticeable from the utilization expression: halving $T_{CASHIER}$ halves $\rho_{CASHIER}$, while supposing half an arrival rate of one class (e.g. of the VIPs, obtained selling less memberships) does not.

Performance index [s]	$\boldsymbol{T_{CASHIER}}$ $\boldsymbol{1min}$	$\boldsymbol{T_{CASHIER}}$ $\boldsymbol{1.5min}$	$\boldsymbol{T_{CASHIER}}$ $\boldsymbol{2min}$	$\boldsymbol{T_{CASHIER}}$ $\boldsymbol{2.5min}$
$W^{VIP}_{CASHIER,0.5}$	0	10.71 ± 0.54	70.9 ± 0.75	165.13 ± 1.21
$W^{NORMAL}_{CASHIER,0.5}$	0	17.91 ± 0.85	203.57 ± 2.63	1680.81 ± 13.53
$R^{VIP}_{CASHIER,0.5}$	62.71 ± 0.34	118.54 ± 0.62	202.64 ± 1	325.07 ± 1.38
$R^{NORMAL}_{CASHIER,0.5}$	65.83 ± 0.41	142.13 ± 0.97	349.74 ± 2.5	1835.67 ± 13.5

Table 1. Median of the waiting and response times expressed in seconds, with a confidence level of 99%.

In terms of probability distributions, the response times for both classes match a Weibull: the distribution for the normal customers is characterized by a shape $k < 1$ (heavy-tailed), while the one of the VIP customers by a shape $k > 1$ (light-tailed): this is coherent with the previous results, since the normal guests experience high response times with a not negligible probability when the utilization is low, while the VIP guests unlikely will experience the same. In both cases, the shapes increase with the utilization and the tails lighten, because of the system being closer to saturation: a higher variability would lead the node to an unstable state, which is not possible if the stability condition is verified.

Furthermore, it resembles the case of an $M/M/1/FCFS$ system, where the response times are exponentially distributed and the population of customers is concentrated inside the classical exponential curve: if the latter is split to represent the different privileges, the result is two curves with different slopes of the tails and different probabilities of experiencing small response times. This is confirmed by the coefficients of variation (< 1 for the VIP customers, > 1 for the normal customers), the histograms (which show a clear Weibull peak for the VIP distribution when the utilization is higher) and the QQ plots [Figure 9-10].

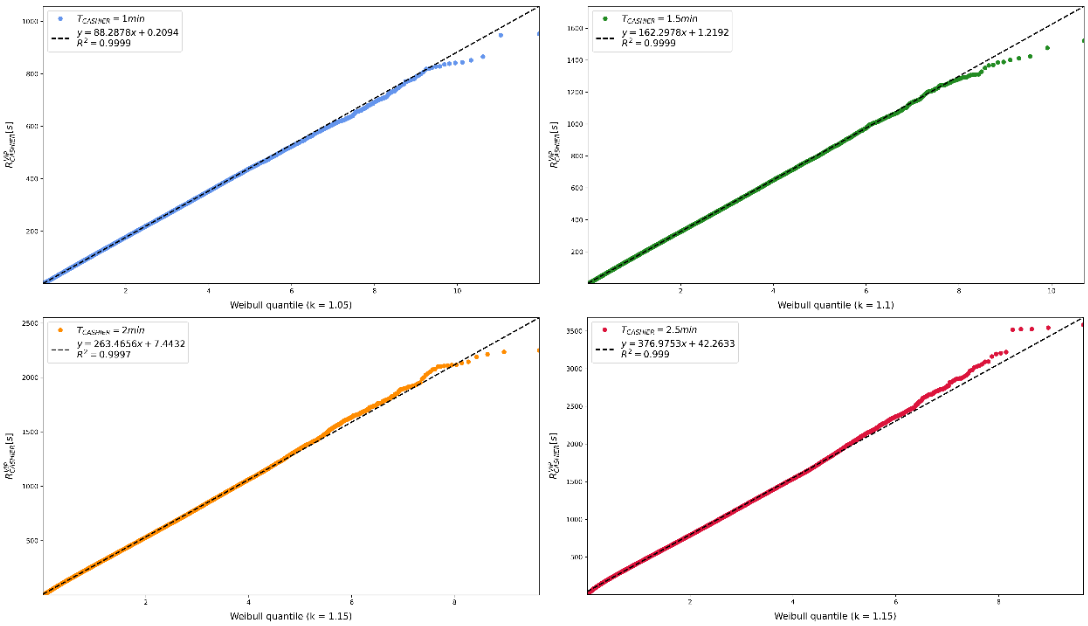

Figure 9 QQ plots of $R^{VIP}_{CASHIER}$ for increasing $T_{CASHIER}$. From left to right, proceeding by row, the Weibull shapes are 1.05, 1.1, 1.15 and 1.15.

