

Un modello di ottimizzazione in IBM OPL CPLEX per la risoluzione dei problemi di network design e di multicommodity flow

Lorenzo LEONI

Università degli studi di Bergamo, Dipartimento di Ingegneria Gestionale,
dell'Informazione e della Produzione

8 giugno 2022

Sommario

Progettare una rete di telecomunicazione e individuare la migliore politica di instradamento dei traffici su di essa non è banale per via dei numerosi vincoli che devono essere presi in considerazione. L'obiettivo di questo progetto è realizzare un modello di ottimizzazione in ambiente IBM OPL CPLEX finalizzato alla risoluzione di questi problemi ispirandosi a una rete realmente esistente, la francese RENATER.

1 Introduzione

La RENATER (REseau NATional de télécommunications pour la technologie, l'Enseignement et la Recherche) è la NREN (National Research and Education Network) francese, una rete nata nel 1993 con l'obiettivo di fornire un servizio di telecomunicazione a elevata affidabilità, velocità e sicurezza tra i diversi poli di ricerca e di istruzione dislocati sul territorio nazionale. Essa all'1 gennaio 2011 era costituita da 43 nodi e da 56 canali, due dei quali a 10 Gbit/s diretti verso la GEANT, la NREN europea, la quale permette alla RENATER di connettersi con le reti internazionali. Per questo progetto il network francese è stato utilizzato come dorsale per la risoluzione dei seguenti problemi di ottimizzazione:

- **network design:** dato un insieme di client e un insieme di server, si vuole determinare presso quali nodi della dorsale installare i dispositivi di instradamento dei traffici (es. router) affinché le sorgenti possano comunicare con le destinazioni. La scelta dei nodi dev'essere fatta con l'obiettivo di minimizzare il costo totale di installazione dei router.
- **multicommodity flow:** dato un insieme di traffici dai client verso i server, si vuole determinare il cammino di costo minimo per ogni flusso, ossia su quali canali di trasmissione ogni traffico dev'essere instradato affinché venga minimizzato il costo totale di accesso, di attraversamento e di uscita dalla dorsale.

1.1 Risorse e toolchain

Le informazioni relative alla RENATER sono state recuperate dal sito Topology Zoo, un portale a cura dell'università di Adelaide che raccoglie dati inerenti numerose topologie di reti provenienti da tutto il mondo. I dati vengono distribuiti nei formati GML e GraphML. Per il

parsing, la creazione delle strutture dati e l'elaborazione dei risultati numerici è stato utilizzato l'ambiente MATLAB, mentre per la realizzazione ed esecuzione del modello di ottimizzazione, come da titolo, IBM OPL Optimization Studio in abbinata al risolutore CPLEX.

2 Descrizione della rete

2.1 Caratteristiche topologiche

Di seguito sono riportate le caratteristiche topologiche della rete:

- il numero di siti candidati all'installazione di un router è pari a 43. Essi coincidono esattamente con i nodi della dorsale, ossia la francese RENATER (figura 1);
- il numero di client (o sorgenti) è pari a 6. Essi non esistono realmente, ma sono delle entità che sono state create ad hoc per immettere traffico nella dorsale;
- il numero di server (o destinazioni) è pari a 6. Anch'essi non esistono nella realtà, ma sono dei nodi fittizi che sono stati creati ad hoc per ricevere il traffico immesso dai client nella dorsale;
- il numero di canali di trasmissione tra i nodi della backbone è pari a 56. Essi coincidono esattamente con i canali della RENATER e possono essere percorsi in entrambe le direzioni. Ogni canale di trasmissione viene modellato tramite una coppia di archi.
- il numero di canali di accesso alla dorsale è pari a 12, un paio per ogni client. Ogni sorgente non può collegarsi a entrambi contemporaneamente;
- il numero di canali di uscita dalla dorsale è pari a 6, uno per ogni server. Tale scelta è dettata dal fatto che i siti candidati sono tra loro geograficamente molto distanti, quindi risulterebbe dispendioso economicamente collegare la destinazione a più punti della dorsale. Tale considerazione, tuttavia, si potrebbe fare anche per i canali di accesso, ma la necessità di simulare la scelta univoca dell'access point per ogni client comporta un rilassamento dell'assunzione di notevole distanza geografica.

