INAF
ISTITUTO NAZIONALE
DI ASTROFISICA
NATIONAL INSTITUTE
FOR ASTROPHYSICS

Rapporti Tecnici INAF
INAF Technical Reports

Number	345
Publication Year	2025-10-15
Acceptance in OA@INAF	2025-11-21T12:29:32Z
Title	Rapporto tecnico sui vantaggi del Traffic Shaping e la sua implementazione presso OAPd
Authors	PETRELLA, Amedeo, SELVESTREL, DANILO
Publisher's version (DOI)	https://doi.org/10.20371/INAF/TechRep/345
Handle	http://hdl.handle.net/20.500.12386/45171

Rapporto tecnico sui vantaggi del Traffic Shaping e la sua implementazione presso OAPd

Amedeo Petrella , Danilo Selvestrel

Introduzione

Questo rapporto tecnico illustra i vantaggi derivanti dall'implementazione della tecnica del *traffic shaping* nelle reti informatiche. Il *traffic shaping* è un metodo di gestione della larghezza di banda che consente di regolare il flusso dei dati in rete, al fine di ottimizzare le prestazioni, garantire la qualità del servizio (QoS) per applicazioni critiche e prevenire la congestione.
L'adozione di questa tecnologia è divenuta indispensabile per l'Osservatorio Astronomico di Padova (OAPd), poiché le crescenti esigenze di banda, derivanti dai nuovi progetti internazionali, stanno portando a un utilizzo quasi continuo delle nostre risorse di rete.

Principi del Traffic Shaping

Il *traffic shaping* opera ritardando alcuni o tutti i pacchetti di dati in modo da conformare il flusso del traffico a un profilo prestabilito. Ciò può includere la limitazione della velocità di trasmissione, la prioritizzazione di specifici tipi di traffico o la garanzia di una larghezza di banda minima per determinate applicazioni. Le tecniche comuni includono:

- **Token Bucket:** utilizzato per regolare la velocità media e limitare i picchi di traffico.
- **Leaky Bucket:** controlla la velocità di uscita dei pacchetti da un buffer.
- **Class-Based Queuing (CBQ):** permette di assegnare diverse priorità a diverse classi di traffico.

Vantaggi chiave del Traffic Shaping

1. Miglioramento della Qualità del Servizio (QoS)

Il *traffic shaping* è essenziale per ottimizzare QoS per applicazioni sensibili alla latenza e alla larghezza di banda. Applicazioni come il VoIP, lo streaming video, le videoconferenze e le

sessioni interattive (SSH o VPN) beneficiano notevolmente della prioritizzazione del traffico. Questo approccio riduce ritardi, *jitter* e perdita di pacchetti, migliorando significativamente l'esperienza utente.

2. Prevenzione della Congestione della Rete

Limitando la velocità di trasmissione di traffico non essenziale o a bassa priorità, il *traffic shaping* impedisce che la rete venga sovraccaricata. Questo è particolarmente utile in scenari con picchi di traffico o in reti con larghezza di banda limitata, come quella presso OAPd che è attualmente di 1 Gbps, dove la congestione potrebbe altrimenti causare rallentamenti significativi per tutti gli utenti.
In ogni caso, anche in presenza di una banda adeguata, un approccio basato sul traffic shaping si rivela fondamentale per ottimizzare la gestione delle risorse di rete. Questa strategia consente di affrontare in modo proattivo e controllato le varie situazioni di congestione che potrebbero verificarsi, garantendo la fluidità del traffico e la priorità dei servizi critici.

3. Utilizzo Ottimizzato delle Risorse di Rete

Attraverso una gestione più efficiente della larghezza di banda disponibile, il *traffic shaping* assicura che le risorse di rete siano utilizzate in modo ottimale. Questo significa che le applicazioni e i servizi più importanti riceveranno la larghezza di banda necessaria, mentre il traffico meno critico sarà gestito in modo da non compromettere le prestazioni generali.

