

Multimedia

Homework 2

Alessandro Trigo

7 Giugno 2024

Indice

1	Introduzione	3
1.1	Utilizzo dello script	3
1.2	Parametri di progetto	4
2	Performance di rete	5
2.1	Numero di link attraversati	5
2.2	Analisi del Round-Trip-Time	6
2.3	Throughput	13
3	Conclusioni	18

Elenco delle figure

1	Tutte le ottenute con 20 istanze ed uno step di 10.	9
2	Grafici in riferimento all'immagine 1.	10
3	Grafici ottenuti con 100 istanze ed uno step di 2.	11
4	Grafici in riferimento all'immagine 3.	11
5	Grafici ottenuti con 250 istanze ed uno step di 1.	12
6	Grafici in riferimento all'immagine 5.	13
7	Regressione lineare del RTT minimo, in riferimento all'immagine 2. . .	16
8	Regressione lineare del RTT minimo, in riferimento all'immagine 4. . .	16
9	Regressione lineare del RTT minimo, in riferimento all'immagine 6. . .	17

1 Introduzione

Il seguente documento ha l'obiettivo di mostrare la procedura attuata per analizzare le prestazioni di una rete, in particolare di un collegamento tra un *host* ed un server seguendo le indicazioni dell'homework. A tal scopo verrà illustrato e spiegato lo script necessario per effettuare questo tipo di analisi.

In particolare, l'analisi si divide in tre principali parti. La prima ha l'obiettivo di recuperare il numero di connessioni che separano l'host con il server (paragrafo 2.1). In secondo luogo al server verrà mandata una elevata quantità di pacchetti con dimensioni diverse al fine di ottenere la latenza tra la sorgente e la destinazione, ovvero il Round Trip Time (vedi 2.2). Infine, con tutti i dati recuperati sarà possibile modellare la distribuzione di latenze con l'obiettivo di ottenere il throughput (sezione 2.3).

1.1 Utilizzo dello script

Le richieste dell'homework sono state soddisfatte attraverso uno script scritto con il linguaggio *Python* ed eseguito sul sistema operativo *Windows 10* in lingua italiana. Tutte le richieste sono state soddisfatte attraverso l'utilizzo del comando `ping` che si occupa di inviare delle richieste ICMP¹ a una determinata destinazione. La risposta del suddetto comando diagnostico è il periodo impiegato dalla destinazione per rispondere, ovvero la **latenza**.

Prima di lanciare lo script, è necessario aprire il file `script.py` e impostare correttamente il percorso dello script modificando la costante `PATH_TO_SCRIPT`.

```
1 | PATH_TO_SCRIPT = os.path.join("multimedia", "hw-2", "script")
```

In secondo luogo è possibile impostare le specifiche di progetto mediante delle costanti che indicano il nome del server su cui inviare le richieste, il numero di istanze da mandare a ogni richiesta ping e l'incremento di lunghezza (step) che ci sarà tra la lunghezza del pacchetto durante l'iterazione i e la successiva $i + 1$.

```
1 | # project specification
2 | SELECTED_SERVER_CITY = "Atlanta"
3 | INSTANCES = 30
4 | STEP_BETWEEN_LENGTHS = 25
```

Tali impostazioni verranno poi brevemente processate con il seguente codice. In particolare si osserva che la costante `SELECTED_SERVER_CITY` verrà utilizzata per ottenere l'effettivo dominio del server e l'abbreviazione del nome della città. In secondo luogo la costante `STEP_BETWEEN_LENGTHS` verrà utilizzata per generare un vettore che parta da 10 e per arrivare a 1471 incrementa il suo valore di esattamente `STEP_BETWEEN_LENGTHS`.

```
1 | SERVERS = {
2 |     "Atlanta" : "atl.speedtest.clouvider.net",
3 |     "New York City" : "nyc.speedtest.clouvider.net",
4 |     "London" : "lon.speedtest.clouvider.net",
5 |     "Los Angeles" : "la.speedtest.clouvider.net",
```

¹ICMP sta per *Internet Control Message Protocol* è un protocollo utilizzato dai dispositivi di una rete per comunicare la presenza di problemi riguardanti la trasmissione dei dati

```

6         "Paris" : "paris.testdebit.info",
7         "Lillie" : "lille.testdebit.info",
8         "Lyon" : "lyon.testdebit.info",
9         "Aix-Marseille" : "aix-marseille.testdebit.info",
10        "Bordeaux" : "bordeaux.testdebit.info"
11    }
12
13    # set choosen options
14    server = SERVERS[SELECTED_SERVER_CITY]
15    city = server.split(".")[0]
16    payload_lengths = range(10, 1471, STEP_BETWEEN_LENGTHS)

```

Usando le impostazioni dell'esempio, lo script — che sarà descritto nel capitolo 2 — manderà le richieste al server di Atlanta `atl.speedtest.clouvider.net`. Lo script poi, al fine di analizzare le performance di rete di tale server, per ogni valore di lunghezza all'interno della lista `payload_lengths`, effettuerà 30 istanze di richiesta. In altre parole si faranno 30 richieste di lunghezza 10 byte al server di Atlanta, poi altre 30 di lunghezza $10 + 25 = 35$ byte, poi altre 30 di lunghezza $35 + 25 = 60$ bytes e così via fino ad arrivare alla lunghezza 1470 bytes, per un totale di 1770 richieste.