2.2 Specifiche tecniche

Si precisa che i valori di seguito riportati non hanno una diretta correlazione con la realtà, ma sono stati scelti tramite delle assunzioni:

- la capacità dei canali di trasmissione tra i nodi della backbone è pari a 155 Mbit/s. Questa informazione è direttamente fornita dal file .gml messo a disposizione da Topology Zoo;
- la capacità dei canali di uscita è pari a 200 Mbit/s. Si ipotizza che i canali di uscita siano più prestazionali rispetto a quelli della dorsale in modo tale che possano gestire l'ingente mole di traffico che va a concentrarsi presso il singolo router di uscita, uno per ogni server.
- il costo di percorrenza di un canale di accesso è compreso tra $25 \text{ u}^1/\text{Mbit s}^{-1}$ e $30 \text{ u}^1/\text{Mbit s}^{-1}$;
- il costo di percorrenza di un canale di trasmissione tra una coppia di nodi della dorsale è compreso tra $15 \text{ u}^1/\text{Mbit s}^{-1}$ e $20 \text{ u}^1/\text{Mbit s}^{-1}$;
- il costo di percorrenza di un canale di uscita è compreso tra $5 \text{ u}^1/\text{Mbit s}^{-1}$ e $10 \text{ u}^1/\text{Mbit s}^{-1}$;

¹unità di costo.

- la capacità dei router è compresa tra 175 Mbit/s e 200 Mbit/s;
- il costo di installazione di un router è compreso tra 5000 u e 7600 u.

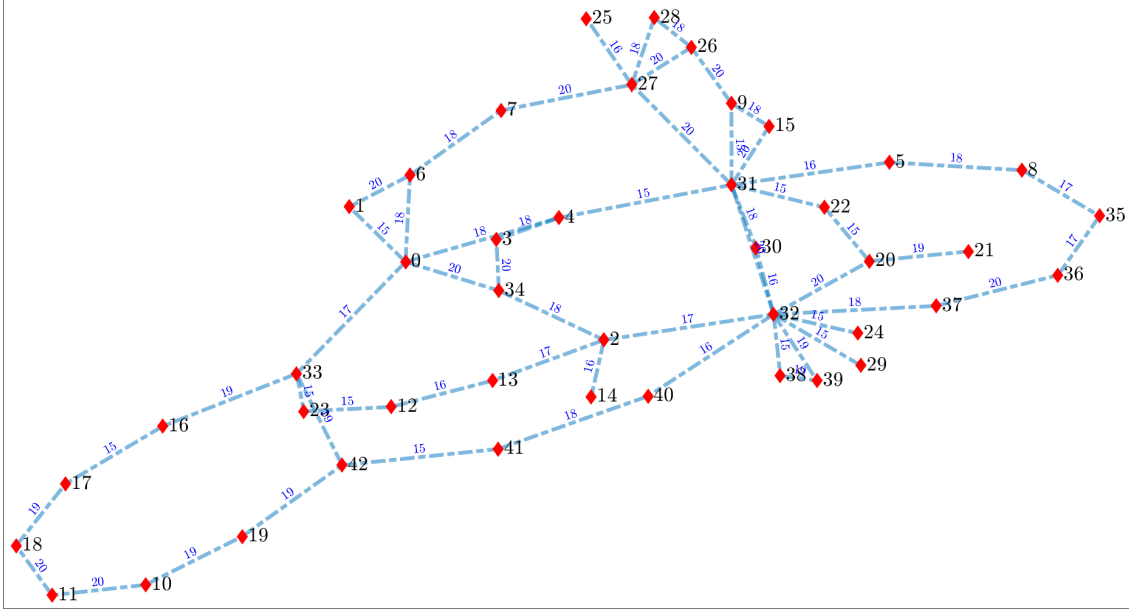


Figura 1: grafo della RENATER. Tale rappresentazione non coincide con la topologia reale, ma i nodi sono stati collocati nello spazio in modo tale da rendere il grafo piacevole alla vista. In blu il costo di attraversamento dei canali, i quali sono tratteggiati poiché la loro abilitazione avviene nell'istante in cui presso entrambi i nodi alle proprie estremità è installato un router.

3 Modello di ottimizzazione

Di seguito è riportato il modello matematico relativo ai problemi di network design e multicommodity flow. Si vuole che l'instradamento venga svolto in modo tale che i traffici non vengano suddivisi, ossia flussi di dati diretti verso la stessa destinazione non possono essere indirizzati da un router su più canali diversi. In questo caso si parla di flussi *unsplittable*, un requisito che richiede un paio di condizioni aggiuntive rispetto al modello base.

