

Indice

1	Introduzione	1
2	Contesto	3
2.1	Software Defined Networking	3
2.2	Intent-based Networking	6
2.3	Controller allo stato dell'arte	8
2.3.1	ONOS	8
2.3.2	ODL	10
3	Teraflow	13
3.1	Componenti	16
3.1.1	Device Level Abstraction	16
3.1.2	Service Level Abstraction	18
3.1.3	Management Level Abstraction	19
3.2	gRPC	20
3.3	P4	22
3.4	Mininet	25
3.4.1	Alcuni comandi fondamentali	27
4	Studio e sperimentazione della gestione di policy in Teraflow	30
5	Conclusioni	35
6	Appendice	36
6.1	Installazione	36

1. Introduzione

L'evoluzione delle reti moderne e l'incremento delle esigenze di prestazioni e affidabilità, insieme con l'aumento della quantità di traffico, hanno portato alla necessità di nuovi paradigmi di gestione e controllo delle reti.

In questo contesto, il Software Defined Networking (SDN)^{2.1} rivoluziona la gestione delle reti attraverso un approccio centralizzato e programmabile, separando il piano di controllo dal piano dati. Questa separazione permette agli amministratori di rete di semplificare configurazioni complesse e di introdurre politiche di rete dinamiche e affidabili.

Tuttavia, le funzioni degli ambienti SDN sono ancora focalizzate principalmente sul controllo del funzionamento dei dispositivi. Ad esempio, si possono impostare regole come *"consenti il traffico tra X e Y solo se appartiene alla VLAN Z"* per gestire il flusso dei dati. Sebbene questo approccio semplifichi la gestione dei dispositivi, non affronta completamente la complessità della gestione delle reti su larga scala.

D'altra parte, l'Intent-Based Networking (IBN) ^{2.2}, o Networking basato su intenti, è un paradigma che mira a separare ulteriormente la complessità di implementazione dal livello di gestione. IBN affronta le sfide emergenti fornendo un'astrazione di alto livello dove gli obiettivi per il comportamento della rete possono essere espressi in modo chiaro e intuitivo. Ad esempio, un intento potrebbe essere: *"consentire alle applicazioni contabili di accedere al server XYZ, ma non consentire l'accesso alle applicazioni di produzione"* oppure *"mantenere sempre un'elevata qualità del servizio e un'elevata larghezza di banda per gli utenti di un certo livello"*. Il framework IBN si occupa di tradurre questi intenti in configurazioni di rete specifiche e di adattare dinamicamente la rete per rispettarli, gestendo automaticamente le complessità sottostanti. Questo approccio consente alla rete di adattarsi dinamicamente sulla base dei cambiamenti in tempo reale, gestendo le difficoltà in modo autonomo.

L'obiettivo della tesi è l'implementazione di un intento in Teraflow, uno dei più recenti controller SDN. Inizialmente è stato necessario comprendere l'architettura del controller, incluse le sue componenti, per capire le loro interazioni con l'ambiente circostante e la

gestione dei servizi e delle politiche. Per capire meglio il funzionamento del software, sono stati eseguiti dei casi di test forniti dagli amministratori per simulare e valutare il comportamento del sistema in contesti applicativi specifici. Successivamente, si è passati alla fase di sperimentazione. Non avendo a disposizione una rete reale, è stato utilizzato l'emulatore di reti Mininet 3.4 come ambiente di test.

Partendo da file esistenti, si è creato il contesto necessario e un servizio tra due end-points. Al servizio è stata associata una politica per mantenere i livelli di loss ratio e latenza sotto una determinata soglia. Alla politica viene associata un'azione nel caso le richieste non fossero più rispettate, in questo caso è stato scelto il ricalcolo del percorso. Per monitorare dinamicamente in tempo reale la qualità del servizio, è stato implementato un probe che, attraverso il ping tra i due end-point, misura le due metriche e invia i relativi valori.

Il resto del documento sarà organizzato come segue.

Nel capitolo 2 verrà esposto il contesto. All'interno ci sarà un'analisi approfondita del Software Defined Networking (SDN) e verranno illustrati alcuni dei principali controller SDN presenti sul mercato. Parallelamente nello stesso capitolo verrà, introdotto anche l'Intent-Based Networking (IBN).