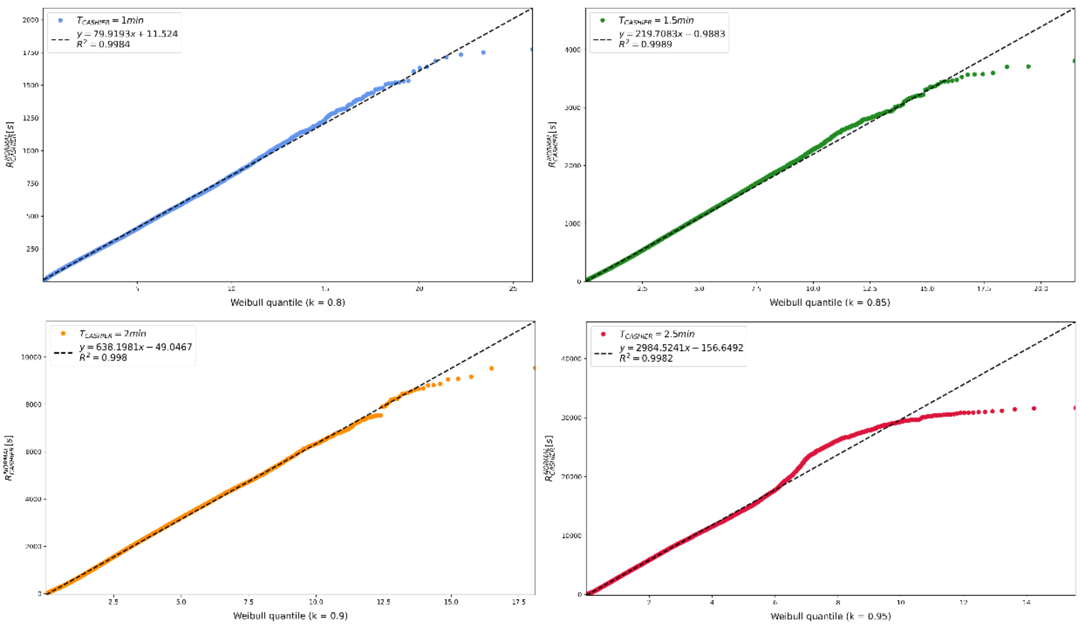

Figure 10 QQ plots of $R^{NORMAL}_{CASHIER}$ for increasing $T_{CASHIER}$. From left to right, proceeding by row, the Weibull shapes are 0.8, 0.85, 0.9 and 0.95.

The QQ plots show a good linearity, except for $R^{NORMAL}_{CASHIER}$ when $T_{CASHIER} = 2.5min$, where the right tail of the response time distribution looks shorter than the Weibull one. This is due to the system

being close to saturation, but still stable: the response time of a normal customer can be high, but not too high, since the latter case belongs to an unstable system.
The same reasoning can be applied to the waiting time distributions, whose PDFs have the shape $f_W(x) = \delta(x)r_0 + 2^{nd}\ term$, where r_0 is the customer arrival-time probability of finding an empty system and the second term depends on the arrival-time probabilities r_n of finding a non-empty system. Filtering the null waiting times, namely removing the Dirac's delta and the discontinuity in zero of the CDFs, the distribution of $2^{nd}\ term$ can be estimated using QQ plots, histograms and coefficients of variation. The result is that the one of $W_{CASHIER}^{VIP}$ match an exponential (or Weibull with $k = 1$), while the second term of $W_{CASHIER}^{NORMAL}$ resembles a Weibull with $k < 1$ (heavy-tailed) and with shape increasing with the utilization.