1.2 Parametri di progetto

All'interno dello script vengono utilizzati due tipi di comando `ping`. Il primo, utilizzato nella prima task, descritta nel paragrafo 2.1, è il comando `ping` di default di Windows: lo script funziona solo e soltanto se la lingua di Windows è in italiano altrimenti non è possibile analizzare correttamente le risposte del comando. Nella seconda task, esplorata nel paragrafo 2.2, viene utilizzato invece il comando `psping`, che è in lingua inglese. Al fine di riuscire ad utilizzare lo script è necessario avere l'eseguibile `psping.exe` all'interno della cartella da cui poi viene eseguito lo script (ovvero la cartella che, attraverso il percorso all'interno di `PATH_TO_SCRIPT`, permette di arrivare al file `script.py`). Si è scelto di utilizzare `psping` in quanto è più prestante del comando `ping` di default. Inoltre nel primo punto non si è utilizzato `psping` in quanto non è presente l'impostazione che permette di ridurre il Time-To-Live del pacchetto.

In secondo luogo, al fine di eseguire i comandi nel terminale, non si è utilizzata la funzione `os.system`, bensì la libreria `subprocess`, che permette di ottenere il risultato direttamente in una stringa senza avere la necessità di scriverlo sul file per poi accedere al file per leggerlo. Questo migliora le prestazioni computazionali dello script, in particolare nella seconda parte dove in genere viene inviata una mole elevata di pacchetti ICMP. È comunque possibile scrivere su file i risultati dell'esecuzione del comando, è sufficiente impostare a `True` la *flag* che si chiama `SAVE_TO_FILE`.

Le analisi che seguiranno saranno effettuate sul server di Los Angeles il cui dominio è `la.speedtest.clouvider.net`. In particolare verranno prese in considerazione tre diverse istanze di esecuzione dello script, le quali avranno un numero crescente di pacchetti da mandare:

1. 20 messaggi ed uno step di 10, quindi un totale di 2 940 pacchetti (figure 1, 2, 7);
2. 100 istanze ed uno step di 2, ammontando a 73 100 blocchi (figure 3, 4, 8);
3. 250 pacchetti ed uno step di 1, per un totale di 365 250 messaggi (figure 5, 6, 9).

2 Performance di rete

Dopo aver introdotto i parametri di progetto adottati, andremo ora ad analizzare le performance di rete. In particolare ci occupiamo di analizzare il numero di connessione che debbono essere attraversate al fine di raggiungere il server a Los Angeles `la.speedtest.clouvider.net`. In secondo luogo, attraverso una massiccia fase di richieste ICMP al server analizzeremo i tempi di latenza, estraendone i minimi, i massimi, la media e la deviazione standard. Infine utilizzeremo i dati raccolti per stimare ed analizzare il throughput verso il server scelto.

2.1 Numero di link attraversati

Il primo valore che andremo a calcolare è il numero di connessioni — dette anche *links* — che ci separa dal server di Los Angeles. In particolare ci occuperemo di recuperare tale valore in due modi diversi. Il primo modo per farlo è tramite l'utilizzo del comando Windows `tracert` che permette di tracciare interamente il percorso che un pacchetto IP² compie al fine di raggiungere la destinazione. Il comando `tracert` si basa su una serie contigua di ping che sono inviati impostando il Time-To-Live a 1 e incrementandolo man mano. In questo modo, nel momento in cui il comando ping non scade, il TTL dello specifico messaggio ICMP corrisponde al numero di link che separa il sorgente dal destinatario. Invocando il comando `tracert` sul server di Los Angeles `la.speedtest.clouvider.net`, si ottiene sul terminale la lista numerata di tutti i link attraversati, come mostrato nel frammento di codice seguente.

```
Traccia instradamento verso la.speedtest.clouvider.net [77.247.126.223]
su un massimo di 30 punti di passaggio:

 1  <1 ms  <1 ms  <1 ms  CLOUD-STORAGE [192.168.1.1]
 2  36 ms  17 ms   8 ms  151.6.141.50
 3  11 ms   7 ms   8 ms  151.6.141.34
 4  14 ms  14 ms  14 ms  151.6.0.68
 5  13 ms  15 ms  12 ms  151.6.1.182
 6  15 ms  14 ms  14 ms  mno-b3-link.ip.twelve99.net [62.115.36.84]
 7  30 ms  29 ms  29 ms  prs-bb1-link.ip.twelve99.net [62.115.135.224]
 8  113 ms 113 ms 127 ms ash-bb2-link.ip.twelve99.net [62.115.112.242]
 9  172 ms 172 ms 171 ms lax-b22-link.ip.twelve99.net [62.115.121.220]
10  173 ms 173 ms 173 ms clouvider-ic-355873.ip.twelve99-cust.net
    [213.248.74.63]
11  176 ms   *    203 ms 77.247.126.1
12  172 ms 171 ms 172 ms 77.247.126.223

Traccia completata.
```

Tale risultato si traduce nel seguente codice Python dove, dopo aver memorizzato il risultato del comando nella stringa `result`, quest'ultima viene suddivisa in righe e viene analizzata in modo tale da restituire l'ultimo numero del link, che coincide con il numero totale di connessioni attraversate per raggiungere il server. Mediante tale