4. Equità nell'Allocazione della Larghezza di Banda

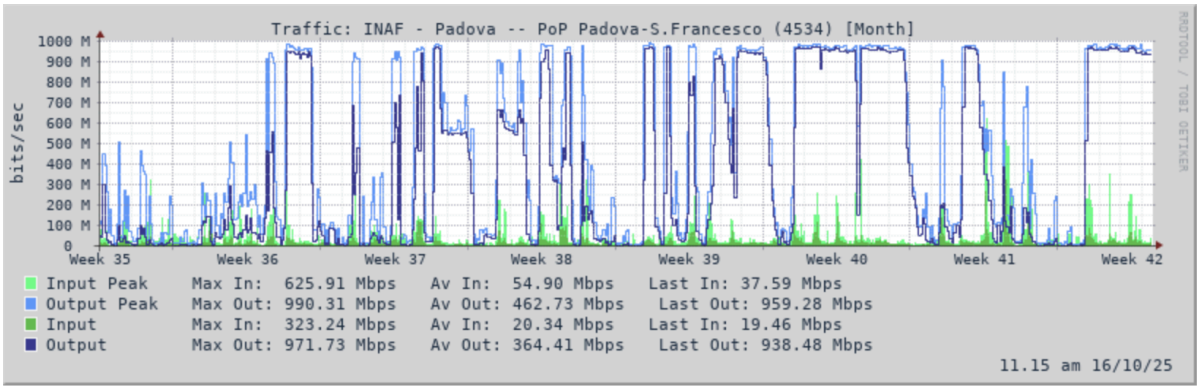
Nelle reti multi-utente, il *traffic shaping* può garantire che nessun singolo utente o applicazione monopolizzi l'intera larghezza di banda disponibile. Questo promuove un'allocazione equa delle risorse, assicurando che tutti gli utenti possano accedere ai servizi di rete in modo efficiente. In particolare, per quanto concerne la rete OAPd, l'adozione di questa tecnica si è resa necessaria a causa di un singolo utente che doveva scaricare diversi TB di dati da Amazon AWS.

Implementazione e Considerazioni

L'implementazione del *traffic shaping* richiede un'attenta pianificazione e configurazione, spesso attraverso dispositivi di rete come router, firewall o *traffic shaper* dedicati. È essenziale analizzare i modelli di traffico esistenti e definire policy chiare per l'allocazione della larghezza di banda. Una configurazione errata può portare a effetti indesiderati, come una limitazione eccessiva del traffico o la compromissione delle prestazioni. Strumenti di rete avanzati permettono un'analisi dettagliata del traffico anche all'interno di protocolli come, ad esempio, HTTPS, che è impiegato sia per la navigazione web che per i download. A questo scopo, presso l’Osservatorio Astronomico di Padova, abbiamo utilizzato il nostro firewall di frontiera, un Fortigate 200F.

Situazione presso OAPd

La seguente immagine evidenzia chiaramente le motivazioni sopra esposte



GINS (GARR Integrated Networking Suite)

Alla luce delle frequenti e notevoli latenze riscontrate quotidianamente, abbiamo deciso quindi di implementare con urgenza il traffic shaping.

Traffic shaping dei firewall Fortigate

Tipologie

FortiGate supporta due approcci fondamentali per il controllo del traffico:

1. **Traffic Shaping con Policing**: quando il traffico supera i limiti di banda configurati, i pacchetti in eccesso vengono scartati (dropped). Il Policing non utilizza code (queues).
2. **Traffic Shaping con Queuing**: quando il traffico supera i limiti di banda configurati, i pacchetti vengono ritardati per il trasporto fino a quando la banda non si libera. Se le code si riempiono, i pacchetti possono comunque essere scartati. Il Queuing utilizza algoritmi come RED o FIFO per l'accodamento e l'algoritmo HTB per il de-accodamento.

Metodi

Il FortiGate offre tre metodi principali per configurare il traffic shaping, che hanno capacità e preferenze diverse:

1. **Profili di Traffic Shaping** (Traffic Shaping Profiles)
2. **Shaper di Traffico** (Traffic Shapers)
3. **Prioritizzazione Globale del Traffico** (Global Traffic Prioritization)

Se tutti e tre i metodi sono configurati, il Profilo di Traffic Shaping ha la preferenza maggiore, seguito dallo Shaper di Traffico, e infine dalla Prioritizzazione Globale.

1. Profili di Traffic Shaping (Traffic Shaping Profiles)

Un profilo di traffic shaping consente la configurazione del traffic shaping utilizzando sia il policing che il queuing. Questo metodo è tipicamente applicato a livello di interfaccia in uscita (egress shaping profile) e basa i suoi limiti su percentuali.