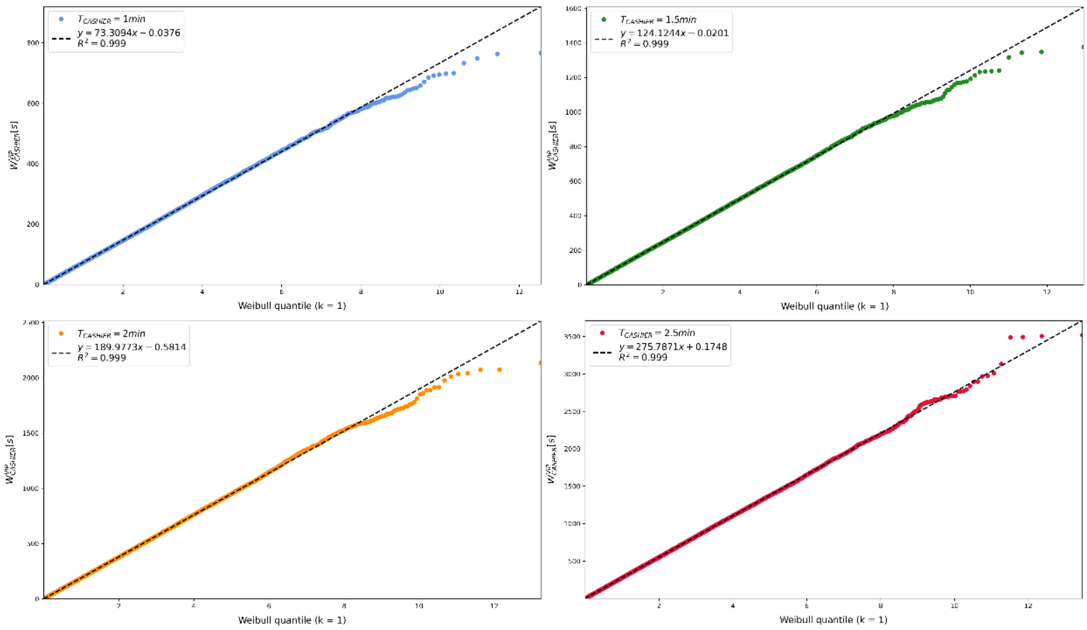

Figure 11 QQ plots of 2^{nd} term of $W_{CASHIER}^{VIP}$ for increasing $T_{CASHIER}$.

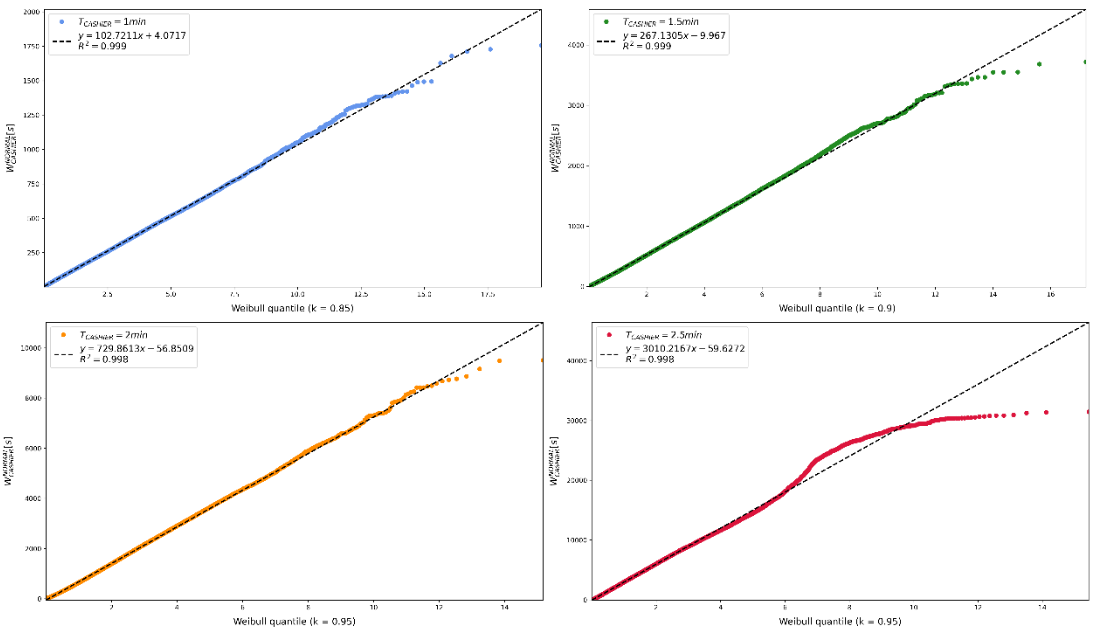

Figure 12 QQ plots of 2^{nd} term of $W_{CASHIER}^{NORMAL}$ for increasing $T_{CASHIER}$. From left to right, proceeding by row, the Weibull shapes are 0.85, 0.9, 0.95 and 0.95.