Nel capitolo 3 verrà descritto in modo approfondito il controller Teraflow su cui si incentra la tesi. Verrà delineata l'architettura e l'interazione necessaria tra le componenti affinché la politica venga rispettata. Saranno inoltre introdotti gli strumenti utilizzati nella sperimentazione, tra cui P4, gRPC e Mininet, spiegando il loro scopo. Nel capitolo 4 viene descritta, dopo aver introdotto la teoria necessaria, la parte pratica e il fulcro di questo lavoro. Infine, nell'ultimo capitolo, verranno presentate le conclusioni.

2. Contesto

In questo capitolo verranno presentate le principali tecnologie alla base di questo lavoro, ovvero il Software Defined Networking (SDN) e l'Intent-Based Networking (IBN).

Verranno inoltre descritti due dei controller SDN più rilevanti, ONOS e OpenDaylight (ODL).

Questa panoramica fornirà il contesto necessario per comprendere meglio i successivi sviluppi trattati nel documento.

2.1 Software Defined Networking

L'architettura tradizionale di rete si basa su dispositivi fisici di interconnessione situati al livello rete dello stack TCP/IP, switch e router, che collegano più host a livello locale e permettono di scambiarsi informazioni. Questi dispositivi, nell'architettura tradizionale con un controller decentralizzato, implementano al loro interno diverse funzioni del piano data e del piano controllo.

Le funzioni del piano dati (data plane) si occupano della ricezione, del processamento e dell'inoltro dei pacchetti, ossia ciò che serve per l'instradamento. La decisione dell'inoltro del dato viene presa in base alla tabella di routing, nel caso dei router, o alla MAC address table nel caso degli switch.

Le tabelle vengono inizializzate e modificate grazie al piano di controllo (control plane) che si occupa di decidere i percorsi per l'instradamento sulla base della destinazione e comunicarli al data plane aggiornando le tabelle di inoltro dei dispositivi.

Nei protocolli di rete tradizionali questi due piani all'interno dei dispositivi sono separati tra loro e svolgono i loro compiti indipendentemente.

Per decidere i percorsi il control plane può scegliere tra due tipi di algoritmi di instradamento con caratteristiche differenti:

- decentralizzati: nessun nodo conosce la topologia di tutta la rete ma ha informazioni solo dai nodi vicini.

- globali: basati sulla conoscenza della topologia di tutta la rete. Questa soluzione porta allo scambio di molti messaggi di broadcast, inoltre ogni nodo si occupa di calcolare soltanto la propria tabella.

Entrambi gli approcci necessitano di un algoritmo di routing in esecuzione su ciascun router che, attraverso un apposito protocollo, scambia messaggi con gli altri componenti della rete per prendere decisioni. Ciò introduce ritardi che non sono necessari non rendendo così la rete adatta alle nuove esigenze delle applicazioni moderne che necessitano di alta dinamicità.

Il Software Defined Networking (SDN) è un'architettura di rete proposta negli ultimi anni dalla Open Networking Foundation (ONF) [1] per rimediare ai problemi di scalabilità e affidabilità dei controller decentralizzati. SDN rappresenta un nuovo paradigma dinamico, gestibile e facilmente adattabile grazie alla separazione del piano di controllo dal piano dati, che risulta direttamente programmabile. Il piano di controllo diventa centralizzato e gestito da un unico controller SDN che prende le decisioni di instradamento e controllo di tutta la rete. Il piano di dati è invece distribuito sui dispositivi di rete che eseguono le istruzioni ricevute dal controller. Di conseguenza è possibile un disaccoppiamento tra hardware e software per la gestione di device con API diverse.

Per poter funzionare, i dispositivi devono essere in grado di comunicare con il controller centrale e riconoscere cambiamenti significativi degni di notifica per una gestione della rete adattabile ai cambiamenti in tempo reale. Questo è possibile tramite l'installazione al loro interno di componenti software con le caratteristiche necessarie detti Control Agents. La base di questo paradigma è quindi un controller remoto che, interagendo con i Control Agents locali, riceve informazioni sui collegamenti e sul traffico in tempo reale ed è in grado di configurare autonomamente i dispositivi collegati sulla base degli eventi notificati. Lo scopo principale è quindi ridurre e semplificare il carico di amministrazione per i singoli dispositivi.

