

Contents

1	Introduzione e obiettivi	1
1.1	Introduzione	1
1.2	Obiettivi	2
2	Modellazione Concettuale	3
2.1	Astrazione	3
2.2	Modellazione	4
2.2.1	Indagini preliminare	4
2.2.2	Udienza preliminare	5
2.2.3	Primo grado	6
2.2.4	Secondo grado	8
2.2.5	Terzo grado	9
2.2.6	Definizione dello stato	10
3	Modello delle specifiche	11
3.1	Distribuzioni	11
3.2	Probabilità di routing	12
3.2.1	Procura	12
3.2.2	Udienza preliminare	13
3.2.3	Ingresso tribunali fisici	14
3.2.4	Uscita Tribunale fisico	15
3.2.5	Uscita corte di appello fisica	15
3.2.6	Uscita corte di appello assise fisica	16
3.3	Tempi medi di servizio	17
4	Implementazione	18
4.1	Architettura	18
4.2	Generatori pseudocasuali	18
5	Verifica e Validazione	19
5.1	Verifica	19
5.1.1	Single Server	20
5.1.2	Multi Server	21

5.1.3	Sottoreti	22
5.2	Validazione	25
6	Simulazione orizzonte temporale infinito	26
6.1	Procura	26
6.2	Udienza preliminare	26
6.3	Primo grado monocratico	28
6.4	Tribunale	28
6.5	Corti	30
7	Simulazione orizzonte finito	31
7.1	Introduzione	31
7.2	Tempi di risposta	31
7.3	Popolazione nel centro	31
7.4	Andamento utilizzazione	33
8	Modello migliorativo	34
8.1	Introduzione	34
8.2	Versione 1	34
8.3	Versione 2	37
8.4	Versione Finale	39
	Bibliography	44

–1–

Introduzione e obiettivi

1.1 Introduzione

Obiettivo di questo progetto è la realizzazione di un modello a reti di code ispirato a un caso di studio reale. Il sistema reale che si prende in analisi è il sistema giudiziario **penale** applicato alla regione Lazio. Tale sistema può essere descritto da un insieme di fasi in sequenza che vengono sintetizzate nell'immagine 1.1. In questo paragrafo verrà descritto il sistema giudiziario italiano, che è applicato allo stesso modo in ogni regione italiana e quindi anche nel Lazio.

Il tutto ha inizio con la denuncia da parte di un cittadino che si traduce in una sequenza di operazioni che possono coinvolgere diverse tipologie di processo. In generale la fase che segue la denuncia è la fase di **indagini preliminari** che viene condotta dalla procura. In questa fase si raccolgono dati, si effettuano interrogatori e altre attività specifiche. Le indagini preliminari si svolgono a cicli di 180 giorni. Terminato il primo ciclo viene fissata un'**udienza preliminare** nella quale il GUP(ovvero giudice dell'udienza preliminare) comunica se le indagini condotte sono sufficienti o è necessario aprire un nuovo ciclo di indagini. Quando il GUP riterrà i dati a sua disposizione sufficienti li analizzerà e stabilirà se la denuncia sussiste (e quindi prosegue il suo percorso nel sistema giudiziario) oppure se è possibile archiviare il caso. Se la denuncia sussiste è possibile attraversare tre macro fasi che sono il **processo di primo grado** , **il processo di secondo grado** e **il processo di terzo grado**. La modellazione di ciascuna di queste fasi verrà descritta nel dettaglio. Giusto per dare un'idea si parla di processo di primo grado nel momento in cui un processo è analizzato per la prima volta da un giudice. Nel processo di secondo grado un altro giudice analizza lo stesso caso per una seconda volta , questo scenario potrebbe verificarsi se ad esempio una persona condannata non accetta l'esito del processo di primo grado e richiede un ricorso. Analogamente si parla di processo di terzo grado se un processo è analizzato per ben tre volte : questo scenario si verifica se un condannato richiede il ricorso sia per il processo di primo grado che per il processo di secondo grado.

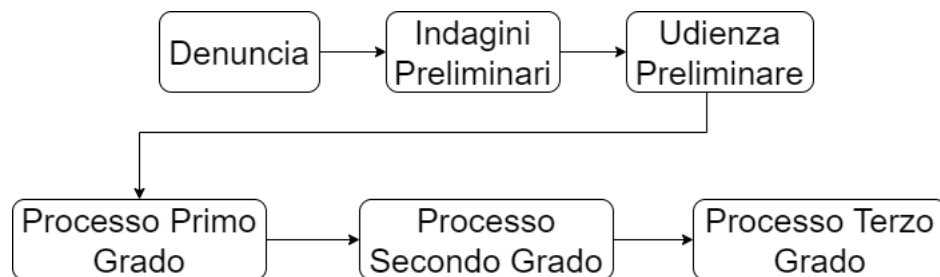


Figure 1.1: L'immagine mostra le fasi che compongono il sistema penale italiano. Queste fasi nel modello reale non sono perfettamente in sequenza tra di loro , possono verificarsi dei collegamenti particolari che verranno discussi in maniera approfondita successivamente. L'immagine vuole essere esclusivamente a scopo illustrativo , per mostrare quali sono le fasi nella loro totalità e per introdurre gli obiettivi del progetto.

1.2 Obiettivi

Attualmente il sistema giudiziario penale nella regione Lazio ,ma in generale in tutta Italia, risulta essere uno dei meno performanti in Europa. Per dare un'idea, la tabella 1.2 mette a confronto le durate medie della varie fasi nel sistema giudiziario italiano e nel sistema giudiziario europeo. Si noti come i tempi medi siano molto lontani tra loro. Addirittura il processo di secondo grado in Italia è circa 6 volte più lento che in Europa.

	Italia	Europa
Indagini Preliminari	323 giorni	160 giorni
Udienza Preliminare	35 giorni	30 giorni
Primo Grado	310 giorni	128 giorni
Secondo Grado	876 giorni	143 giorni
Terzo Grado	191 giorni	143 giorni

Alla luce di questi dati gli obiettivi dello studio sono:

- (i) creare un modello a reti di code e studiarlo tramite simulazione next event. Il modello base deve contenere centri che modellano le vari fasi di indagini preliminari, udienza preliminare ecc.. ; tali centri devono avere un tempo di risposta medio che si avvicina quanto più possibile ai valori medi italiani mostrati in tabella.
- (ii) creare un modello migliorativo partendo dal modello base in cui i tempi di risposta medi dei centri converga ai tempi di risposta medi europei.

–2–

Modellazione Concettuale

2.1 Astrazione

Per la realizzazione di questo modello è fondamentale ricorrere ad alcune astrazioni in modo da semplificare la complessità del dominio applicativo.

1. La prima astrazione fondamentale è che una denuncia corrisponde ad un job. Nel momento in cui verrà sporta una nuova denuncia verrà generato un **job** che entrerà nel sistema e verrà opportunamente elaborato. Da questa modellazione ne consegue che **il tasso di ingresso medio** dei job nel sistema corrisponde al numero medio di denunce che vengono sporte ogni anno.
2. La seconda astrazione riguarda concetto di **servente**. I serventi in questo modello sono sia le persone (come i giudici), sia le infrastrutture fisiche che ospitano le sentenze presiedute dai giudici stessi. Per quanto riguarda le infrastrutture fisiche si considerano come serventi:
 - tribunale fisico(T) : accoglie processi di primo e secondo grado;
 - Corte di Appello (CA) : accoglie processi di primo e secondo grado;
 - Corte di Assise fisica (CAS) : accoglie solo processi di secondo grado;
 - Corte di Cassazione fisica (CC) : accoglie solo processi di terzo grado;

Si vogliono considerare anche le infrastrutture fisiche come serventi in quanto ,avendo in Italia e di conseguenza nel Lazio un numero di strutture carenti , è verosimile pensare che un giudice ,per comunicare l'esito della condanna, debba attendere che una struttura fisica possa accogliere la sentenza. Si tratta ,quindi, dell'introduzione di un ritardo che potrebbe non essere trascurabile e che è opportuno modellare.

-
3. La terza astrazione riguarda il concetto di **tempo di esecuzione**. Per tempo di esecuzione di un job da parte di un giudice si intende il tempo impiegato da una o più figure giuridiche per studiare un caso e trovare una soluzione alla controversia. Dunque ci si riferisce a tutto il lavoro che viene prima di una sentenza. Il lavoro del giudice ha inizio nel momento in cui egli apre il fascicolo ed analizza la pratica di una determinata denuncia e termina nel momento in cui ha realizzato il verbale contenente l'esito della sentenza che dovrà essere comunicata in una opportuna infrastruttura fisica. La definizione di tempo di esecuzione è molto aleatoria, tuttavia per modellare questi aspetti lo studio verrà guidato da dati statistici [1]. In particolare, verranno sfruttati dati offerti dal sito ufficiale del Ministero della Giustizia. Per quanto riguarda le infrastrutture fisiche il tempo di esecuzione indica effettivamente il tempo necessario per comunicare l'esito del processo alle persone denunciate.

2.2 Modellazione

2.2.1 Indagini preliminare

La prima fase che si intende modellare è quella della indagini preliminari. Tale fase nel dominio applicativo è gestita dalla procura. Come assunzione si ipotizzi la presenza di una procura per ogni città nel Lazio. Quindi un totale di 5 procure: Roma, Frosinone, Viterbo, Rieti e Latina. Si modella la procura come un unico server. In totale 5 serveri singoli che lavorano in parallelo ma che ricevono in ingresso flussi indipendenti. Si parla di flussi indipendenti perchè ogni città ha il proprio pool di cittadini che possono dar luogo ad una denuncia. Ad esempio una denuncia di un cittadino romano è gestita dalla procura di Roma. Queste osservazioni portano a modellare la procura della regione Lazio come un multiserver indipendente. La modellazione è visibile nell'immagine 2.1.

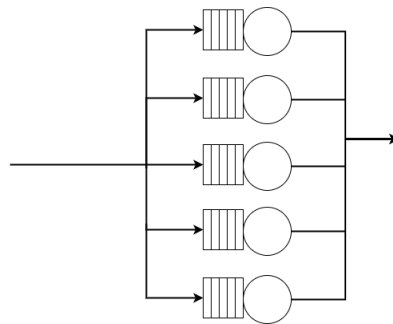


Figure 2.1

2.2.2 Udienza preliminare

Terminata la fase di indagini preliminari da parte della procura il testimone è passato al GUP (Giudice dell'udienza preliminare), il quale procede con l'udienza preliminare. Per questa modellazione è richiesto uno sforzo di astrazione. Per semplificare il modello si considera il pool di tutti i giudici del Lazio che possono presidiare l'udienza preliminare come un unico single server. In pratica il single server è come se nascondesse al suo interno una struttura ben più complessa che è considerata in questo modello in maniera black box. Una possibile alternativa potrebbe essere quella di modellare l'udienza preliminare come un multiserver dove ogni servente è un giudice. Tuttavia questo avrebbe dato luogo a un sistema eccessivamente complesso.

Si noti come il flusso in uscite delle procure si ricongiunge nell'unico single server dell'udienza preliminare poichè il lavoro dei GUP non è "locale". Un giudice che ha sede legale a Roma potrebbe ricevere un incarico di lavoro anche a Latina. In questo caso inoltre si è ritenuto non necessario considerare un'infrastruttura fisica che accolga la sentenza dell'udienza preliminare. Infatti il Gup quando termina il proprio lavoro comunica il risultato in un incontro che si svolge solitamente in 15 minuti e in sedi private. Proprio a causa della brevità di questo incontro si è preferito non modellarlo in quanto avrebbe introdotto ritardi trascurabili.

Terminata anche la fase di udienza preliminare il flusso può seguire una ripartizione. Il GUP può ritenere che dati in suo possesso non siano sufficienti per completare l'udienza preliminare e quindi richiede alla procura di poter effettuare ulteriori indagini. Dal punto di vista della modellazione del sistema si viene a creare un feedback che connette il server dell'udienza preliminare con il server delle indagini preliminari. Se invece il giudice ritiene che i dati in suo possesso siano adeguati allora è possibile passare per il processo di primo grado. Il modello parziale che si è venuto a creare è descritto nell'immagine 2.2.

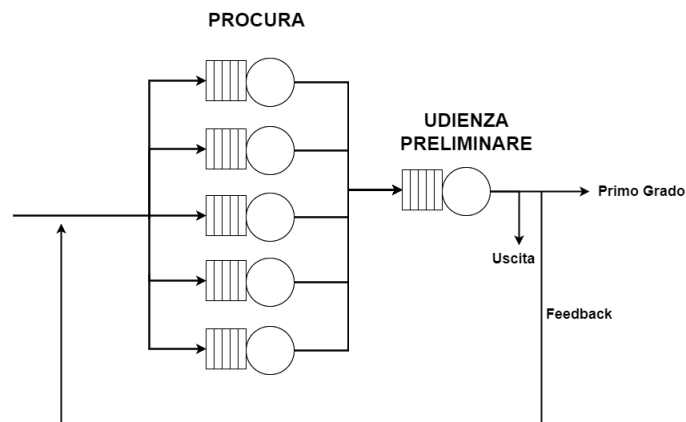


Figure 2.2

2.2.3 Primo grado

Nel processo di primo grado i job possono essere classificati come poco gravi, a gravità media e molto gravi. Un esempio di job poco grave potrebbe essere una denuncia per furto, un job a gravità media potrebbe essere una denuncia per aggressione mentre un job a gravità alta potrebbe essere una denuncia per omicidio. Ciascuna di questa tipologia di job è gestita in maniera opportuna.

- I job poco gravi vengono gestiti dal dibattimento monocratico. Ovvero tutta la pratica della denuncia è gestita da un unico giudice, che effettua le analisi e le valutazioni del caso. Ancora una volta il pool di giudici che possono presidiare il dibattimento monocratico verranno modellati come un unico single server e si considera il sottosistema in modo black box. Terminato il lavoro da parte del singolo giudice devono essere comunicati i risultati. Dunque ci sarà una sentenza in un tribunale fisico che non è organizzata immediatamente, ma avverrà al trascorrere di un certo numero di giorni che dipendono da quante sentenze si sono accumulate nel tempo. Questa attesa per poter ottenere l'esito della condanna può essere modellata con un insieme di 5 single server paralleli tra loro. Si è preferito utilizzare un single server multiplo piuttosto che un multi server perché i tribunali fisici sono "locali". Un cittadino romano verrà convocato dal tribunale fisico di Roma. Ci sono infatti 5 tribunali, uno per ogni città.

