Online movie streaming project

Optimization and Simulation Course (MATH-600)

Teaching Assistants: CCB,PI,LR **Professor:** MB

Spring Semester 2024

Introduction

We consider an online movie streaming service provided in Switzerland. The provider plans to offer happy hour to its customers in the future, during which the price of the ten most popular movies will be reduced. The provider stores all the movies in the main storage node (MSN) outside Switzerland. It also has two additional storage nodes (ASN) in Switzerland, in Lausanne and Locarno. The ASNs have limited capacity. The time it takes to serve a movie from a storage node to a customer depends on the distance between them (it increases with an increase of the distance). The customers are considered at an aggregated level. They are grouped according to the main languages spoken in Switzerland: French, Swiss German and Italian.

Given the description of groups, movies and storage nodes you have to decide which movies to put on which additional storage node, so as to minimize waiting time for all requests.

The aim of the "Simulation Project" is to develop a discrete event simulation that represents the system and to evaluate the performance of two solutions.

During the "Optimization Project", the discrete event simulation is expanded, and a good solution is identified by an optimization algorithm.

Develop the discrete event simulation with a modular structure. It should be possible to modify the various components, such as the activity pattern, service times, etc., during the "Optimization Project".

Project description

We consider the system illustrated in Figure 1. It consists of

- 1 MSN: unlimited capacity, located outside Switzerland
- 2 ASNs, both with the capacity of 3500 megabytes [MB]: ASN1 located in Lausanne and ASN2 located in Locarno
- 10 movies: movie #0,...,movie #9
- 3 customer groups: G1-French, G2-Swiss German, G3-Italian.

The MSN stores all the movies existing in the system. Additionally, each movie can be stored on 0, 1 or both ASNs. The ASNs serve customers on a First-Come-First-Served (FCFS) basis.

Each movie is characterized by the size in MB. The sizes of the movies are given in Table 1.

Table 1: Size of the movies [MB]

Movie	#0	#1	#2	#3	#4	#5	#6	#7	#8	#9
Size [MB]	850	950	1000	1200	800	900	1000	750	700	1100

The movies are associated with their popularity within each group, which is reported in Table 2.

Group/Movie	#0	#1	#2	#3	#4	#5	#6	#7	#8	#9
G1	$\omega_{1-0} = 2$	$\omega_{1-1} = 4$	$\omega_{1-2} = 9$	$\omega_{1-3} = 8$	$\omega_{1-4} = 1$	$\omega_{1-5} = 3$	$\omega_{1-6} = 5$	$\omega_{1-7} = 7$	$\omega_{1-8} = 10$	$\omega_{1-9} = 6$
G2	$\omega_{2-0} = 6$	$\omega_{2-1} = 1$	$\omega_{2-2} = 3$	$\omega_{2-3} = 4$	$\omega_{2-4} = 7$	$\omega_{2-5} = 9$	$\omega_{2-6} = 2$	$\omega_{2-7} = 5$	$\omega_{2-8} = 8$	$\omega_{2-9} = 10$
G3	$\omega_{3-0} = 4$	$\omega_{3-1} = 7$	$\omega_{3-2} = 3$	$\omega_{3-3} = 6$	$\omega_{3-4} = 1$	$\omega_{3-5} = 10$	$\omega_{3-6} = 2$	$\omega_{3-7} = 9$	$\omega_{3-8} = 8$	$\omega_{3-9} = 5$

Table 2: Group/Movie popularity

All groups are connected to the MSN. Additionally, the G1 is connected to the ASN1, the G2 is connected to the ASN1 and ASN2, and the G3 is connected to the ASN2. The groups are characterized by their activity pattern, expressed in the number of requests per second. The provider has estimated the group activity pattern based on historical data for a representative hour, and it is given in Table 3.

Each group is associated with the time it takes to send a request to the MSN and to the ASNs, and the time it takes to serve a movie from the MSN and

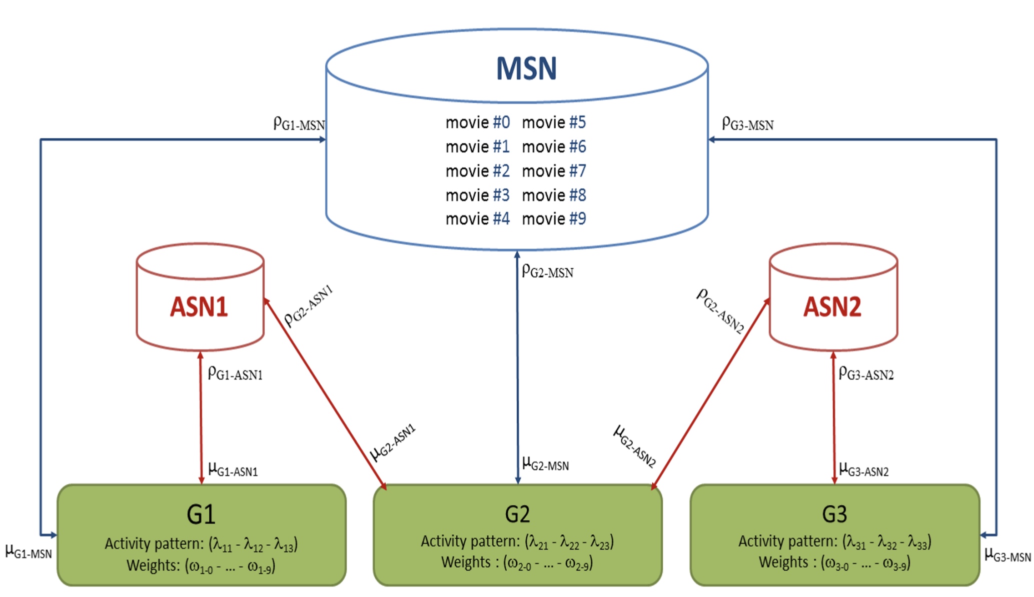

Figure 1: Online movie streaming - storage

Groups	First interval: 20 min	Second interval: 20 min	Third interval: 20 min
G1	$\lambda_{11} = 0.8$	$\lambda_{12} = 1.2$	$\lambda_{13} = 0.5$
G2	$\lambda_{21} = 0.9$	$\lambda_{22} = 1.3$	$\lambda_{23} = 0.3$
G3	$\lambda_{31} = 0.7$	$\lambda_{32} = 1.5$	$\lambda_{33} = 0.4$