Caratteristiche principali:

- **Classificazione del Traffico**: è possibile definire fino a 30 classi di traffico (ID di classe da 2 a 31), ognuna con i propri limiti e priorità. Per i processori NP6, NP6Lite (SoC3) o NP6Xlite (SoC4), il limite di ID di classe per il traffico in uscita è 2-15.
- **Limiti di Banda basati su percentuali**: la banda garantita (guaranteed bandwidth) e la banda massima (maximum bandwidth) sono configurate come una percentuale della banda in uscita dell'interfaccia (outbandwidth).
    - La somma di tutta la banda garantita di tutte le classi non può superare il 100% dell'outbandwidth. La banda garantita viene sempre rispettata.
    - La banda massima definisce la percentuale massima di outbandwidth utilizzabile da una classe.
- **Prioritizzazione**: il traffico può essere collocato in cinque livelli di priorità: top, critical, high, medium, low.
    - Nota: I processori NP7 e NP7Lite (SOC5) ignorano l'opzione di priorità nel profilo di shaping.
- **Queuing Avanzato**: quando il queuing è abilitato, è possibile configurare opzioni aggiuntive per classe, come il Weighted Random Early Detection (WRED) e il controllo del burst.

2. Shaper di Traffico (Traffic Shapers)

I *Traffic Shapers* sono un metodo per il **policing del traffico**, un approccio allo shaping in cui il traffico che eccede i limiti di banda configurati viene **scartato (dropped)**.

Questi shapers vengono applicati al traffico desiderato attraverso l'uso delle **politiche di shaping del traffico** (*traffic shaping policies*). Le politiche di shaping stabiliscono le regole che associano il traffico specifico a un determinato *shaper* o a una *classe*.
È fondamentale notare che i limiti di banda (massima e garantita) per i Traffic Shapers sono definiti come un tasso (rate), misurato in unità come Kbps o Mbps. Questa configurazione si distingue dai Traffic Shaping Profiles, i quali definiscono i limiti di banda in termini di percentuale della banda di uscita (outbandwidth).

Tipi di Shaper di Traffico:

A. Shaper Condiviso (Shared Traffic Shaper)

- **Utilizzato** per indicare priorità e limiti di banda garantita e massima **per un tipo specifico di traffico**.

- **Limiti di Banda basati su percentuale**: banda garantita e banda massima sono configurate come un tasso specifico (Kbps, Mbps, ecc.).
    - Il FortiGate non impone un limite rigido sulla somma della banda garantita, per cui gli amministratori devono assicurarsi di non superare la banda totale in uscita dell'interfaccia.
- **Prioritizzazione**: classi di priorità high (2), medium (3) o low (4). Il traffico al di sotto della banda garantita ottiene automaticamente il livello critical (1).
- **Applicazione**: può essere configurato per applicare lo shaping per ogni singola politica (per-policy) o per tutte le politiche che utilizzano lo shaper (all policies), che condividono la banda.
- **Direzione**: uno shaper condiviso si applica al traffico in direzione forward (upload); uno shaper inverso (reverse shaper) può essere definito per il traffico in direzione egress-to-ingress (download).
- **Multi-Stage Marking**: supporta la marcatura dinamica DSCP e VLAN CoS basata su diversi livelli di velocità del traffico.

B. Shaper Per-IP (Per-IP Traffic Shaper)

- **Utilizzato per limitare** il comportamento di ogni **singolo indirizzo IP**, prevenendo che un utente consumi tutta la banda disponibile.
- **Limiti di Banda**: applica il limite massimo di banda configurato (su tasso/rate) a ciascun IP individuale.
- **Limiti di Sessione**: consente anche di definire il numero massimo di sessioni concorrenti per un indirizzo IP.
- **Direzione**: a differenza degli shaper condivisi, gli shaper Per-IP applicano il limite di velocità sia alle operazioni di upload che di download.

3. Prioritizzazione Globale del Traffico (Global Traffic Prioritization)

Questo è l'approccio più semplice, focalizzato solo sulla prioritizzazione.

Caratteristiche principali:

- **Meccanismo**: supporta solo il policing.
- **Prioritizzazione basata su ToS/DSCP**: Il traffico viene classificato come high (2), medium (3) o low (4) in base al valore del campo ToS (Type of Service) o DSCP (Differentiated Services Code Point) nell'header IP.
- **Limiti di Banda**: non supporta la configurazione di limiti di banda garantita o massima.
- **Prerequisito**: affinché la prioritizzazione globale abbia effetto, è necessario definire l'outbandwidth sull'interfaccia, anche se non viene applicato alcun profilo di shaping in uscita.