I job di gravità bassa, dopo essere passati per il dibattimento monocratico e per il tribunale fisico, possono trovarsi di fronte due possibili strade: o il giudice ritiene conclusa la causa e il job lascia il sistema, in alternativa il denunciato non accetta la sentenza e richiede il ricorso. In questo caso si parla di processo di secondo grado. Il modello parziale è descritto nell'immagine 2.3.

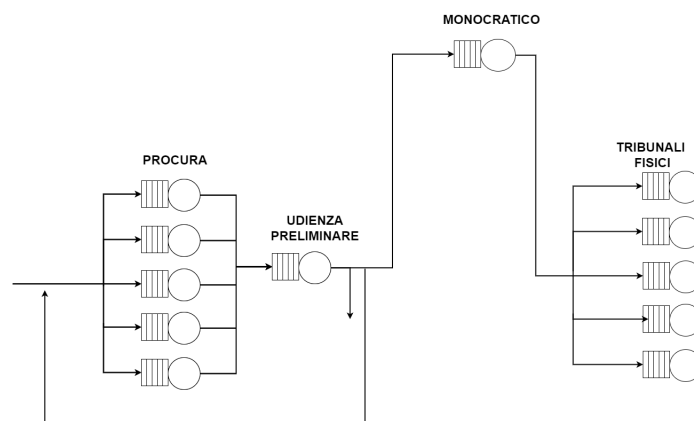


Figure 2.3

- I job a gravità media invece , data la maggiore importanza , verranno analizzati da un gruppo di 3 diversi giudici definito *tribunale*. Per evitare incomprensioni nella lettura del documento, quando si utilizzerà il termine "tribunale" si farà riferimento all'insieme di giudici(quindi un insieme di persone), quando invece si utilizzerà il termine "tribunale fisico" si farà riferimento all'infrastruttura fisica che accoglie la sentenza. Si modella anche in questo caso il pool di giudici che possono presidiare un processo di primo grado a gravità media come un single server. Anche per i job di gravità media il risultato del lavoro dei giudici verrà comunicato in una sentenza organizzata in un tribunale fisico. Per la precisione, gli stessi tribunali fisici utilizzati dal dibattito monocratico. Inevitabilmente in questa fase si viene a creare una fusione di due flussi verso il tribunale fisico. Si osservi l'immagine2.4.

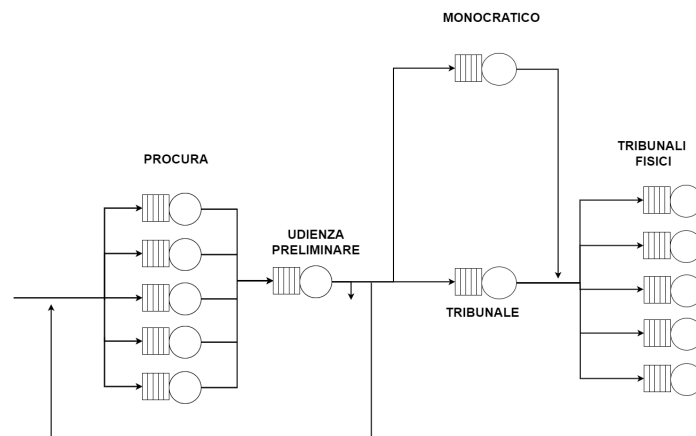


Figure 2.4

- I job molto gravi ,infine ,vengono gestiti da un gruppo di 8 giudici definito *corte di appello*, che lavora alla stessa denuncia. Valgono le stesse osservazioni fatte in precedenza. Per cui, anche questa fase può essere modellata come un unico single server. A differenza dei job di gravità bassa e media , i risultati della sentenza verranno comunicati in una struttura definita corte di appello fisica che è una struttura separata rispetto a quelle del tribunale fisico. In particolare ogni regione ha la propria corte di appello fisica. Quindi, il flusso in uscita dal single server è collegato a un altro single server. Si osservi l'immagine2.5.

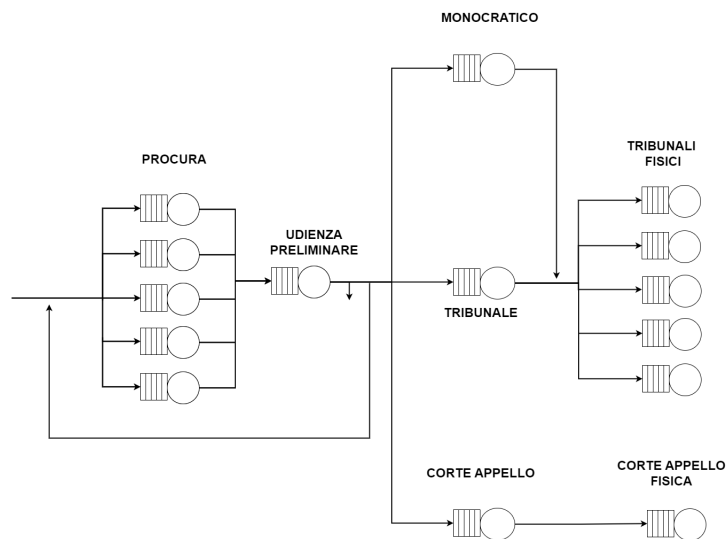


Figure 2.5

2.2.4 Secondo grado

Per quanto riguarda il secondo grado è opportuno seguire ancora la distinzione tra job di gravità media, bassa e alta. I job a gravità bassa dopo essere passati per il dibattimento monocratico e il tribunale fisico vengono presi in carico da un gruppo 3 di giudici. Qui il sistema inizia a complicarsi pochè si vengono a creare degli intrecci particolari. I giudici che possono lavorare a questi processi di secondo grado di gravità bassa sono esattamente gli stessi che si occupano dei processi di primo grado di media gravità. Poichè al termine del processo di secondo grado la sentenza avverrà sempre in un tribunale fisico, per modellare questa fase è sufficiente aggiungere un feedback che va dai tribunali fisici all'ingresso del tribunale.

Per quanto riguarda i processi(job) di secondo grado a gravità media , dopo essere usciti dai tribunali fisici vengono presi in analisi dalla *corte di appello*, che è anche incaricati di lavorare ai processi di primo grado gravi.

Infine, i job di gravità alta per il processo di secondo grado verranno analizzati da un gruppo di 8 giudici definito *corte di assise*. Anche in questo caso valgono i ragionamenti fatti in precedenza: la corte di assise fisica, in cui si espongono i risultati della sentenza, è unica nella regione. Pertanto anche in questo caso ci sarà una sequenza di due single server. Si osservi l'immagine2.6. Le aggiunte rispetto all'immagine 2.5 sono evidenziate in rosso.

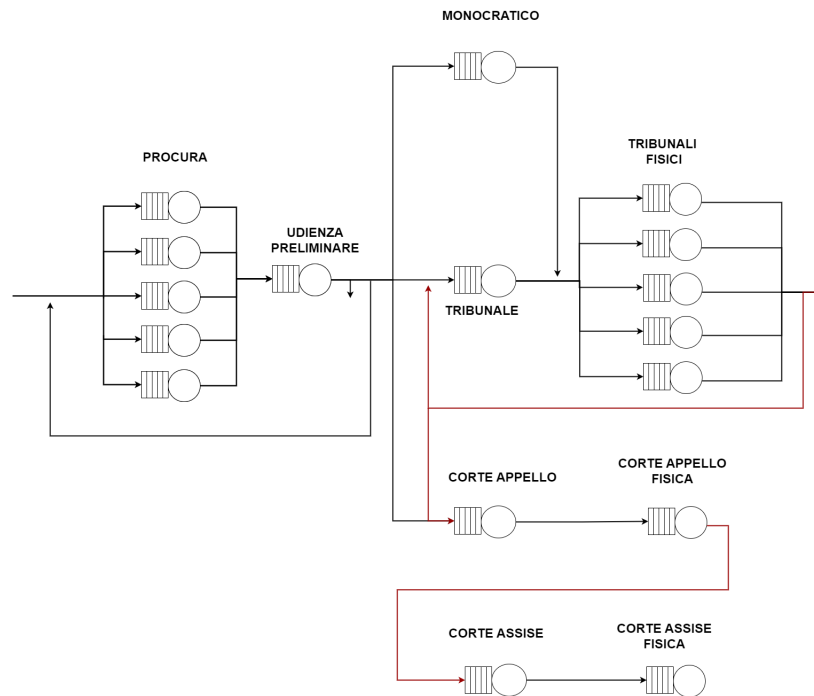


Figure 2.6

2.2.5 Terzo grado

La modellazione del terzo grado di processo è abbastanza semplice. Tutti i job provenienti da un processo di secondo grado convergeranno nella stessa e identica struttura. Il processo di terzo grado è gestito da un insieme di 5 giudici definito *corte di cassazione*. Il gruppo di 5 giudici dopo aver terminato il proprio lavoro comunicherà i risultati in una sentenza che verrà tenuta in una struttura fisica (la corte di cassazione fisica che è unica in Italia). Ne consegue che nel modello ci sarà una sequenza di due single server. Si osservi l'immagine 2.7.

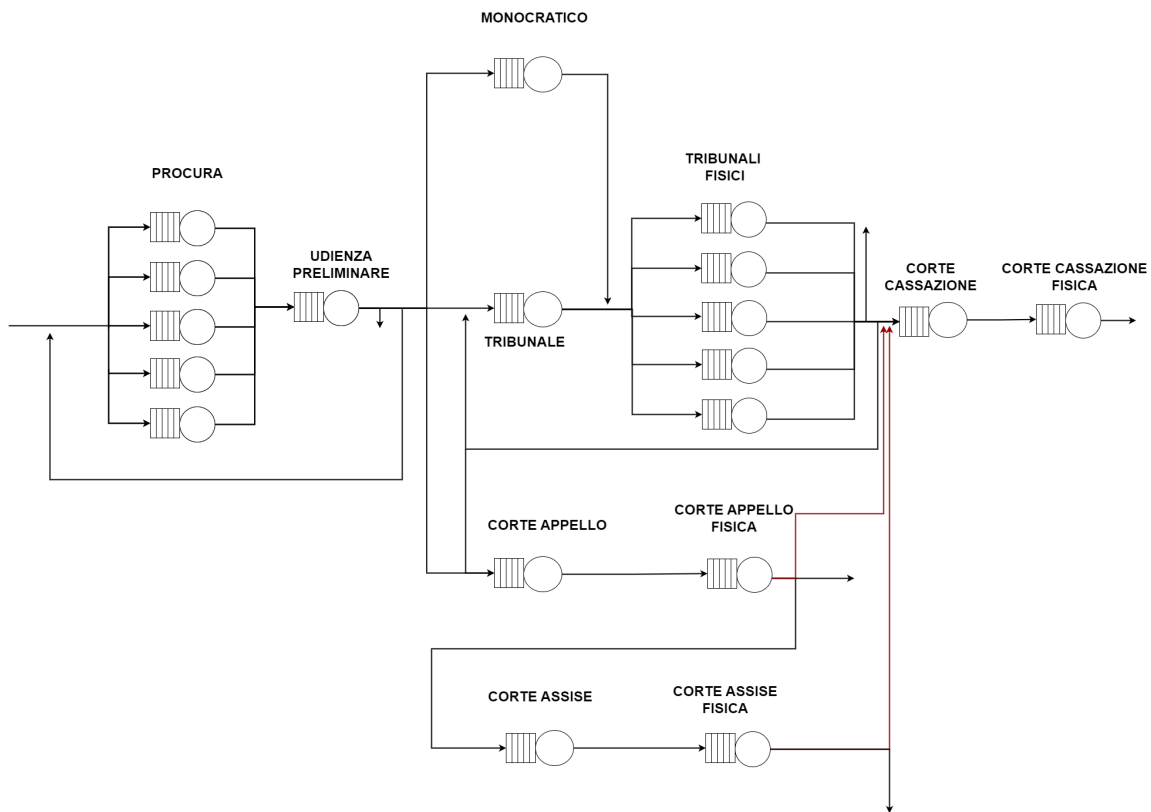


Figure 2.7

2.2.6 Definizione dello stato

Si definisce lo stato del sistema come l'insieme degli stati di ogni servente. In particolare nello stato di ogni servente verrà memorizzato il numero di job in attesa, il numero di job in servizio, lo stato del servente (libero o occupato), il numero di arrivi e il numero di partenze. Per quanto riguarda gli eventi se ne distinguono due tipologie: arrivo di una nuova denuncia, che corrisponde all'ingresso di un job nel sistema, e il completamento di un servente. Infine i job possono essere di tre tipologie: poco gravi, gravità media e gravità alta.

–3–

Modello delle specifiche

3.1 Distribuzioni

Per l'udienza preliminare, le indagini preliminari, il primo grado monocratico, il tribunale, la corte di cassazione e corte di appello, si è preferito utilizzare come distribuzione quella di Pareto. La decisione è dovuta principalmente al fatto che la Pareto è heavy tail. Questa caratteristica rispecchia perfettamente una proprietà del carico in ingresso al sistema: potrebbe verificarsi uno scenario in cui poche denunce entranti nel sistema, in un intervallo di tempo relativamente breve, possono tenere impegnati i centri per molto tempo a causa della loro complessità. Invece per quanto riguarda l'infrastruttura fisica del tribunale e delle cassazioni si è preferito utilizzare una distribuzione deterministica. Una sentenza in un'infrastruttura fisica dura mediamente 5-6 ore che vengono però suddivise tra mattina e pomeriggio. Ne consegue che una qualsiasi infrastruttura fisica verrà mantenuta occupata un giorno intero per comunicare il risultato di un'unica sentenza. E' anche vero che questa assunzione è una forte semplificazione. Ogni tribunale fisico non ha una sola aula. Tuttavia, non trovando dati relativi alla dimensione dei tribunali nel Lazio, si assume che ogni tribunale abbia solo un'aula. Da queste osservazioni è possibile concludere che il tempo di servizio di tali centri è deterministico e dura esattamente un giorno. La tabella 3.1 sintetizza le distribuzioni scelte. I parametri delle distribuzioni verranno scelti nei paragrafi seguenti.