²IP è l'acronimo di *Internet Protocol*

funzione si ottiene che il numero di links necessari per raggiungere il server di Los Angeles corrisponde a 12.

```
1 def get_links_from_tracert(server: str) -> int:
2
3     # get number of links from tracert
4     cmd = f"tracert {server}"
5     result = subprocess.run(cmd, shell=True,
6                             stdout=subprocess.PIPE, stderr=subprocess.PIPE,
7                             text=True)
8
9     last_link_line = result.stdout.split("\n")[-4]
10
11     return int(last_link_line.split(" ")[1])
```

Per contare il numero di connessioni, oltre all'utilizzo del comando `tracert`, è possibile impiegare direttamente il comando `ping`. In particolare, è sufficiente iterare attraverso una serie di ping, diminuendo a ogni iterazione il parametro TTL (Time-To-Live), che rappresenta il numero massimo di link che il pacchetto può attraversare prima di essere eliminato. La seguente funzione in Python itera sui valori di TTL da 20 a 0. a ogni iterazione viene inviata una richiesta ping predefinita, con 4 pacchetti di dimensione di 32 byte. Quando il ping non ottiene una risposta, significa che il valore di TTL è diminuito al punto da non essere più sufficiente per raggiungere la destinazione finale. Pertanto, viene restituito il valore di TTL incrementato di 1, che è il minimo TTL necessario per raggiungere il server. Di fatto questo appriccio è l'esatto opposto dell'approccio che adotta `tarcert`: in questo caso infatti al posto di incrementare il TTL fino a che non si riceve una risposta, lo si decrementa fino a che il messaggio ICMP scade, non riuscendo quindi a raggiungere la destinazione.

```
1 def get_links_from_ping(server: str) -> int:
2
3     for ttl in range(20, 0, -1):
4
5         cmd = f"ping {server} -i {ttl}"
6         result = subprocess.run(cmd, shell=True,
7                                 stdout=subprocess.PIPE, stderr=subprocess.PIPE,
8                                 text=True)
9
10        if "TTL scaduto durante il passaggio" in result.stdout:
11            return (ttl + 1)
```

In tutti e due i casi, come risultato, il numero di connessioni che separa il sorgente dal server di Los Angeles — `la.speedtest.clouvider.net` — è 12.

2.2 Analisi del Round-Trip-Time

La seconda parte, volta ad analizzare il *RTT*, necessita uno script Python più complesso, anche volto a migliorare le prestazioni computazionali dello script. Si è quindi scelto di utilizzare un approccio **multi-threading** ed il comando `psping`, che è più

prestante rispetto al comune ping, inoltre permetto di ottenere delle latenze più precise (fino al centesimo di millisecondo).

Per permettere l'utilizzo di diversi thread, è necessario racchiudere le richieste ping in una funzione, chiamata `ping_server`. In questa funzione la richiesta di ping ha diverse proprietà specificate:

- `-n`, che specifica il numero di istanze per ogni richiesta ping;
- `-l`, indica la lunghezza del *payload* di ciascuna richiesta;
- `-i`, specifica l'intervallo di secondo tra un ping e l'altro; viene impostato a zero per aumentare la rapidità dello script;
- `-w`, che indica il numero di *warmup request* da fare prima di iniziare la sequenza di ping.

Dopo aver eseguito il comando, il risultato viene analizzato e ne sono estratte le latenze, che sono aggiunte in una lista (o vettore). Infine sono ritornati due valori, la lunghezza corrente e la lista con tutte le latenze.

```
1 def ping_server(server, instances, length):
2
3     cmd = f"psping -n {instances} -l {length} -i 0 -w 0 {server}"
4     result = subprocess.run(cmd, shell=True,
5                             stdout=subprocess.PIPE, stderr=subprocess.PIPE,
6                             text=True)
7
8     millisecs_vector = []
9     lines = result.stdout.split("\n")
10
11     for line in lines:
12         if "Reply from" in line:
13             duration = float(line.split(": ")[1].split("ms")[0])
14             millisecs_vector.append(duration)
15
16     return length, millisecs_vector
```

La funzione descritta sopra viene utilizzata nel seguente script dove, mediante il pacchetto `ThreadPoolExecutor`, viene eseguita da ciascun thread. In questo modo molte più richieste sono mandate contemporaneamente migliorando notevolmente le prestazioni del programma. Senza l'utilizzo di un algoritmo multithreading, ottenere tempistiche di esecuzione superiori a una o due ore era inevitabile, anche con un numero limitato di istanze (ad esempio, 50 istanze con incrementi di 75). Infine, la coppia restituita dalla funzione `ping_server`, viene aggiunta al dizionario `stats` dove sono immagazzinate tutte le latenze per poi essere analizzate.

```
1 stats = {}
2
3 # using ThreadPoolExecutor to improve computational capabilities
4 with concurrent.futures.ThreadPoolExecutor(max_workers=100) as
5     executor:
6
7     futures = []
```

```

7     for length in payload_lengths:
8         # submit a task to the executor to ping the server with the
          specified parameters
9         future = executor.submit(ping_server, server,
          INSTANCES, length)
10        futures.append(future)
11
12        # iterate over completed futures as they become available
13        for future in concurrent.futures.as_completed(futures):
14            length, millisecs_vector = future.result()
15            stats[length] = millisecs_vector
16
17        # order stats by length
18        stats = dict(sorted(stats.items()))