3.1 Insiemi

- *CSs*: insieme dei siti candidati (Candidate Sites) all'installazione dei router;
- *TPs*: insieme dei client (Test Points);
- *DNs*: insieme dei server (Destination Nodes);
- *INs*: insieme dei canali di accesso (INputs) alla dorsale;
- *OUTs*: insieme dei canali di uscita (OUTputs) dalla backbone;
- *TRASFs*: insieme dei canali di trasferimento (TRASFers) all'interno della dorsale.

3.2 Dati

- $Cap^R[CSs]$: vettore contenente le capacità dei router, una per ogni CS ;
- $Costo^R[CSs]$: vettore contenente i costi di installazione di un router, uno per ogni CS ;
- $Sorg^I[INs]$: vettore contenente i nodi di partenza, uno per ogni canale IN ;
- $Dest^I[INs]$: vettore contenente i nodi di arrivo (access point), uno per ogni canale IN ;
- $Costo^I[INs]$: vettore contenente i costi di attraversamento, uno per ogni canale IN ;
- $Sorg^O[OUTs]$: vettore contenente i nodi di partenza, uno per ogni canale OUT ;
- $Dest^O[OUTs]$: vettore contenente i nodi di arrivo, uno per ogni canale OUT ;
- $Cap^O[OUTs]$: vettore contenente le capacità dei canali OUT ;
- $Costo^O[OUTs]$: vettore contenente i costi di attraversamento, uno per ogni canale OUT ;
- $Sorg^T[TRASFs]$: vettore contenente i nodi di partenza, uno per ogni canale $TRASF$;
- $Dest^T[TRASFs]$: vettore contenente i nodi di arrivo, uno per ogni canale $TRASF$;
- $Cap^T[TRASFs]$: vettore contenente le capacità dei canali $TRASF$;
- $Costo^T[TRASFs]$: vettore contenente i costi di attraversamento, uno per ogni canale $TRASF$;
- $Traf[INs][OUTs]$: matrice dei traffici, uno per ogni coppia client-server.

3.3 Variabili decisionali

- $x[INs]$: variabile binaria che indica se un TP è stato assegnato o meno a uno specifico CS ;
- $z[CSs]$: variabile binaria che indica se un router è stato installato o meno presso uno specifico CS ;
- $w[OUTs]$: variabile binaria che indica se un CS è connesso o meno a uno specifico DN ;
- $y[TRASFs][DNs]$: variabile binaria che indica se attraverso un canale di trasferimento $TRASF$ scorre o meno un traffico diretto verso uno specifico DN ;
- $f^{d1}[TRASFs][DNs]$: variabile reale positiva che rappresenta il traffico instradato su un canale di trasferimento $TRASF$ diretto verso uno specifico DN ;
- $f^{d2}[OUTs]$: variabile reale positiva che rappresenta il traffico instradato da un CS verso uno specifico DN .

3.4 Funzione obiettivo

Da *minimizzare* la somma delle seguenti quattro componenti:

1. il costo totale di installazione dei router, ossia:

$$\sum_{j \in CSs} Costo_j^R \cdot z_j;$$

2. il costo totale dovuto al trasferimento dei traffici all'interno della dorsale, ossia:

$$\sum_{a \in TRASFs} \sum_{k \in DN_s} Costo_a^T \cdot f_{a,k}^{d1};$$

3. il costo costo totale dovuto all'accesso dei *TP* ai rispettivi access point, ossia:

$$\sum_{a \in INs} \sum_{k \in DN_s} Costo_a^I \cdot Traf_{Sorg_a^I, k} \cdot x_a;$$

4. costo totale dovuto alla trasmissione dei traffici verso i rispettivi *DN*, ossia:

$$\sum_{a \in OUTs} Costo_a^O \cdot f_a^{d2} \cdot w_a.$$

3.5 Vincoli

Da rispettare la seguente serie di vincoli:

1. vincoli di univocità della connessione di accesso alla dorsale, ossia:

$$\sum_{a \in INs | Sorg_a^I = i} x_a = 1, \forall i \in TP_s;$$

2. vincoli di coerenza relativi all'accesso alla backbone, ossia:

$$x_a \leq z_{Dest_a^I}, \forall a \in INs;$$

3. vincoli di coerenza relativi all'uscita dalla dorsale, ossia:

$$w_a \leq z_{Sorg_a^O}, \forall a \in OUTs;$$

4. vincoli di conservazione dei traffici ai nodi della backbone, ossia:

$$\begin{aligned} & \sum_{a \in INs | Dest_a^I = j} Traf_{Sorg_a^I} \cdot x_a + \sum_{a \in TRASFs | Dest_a^T = j} f_{a,k}^{d1} - \sum_{a \in TRASFs | Sorg_a^T = j} f_{a,k}^{d1} - \\ & - \sum_{a \in OUTs | Sorg_a^O = j, Dest_a^O = k} f_a^{d2} = 0, \forall j \in CSs, k \in DN_s \end{aligned}$$