Servente	Distribuzione
Indagini preliminari	Pareto
Udienza preliminare	Pareto
Dibattimento monocratico	Pareto
Dibattimento collegiale	Pareto
Corte Appello	Pareto
Corte Assise	Pareto
Corte Cassazione	Pareto
Tribunale fisico	Deterministica
Corte Appello fisica	Deterministica
Corte Assise fisica	Deterministica
Corte Cassazione fisica	Deterministica

3.2 Probabilità di routing

Per definire le probabilità di routing si utilizzano dati statistici. Si calcolano i parametri di interesse usando i dati statistici relativi all'Italia nella sua totalità. Si assumerà successivamente che i parametri calcolati per l'Italia siano esattamente gli stessi per la regione Lazio.

3.2.1 Procura

La procura, come detto in precedenza, lavora in modo locale. Ogni procura prenderà in carico le denunce provenienti dalla città in cui è collocata. Per definire le percentuali di routing dei job in ingresso al sistema si è deciso di analizzare la popolazione delle città. In particolare, la probabilità di un job di entrare in una procura specifica sarà data dal rapporto della popolazione della città fratto quella della popolazione del Lazio. Questa formula deriva dal fatto che in una città più popolosa è più probabile avere un maggior numero di denunce. Formalizzando quindi:

$$p(P_{citta}) = N_{citta} / N_{Lazio}$$

Dove $p(P_{citta})$ è la probabilità di entrare nella procura di una determinata città, N_{Lazio} la popolazione del Lazio e N_{citta} la popolazione di una città specifica. I dati utilizzati sono racchiusi nella seguente tabella:

Città/Regione	Popolazione	Percentuale
Lazio	5.714.882	-
Roma	4.216.874	75%
Latina	567.439	10%
Frosinone	470.689	8%
Viterbo	308.737	5%
Rieti	151.143	2%

3.2.2 Udienza preliminare

In uscita all'udienza preliminare il flusso può subire diverse strade. Un 58 % torna in feedback. Si ottiene questo dato perchè la somma dei tempi di risposta dell'udienza e delle indagini preliminari è 358 giorni. Per legge le indagini si svolgono a cicli di 180 giorni. Al termine di ogni ciclo si passa per l'udienza preliminare. Terminata l'udienza è possibile richiedere un'ulteriore ciclo di indagini. Complessivamente quindi un ciclo dura 180 +30 giorni, ovvero 210. In proporzione 210 giorni è il 58 % di 358. Questo vuol dire che mediamente la metà dei job subisce una retrazione in feedback e ripete il ciclo di indagini, il restante 42 % passa al primo grado. Formalizzando:

$$p(F_{UP}) = (C_I + \overline{D_{UP}}) / \overline{D_{UI}}$$

$$p(\neg F_{UP}) = 1 - p(F_{UP})$$

Dove $p(F_{UP})$ è la probabilità di feedback, C_I è la durata media di un ciclo di indagini, $\overline{D_{UP}}$ è la durata media di permanenza all'interno dell'udienza preliminare, $\overline{D_{UI}}$ la durata media della somma dell'udienza preliminare più indagini. Invece $p(\neg F_{UP})$ è la probabilità di non finire in feedback. Da cui si ottiene:

$$p(F_{UP}) = (180 + 30) / 358 \approx 0.58$$

$$p(\neg F_{UP}) = 1 - 0.58 \approx 0.42$$

Di questo 42% che supera le indagini preliminari il 65% esce dal sistema a causa dell'archiviazione. Quindi, normalizzando, il 31% dei job che abbandonano l'udienza escono dal sistema. Il restante flusso è ripartito nel primo grado. In particolare il 16% sono job poco gravi, 0.85% sono job di gravità media e il restante 0.15% sono job gravi. Formalizzando:

$$p(A) = p(\neg F_{UP}) * p(A_t) = 0.42 * 0.60 \approx 0.31$$

$$p(P_b) = (1 - p(\neg F_{UP}) - p(A)) * p(P_{bt}) \approx 0.16$$

$$p(P_m) = (1 - p(\neg F_{UP}) - p(A)) * p(P_{mt}) \approx 0.0085$$

$$p(P_a) = (1 - p(\neg F_{UP}) - p(A)) * p(P_{at}) \approx 0.0015$$

Dove $p(A)$ è la probabilità di archiviazione sui job uscenti dall'udienza preliminare, $p(A_t)$ è la probabilità di archiviazione dei job che escono dall'udienza ma che non tornano in feedback. $p(P_b)$ è la probabilità di entrare in un processo di primo grado di gravità bassa calcolata su tutti i job uscenti dall'udienza preliminare, $p(P_{bt})$ è la probabilità di entrare in un processo di primo grado di gravità bassa calcolata sui job che non tornano in feedback. $p(P_m)$ è la probabilità di entrare in un processo di primo grado di gravità media calcolata su tutti i job uscenti dall'udienza preliminare, $p(P_{mt})$ è la probabilità di entrare in un processo di primo grado di gravità media calcolata sui soli job che non tornano in feedback. $p(P_a)$ è la probabilità di entrare in un processo di primo grado di gravità alta calcolata su tutti i job uscenti dall'udienza preliminare, $p(P_{at})$ è la probabilità di entrare in un processo di primo grado di gravità alta calcolata sui soli casi che non tornano in feedback.

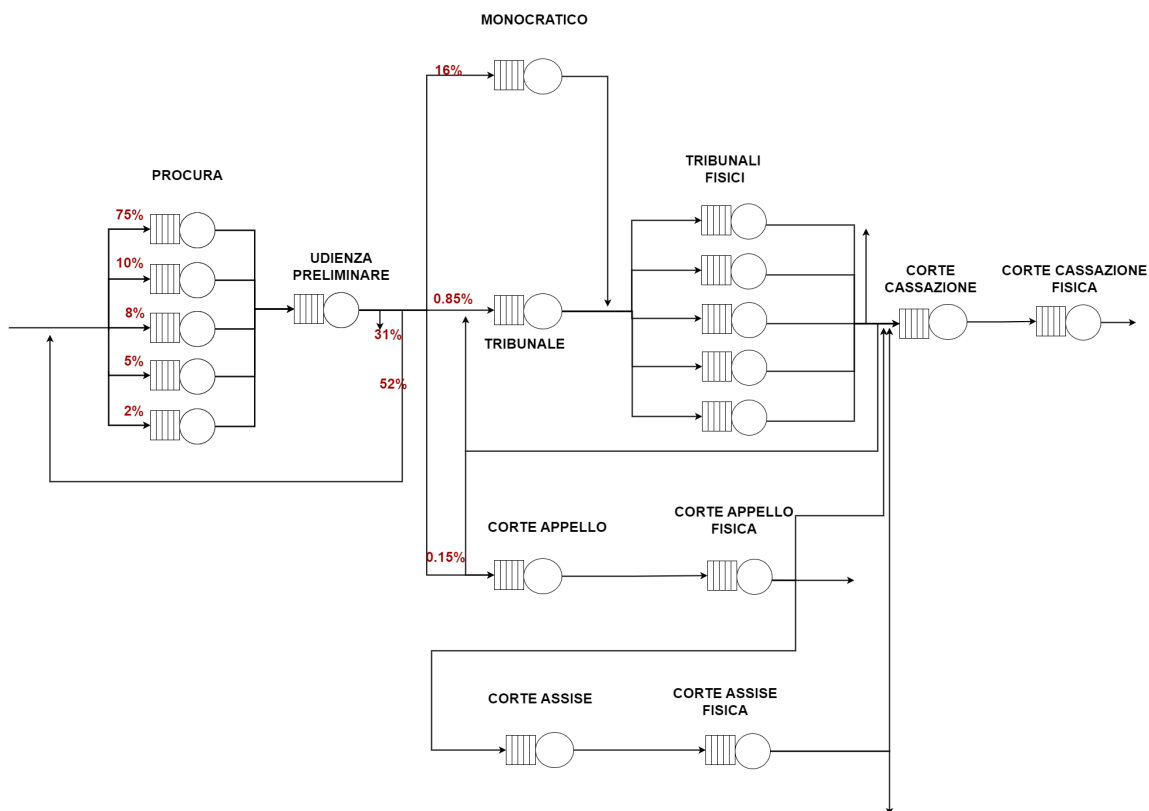


Figure 3.1: L'immagine mostra come è stato modellato il sottosistema formato da indagini preliminari e udienza preliminare.

3.2.3 Ingresso tribunali fisici

Il flusso in uscita dal dibattimento monocratico si combina con quello proveniente dal tribunale. A questo punto il flusso unificato deve essere suddiviso nei 5 tribunali fisici. La probabilità di routing nei 5 diversi tribunali fisici è calcolata come visto in precedenza nella fase di procura essendoci un tribunale fisico per ogni città. I dati relativi alle popolazioni sono sempre descritti dalla tabella 3.2.1.

3.2.4 Uscita Tribunale fisico

All'uscita del tribunale fisico lo 0.01% del flusso va in terzo grado e quindi alla corte di cassazione. Questo parametro è in realtà frutto di un'approssimazione. Gli ingressi totali nel terzo grado sono circa 100 job annui in tutta Italia. Le uscite complessive dai tribunali fisici sono circa 1 milione all'anno. Il valore corretto di probabilità di routing dovrebbe essere 10^{-6} . Si tratta di un valore molto molto basso che rischia di non far giungere al centro della corte di cassazione un numero sufficiente di job che rendano il centro analizzabile dal punto di vista statistico.

Continuando l'analisi, un 7% del flusso entra in secondo grado con gravità bassa: dai dati statistici si osserva che dei 1,3 milioni di job entranti nel tribunale 100000 sono job di secondo grado di bassa gravità. In percentuale quindi il 7% dei casi in uscita dai tribunali fisici va nella corte di appello. Sempre dai dati statistici si vede come circa 1% dei casi entranti nel tribunale siano a gravità media. Formalizzando:

$$p(S_m) = N_{S_m}/N_T$$

$$p(S_m) = 100000/1333484 \approx 0.07$$

E anche

$$p(S_a) = N_{S_a}/N_T \approx 0.01$$

Dove $p(S_m)$ è la probabilità di entrare nel secondo grado con gravità bassa, N_{S_m} numero di job uscenti dal tribunale fisico e entrati nel secondo grado con gravità bassa, N_T numero totali di job uscenti dai tribunali fisici. $p(S_a)$ è la probabilità di entrare nel secondo grado con gravità media, N_{S_a} numero di job uscenti dal tribunale fisico e entrati nel secondo grado con gravità media. La probabilità di uscire dal sistema è invece $1 -$ la probabilità di prendere le altre strade.

$$p(U) = 1 - p(F_{DM}) - p(S_m) - p(S_a) = 1 - 0.07 - 0.01 - 0.0001 = 0.9299$$

3.2.5 Uscita corte di appello fisica

In uscita dalla corte di appello fisica un 14% del flusso va in corte di assise. Dai dati statistici si osserva che il numero di ingressi in assise è circa 600 ogni anno, quelli in uscita dalla corte di appello sono circa 4354. In proporzione quindi il 14% è entrante in corte di assise. Il restante flusso si ripartisce tra terzo grado e uscente. Per il terzo grado si usa ancora un'approssimazione dello 0.01%. Tutto il resto è uscente. Formalizzando:

$$p(CA) = N_{CA}/N_T = 600/4354 \approx 0.14$$

$$p(CC) \approx 0.0001$$

$$p(U) = 1 - p(CA) - p(CC) \approx 0.8599$$

Dove $p(CA)$ è la probabilità di entrare in corte di assise, $p(CC)$ è la probabilità di entrare in corte di cassazione e $p(U)$ è la probabilità di uscire dal sistema.

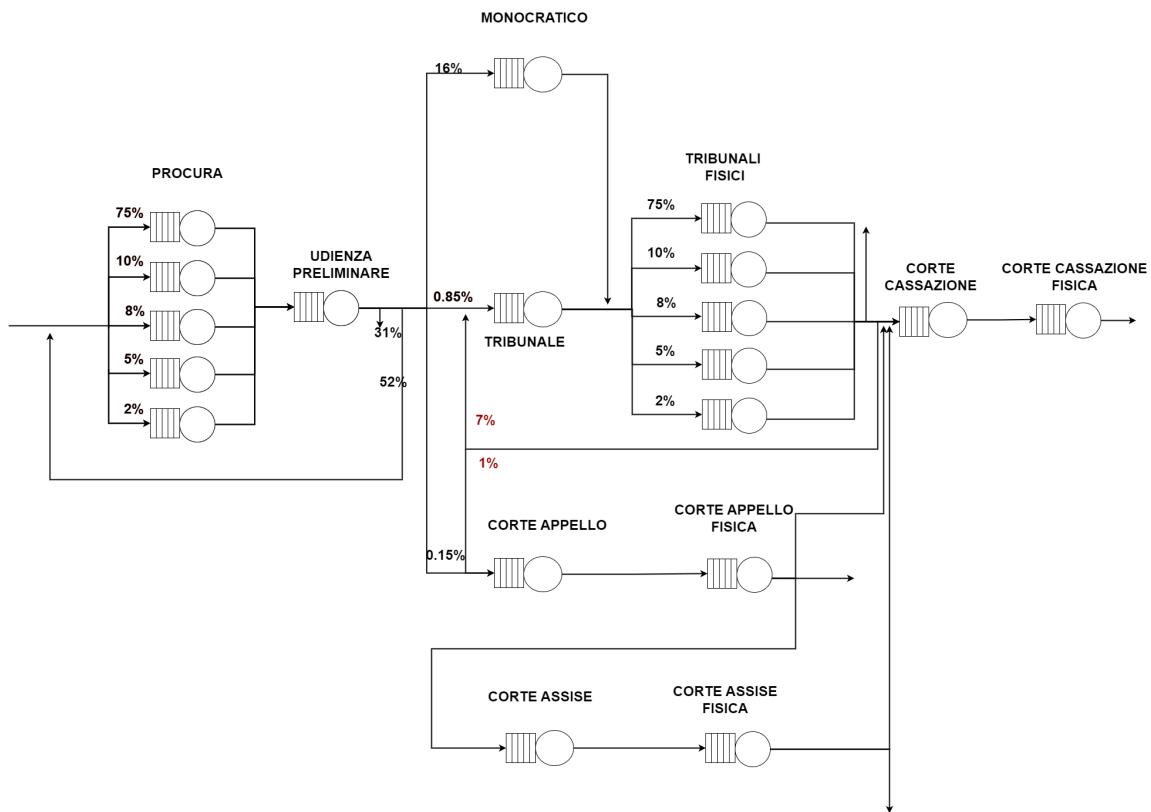


Figure 3.2

3.2.6 Uscita corte di appello assise fisica

Il flusso in uscita può subire una sola biforcazione. Un job o esce dal sistema oppure entra in corte di cassazione. Poichè ,ancora una volta mancano dati sulla corte di cassazione, si ipotizza che lo 0.01% vada in corte di cassazione e tutto il resto sia uscente.