```

Al fine di calcolare tutti i massimi, i minimi, le medie e le varianze è sufficiente iterare lungo il dizionario creato. Lo seguente script si occupa proprio di questo, anche se in modo poco efficiente ma perlomeno leggibile.

```

1     max_values = {}
2     min_values = {}
3     average_values = {}
4     standard_deviations = {}
5
6     for key, value in stats.items():
7         max_values[key] = max(value)
8         min_values[key] = min(value)
9         average_values[key] = sum(value) / len(value)
10        standard_deviations[key] = math.sqrt(sum((x -
          average_values[key]) ** 2 for x in value) / len(value))

```

Risultati

Dopo aver eseguito gli script, è quindi possibile mostrare tutte le latenze raccolte, i loro massimi, i minimi, le latenze medie per ciascuna lunghezza e la loro varianza. Sono presentati di seguito tre immagini contenenti i risultati delle richieste; ciascuna delle tre immagini fa riferimento a una diversa scelta di istanze e di step. Infine, si osserva che nei grafici seguenti la lunghezza L dei pacchetti è espressa in bit attraverso la seguente formula di traduzione:

$$L = 8 \cdot (L_{\text{byte}} + 28)$$

L'immagine 1 di seguito rappresenta tutte le latenza raccolte durante l'esecuzione dei ping con 20 istanze per richiesta e con uno step di 10 tra una lunghezza e l'altra, per un totale di 2940 ping effettuati in circa un minuti. Si può osservare che in genere tutte le latenze sono attorno ai 170 millisecondi, con alcuni picchi tra i 750 e gli 800 bytes. Non a caso, nei grafici rappresentati in figura 2, si osserva infatti che i picchi sono rispecchiati nella rappresentazione dei massimi (in alto a destra), delle medie (in basso a sinistra) e delle varianze (in basso a destra). Si osserva inoltre che i

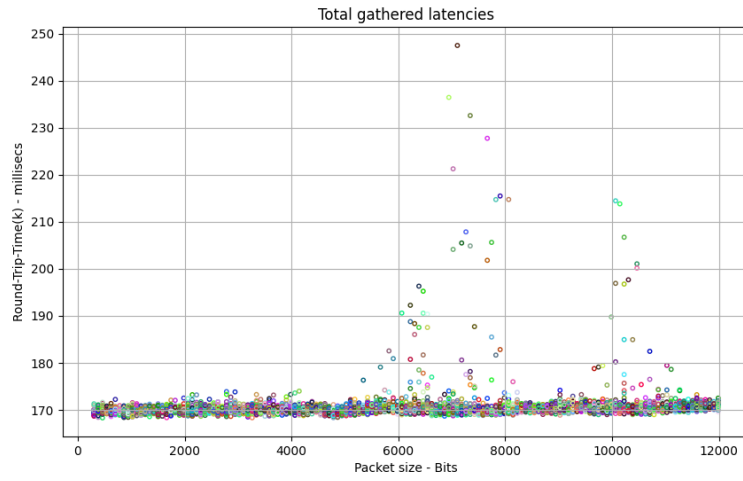


Figura 1: Tutte le ottenute con 20 istanze ed uno step di 10.

minimi variano tra 168 e 170 millisecondi e che crescono al crescere della dimensione del pacchetto; questo aspetto sarà dettagliatamente esplorato nel prossimo paragrafo, dedicato al *throughput* (2.3).

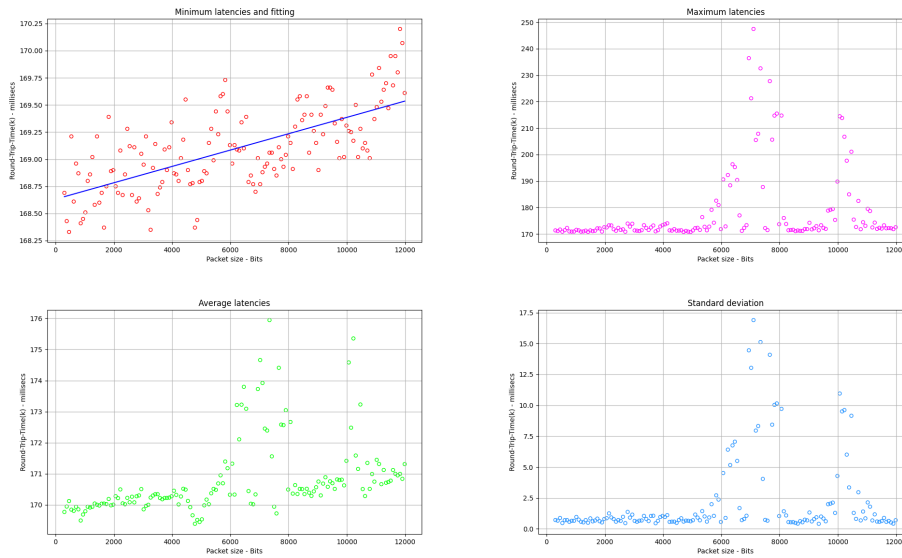


Figura 2: Grafici in riferimento all'immagine 1.