Table 3: Activity patterns for a representative hour[the number of requests/second]

from the ASNs to a group, both in seconds. The time needed to send a request is given in Table 4. The reported times indicate how close a storage node is to each group. For instance, ASN1 is closer to the G2 than ASN2, and ASN2 is closer to the G2 than MSN.

Each storage node needs certain time to handle a request. This time follows an exponential distribution with the mean of 0.5 second, which is independent of movie characteristics, groups or storage nodes. During this time, the storage node cannot handle other requests. Once the request is handled by a storage node, additional time is needed to serve a movie from a storage node to a group. A user can start watching the movie immediately after the time needed to serve it. This time consists of a deterministic part and an extra random part. The deterministic part depends on the movie size as listed in Table 5. The random part does not depend on movie characteristics, groups or storage nodes. It is uniformly distributed between 0.3 and 0.7 seconds.

Group	MSN	ASN1	ASN2
G1	$\rho_{\text{G1-MSN}} = 0.5$	$\rho_{\text{G1-ASN1}} = 0.2$	/
G2	$\rho_{\text{G2-MSN}} = 0.5$	$\rho_{\text{G2-ASN1}} = 0.3$	$\rho_{\text{G2-ASN2}} = 0.4$
G3	$\rho_{\text{G3-MSN}} = 0.5$	/	$\rho_{\text{G3-ASN2}} = 0.2$

Table 4: Time needed to send a request [in seconds]

Each group request is sent to the closest storage node which contains the

Movie size [MB]	Group	MSN	ASN1	ASN2
700-900	G1	$\mu_{\text{G1}-\text{MSN}} = 9$	$\mu_{\text{G1}-\text{ASN1}} = 3$	/
	G2	$\mu_{\text{G2}-\text{MSN}} = 8$	$\mu_{\text{G2}-\text{ASN1}} = 4$	$\mu_{\text{G2}-\text{ASN2}} = 5$
	G3	$\mu_{\text{G3}-\text{MSN}} = 10$	/	$\mu_{\text{G3}-\text{ASN2}} = 4$
900-1100	G1	$\mu_{\text{G1}-\text{MSN}} = 12$	$\mu_{\text{G1}-\text{ASN1}} = 4$	/
	G2	$\mu_{\text{G2}-\text{MSN}} = 11$	$\mu_{\text{G2}-\text{ASN1}} = 5$	$\mu_{\text{G2}-\text{ASN2}} = 6$
	G3	$\mu_{\text{G3}-\text{MSN}} = 13$	/	$\mu_{\text{G3}-\text{ASN2}} = 5$
1100-1500	G1	$\mu_{\text{G1}-\text{MSN}} = 15$	$\mu_{\text{G1}-\text{ASN1}} = 5$	/
	G2	$\mu_{\text{G2}-\text{MSN}} = 14$	$\mu_{\text{G2}-\text{ASN1}} = 6$	$\mu_{\text{G2}-\text{ASN2}} = 7$
	G3	$\mu_{\text{G3}-\text{MSN}} = 16$	/	$\mu_{\text{G3}-\text{ASN2}} = 6$

Table 5: Time needed to serve a movie – deterministic part [in seconds]

requested movie. The provider of the service considers *user satisfaction* to quantify the performance of the system. The satisfaction is a function of waiting time, and it decreases with an increase of waiting time.

We assume that at the beginning of the happy hour the system is empty. All the requests made within the considered hour are served. The requests arriving after the happy hour are ignored.

Simulation

For the simulation project, you are requested to:

- Develop a discrete event simulation to represent the described system.
- Define the indexes used to quantify the quality of the service
    - Remember that extreme cases are also important; evaluate other indexes in addition to the mean. Moreover, the waiting time is an important factor for how your service is perceived.

  - Report the mean squared error of your estimation using bootstrapping when necessary.
  - Use variance reduction techniques to reduce the computational time. The computational time is important in the optimization part.

- Evaluate the quality of the system during happy hour with two different configurations
  - Movies #2, #3 and #9 are stored on both ASN1 and ASN2.
  - Movies #3, #5, #7 and #8 are stored on both ASN1 and ASN2.
  - (Or other scenarios which show the performance of simulation model well.)
- Make any necessary assumptions.

Optimization

The performance of the service can be improved by a smarter choice of the movies to be stored on the ASNs.

For the optimization project, you are requested to:

- Identify the decision variables of the problem.
- Define the objective function.
- Design an optimization algorithm and apply it to solve the problem. The value of the objective function is evaluated using simulation.
- Like in the simulation project, the objective function can reflect various policies: whether they want to optimize over the average, best, worst, or certain percentile of the objective function distribution. Decide what your position is and justify it, or present results for several alternatives.
- Use your creativity and design a new storing strategy that leads to better performance.