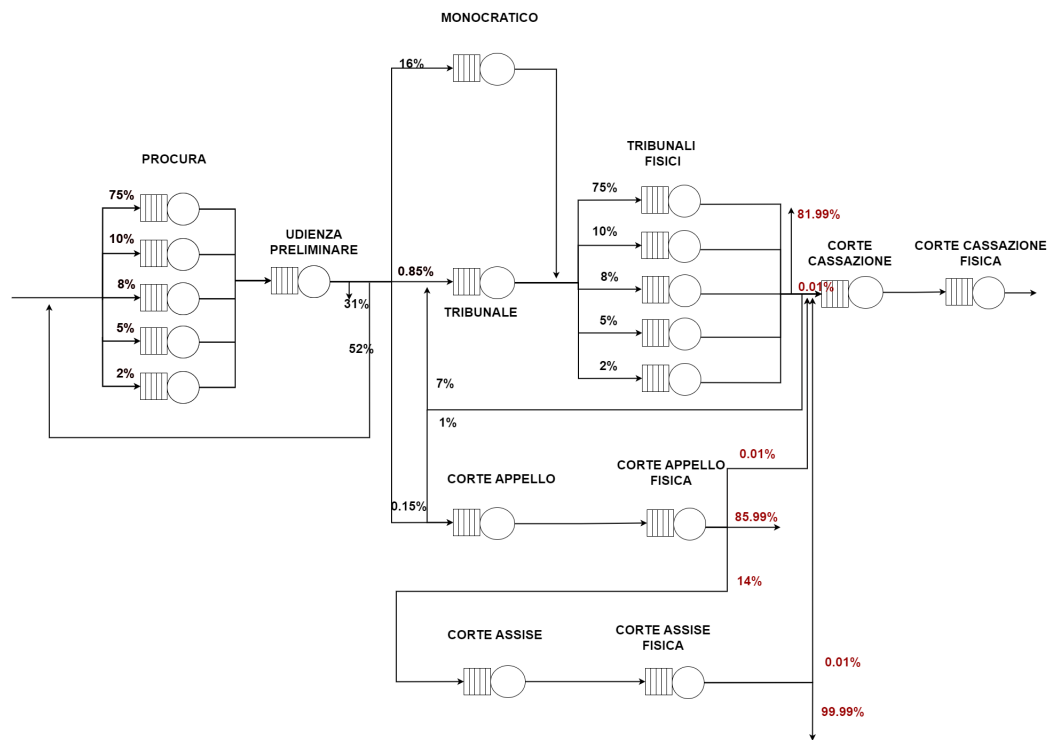


Figure 3.3

3.3 Tempi medi di servizio

Per individuare i tempi medi di servizio sono stati condotti alcuni esperimenti poichè non è stato possibile trovare dati che potessero essere utili a tal fine. In particolare gli esperimenti sono stati condotti dopo la fase di validazione. L'idea era quella di partire con un sistema vuoto e provare diversi valori dei tempi medi di esecuzione in modo da far convergere le statistiche dei serventi ai valori medi reali. La tabella 8.4 mostra tutti i valori trovati al termine degli esperimenti. Per tutte le infrastrutture fisiche che hanno una distribuzione deterministica si sceglie come tempo medio di servizio 1 giorno.

Centro	E[S]
Procura Roma	0.8 giorni
Procura Frosinone	2.15 giorni
Procura Latina	3.5 giorni
Procura Viterbo	2.55 giorni
Procura Rieti	3.05 giorni
Udienza	0.375 giorni
Monocratico	1.79 giorni
Tribunale	23 giorni
Appello	80 giorni
Assise	250 giorni

-4-

Implementazione

4.1 Architettura

Il modello è stato implementato utilizzando il linguaggio di programmazione Java. Il linguaggio è stato preferito al C esclusivamente perchè basato su un paradigma di programmazione ad oggetti. Questa caratteristica semplifica notevolmente l'implementazione e permette la creazione di un software quanto più generico possibile che si adatta ad ogni modello a rete di code. Il software si baserà infatti su un file di configurazione in cui verrà descritta la topologia della rete da simulare. In particolare nel file di configurazione devono essere specificati quali e quanti sono i centri, come sono collegati tra di loro, il tasso di servizio medio per ogni centro e la distribuzione del tempo di servizio. L'architettura implementativa del modello si compone di più strati. Il layer più esterno è il *Network* che converte il file di configurazione in un oggetto della classe *Network* contenente i nodi della rete e i collegamenti tra di essi. Il secondo layer è quello di *Handle*. Ogni nodo della rete mantiene il riferimento ad un "handle" che si occupa esclusivamente della gestione e la generazione degli eventi. Infine l'ultimo layer è quello che nel concreto calcola i tempi di servizio e aggiorna le variabili e strutture dati che permettono di calcolare i tempi di attesa medi e i numero medi di job. Ogni handler avrà infatti il riferimento a un *server* che ha la responsabilità di implementare la logica di un single server o multiserver. Proprio in questo layer vengono utilizzati i generatori pseudorandom.

4.2 Generatori pseudocasuali

Per la generazione dei tempi di servizio è stato utilizzato un generatore pseudocasuale che è in realtà la conversione in java del file C *Rngs.c*. Si tratta di un generatore multistream i cui parametri utilizzati per la generazione dei valori sono : *MODULUS* = 2147483647, *MULTIPLIER* = 48271, *A256* = 22925, seme *DEFAULT* = 123456789, *STREAMS* = 256. Sono previste diverse funzionalità sia per piantare un seme, cambiare stream sia per generare nuovi valori secondo una distribuzione normale, pareto esponenziale ecc.

–5–

Verifica e Validazione

5.1 Verifica

L'obiettivo di questa fase è verificare che il software rispetti le specifiche di progetto e produca risultati in accordo con la teoria. Poichè il programma che è stato implementato permette di simulare una qualsiasi topologia di rete, non ha molto senso verificare che il software si comporti in maniera corretta con i soli parametri individuati nella fase di modellazione delle specifiche. L'approccio seguito si basa su una tecnica di test definita *partizionamento del dominio*. In pratica si cerca di identificare tutti i possibili sottosistemi che è possibile simulare e si va a dimostrare che in tutte le configurazioni il modello implementativo produce risultati concordi con la teoria. Di conseguenza partendo da una base di software verificata in questo modo, si può considerare e simulare una rete complicata a piacere, ma c'è comunque una ragionevole confidenza che la simulazione restituisca risultati corretti. Si distinguono quindi due casistiche:

- Rete formata da un solo centro
- Rete formata da più centri

Nel caso in cui la rete sia formata da un solo centro si distinguono i due casi in cui il centro è un multiserver o un single server. Nel caso di un **single server** si vuole verificare che la simulazione generi i valori di tempi di attesa, tempi di risposta, popolazione media in coda, popolazione media nel sistema concordi con le equazioni teoriche. In più si vuole verificare che la legge di Little venga rispettata ma anche che il numero medio di job nel sistema sia la somma del numero medio di job nel centro più quelli in coda. Tutte queste condizioni devono essere verificate anche al variare della distribuzione del tempo di servizio. Per il **multiserver** si vogliono verificare le stesse cose, in aggiunta però, si vuole anche verificare se la concordanza tra valori teorici e simulati avvenga anche con un numero crescente di serventi. Si noti come per tutti i test eseguiti i valori teorici sono calcolati attraverso la KP, mentre i valori simulati sono ottenuti attraverso una simulazione a lungo orizzonte temporale che sfrutta il metodo delle batch means.

In particolare sono stati scelti $k=64$ e $b=1028$. Gli intervalli di confidenza sono calcolati utilizzando un valore di confidenza al 95% e sfruttando la distribuzione di Students.

5.1.1 Single Server

Si consideri un single server con le seguenti caratteristiche: **numero job iniziale** pari a 0, **tempo di servizio medio** pari a 4 j/s, **tempo di interarrivo** pari a 5 j/s.

Di seguito vengono riportate le tabelle dei confronti dei risultati teorici e dei risultati simulati. La prima tabella mostra i risultati ottenuti con una **distribuzione esponenziale**.

	Analitico	Simulato
E[S]	4	3.9986 ± 0.0218
E[TQ]	16	15.8358 ± 0.3281
E[T]	20	19.8344 ± 0.4111
E[NQ]	3,2	3.1650 ± 0.0391
U	0.8	0.7986 ± 0.0218
E[N]	4	3.9636 ± 0.1237
$1/\lambda$	5	$5.0070 \pm 0.833 * 10^{-2}$

Utilizzando come distribuzione la **deterministica** si ottiene:

	Analitico	Simulato
E[S]	4	3.9999 ± 0.0254
E[TQ]	8	8.0221 ± 0.3181
E[T]	12	12.0221 ± 0.4172
E[NQ]	1.6	1.6056 ± 0.0117
U	0.8	0.8000 ± 0.0201
E[N]	2.4	2.4057 ± 0.0127
$1/\lambda$	5	$4.9997 \pm 0.723 * 10^{-2}$

Infine utilizzando come distribuzione la **pareto** con $\alpha=2.100$ e $k=2.0952$ si ottiene:

	Analitico	Simulato
E[S]	4	4.0494 ± 0.1328
E[TQ]	223.7545	225.4287 ± 2.9032
E[T]	227.7545	229.4781 ± 2.7431
E[NQ]	43.5023	42.6724 ± 0.8891
U	0.8	0.8047 ± 0.02022
E[N]	44.3023	43.4771 ± 0.08791
$1/\lambda$	5	$5.0294 \pm 0.813 * 10^{-2}$

In tutti i casi si può facilmente osservare come i valori teorici rientrino nell'intervallo di confidenza individuato tramite la simulazione, ma anche come la legge di Little e i vari controlli di consistenza siano rispettati.

5.1.2 Multi Server

Si consideri ora un multi server con le seguenti caratteristiche: **numero job iniziale** pari a 0 , **tempo di interarrivo** pari a 5 j/s, **distribuzione** esponenziale.

Di seguito vengono riportate le tabelle dei confronti dei risultati. La prima tabella mostra i risultati ottenuti con un numero di serventi pari a 1 e un tempo di servizio pari a 4j/s. L'idea è verificare che ci sia la convergenza a un single server.

	Analitico	Simulato
E[S]	4	3.9978 \pm 0.0219
E[TQ]	16	15.7958 \pm 0.3271
E[T]	20	19.9346 \pm 0.4151
E[NQ]	3,2	3.1751 \pm 0.0394
U	0.8	0.7996 \pm 0.0318
E[N]	4	3.9655 \pm 0.2237
1/ λ	5	5.0055 \pm 0.837 * 10 ⁻²

Nel secondo caso si fa crescere in numero di serventi a 3. E il tempo di servizio per ogni servente si sceglie pari a 10.

	Analitico	Simulato
E[Si]	10	10.0235 \pm 0.1237
E[TQ]	4.5543	4.6468 \pm 0.2349
E[T]	14.5543	14.6704 \pm 0.5432
E[NQ]	0.9216	0.9313 \pm 0.1289
U	0.6666	0.6691 \pm 0.0994
E[N]	2.9196	2.9388 \pm 0.2354
1/ λ	5	4.9936 \pm 0.0099

Nel terzo caso si fa crescere il numero di serventi a 5. E il tempo di servizio per ogni servente si sceglie pari a 15.

	Analitico	Simulato
E[Si]	15	14.9895 \pm 0.2376
E[TQ]	1.7545	1.7748 \pm 0.1342
E[T]	16.7545	16.7644 \pm 0.0990
E[NQ]	0.3555	0.3546 \pm 0.5467
U	0.6	0.5985 \pm 0.0099
E[N]	3.3555	3.3475 \pm 0.1288
1/ λ	5	5.0086 \pm 0.0190

Questo approccio serve per verificare il sistema per induzione. Se il software simula correttamente un multiserver con 1,3,5.. serventi è altrettanto probabile che simuli correttamente multiserver con centinaia di serventi.

5.1.3 Sottoreti

Si procede ad analizzare le possibili sottoreti che si possono creare combinando tra loro più server. Si analizzano in particolare 3 casi :

- 2 server in sequenza
- 3 server senza feedback
- 3 server con feedback

Ancora un volta si ragiona per induzione. Se si riesce a dimostrare che questi casi base funzionano correttamente , è verosimile che anche una rete molto più complessa formata da centinaia di centri funzioni in maniera corretta , dato che ogni rete può essere vista come la composizione di queste sottoreti di base. Si noti come i valori teorici siano stati calcolati a partire dalle equazioni di flusso combinate con la KP. La distribuzione di tutti i centri considerati nel test è esponenziale. Il primo caso che si analizza è il seguente:

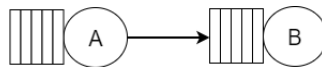


Figure 5.1

Simulando si ottengono i seguenti risultati:

Server A	Analitico	Simulato
E[S]	4	3.9986 ± 0.0218
E[TQ]	16	15.8358 ± 0.3281
E[T]	20	19.8344 ± 0.4111
E[NQ]	3,2	3.1650 ± 0.0391
U	0.8	0.7986 ± 0.0218
E[N]	4	3.9636 ± 0.1237
1/ λ	5	5.0070 $\pm 0.833 * 10^{-2}$

Server B	Analitico	Simulato
E[S]	4	3.9886 ± 0.0218
E[TQ]	16	15.9358 ± 0.3281
E[T]	20	19.8844 ± 0.4111
E[NQ]	3,2	3.1780 ± 0.0391
U	0.8	0.7989 ± 0.0218
E[N]	4	3.9886 ± 0.1237
1/ λ	5	5.0070 $\pm 0.821 * 10^{-2}$

Nel secondo caso si considera la seguente rete.