4.1.3 Queue occupancy

In order to study the queue occupancy for both classes, it is possible to proceed with a $2^k r$ factorial analysis as done in the previous section, this time with performance indexes $E[N_{q,CASHIER}^{VIP}]$ and $E[N_{q,CASHIER}^{NORMAL}]$; the factor intervals are the same, hence T_{VIP}, $T_{NORMAL} \in [5.5min, 10min]$, $T_{CASHIER} \in [1min, 2.5min]$. Applying a logarithmic transformation to meet the hypotheses of the method, the resulting models have negligible errors ($\cong 1\%$) and show that:

- 76.35% and 20.79% of the variation of $E[N_{q,CASHIER}^{VIP}]$ are due respectively to $T_{CASHIER}$ and T_{VIP}, with negligible impact due to other factors or interactions between them (at most $\cong 2\%$);
- 75.06% and 13.91% of the variation of $E[N_{q,CASHIER}^{NORMAL}]$ are due respectively to $T_{CASHIER}$ and T_{NORMAL}, with negligible impact due to other factors or interactions between them (at most $\cong 7\%$ due to T_{VIP});

Thus, the occupancy of a queue can be studied varying the inter-arrival times of the related customer class in the set $\{5.5min, 7.5min, 10min\}$ and varying $T_{CASHIER}$ in the set $\{1min, 1.5min, 2min, 2.5min\}$ [Figure 11].

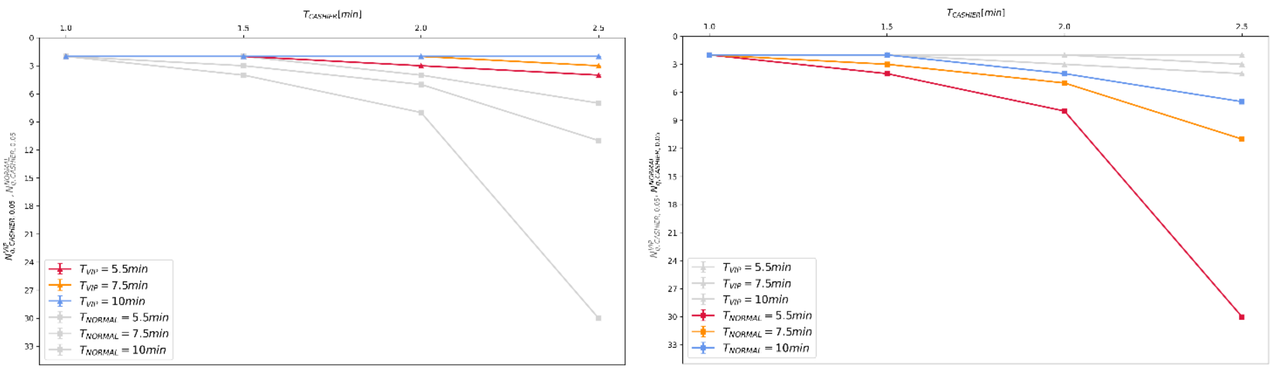

Figure 13 The behaviour of the 0.95 quantile of $N_{q,CASHIER}^{VIP}$ (left) and $N_{q,CASHIER}^{NORMAL}$ (right) varying T_{VIP}, T_{NORMAL} and $T_{CASHIER}$. The confidence level is 99%.

Both queues grow with the utilization, but asymmetrically depending on the customer class: VIP queue is less crowded, more stable when the utilization is higher and varies gracefully in length with the variation of T_{VIP}; the normal queue increases quickly together with the utilization and it is sensible to the variation of T_{NORMAL}, which can cause significative peaks. This can be justified by the normal queue being emptied only when the VIP counterpart is empty.

The probability distributions show an index of dispersion in all the scenarios higher than 1, with all the histograms showing an early peak [Figure 12].