Individuazione del metodo

In prima istanza abbiamo pensato di applicare uno shaper per IP in modo da limitare un singolo client, ma questo approccio si è rivelato insufficiente per i seguenti motivi:

1. funziona con un singolo client. Se ce ne fossero altri (come poi è successo) si dovrebbe applicare lo shaper a ciascun IP
2. il limite di banda è assoluto e quindi non si adatta dinamicamente alla reale occupazione di banda

Abbiamo quindi valutato l'utilizzo di una Shaper condiviso, ma anche qui il limite assoluto sulla banda si è dimostrato troppo stringente.
Infine abbiamo quindi optato per i Traffic Shaping Profiles che meglio si adattavano alle nostre esigenze.

Difficoltà

La principale sfida incontrata nell'implementazione del traffic shaping è stata la corretta comprensione dei flussi di traffico esistenti e l'identificazione dei punti ottimali per l'applicazione dei profili e delle policy associate. Questo processo ha richiesto un'analisi approfondita del comportamento della rete e delle esigenze specifiche degli utenti, al fine di garantire che il traffic shaping non limitasse in modo improprio le attività critiche, ma ottimizzasse efficacemente l'utilizzo della larghezza di banda disponibile.

Considerando che il traffico di maggiore interesse è quello di download, si è reso necessario identificare con precisione i seguenti elementi chiave:

1. **Origine del traffico:** nella stragrande maggioranza dei casi analizzati, il traffico ha origine da un client situato sulla nostra LAN interna. Questi client stabiliscono una connessione TCP con un server remoto, avviando così il processo di download. È fondamentale comprendere il numero e la tipologia di questi client, nonché i servizi o le applicazioni che generano la maggiore quantità di traffico di download, per poter applicare policy mirate ed efficaci.
2. **Interfaccia di ingresso:** l'interfaccia di ingresso per il traffico di download è stata univocamente identificata come quella WAN (Wide Area Network). All'interno della nostra architettura di rete, questa interfaccia è stata esplicitamente etichettata e denominata "Untrusted", indicando il punto di confine tra la nostra rete interna e la rete Internet esterna. Questa chiara identificazione è cruciale per la corretta configurazione delle regole di traffic shaping sui dispositivi di rete pertinenti, come firewall o router edge, dove il traffico in ingresso dalla WAN può essere ispezionato e gestito prima di raggiungere la LAN interna.
3. **Larghezza di banda:** l'OAPd (Osservatorio Astronomico di Padova) è dotato di una connessione internet in fibra ottica di tipo simmetrico, che garantisce una banda dedicata e costante di 1 Gbps (Gigabit al secondo). Sebbene la capacità teorica sia elevata, l'esperienza ha dimostrato che saturare completamente la banda disponibile può portare a un degrado delle prestazioni complessive della rete, influenzando negativamente anche il traffico non soggetto a shaping. Per mitigare questo rischio e mantenere un margine di operatività ottimale, si è deciso di applicare una soglia inferiore per il traffic shaping, fissata a 921600 Kbps (Kilobit al secondo), equivalente a circa 900 Mbps. Questa scelta consente di sfruttare la maggior parte della capacità disponibile, prevenendo al contempo situazioni di congestione e garantendo una

migliore esperienza utente per tutte le applicazioni e i servizi che utilizzano la connessione internet. La differenza tra la banda garantita e la soglia applicata offre un "cuscinetto" di capacità che può essere utilizzato per picchi di traffico imprevisti o per servizi che richiedono una latenza estremamente bassa.

Implementazione su Fortigate

Il processo si articola in quattro passi principali:

1. **Definire le classi** per i diversi tipi di traffico.
2. **Creare un Profilo di Traffic Shaping** che assegna priorità e limiti di banda a ciascuna classe.
3. **Creare delle Policy di Traffic Shaping** per classificare il traffico in entrata e assegnarlo alla classe corretta.
4. **Applicare il profilo all'interfaccia** WAN per il traffico in ingresso.