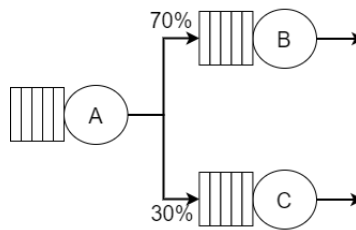


Figure 5.2

Simulando si ottengono i seguenti risultati:

Server A	Analitico	Simulato
E[S]	4	3.9986 ±0.0218
E[TQ]	16	15.8358 ±0.3281
E[T]	20	19.8344 ±0.4111
E[NQ]	3,2	3.1650 ±0.0391
U	0.8	0.7986 ±0.0218
E[N]	4	3.9636 ±0.1237
1/λ	5	5.0070 ±0.833 * 10 ⁻²

Server B	Analitico	Simulato
E[S]	4	3.9756 ±0.2376
E[TQ]	5.8755	4.9551 ±0.9976
E[T]	9.8755	8.9307 ±0.9809
E[NQ]	0.7085	0.6928 ±0.0999
U	0.5600	0.5556 ±0.0732
E[N]	1.3185	1.2485 ±0.2098
1/λ	7.1428	7.1548 ±0.2254

Server C	Analitico	Simulato
E[S]	4	3.9989 ±0.0990
E[TQ]	1.3550	1.2855 ±0.1234
E[T]	5.3550	5.2844 ±0.2300
E[NQ]	0.0825	0.0771 ±0.0099
U	0.2400	0.2397 ±0.0679
E[N]	0.3225	0.3168 ±0.1209
1/λ	16.6666	16.6814 ±0.5443

Si verifica infine l'ultimo caso.

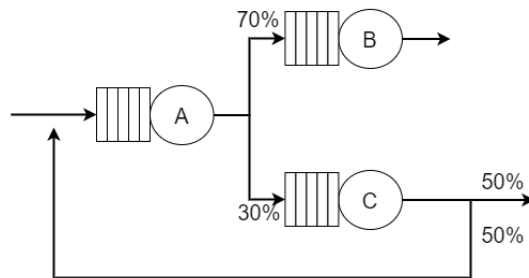


Figure 5.3

Simulando si ottengono i seguenti risultati:

Server A	Analitico	Simulato
E[S]	4	4.0019 ±0.0126
E[TQ]	63.8509	63.8529 ±0.0345
E[T]	67.8509	67.8549 ±0.1654
E[NQ]	15.0455	15.0452 ±0.2109
U	0.9411	0.94027 ±0.0034
E[N]	15.9866	15.9854 ±0.0456
1/λ	4.2500	4.2561 ±0.0232

Server B	Analitico	Simulato
E[S]	4	3.9812 ±0.1543
E[TQ]	7.3977	7.3954 ±0.2366
E[T]	11.3977	11.3767 ±0.0889
E[NQ]	1.2133	1.2154 ±0.0778
U	0.6666	0.6541 ±0.0998
E[N]	1.8799	1.8696 ±0.1278
1/λ	6	6.0859 ±0.1002

Server C	Analitico	Simulato
E[S]	4	3.9967 ±0.0349
E[TQ]	1.5911	1.5965 ±0.2355
E[T]	5.5911	5.5932 ±0.0235
E[NQ]	0.1100	0.1129 ±0.0090
U	0.2826	0.2823 ±0.0089
E[N]	0.3926	0.3952 ±0.0127
1/λ	14.1500	14.1557 ±0.1776

5.2 Validazione

Per quanto sia stato possibile trovare dei dati per la realizzazione delle specifiche del progetto, è stato più complicato trovare dati utili per la validazione. Questa fase si limiterà quindi a verificare che cambiando anche di poco alcuni parametri del sistema, il modello mostri un comportamento verosimile che rispecchi la realtà. I test che verranno eseguiti sono sostanzialmente due:

- Aumentare leggermente il tasso di ingresso dei job nel sistema e verificare se c'è un aumento dei tempi di risposta nei centri;
- Diminuire leggermente il tasso di ingresso dei job nel sistema e verificare che ci sia una diminuzione dei tempi di risposta.

Per quanto riguarda il primo test si aumenta il tasso di ingresso del sistema da 1 job/giorno a 1,1 job/giorno. La tabella mostra i risultati ottenuti con questa modifica in alcuni dei centri del sistema. Si può osservare come ci sia stata un aumento dei tempi di risposta di ogni centro. Tale aumento non è spropositato ma è comunque verosimile.

Centro	E[T] versione base	E[T] con tasso modificato
<i>Udienza</i>	33 giorni	112 giorni
<i>Monocratico</i>	312 giorni	358 giorni
<i>Tribunale</i>	868 giorni	1037 giorni
<i>Appello</i>	392 giorni	417 giorni
<i>Assise</i>	657 giorni	658 giorni

Infine nell'ultimo test si diminuisce il tasso di ingresso del sistema da 1 job/giorno a 0.9 job/giorno. I risultati sono mostrati nella seguente tabella. Si nota anche in questo caso una diminuzione verosimile di tutti i tempi di risposta dei centri.

Centro	E[T] versione base	E[T] con tasso modificato
<i>Udienza</i>	33 giorni	13 giorni
<i>Monocratico</i>	312 giorni	138 giorni
<i>Tribunale</i>	868 giorni	602 giorni
<i>Appello</i>	392 giorni	375 giorni
<i>Assise</i>	657 giorni	646 giorni

–6–

Simulazione orizzonte temporale infinito

L'obiettivo del progetto è quello di far convergere i tempi di risposta medi dei vari centri ai tempi di risposta medi europei. Appare evidente che, poichè l'obiettivo richiede il raggiungimento di una convergenza, ha molto più senso realizzare una simulazione a orizzonte temporale infinito. Nello specifico è stata condotta un'unica simulazione a lungo orizzonte temporale utilizzando il metodo delle batch means. I parametri utilizzati sono $k=128$ e $b=8192$. I risultati riportati in seguito fanno riferimento al modello base.

6.1 Procura

Il modello base prevede che la media dei tempi di attesa in Italia di tutte le procure sia 323 giorni. Dall'immagine 6.1 si può notare come la procura di Roma che riceve in ingresso un flusso maggiore di job non riesce a smaltire agevolmente tutto il carico e di conseguenza registra tempi di risposta medi che superano il valore di riferimento di 323 giorni. Le altre procure invece dovendo rispondere ad un carico di lavoro nettamente inferiore riescono a restare sotto il valore di riferimento. Si osservi però come sia la procura di Roma che quella di Frosinone non convergono perfettamente al valore medio. Sembrerebbe quasi che i tempi di risposta possano poi esplodere considerando un orizzonte temporale ancora maggiore.

6.2 Udienza preliminare

Segue poi l'udienza preliminare. Per legge l'udienza preliminare non può superare i 35 giorni di tempo di risposta medio. Nell'immagine 6.2 si osserva tale convergenza.

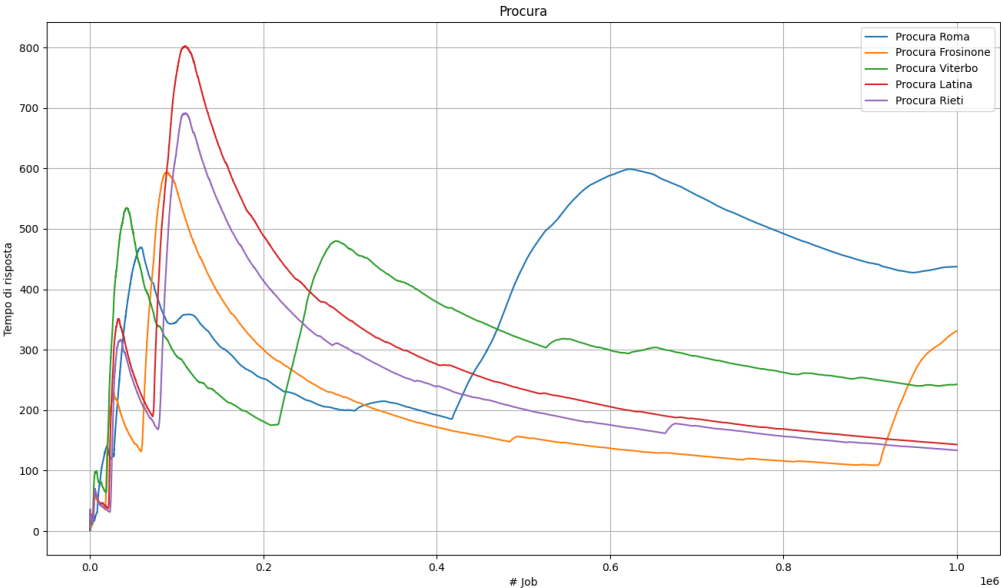


Figure 6.1

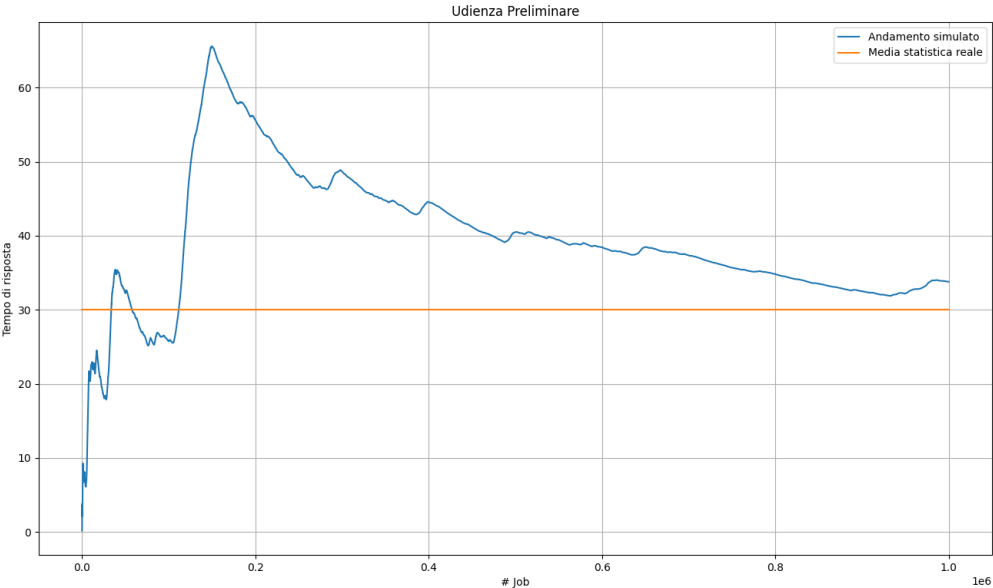


Figure 6.2

6.3 Primo grado monocratico

Il primo grado monocratico registra un tempo di risposta medio di 310 giorni. La simulazione genera invece un tempo di risposta medio che si mantiene leggermente al di sotto. Questo può essere notato nell'immagine 6.3

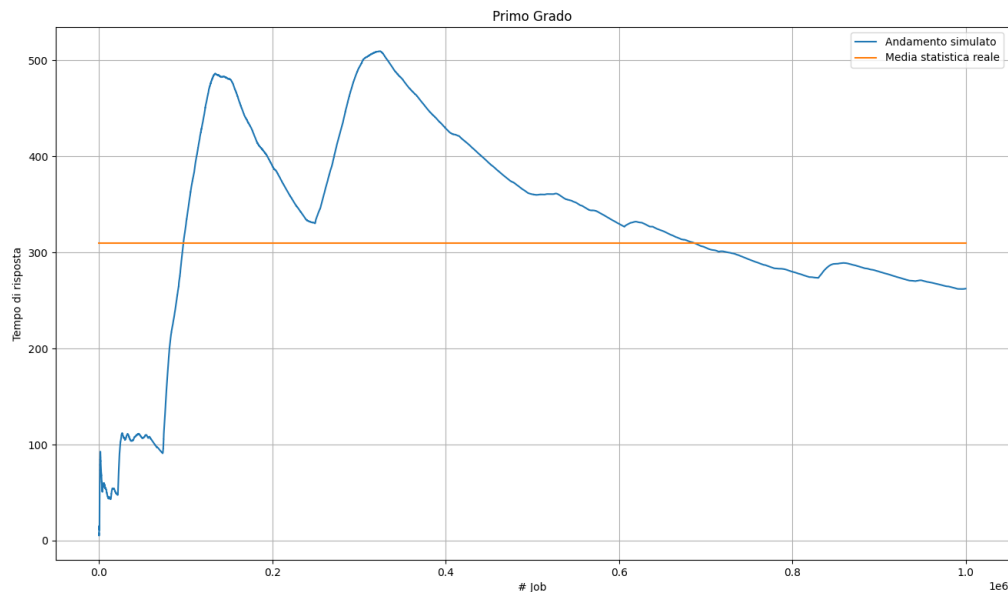


Figure 6.3

6.4 Tribunale

Il Tribunale ha un tempo medio di risposta di 876 giorni. L'immagine 6.4 mostra come il modello si avvicini a questo valore empirico anche se non è stata raggiunta una perfetta convergenza.

Nell'immagine 6.5 è mostrato anche l'andamento del tribunale fisico di Roma. Stranamente, rispetto a quanto si potesse pesare, si osserva come il tempo di risposta medio si mantiene attorno ad 1 giorno. Si ricordi che il tempo di esecuzione medio di tale servente è esattamente 1 giorno. Questo risultato evidenzia un aspetto importante: sembrerebbe che non sono le infrastrutture fisiche a causare un eccessivo ritardo nei tempi di risposta. Infatti i tempi di attesa medi per questo centro sono praticamente nulli. Da qui si evince che per poter aver un miglioramento di prestazioni bisognerà agire su altri centri e/o eventualmente su altre politiche di scheduling.