Aumentando il numero di ping effettuati, in particolare aumentando il numero di istanze a 100 e riducendo l'incremento delle lunghezze a 2, otteniamo un grafico come quello della figura 3, che contiene 73 100 valori di latenza, ottenuti in circa 6 minuti. Anche in questo caso possiamo notare dei picchi che sono distribuiti lungo tutte le lunghezze. In particolare, i più elevati si aggirano attorno a 600 bytes di lunghezza del payload. Osservando infatti i grafici nell'immagine 4, notiamo una forte varianza (fino a 30 millisecondi) proprio in queste zone. In questa immagine si rende ancora più evidente la crescita dei minimi e anche delle medie con il crescere della lunghezza dei pacchetti.

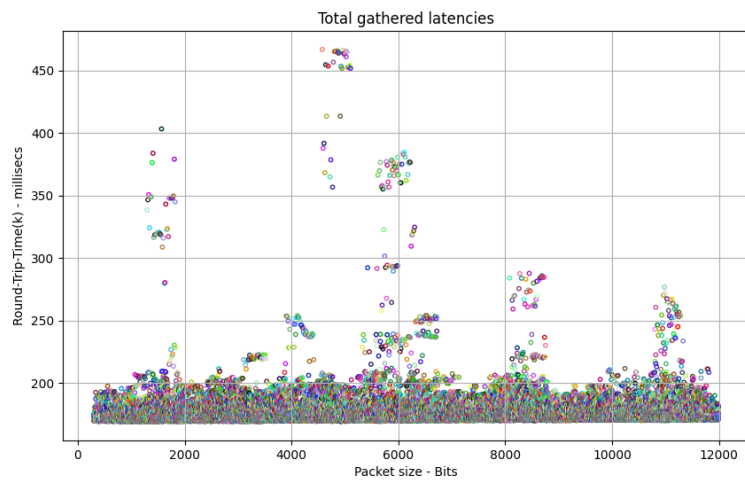


Figura 3: Grafici ottenuti con 100 istanze ed uno step di 2.

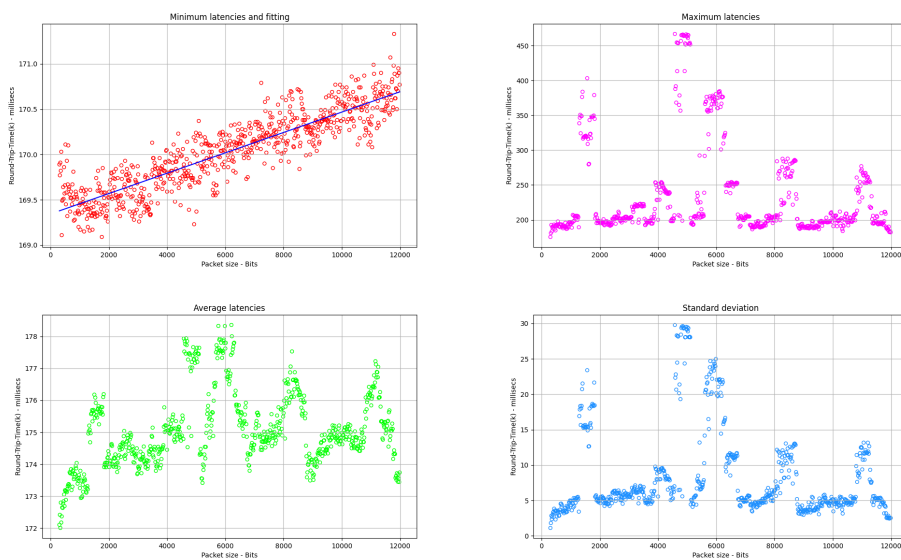


Figura 4: Grafici in riferimento all'immagine 3.

Infine, nell'immagine 5 sono state processate un totale di 365 250 richieste in circa un quarto d'ora, risultato di 250 istanze per lunghezza ed uno step di 1 tra una lunghezza e l'altra. Si può infatti notare che il numero di dati in quest'ultimo grafico è molto più elevato rispetto ai due precedenti garantendo quindi che i dati analizzati sono più simili a quelli reali. In questo caso è interessante osservare che sia la media che la

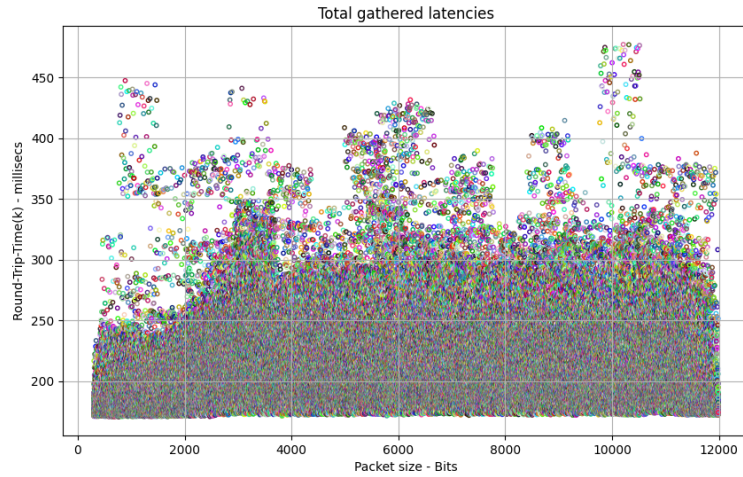


Figura 5: Grafici ottenuti con 250 istanze ed uno step di 1.

varianza sono elevate per la maggior parte delle lunghezze per decrescere solo quando il pacchetto è di dimensione ridotta oppure molto elevata (figura 6). Anche il grafico che rappresenta il minimo non segue più una retta definita (a differenza di quello mostrato nei grafici 4), ma dopo i 350 bytes si hanno minimi che arrivano fino a 178 millisecondi.