5. vincoli di capacità relativi ai canali di trasferimento, ossia:

$$\sum_{k \in DN_s} f_{a,k}^{d1} \leq Cap_a^T \cdot z_{Sorg_a^T}, \forall a \in TRASFs,$$

$$\sum_{k \in DN_s} f_{a,k}^{d1} \leq Cap_a^T \cdot z_{Dest_a^T}, \forall a \in TRASFs.$$

Da notare che essi vengono abilitati soltanto se a entrambe le estremità dell'*a*-esimo *TRASF* è installato un router;

6. vincoli di capacità relativi agli access point, ossia:

$$\sum_{a \in INs | Dest_a^I = j} \sum_{k \in DNs} Traf_{Sorg_a^I, k} \cdot x_a \leq Cap_j^R, \forall j \in CSs;$$

7. vincoli di capacità relativi ai canali di uscita dalla dorsale, ossia:

$$f_a^{d2} \leq Cap_a^O \cdot w_a, \forall a \in OUTs;$$

Da notare che essi vengono abilitati soltanto se l' a -esimo OUT è attivo, ossia è percorso da un traffico;

8. **vincoli di coerenza relativi alla non suddivisione dei traffici**, ossia:

$$f_{a,k}^{d1} \leq Cap_a^T \cdot y_{a,k}, \forall a \in TRASFs, k \in DNs,$$

$$y_{a,k} \leq f_{a,k}^{d1};$$

Il fatto che l' a -esimo $TRASF$ venga percorso da un traffico diretto verso il k -esimo DN comporta l'abilitazione del vincolo, ossia $y_{a,k} = 1$. La seconda condizione, invece, impone $y_{a,k} = 0$ se l' a -esimo $TRASF$ non fosse percorso da un traffico diretto verso il k -esimo DN ;

9. **vincoli di non suddivisione dei traffici**, ossia:

$$\sum_{a \in TRASFs | Sorg_a^T = j} y_{a,k} \leq 1, \forall j \in CSs, k \in DNs.$$

Un traffico diretto verso uno specifico DN dev'essere instradato da ogni router installato presso un determinato CS verso uno e un solo canale di trasferimento.

4 Discussione sui risultati

I problemi di network design e di multicommodity flow sono stati risolti tramite il modello di ottimizzazione precedentemente descritto per un numero k di traffici (o connessioni) pari a 2, 5, 10, 15 e 20. Per ogni configurazione di flusso è stato provato il modello sia con sia senza suddivisione dei traffici in modo tale da poter individuare le differenze prestazionali tra le due politiche di instradamento.

4.1 Esempio con $k = 2$

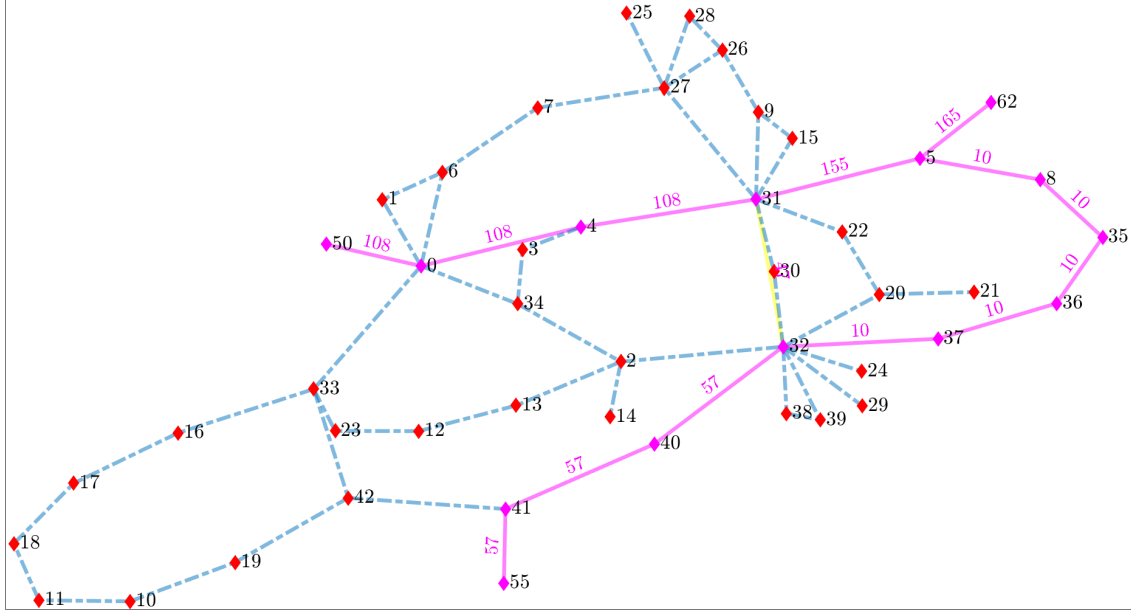
In figura 2 è possibile osservare come i flussi vengono instradati sulla RENATER a seconda del fatto che i traffici diretti verso la stessa destinazione possano essere o meno suddivisi dai router della backbone. Se i flussi sono *splittable* (figura 2(a)), allora 47 dei 57 Mbit/s di traffico dal client 55 al server 62 vengono indirizzati verso il canale (31, 5) attraverso l'arco (32, 31) per minimizzare il costo totale di attraversamento della dorsale. Ovviamente non possono essere indirizzati più di 57 Mbit/s poiché la capacità dei canali di trasferimento della RENATER è pari a 155 Mbit/s. Il costo totale vale:

$$C_{tot}^{split} = c_{router}^{split} + c_{trasf}^{split} + c_{in}^{split} + c_{out}^{split} = 81587 + 9822 + 4182 + 990 = 96581.$$

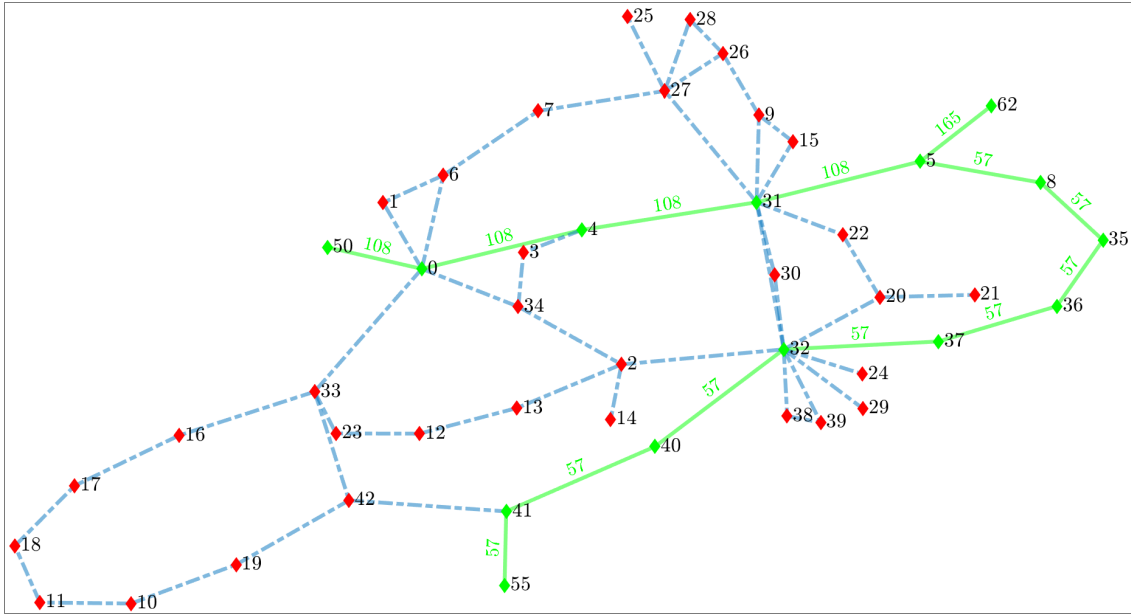
Se i flussi sono *unsplittable*, invece, (figura 2(b)), il traffico $50 \rightarrow 62$ non può essere suddiviso dal router installato presso il nodo 32, quindi tutti i 57 Mbit/s percorrono lo stesso cammino per arrivare a destinazione. Ciò comporta un aumento del costo totale:

$$C_{tot}^{unsplit} = c_{router}^{unsplit} + c_{trasf}^{unsplit} + c_{in}^{unsplit} + c_{out}^{unsplit} = 81587 + 12360 + 4182 + 990 = 99119.$$

Confrontando i valori della variabile obiettivo C_{tot} ottenuti si può notare che la differenza risiede nel costo totale di attraversamento della dorsale c_{trasf} . Infatti, se i traffici non possono essere suddivisi, allora tutti i 57 Mbit/s del flusso $55 \rightarrow 62$ devono percorrere il cammino più lungo e costoso. Da sottolineare, infine, che il costo totale di installazione dei router c_{router} tiene conto anche delle connessioni tra la backbone e i rimanenti quattro client che, nonostante non generino traffico in questo contesto, devono essere connessi a uno e a un solo access point.