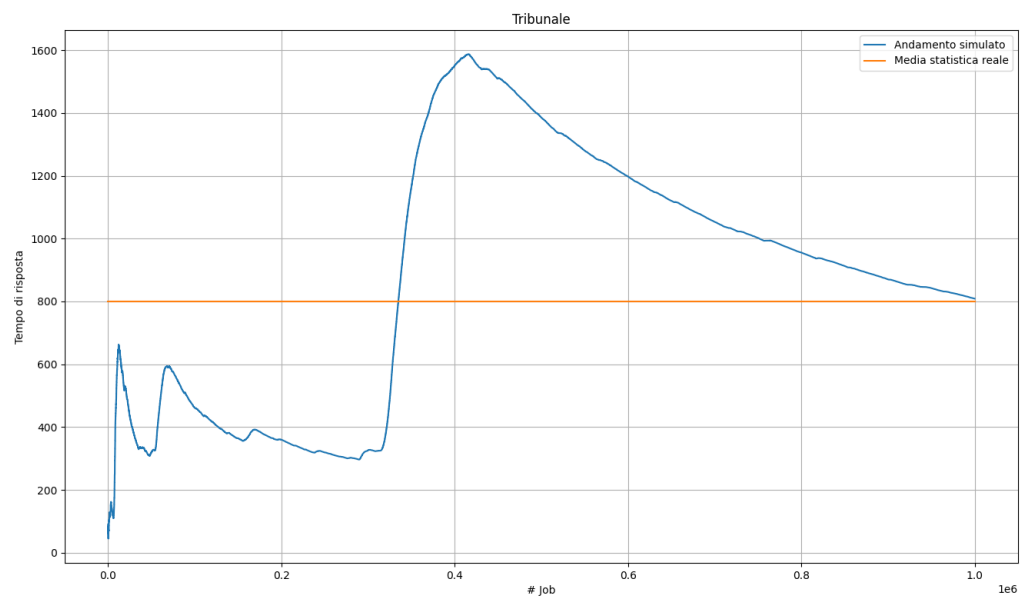


Figure 6.4

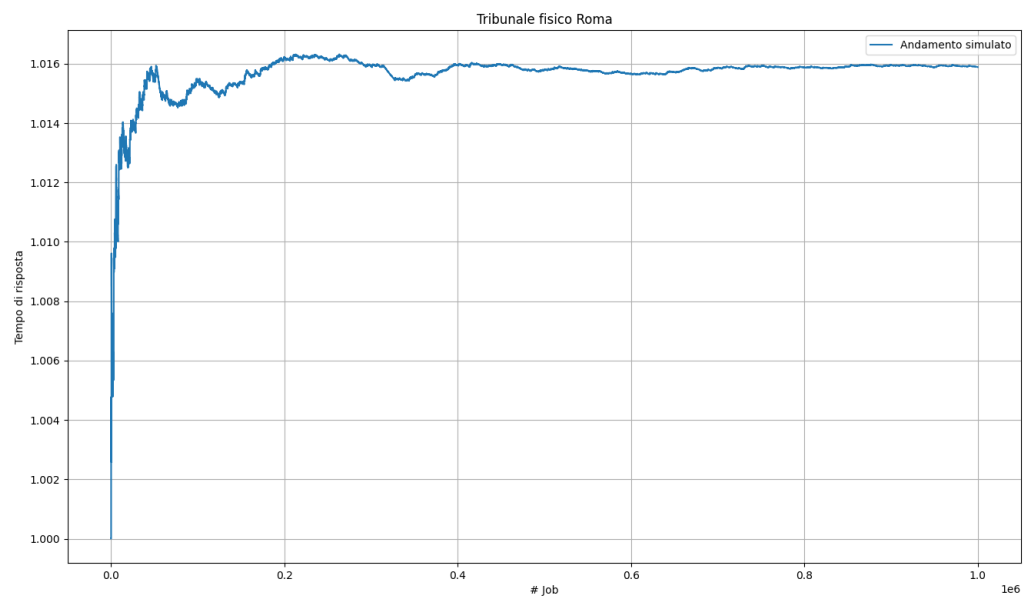


Figure 6.5

6.5 Corti

Nell'immagine 6.6 vengono presentate contemporaneamente sia la corte di appello che la corte di assise. La corte di appello riceve prevalentemente flusso di primo grado. Per cui i tempi di risposta medi si avvicinano a 310 giorni. La corte di assise è esclusivamente di secondo grado. Quindi i suoi tempi di risposta si avvicinano a 876 giorni.

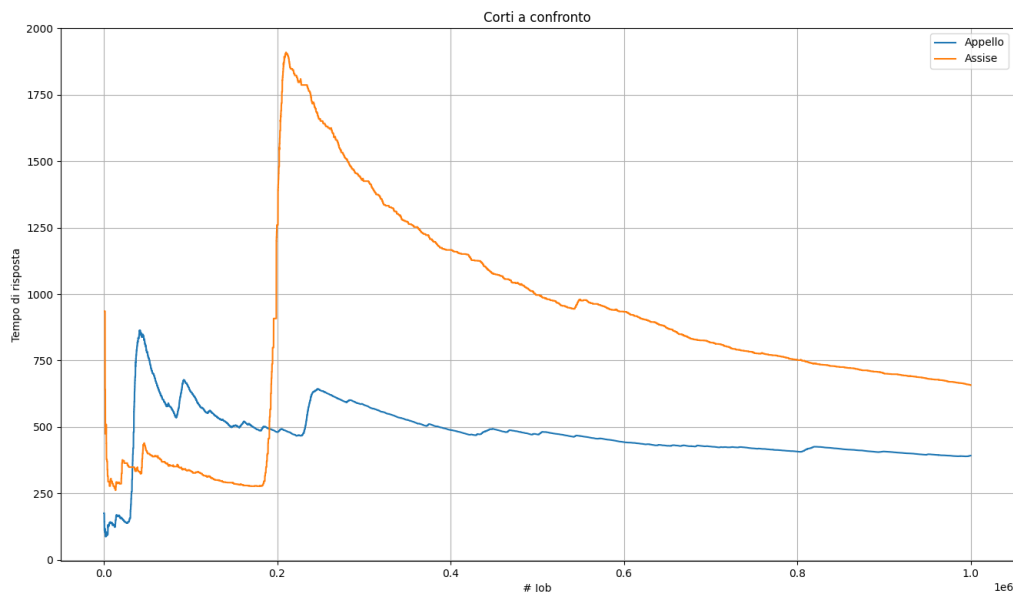


Figure 6.6

Per la corte di cassazione invece deve essere fatto un discorso più approfondito. Mediamente ogni anno giungono in Italia 136 nuovi job in corte di cassazioni. Nella sola regione del Lazio ci si aspetta mediamente di avere 1/20 di questo valore ovvero circa 7 job all'anno. Questo valore così basso fa in modo che la percentuale di routing verso il la cassazione sia talmente bassa da non riuscire ad effettuare un'analisi approfondita e corretta da un punto di vista statistico. Poichè inoltre la struttura non prevede in ingresso solo il flusso del Lazio ma di tutta Italia si preferisce escluderla da questa analisi.

–7–

Simulazione orizzonte finito

7.1 Introduzione

Come descritto anche in precedenza questa simulazione non è strettamente necessaria per raggiungere gli obiettivi prefissati. Sono stati comunque implementate delle simulazioni a breve orizzonte temporale. In questa simulazione si prenderà in considerazione un unico centro del sistema e si effettueranno più run di simulazione facendo variare il seme utilizzato per la generazione dei valori pseudocasuali. Si vuole quindi osservare come si comporta un centro specifico quando si presentano scenari di carico differenti. Si sceglie come centro quello della procura di Roma. Ma questa analisi può essere ripetuta allo stesso modo per tutti gli altri centri.

7.2 Tempi di risposta

L'immagine 7.1 mostra come variano i tempi di risposta in un orizzonte di simulazione molto breve al variare del seme. In questo esperimento sono stati considerati 3 semi e 10000 job entranti nel sistema. Si osservi come gli andamenti al variare del seme siano leggermente diversi tra loro, tuttavia, si comincia ad intravedere una convergenza ad un valore comune.

7.3 Popolazione nel centro

Nell'immagine 7.2 è mostrato invece come varia il numero di job all'interno del centro all'aumentare del tempo di simulazione. Si osserva come il grafico sia uguale a quello mostrato nell'immagine 7.1 ma con un cambiamento di scala. Questo non corrisponde ad un errore infatti la legge di Little afferma che è possibile ricavare il tempo di risposta dal numero medio di job nel sistema tramite una relazione lineare. Quindi è del tutto normale aspettarsi gli stessi grafici ma con un leggero cambiamento di scala.

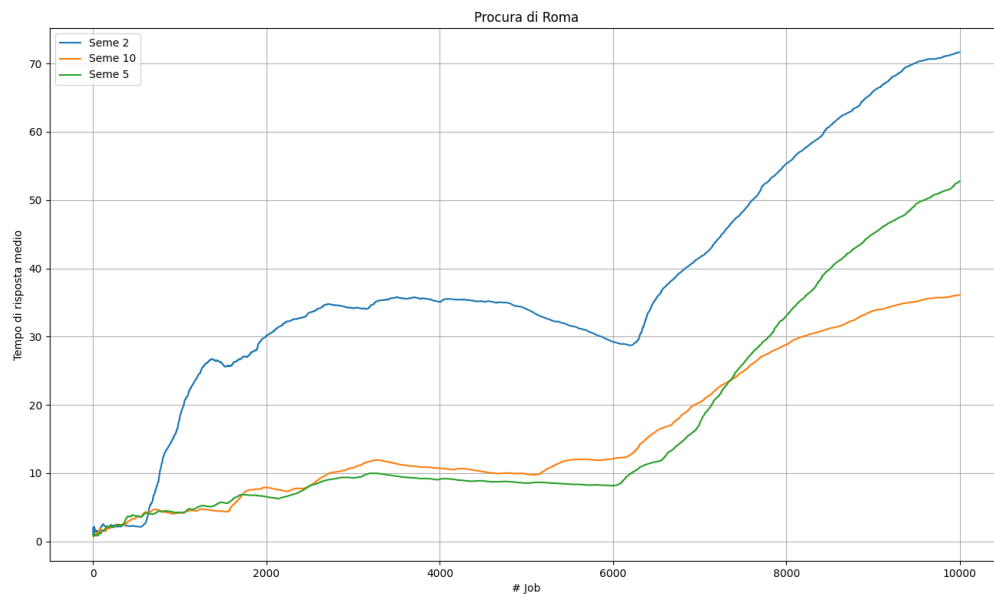


Figure 7.1

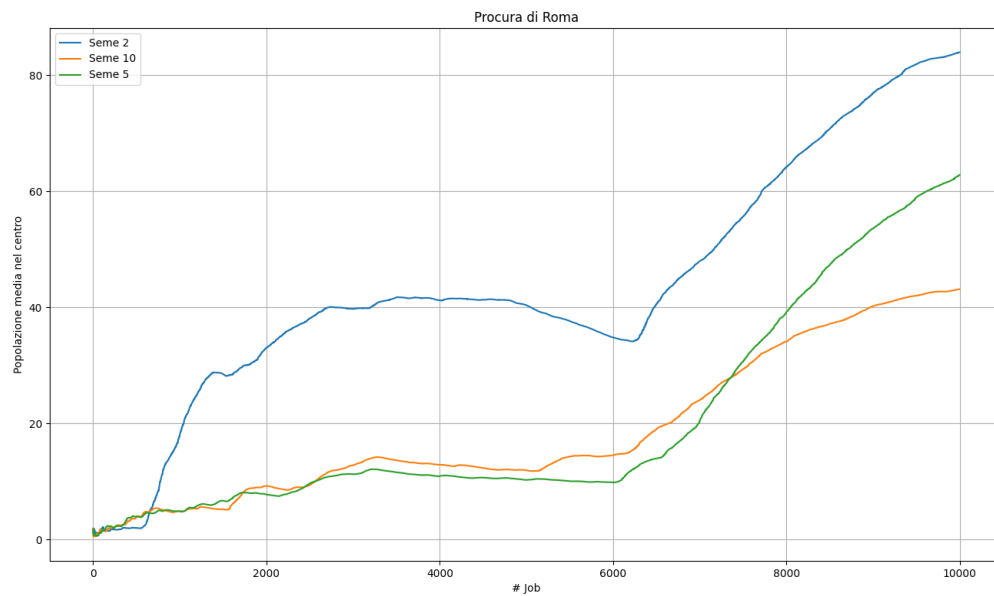


Figure 7.2

7.4 Andamento utilizzazione

Infine nell'ultimo grafico nell'immagine 7.3 è mostrato l'andamento dell'utilizzazione del centro al variare del seme. Si osserva come questo grafico sia strettamente correlato ai grafici precedenti. Nel momento in cui l'andamento in blu tende ad aumentare si osserva negli altri grafici un'impennata delle curve omologhe (sempre in blu). Questo è facilmente interpretabile: all'aumentare dell'utilizzazione il sistema va in sovraccarico e inevitabilmente aumentano i tempi di risposta e il numero medio di job nel centro.

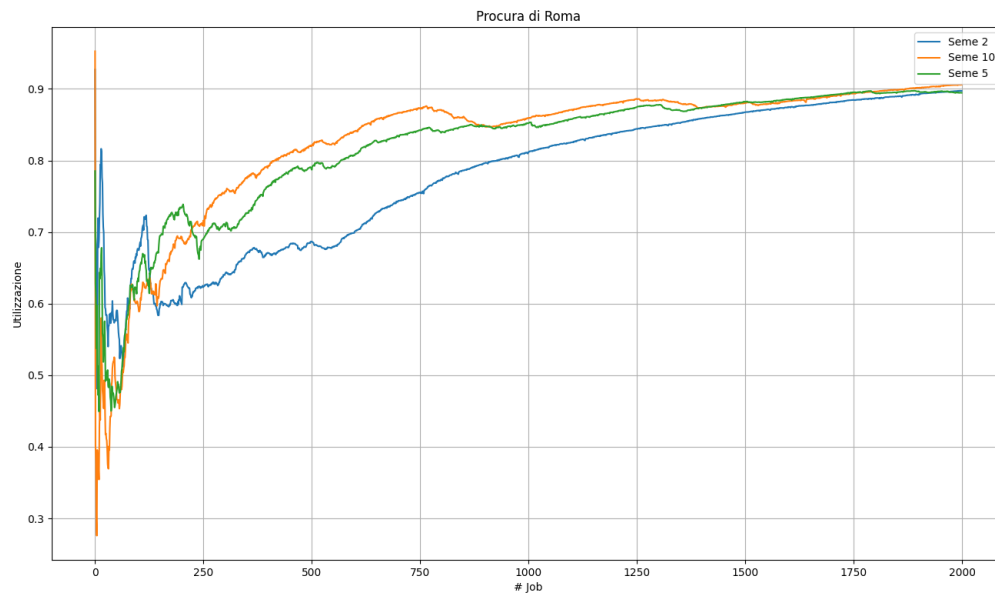


Figure 7.3

–8–

Modello migliorativo

8.1 Introduzione

Il modello migliorativo è stato realizzato partendo dal modello base. Non verrà condotta una nuova analisi del modello concettuale e delle specifiche. Si andranno invece ad apportare delle modifiche puntuali determinate dall'analisi dei dati ottenuti dalla simulazione del modello di base. In particolare si intende apportare due tipologie di modifiche:

- Modificare le politiche di scheduling
- Agire sul tempo di esecuzione medio dei server

Appare evidente che diminuire il tempo di esecuzione medio di un servente comporti un miglioramento delle prestazioni. Questa diminuzione del tempo di esecuzione deve essere però interpretata in maniera opportuna. Si ricordi infatti che alcuni centri, come quello del primo grado monocratico, è un single server che implementa in modo black box una struttura decisamente più complicata che organizza il lavoro di un certo numero di giudici. Agire sul tempo di esecuzione vuol dire implicitamente calcolare l'aumento di giudici che deve esserci nel sistema per ottenere quell'aumento di velocità.