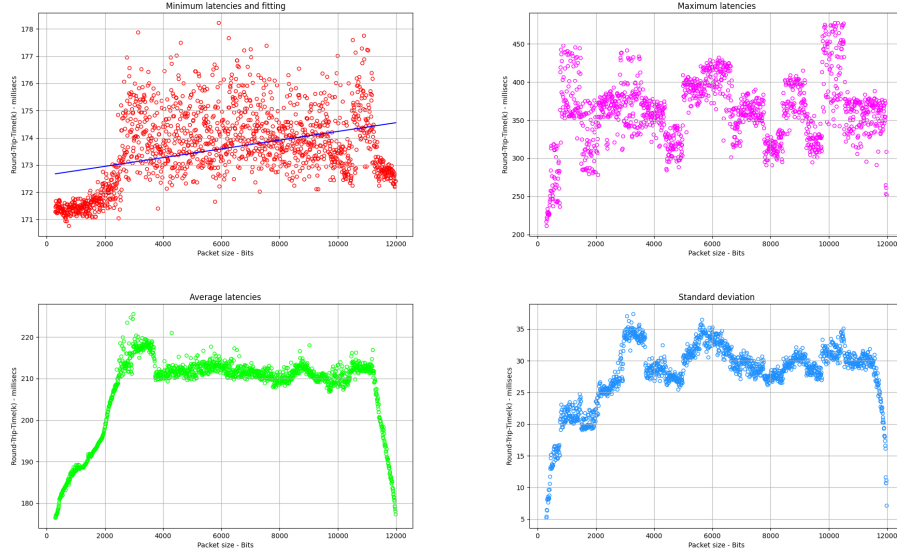


Figura 6: Grafici in riferimento all'immagine 5.

Le analisi relativi ai grafici mostrati, in particolare in riferimento alla drastica crescita di latenza (minima e media) all'inizio dell'invio dei messaggi ping e la drastica decrescita verso la fine dei dati in figura 6 saranno discusse in maniera dettagliata nel capitolo 3.

2.3 Throughput

Attraverso tutti i dati raccolti è possibile riuscire a stimare il *throughput* \mathcal{S} . Questa grandezza è il rapporto tra la quantità di dati trasmessi e il tempo di trasmissione: tale valore serve per stimare la capacità effettiva del canale di trasmissione.

Si fornisce una breve spiegazione del ragionamento matematico che le supporta le scelte di calcolo del valore del throughput. Iniziando dalla definizione del Round Trip Time di un pacchetto k di lunghezza L in bit ($\text{RTT}(L, k)$), questa è la somma dei ritardi d_i causati dai collegamenti i -esimi ($i \in [1, n]$); tale latenza è a sua volta definita come la somma di 4 valori di latenza d :

$$d_i = d_{i,\text{processing}} + d_{i,\text{queue}} + d_{i,\text{transmission}} + d_{i,\text{propagation}}$$

Dove tali latenze indicano rispettivamente:

- ◇ $d_{i,\text{processing}}$, che indica il ritardo impiegato per controllare gli errori e calcolare l'instradamento del pacchetto — in genere è un ritardo trascurabile;
- ◇ $d_{i,\text{queue}}$, ovvero il tempo trascorso del pacchetto k all'interno di un buffer prima di essere elaborato — tale valore varia nel tempo, a seconda di quanto sia pieno il buffer, e verrà indicato con la notazione $q_i(k)$;

- ◇ $d_{i,\text{transmission}}$, è la durata del segnale che rappresenta il pacchetto — può essere matematicamente modellato come il rapporto tra la dimensione L del pacchetto k e dal throughput R_i del link i ;
- ◇ $d_{i,\text{propagation}}$, rappresentato con τ_i , indica il tempo fisico di trasmissione del segnale lungo il medium.

Dunque il valore di d_i diventa:

$$d_i = q_i(k) + \frac{L}{R_i} + \tau_i$$

E il calcolo del Round Trip Time totale risulta quindi essere:

$$\begin{aligned} \text{RTT}(L, k) &= \sum_{i=1}^n d_i = \sum_{i=1}^n \left[q_i(k) + \frac{L}{R_i} + \tau_i \right] \\ &= \sum_{i=1}^n q_i(k) + L \sum_{i=1}^n \frac{1}{R_i} + \sum_{i=1}^n \tau_i \end{aligned}$$

A questo punto cerchiamo di riassumere le tre sommatorie come segue:

$$Q(k) = \sum_{i=1}^n q_i(k) \quad \alpha = \sum_{i=1}^n \frac{1}{R_i} \quad T = \sum_{i=1}^n \tau_i$$