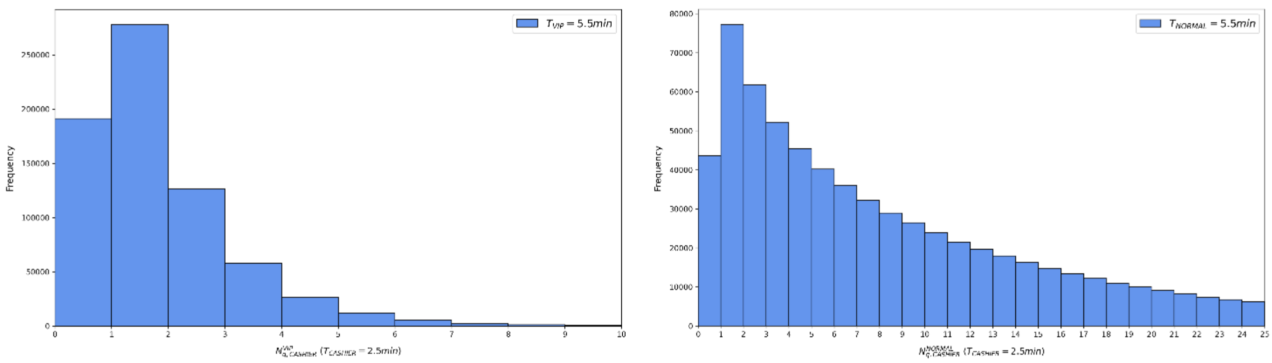

Figure 14 Histograms for $N_{q,CASHIER}^{VIP}$ and $N_{q,CASHIER}^{NORMAL}$, showing an early peak in the EPMFs.

This last result excludes the possibility of a geometric distribution, for both queues: the guess is that $N_{q,CASHIER}^{VIP}$ and $N_{q,CASHIER}^{NORMAL}$ follow a distribution akin to a discrete Weibull, with a duality similar to the one between an exponential and a geometric distribution. QQ plots performed with a rough discretization of a Weibull, with shape $k > 1$ and by means of the operator $\lfloor \cdot \rfloor$, show a good linearity, but they provide less confidence in the result compared to the cases of waiting and response times.

The dimensioning of the queues can be done aiming at minimizing the customer loss rate with the highest utilization, i.e. in the case of $T_{VIP}, T_{NORMAL} = 5.5min$ and $T_{CASHIER} = 2.5min$: a possible sizing consists in a 99% dimensioning of the VIP queue, resulting in $K_{q,CASHIER}^{VIP} = 6$, and a 95% dimensioning of the normal queue, resulting in $K_{q,CASHIER}^{NORMAL} = 30$. The latter prevents a too costly dimensioning in the worst case (the 0.99 quantile of $N_{q,CASHIER}^{NORMAL}$ is 46).

4.2 Seating node

The seating node can be studied leveraging Burke's theorem, in order to estimate the distribution of arrivals to the tables and to study their behaviour in isolation with respect to the cashier. Burke's theorem is valid despite having two customer classes in the cashier node, because the queue discipline is irrelevant, being the output and not the delay distribution of interest (Burke, P. J. *The Output of a Queuing System*, 1956), but requires infinite queues by hypothesis. Since the previously proposed dimensioning aims at minimizing the customer loss, the expected inter-departure times from the cashier should not differ greatly from being exponentially distributed: taking as reference the case of highest $\rho_{CASHIER}$ (highest losses), obtained with $T_{VIP}, T_{NORMAL} = 5.5min$ and $T_{CASHIER} = 2.5min$, the QQ plot, coefficient of variation and mean value show that the arrivals to the seating node are approximately exponential with a rate $\lambda_{VIP} + \lambda_{NORMAL}$ [Figure 13].

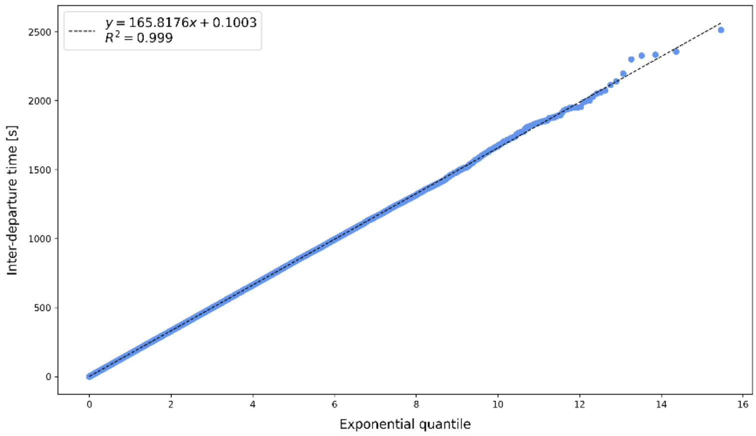

Figure 15 QQ plot of the inter-departure times of the cashier against the exponential quantiles.