8.2 Versione 1

Il primo miglioramento che può essere realizzato osservando il modello a rete di code è sulle indagini preliminari. Le indagini preliminari vengono infatti eseguite dalla procura che è modellata come un insieme di single server paralleli. Dalla teoria emerge che questa politica di scheduling è decisamente più lenta sia di un multiserver sia di un single server a parità di media. Si decide quindi di sostituire i single server con stream paralleli con un unico multiserver che conserva ovviamente la stessa media. Nel modello reale questa modifica può essere interpretato nel seguente modo: se un utente effettua una denuncia a Latina, la procura di Roma può intervenire ugualmente spostandosi nella regione. Non c'è più una divisione settoriale e locale. In questo modo si riesce a distribuire il carico

in maniera più equa. Sicuramente questo comporta maggiori costi per le procure ma sicuramente un miglior servizio.

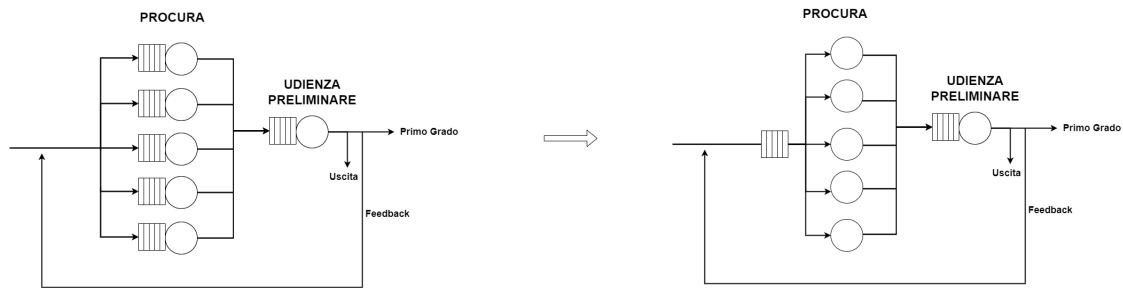


Figure 8.1

L'immagine 8.1 mostra le modifiche effettuate. Il resto del modello a rete di code rimane esattamente lo stesso. Questa modifica ha portato con sé alcune importanti conseguenze. La prima ovviamente è stata un miglioramento delle prestazioni e dei tempi di risposta per il centro che implementa la procura.

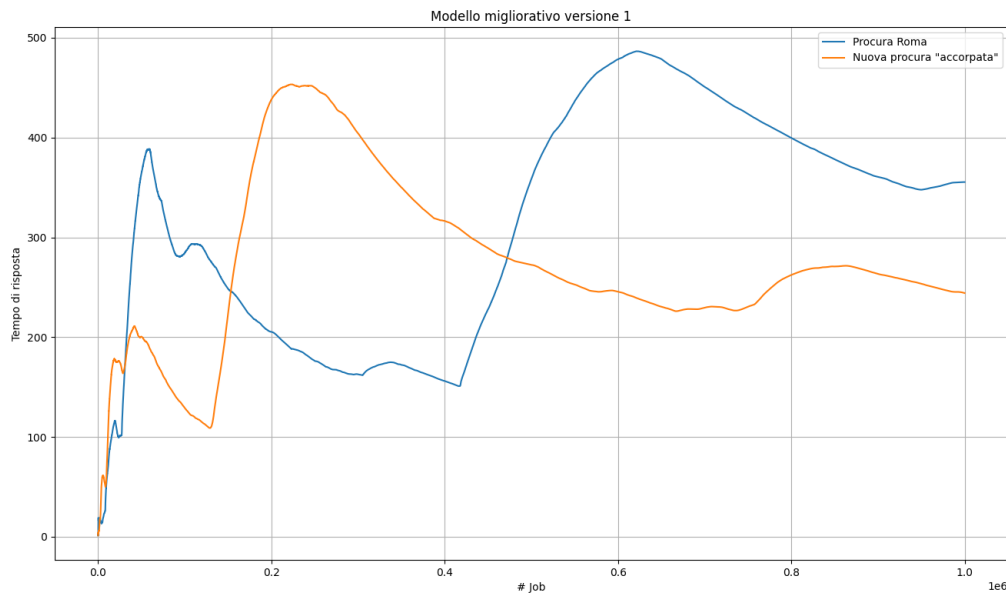


Figure 8.2

L'immagine 8.2 mostra il confronto tra la procura di Roma e la nuova procura modellata dal multiserver. Effettivamente si osserva un miglioramento sui tempi di risposta medi. C'è stata una velocizzazione di circa 100 giorno. Una seconda grande conseguenza però è stata che questa modifica ha messo in crisi l'intero sistema. Tutti gli altri centri sono

esplosi facendo aumentare i tempi di risposta. Questo è accaduto perché la procura ha velocizzato la frequenza con cui restituisce i job in output, causando un sovraccarico degli altri centri che già di base registravano valori di utilizzazione molto alti.

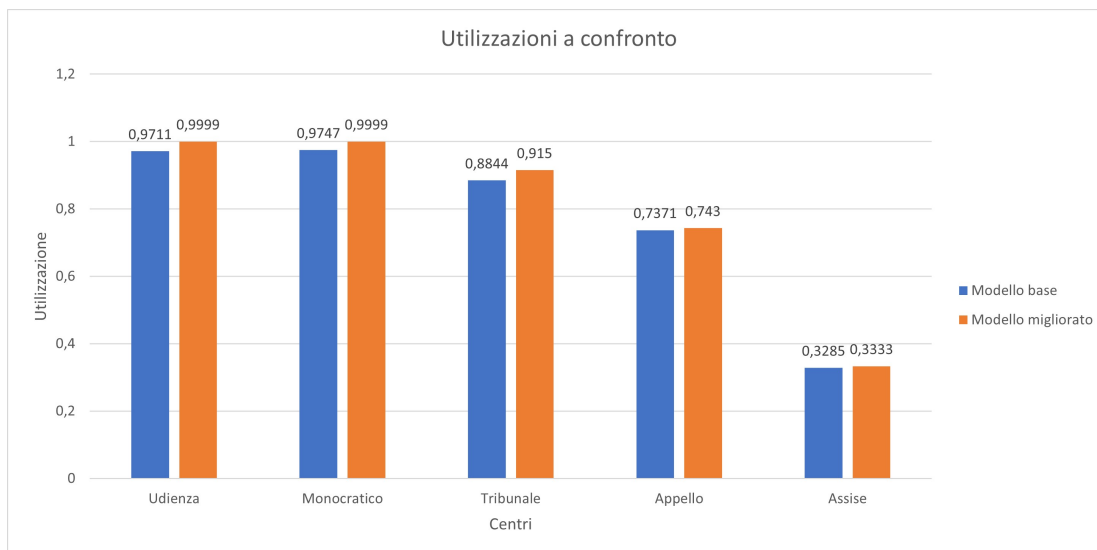


Figure 8.3

L'immagine 8.3 mostra quantitativamente il peggioramento delle prestazioni mettendo a confronto i valori di utilizzazione dei maggiori centri del sistema prima e dopo aver apportato la modifica. Questo peggioramento può essere osservato in maniera più netta nell'immagine 8.4 che mette a confronto le prestazioni del dibattimento monocratico prima e dopo la modifica. Questa forte aumento dei tempi di risposta si registra anche in tutti gli altri centri.

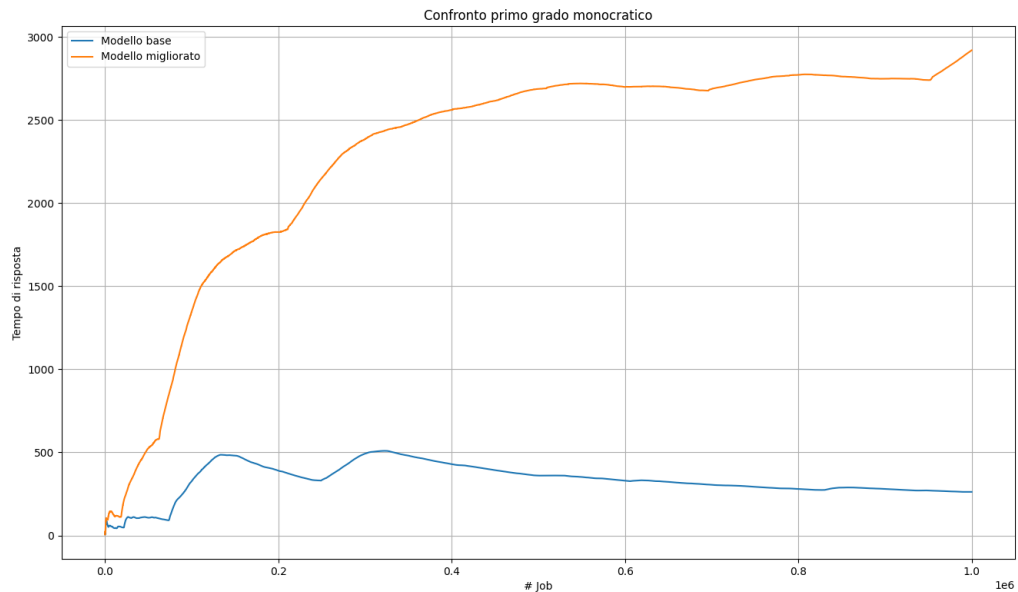


Figure 8.4

8.3 Versione 2

Nella versione migliorativa numero 2 si parte dalla rete individuata nel modello migliorativo numero 1 (quella con la procura modellata da un multiserver) e si determinano il numero di giudici necessari per tornare alle prestazioni del modello base. Si modificano i tempi di esecuzione dell'udienza preliminare, del dibattimento monocratico, del tribunale, della corte di appello. I tribunali fisici non vengono invece modificati. Come detto in precedenza, la diminuzione del tempo di servizio corrisponde ad un aumento del numero di giudici. Le modifiche sono riportate in tabella con conseguente aumento percentuale del numero di giudici:

Centro	E[S] base	E[S] migliorato	Incremento giudici
<i>Udienza</i>	0.375	0.315	16%
<i>Monocratico</i>	1.79	1.48	17%
<i>Tribunale</i>	23	15	34%
<i>Appello</i>	80	75	6%
<i>Assise</i>	250	250	0%

A seguito di queste modifiche tutti i server sono ritornati ai loro valori medi. Un esempio è l'immagine 8.5.

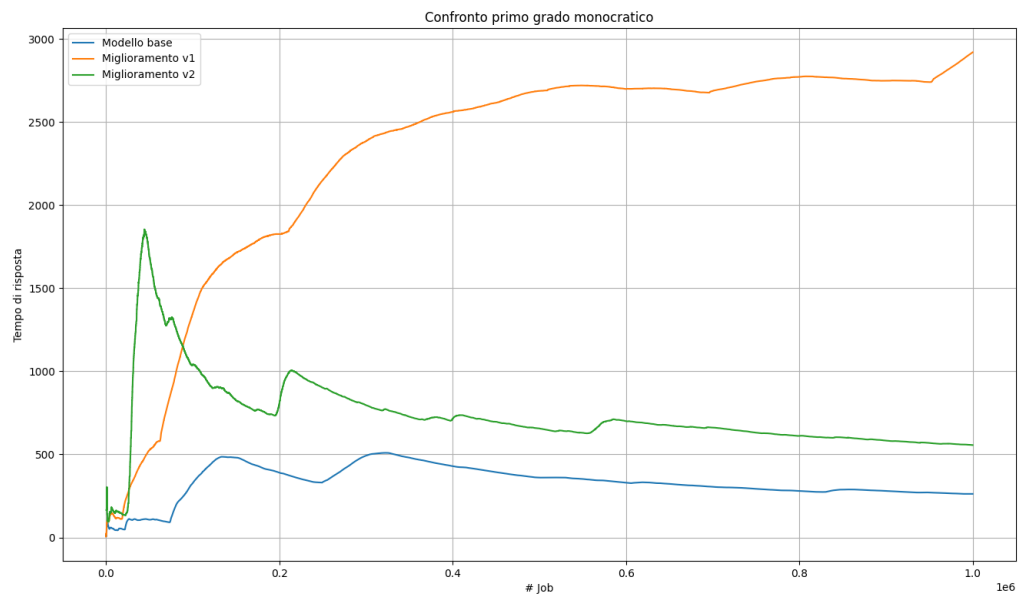


Figure 8.5

Si osservi invece il confronto tra le utilizzazioni nell'immagine 8.6.

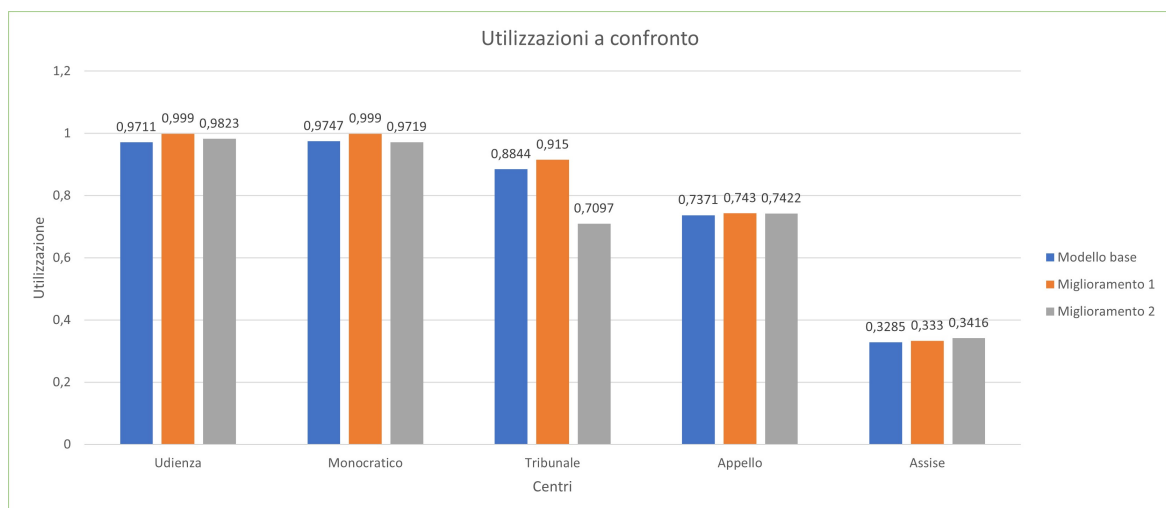


Figure 8.6

8.4 Versione Finale

Nella versione finale si cerca di capire quanti giudici è necessario aggiungere per allineare il sistema giudiziario italiano nel Lazio a quello europeo. Le modifiche apportate sono le descritte nella tabella seguente.