E conseguentemente il calcolo del Round Trip Time diventa:

$$\text{RTT}(L, k) = Q(k) + \alpha L + T$$

Per riuscire a ridurre completamente il Round Trip Time a una retta, è necessario gestire la variabilità del ritardo di coda del pacchetto $Q(k)$. Si osserva che, con un elevato numero di istanze K — come quelle scelte in questo caso, dove $K = 250$ — la probabilità che almeno una volta i buffer dei link siano liberi aumenta. Di conseguenza, di tutti gli RTT, prendere la minima latenza per ogni pacchetto $k \in \{1, \dots, K\}$, garantirebbe che $Q(k) = 0$. Si ottiene quindi un'approssimazione lineare del RTT:

$$\text{RTT}_{\min}(L) = \min_{k \in \{1, \dots, K\}} \text{RTT}(L, k) \approx \alpha L + T$$

A questo punto si hanno tutte le informazioni necessarie per ricavare il throughput \mathcal{S} in quanto, attraverso tutti i dati ottenuti nel paragrafo 2.2, è possibile prendere il minimi RTT per ciascuna lunghezza $L \in \{1, \dots, 1470\}$ e, attraverso una regressione lineare (detto anche interpolazione o *fitting*), è possibile ottenere sia il coefficiente α che la somma dei ritardi di trasmissione T . Una volta recuperato il coefficiente α , è sufficiente invertire la formula definita in precedenza per calcolare il throughput come segue:

$$\alpha = \sum_{i=1}^n \frac{1}{R_i} = \frac{n}{\mathcal{S}} \quad \longrightarrow \quad \mathcal{S} = \frac{n}{\alpha}$$

Si osserva che la relazione soprastante prende come ipotesi quella in cui tutti i link abbiano throughput uguali. Nell'eventualità in cui sia presente un link con throughput nettamente minore rispetto agli altri, si ha un effetto di collo di bottiglia —

bottleneck. In questo caso, si ha una riduzione del throughput sia durante l'andata del pacchetto ping che al ritorno, ciò determina una riduzione del throughput globale, approssimabile con il calcolo del throughput nell'ipotesi in cui $n = 2$:

$$\mathcal{S}_{\text{bottleneck}} = \frac{n}{\alpha} \Big|_{n=2} = \frac{2}{\alpha}$$

I risultati matematici ottenuti si traducono in un numero estremamente ridotto di linee di codice. Prima di tutto è necessario creare la lista di lunghezza trasformata in bit, come menzionato all'inizio del paragrafo, attraverso la formula $L = 8 \cdot (L_{\text{byte}} + 28)$. In secondo luogo, tramite una regressione lineare, fornita dalla libreria `sklearn.linear_model`, è possibile recuperare il coefficiente α e quindi calcolare i due throughput: \mathcal{S} ed $\mathcal{S}_{\text{bottleneck}}$.

```

1 |     payload_lengths_bit = [8 * (length + 28) for length in
   |                           payload_lengths]
2 |
3 |     # use linear regression to retrieve alpha, need to transform list into
   |     # np.arrays
4 |     reg = LinearRegression().fit(
5 |         np.array(payload_lengths_bit).reshape(-1, 1), # transpose of
   |         # payload lengths
6 |         np.array(list(min_values.values()))
7 |     )
8 |
9 |     alpha = reg.coef_[0]
10 |
11 |     # compute throughput
12 |     throughput_identical_link = 2 * tracert_links / alpha
13 |     throughput_bottleneck = 2 / alpha

```

Risultati

Eseguendo il codice descritto con le tre diverse istanze di esecuzione, è possibile calcolare i due casi di throughput e analizzare meglio il *fit* sui minimi ottenuti per ciascuna lunghezza L . Iniziando dall'esecuzione con 20 istanze ed uno step di 10, mediante l'interpolazione lineare si ottiene che $\alpha \approx 7.5 \cdot 10^{-5}$ (figura 7) e quindi:

$$\mathcal{S} = 318511 \text{ bps} \approx 319 \text{ Kbps}$$

$$\mathcal{S}_{\text{bottleneck}} = 26542 \text{ bps} \approx 27 \text{ Kbps}$$

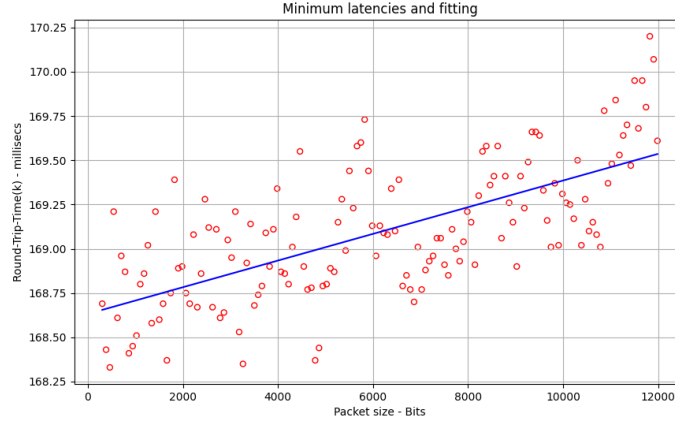


Figura 7: Regressione lineare del RTT minimo, in riferimento all'immagine 2.