Passo 1: Creazione delle Classi di Traffico

Abbiamo creato quattro classi distinte:

- **Critico** (ID: 10): per servizi essenziali come DNS e protocolli di telefonia (SIP).
- **Interattivo** (ID: 11): per traffico interattivo come SSH, RDP, VPN e videoconferenze (Zoom, Teams, etc.).
- **Normale** (ID: 12): per la navigazione web generica (HTTP/HTTPS).
- **Basso** (ID: 13): per tutto il resto, inclusi i download pesanti.

```
config firewall traffic-class
    edit 10
        set class-name "Critico"
    next
    edit 11
        set class-name "Interattivo"
    next
    edit 12
        set class-name "Normale"
    next
    edit 13
        set class-name "Basso"
    next
end
```

Passo 2: Creazione del Profilo di Traffic Shaping

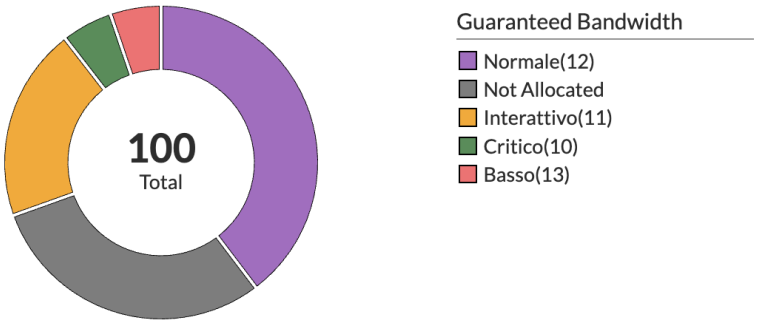
Il profilo, denominato **Download-QoS**, è così strutturato:

Traffic Shaping Classes

+ Create new | Edit | Delete | Set as default

Default	Class ID	Guaranteed Bandwidth	Maximum Bandwidth	Priority
Yes	Basso (13)	5%	80%	Low
	Critico (10)	5%	100%	Critical
	Interattivo (11)	20%	100%	High
	Normale (12)	40%	100%	Medium

Guaranteed Bandwidth Usage



Da notare che il profilo di default è quello “Basso”.

Logica di questa configurazione:

- **Guaranteed Bandwidth**: ogni classe ha una porzione di banda sempre riservata. La somma totale è 70%, ben al di sotto del 100% massimo consentito.
- **Maximum Bandwidth**: le classi prioritarie possono usare fino al 100% della banda se disponibile. La classe a bassa priorità è limitata all'80% per evitare che saturi completamente la linea anche quando non c'è altro traffico.
- **Priority**: quando la banda totale richiesta supera quella disponibile (congestione), dopo aver soddisfatto la banda garantita, la banda rimanente viene distribuita partendo dalle classi con priorità più alta (critical, poi high, etc.).

Passo 3: Creazione delle Policy di Traffic Shaping

Ora dobbiamo dire al FortiGate come classificare il traffico. Abbiamo quindi creato delle policy che associano specifici servizi o applicazioni alle classi definite.

La criticità di questo passo è quella descritta sopra: definire i corretti flussi del traffico. Inoltre è fondamentale ordinare le policy da quella più specifica a quella più generica. Qui di seguito le varie policy secondo l'ordine di applicazione.

```
config firewall shaping-policy
    edit 1
        set uuid a5f8a71a-a9ae-51f0-ca24-5a7cee62d0d4
        set name "Traffico Critico"
        set service "DNS" "SIP"
        set dstintf "Untrusted"
        set class-id 10
        set srcaddr "all"
        set dstaddr "all"
    next
end

config firewall shaping-policy
    edit 2
        set uuid 29506094-a9af-51f0-fbd9-68aaa23f5677
        set name "Traffico Interattivo"
        set service "SSH" "custom-ssh" "RDP" "IKE"
          set application 43541 16354 47385 48983 16350 27948
17678 43273
        set dstintf "Untrusted"
        set class-id 11
        set srcaddr "all"
        set dstaddr "all"
    next
end

config firewall shaping-policy
    edit 5
        set uuid a82ed77c-a9c0-51f0-f5ca-313afe957109
        set name "QoS-AWS-Bassa-Priorita"
        set internet-service enable
        set internet-service-name "Amazon-AWS"
        set dstintf "Untrusted"
        set class-id 13
        set srcaddr "all"
    next
end
```

```
config firewall shaping-policy
    edit 3
        set uuid 6cd7726c-a9af-51f0-43d3-627cdf091d76
        set name "Traffico Web Normale"
        set service "HTTP" "HTTPS" "quic"
        set dstintf "Untrusted"
        set class-id 12
        set srcaddr "all"
        set dstaddr "all"
    next
end

config firewall shaping-policy
    edit 4
        set uuid c61e56b0-a9b4-51f0-86c5-31b7ad621d27
        set name "Catch-All"
        set service "ALL"
        set dstintf "Untrusted"
        set class-id 13
        set srcaddr "all"
        set dstaddr "all"
    next
end
```