Centro	E[S] base	E[S] migliorato	Incremento giudici
<i>Udienza</i>	0.375	0.315	16%
<i>Monocratico</i>	1.79	1.41	21%
<i>Tribunale</i>	23	9	60%
<i>Appello</i>	80	53	33%
<i>Assise</i>	250	100	60%

Si osservi invece il confronto tra le utilizzazioni nell'immagine 8.7. Si noti come in questa ultima versione ci sia stato un netto miglioramento dell'utilizzazione in quasi tutti i centri. La discesa in percentuale maggiore del valore dell'utilizzazione si registra per il tribunale. Si passa infatti da un valore di 0.8844 del modello di base a 0.427 nel modello finale. Questo risultato è stato raggiunto a fronte di un aumento non indifferente del 60% del numero di giudici che possono presiedere una sentenza in un tribunale fisico.

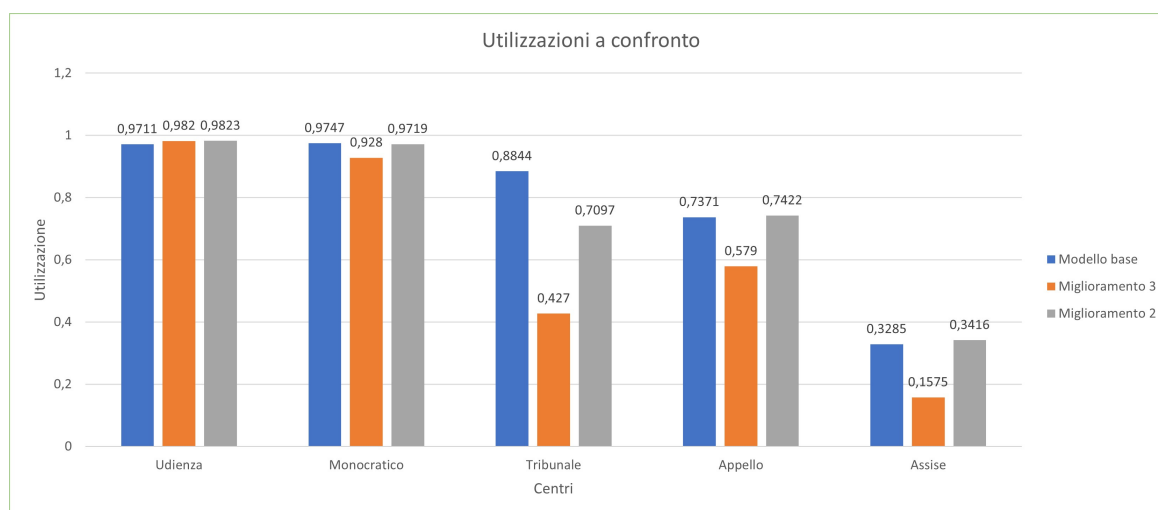


Figure 8.7

Per completezza vengono riportati di seguito tutti i grafici che mostrano la convergenza ai tempi medi europei.

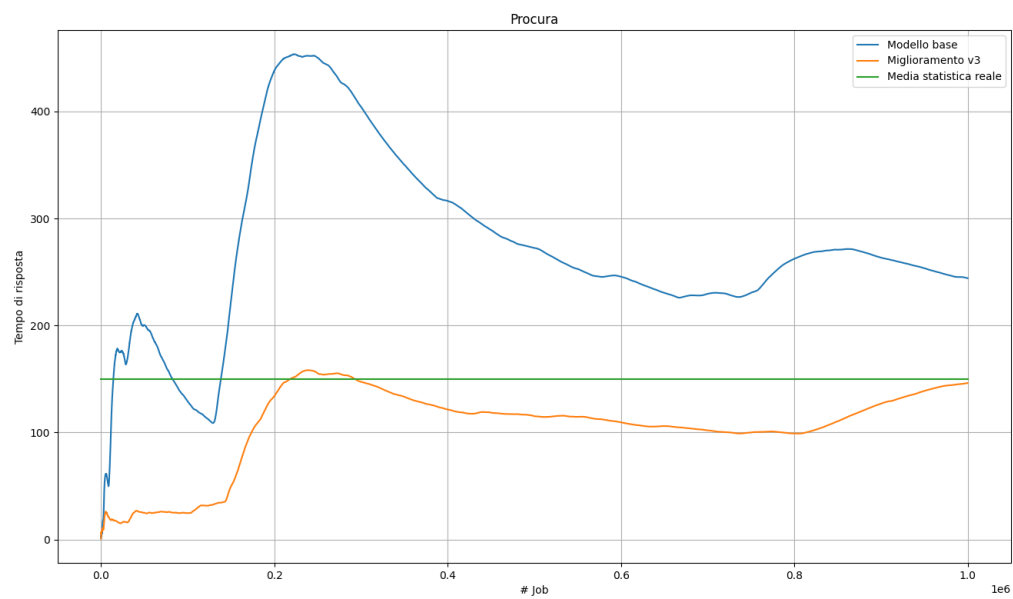


Figure 8.8

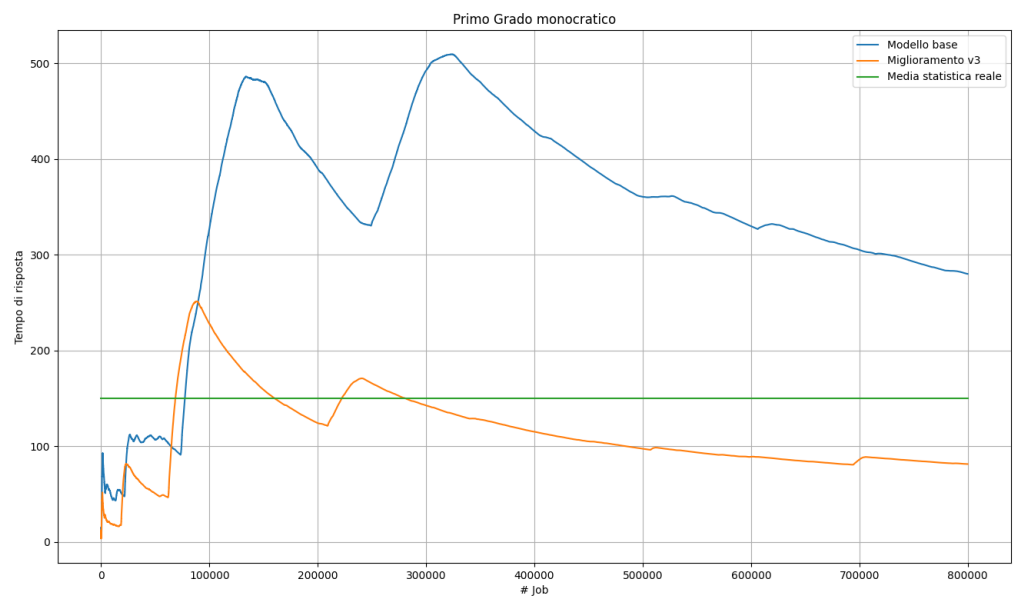


Figure 8.9

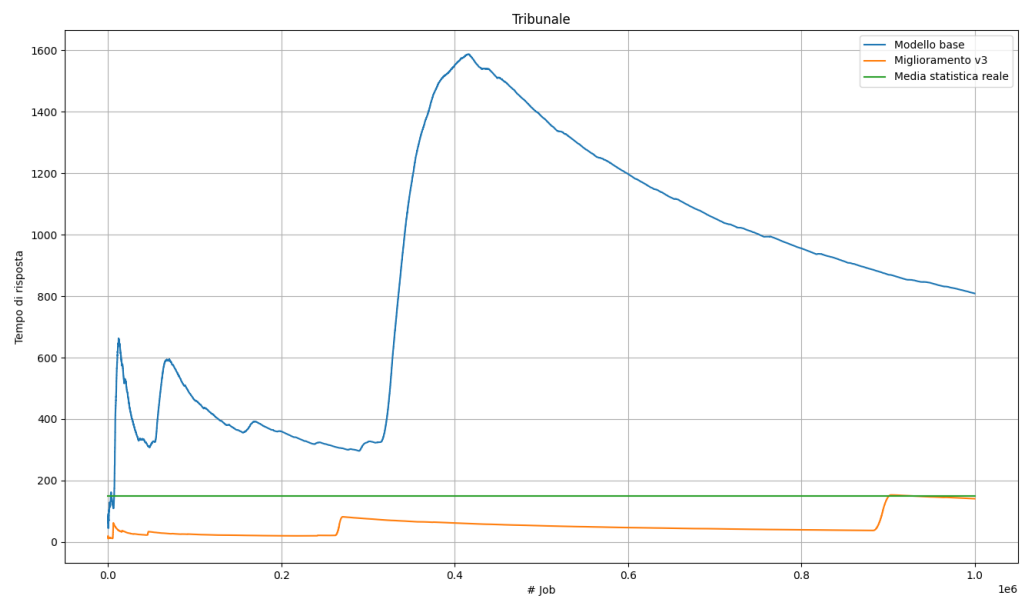


Figure 8.10

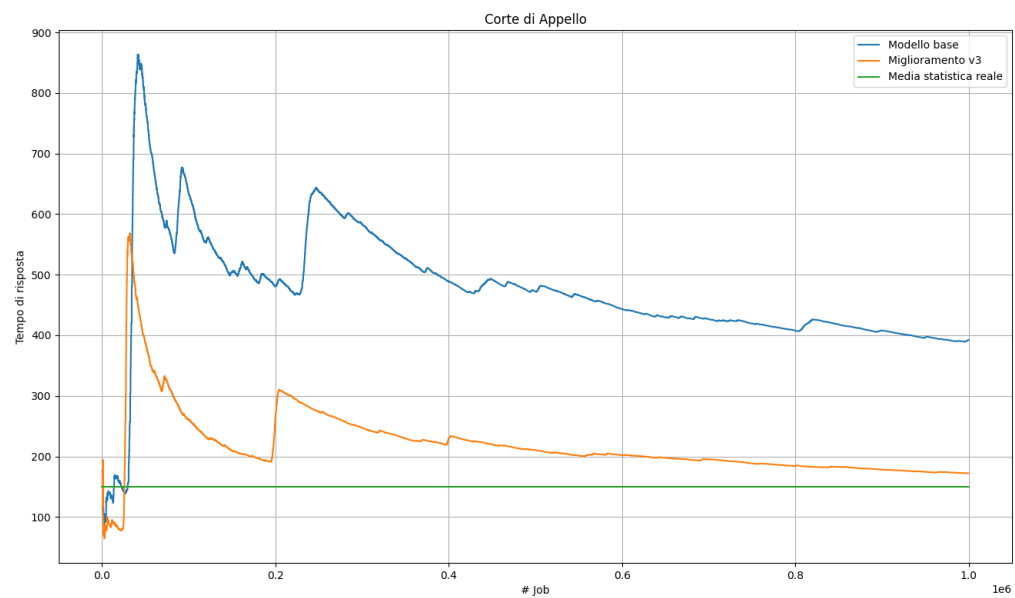


Figure 8.11

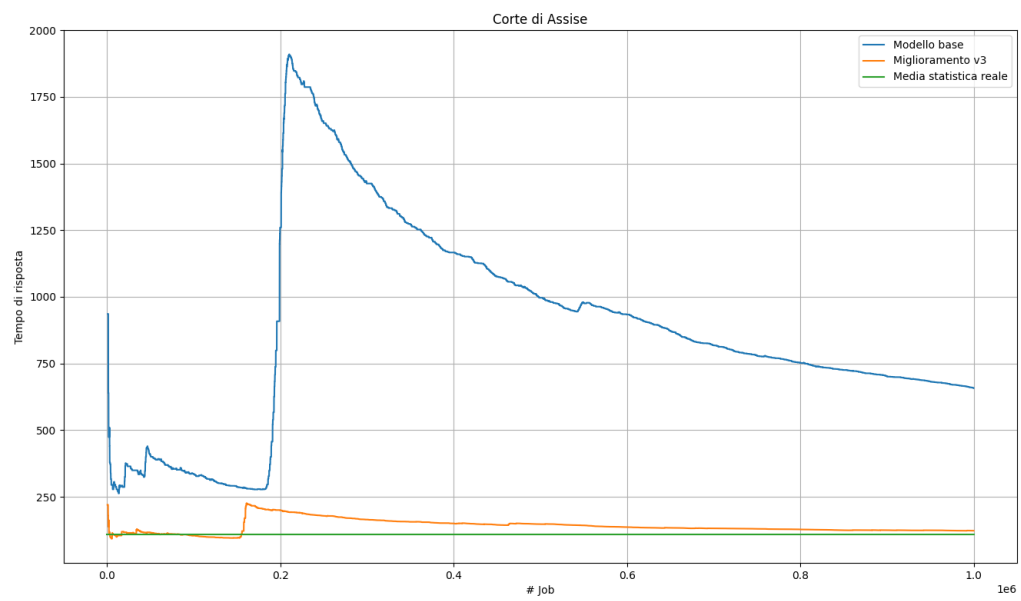


Figure 8.12

Conclusioni

L'analisi condotta ha permesso di individuare alcune criticità nel sistema giudiziario nel Lazio. La lentezza del sistema non è da imputare ad una carenza di infrastrutture fisiche. Il problema sembrerebbe essere dovuto ad una mancanza di giudici che non riescono a smaltire velocemente tutto il carico in ingresso. Questa affermazione trova conferma nel fatto che sono state apportate anche delle modifiche al modello concettuale originale ma questo non ha portato dei miglioramenti notevoli. Addirittura una modifica strutturale del modello ha causato un aumento dei tempi di risposta generali determinando il collasso del sistema. Quindi per raggiungere il secondo obiettivo di questa sperimentazione si è reso necessario agire sul tempo di servizio dei serventi. E' emerso che per far convergere i tempi di risposta delle varie fasi in Italia e nel Lazio ai tempi di risposta europei è necessario un forte sforzo di assunzioni.

Bibliography

- [1] Sito ufficiale Ministero della giustizia. “[https://www.giustizia.it/giustizia/it/mg₁₁4.page](https://www.giustizia.it/giustizia/it/mg_1_1_4.page)”. In: (2023).