Eseguendo lo script impostando $K = 100$ e lo step a 2 si giunge ad un coefficiente $\alpha \approx 1.12 \cdot 10^{-4}$ (immagine 8), maggiore rispetto a quello precedente che quindi riduce il valore del throughput a:

$$\mathcal{S} = 213508 \text{ bps} \approx 214 \text{ Kbps} \quad \mathcal{S}_{\text{bottleneck}} = 17792 \text{ bps} \approx 18 \text{ Kbps}$$

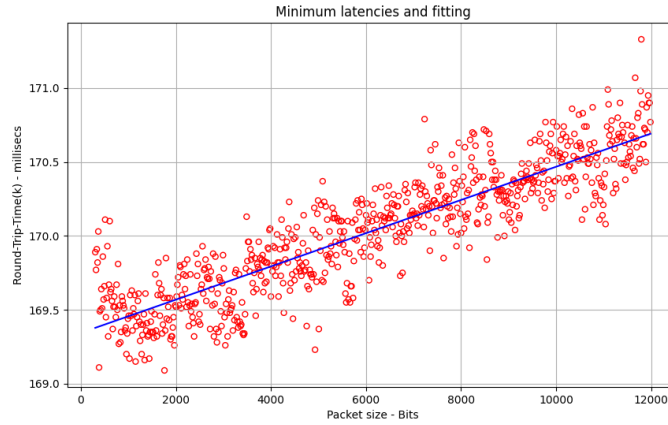


Figura 8: Regressione lineare del RTT minimo, in riferimento all'immagine 4.

Infine impostando le istanze ad un massimo di 250 e lo step minimo di 1, si ottiene che il valore del coefficiente $\alpha \approx 1.61 \cdot 10^{-4}$ è ancora maggiore (grafico 9), diminuendo ancora di più il valore del throughput:

$$\mathcal{S} = 149382 \text{ bps} \approx 149 \text{ Kbps} \quad \mathcal{S}_{\text{bottleneck}} = 12448 \text{ bps} \approx 12 \text{ Kbps}$$

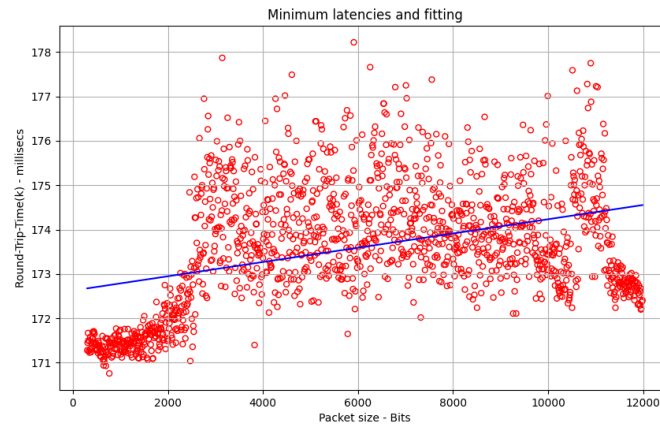


Figura 9: Regressione lineare del RTT minimo, in riferimento all'immagine 6.

Analizzando i risultati ottenuti si osserva che man mano che il numero di istanze di pacchetti inviati aumento, il valore del throughput diminuisce: questo aspetto verrà discusso ed analizzato in maniera dettagliata nel paragrafo seguente (3)

3 Conclusioni

La tabella 1 riassume i valori ottenuti analizzando le latenze ottenute inviando una elevata mole di pacchetto al server di Los Angeles. La tabella in particolare si concentra sul mostrare la netta differenza tra il throughput con link di pari prestazioni \mathcal{S} e il throughput nel caso in cui sia presente un link molto poco prestante che fa da collo di bottiglia $\mathcal{S}_{\text{bottleneck}}$. Si osserva inoltre la curva discendente dei throughput rispetto all'incrementare delle istanze K . La decrescita del throughput \mathcal{S} in virtù dell'aumento

Tabella 1: Risultati finali

K	\mathcal{S}	$\mathcal{S}_{\text{bottleneck}}$
20	319 Kbps	27 Kbps
100	214 Kbps	18 Kbps
250	149 Kbps	12 Kbps

del numero delle istanze può essere dovuta a diversi motivi. Il primo è relativo a una possibile generazione di **congestione** dovuta all'elevato numero di pacchetti inviati in un breve periodo di tempo. In particolare, prendendo come riferimento $K = 250$ — osservando le figure 6, 9 — si nota che quando l'invio dei pacchetti è a regime, ovvero nel momento in cui tutti i thread sono in esecuzione e stanno invocando il comando `psping`, sia il ritardo minimo che la latenza media aumentano drasticamente se comparati con la fase iniziale e finale. Si osserva che questa drastica crescita nel momento in cui il codice è a regime non accade nel momento in cui $K = 20$ o $K = 100$, come mostrato nelle figure 2, 4. Ciò nonostante il valore del throughput \mathcal{S} decresce comunque tra le due casistiche dove non ci sono indizi di possibili congestioni.

Pertanto, oltre alla possibile congestione generata dall'elevato numero di pacchetti inviati, un altro fattore che contribuisce alla decrescita del throughput calcolato all'aumentare del numero di messaggi ping è di natura statistica. Infatti, inviando un numero maggiore di messaggi, si raccoglie un *dataset* più ampio, che può fornire una stima migliore del throughput medio. I risultati ottenuti con 365 250 pacchetti sono probabilmente più rappresentativi rispetto a quelli ottenuti con soli 2 940 messaggi. Inoltre, inviando un numero maggiore di ping, si minimizza l'impatto di valori anomali nei dati. Eventuali problemi occasionali di rete o ritardi del server possono avere un effetto più significativo sul calcolo del throughput quando si utilizza un campione di dimensioni ridotte.