Given these observations, the seating node is an $M/M/N_{SEAT}/FCFS$ system, for which the stability condition is:

$$\rho_{SEATING} = \frac{\lambda_{VIP} + \lambda_{NORMAL}}{N_{SEAT} \cdot t_{EATING}} < 1$$

Theoretical results showing the interaction between $\lambda = \lambda_{VIP} + \lambda_{NORMAL}$, t_{EATING} and N_{SEAT} are available for such a system, including expressions for the probability p_n of having n customers in the node, $E[N_{q,SEATING}]$, $E[W_{SEATING}]$ and $E[R_{SEATING}]$, both with infinite and finite queues.

A formula to dimension the waiting queue according to a desired customer loss probability can be formally derived, considering a finite memory system $M/M/N_{SEAT}/K/FCFS$.

Called $u = \frac{\lambda_{VIP}+\lambda_{NORMAL}}{t_{EATING}}$, the loss probability is equal to:

$$p_{LOSS} = \frac{u^K}{N_{SEAT}!\, {N_{SEAT}}^{K-N_{SEAT}}} p_0$$

$$\begin{cases} p_0 = \dfrac{1}{\sum_{j=0}^{N_{SEAT}-1} \dfrac{u^j}{j!} + \dfrac{u^{N_{SEAT}}}{N_{SEAT}!} \dfrac{1-(u/N_{SEAT})^{K-N_{SEAT}+1}}{1-(u/N_{SEAT})}} & \dfrac{u}{N_{SEAT}} \neq 1 \\ p_0 = \dfrac{1}{\sum_{j=0}^{N_{SEAT}-1} \dfrac{u^j}{j!} + \dfrac{u^{N_{SEAT}}}{N_{SEAT}!} (K - N_{SEAT} + 1)} & \dfrac{u}{N_{SEAT}} = 1 \end{cases}$$

Considering a worst case of $T_{VIP}, T_{NORMAL} = 5.5min$ and $T_{EATING} = 15min$, and starting with $N_{SEAT} = 6$, which is the minimum number of seats that grants a stable system with an infinite queue, the best way to minimize the loss probability is to increase the number of seats at the expense of increasing the queue length [Figure 14].

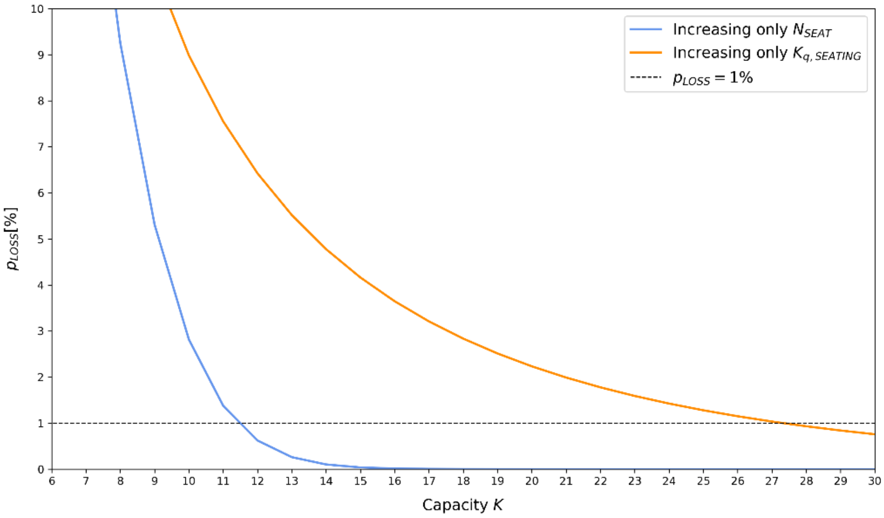

Figure 16 Behaviour of p_{LOSS} increasing only N_{SEAT} or increasing only $K_{q,SEATING}$.

The plot shows that it is possible to achieve $p_{LOSS} < 1\%$ with just $N_{SEAT} = 12$, instead that with $K_{q,SEATING} = 28$. This strategy grants that a customer never waits for a table after being served at the cashier node ($K_{q,SEATING} = 0$) and a response time dependent of only the eating time, so that $E[R_{SEATING}^{VIP}] = E[R_{SEATING}^{NORMAL}] = T_{EATING}$.