(a)



(b)

Figura 2: instradamento di due traffici ($50 \rightarrow 62$ di valore 108 Mbit/s e $55 \rightarrow 62$ di valore 57 Mbit/s) sulla dorsale con (a) e senza (b) suddivisione dei flussi ai nodi. In giallo il canale e il relativo traffico di valore 47 Mbit/s che differenzia le due politiche di instradamento.

4.2 Confronti numerici

Tramite l'analisi dei risultati numerici ottenuti (tabella 1) è possibile osservare che la suddivisione dei flussi determina un aumento delle risorse di rete utilizzate per instradare i k traffici, sia in termini di utilizzo medio dei canali di trasferimento attivi $\bar{\mu}_{canali}$ sia di risorse complessive $\mu_{backbone}$. Ciò è dovuto all'incremento del numero di canali della dorsale che devono sostenere dei traffici elevati a causa dell'impossibilità di suddividere i flussi, ossia di indirizzare una parte dei traffici sui cammini di costo minimo al fine limitare l'utilizzo di risorse a una sola parte della backbone. Il numero medio di canali di trasferimento \bar{n} che un flusso deve percorrere per giungere a un server, invece, tende in generale a diminuire con la politica di instradamento suddiviso poiché, fissata una destinazione, viene meno la percorrenza di quegli archi necessari per ricongiungersi ai cammini di costo minimo (es. arco (32, 31) dell'esempio con $k = 2$). Tuttavia, bisogna altresì mettere in evidenza che \bar{n} dipende anche da come vengono scelti i traffici: se nella dorsale viene immesso un numero notevole di flussi destinati a server poco distanti dai rispettivi client, allora mediamente i cammini risulteranno brevi. Da notare, inoltre, che il numero di router installati rimane pressoché indipendente dalla politica impiegata; anch'esso dipende in parte, come per \bar{n} , dalla scelta dei flussi. Infine, il costo totale C_{tot} aumenta notevolmente senza splitting perché alcuni traffici devono percorrere per intero i cammini più lunghi e costosi.

k	<i>Instradamento</i>	$\bar{\mu}_{canali}$	$\mu_{backbone}$	\bar{n}	<i>Num. router installati</i>	C_{tot} (<i>ottimo</i>)
2	Splitting	34.13 %	3.35 %	5,5	13	96581 u
2	Unsplitting	46.65 %	4.16 %	5	13	99119 u
5	Splitting	35.76 %	6.71 %	4,67	16	130838 u
5	Unsplitting	40.09 %	7.52 %	4,5	16	133376 u
10	Splitting	44.65 %	11.56 %	3,75	20	179213 u
10	Unsplitting	48.82 %	13.51 %	3,92	20	183306 u
15	Splitting	63.71 %	18.20 %	3,78	22	220281 u
15	Unsplitting	66.79 %	19.08 %	3,44	23	226791 u
20	Splitting	63.35 %	19.80 %	3,39	25	240880 u
20	Unsplitting	67.93 %	20.62 %	2,93	25	243442 u

Tabella 1: confronto tra indici che descrivono le prestazioni del modello di ottimizzazione applicato alla dorsale al variare del numero k di traffici e della politica di instradamento. $\bar{\mu}_{canali}$ rappresenta l'utilizzo medio dei canali attivi, ossia quelli che effettivamente trasportano flusso, $\mu_{backbone}$ l'utilizzo delle risorse della RENATER e \bar{n} il numero medio di canali di trasferimento che un traffico deve percorrere per giungere a destinazione.

5 Conclusioni

Alla luce dei risultati ottenuti è evidente come la suddivisione dei traffici sia necessaria per ottimizzare l'utilizzo delle risorse della dorsale e minimizzare il costo totale di instradamento e di installazione dei router, tuttavia lo splitting non è semplice da implementare a livello software e non tutti i protocolli di comunicazione lo prevedono, quindi è bene risolvere i problemi di network design e di multicommodity flow anche non considerando questa politica di instradamento.