Le policy condividono una struttura simile: il campo "srcaddr", indirizzo sorgente, è impostato su "all" per tutte e cinque, mentre il campo di destinazione, "dstaddr", è "all" tranne che per la policy "QoS-AWS-Bassa-Priorita". In quest'ultimo caso, gli indirizzi di destinazione sono definiti tramite "internet-service-name Amazon-AWS", ovvero tutti gli indirizzi Amazon come specificato nell'Internet Service Database (ISDB) di FortiGate. L'ISDB è un database costantemente aggiornato da Fortinet, contenente tutti gli indirizzi IP pubblici utilizzati dai principali servizi Internet, inclusa l'ampia infrastruttura di Amazon AWS.

Un altro elemento comune a tutte le policy è l'interfaccia di destinazione, "dstintf", che è l'interfaccia "Untrusted". L'interfaccia sorgente non è specificata, poiché il traffico soggetto alle policy ha origine da qualsiasi interfaccia ma è diretto verso la WAN (Untrusted).

Analisi delle policy

- **Traffico Critico**: questa policy ha class-id=10 quindi quella prima specificata come “Critico”. Il traffico critico è quindi quello DNS e SIP (VoIP).

- **Traffico Interattivo**: questa policy ha class-id=11, cioè quella definita per il traffico “Interattivo”. Oltre ai protocolli "SSH" "RDP" "IKE" (IPSec), sono compresi alcuni di videoconferenza (Teams, Meet, Zoom, ...)
- **QoS-AWS-Bassa-Priorità**: è stata implementata per gestire il traffico che attualmente satura la banda internet dell'OAPd e ad essa è applicata la class-id 13, quella per gestire il traffico a priorità più bassa. Questa policy precede la policy "Traffico Web Normale" per intercettare il traffico AWS, che si manifesta come normale traffico HTTPS.
- **Traffico Web Normale**: tutto il traffico internet HTTP(S) e QUIC (HTTPS su UDP). La class-id è la 12, per il traffico “Normale”.
- **Catch-All**: tutto il resto del traffico. Bassa priorità per tutto il traffico non intercettato dalle precedenti policy.

Passo 4: Applicazione alla Porta WAN

L'ultimo passo è applicare il profilo "Download-QoS" all'interfaccia WAN per il traffico in download. Ogni porta di rete dei firewall Fortigate prevede due direzioni di flusso: *ingress* e *egress*. In questo caso, essendo traffico che da internet entra nella rete OAPd, è un traffico *ingress*.
Questi i principali parametri di configurazione:

```
config system interface
  edit "port20"
    ...
    set inbandwidth 921600
    set ingress-shaping-profile "Download-QoS"
    ...
  next
end
```

- **inbandwidth** definisce la larghezza di banda in ingresso (fissata, come sopra specificato, a 900 Mbps),
- **ingress-shaping-profile** applica, in ingresso, alla porta di rete il profilo "Download-QoS"

Conclusione

L'implementazione del *traffic shaping* presso l'Osservatorio Astronomico di Padova, attraverso l'adozione dei Profili di Traffic Shaping sui firewall FortiGate, ha permesso di affrontare efficacemente le sfide legate alla congestione della rete e alla gestione della larghezza di banda. Le diverse classi di traffico (Critico, Interattivo, Normale, Basso), associate a specifiche policy e configurazioni di banda garantita e massima, hanno assicurato una Quality of Service (QoS) ottimale per le applicazioni critiche e una distribuzione equa delle risorse di rete. Nonostante le iniziali difficoltà nell'identificazione dei flussi di traffico e nella definizione delle policy, l'approccio

dettagliato ha consentito di mitigare i problemi di latenza e di garantire una navigazione e un utilizzo dei servizi più fluidi, anche in presenza di carichi di download elevati. Questa soluzione si è dimostrata robusta e adattabile alle esigenze dinamiche della rete dell'OAPd, contribuendo a migliorare significativamente l'efficienza operativa.