Come si può notare dalla figura 2.1 la rete viene suddivisa in tre livelli: Infrastructure layer, Control layer e Application layer. Partendo dal livello più basso troviamo l'infrastruttura di rete il cui unico compito è implementare il piano dati, ossia la parte che supporta un protocollo condiviso per comunicare con il controller e gestire i pacchetti sulla base del-

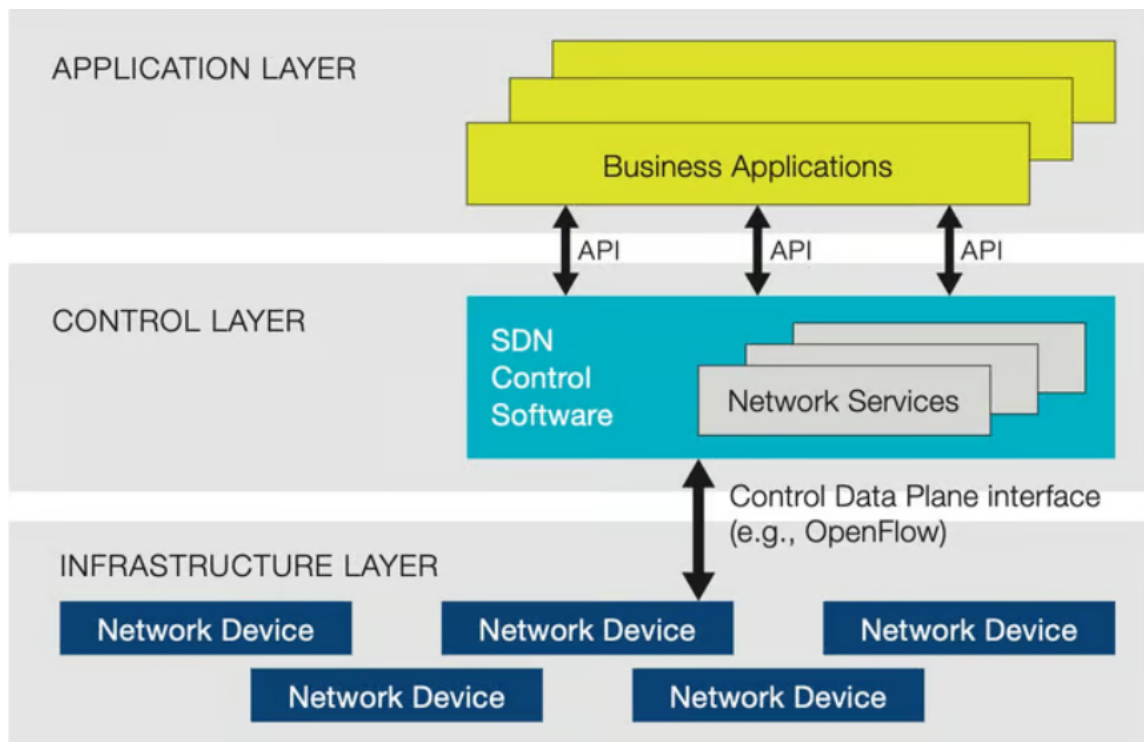


Figura 2.1: struttura di una rete SDN

le configurazioni imposte da quest'ultimo. Questa divisione consente di evitare algoritmi di routing per il forwarding dei pacchetti all'interno dei dispositivi di rete visto che sarà gestito direttamente dai livelli sovrastanti [2].

Nel control layer si trova il controller SDN che, tramite API northbound (NBI) e south-bound (SBI), permette di comunicare con gli altri due livelli. Le API NBI consentono al controller di interfacciarsi con le applicazioni e i servizi situati nel livello superiore, mentre le API SBI, tipicamente implementate tramite OpenFlow, permettono al controller di comunicare con i dispositivi di rete nel livello inferiore.

Questo livello consente il monitoraggio e l'implementazione del piano dati. E' in grado di configurare e di gestire rapidamente le risorse di rete tramite programmi dinamici e automatizzati. Consente infatti di imporre regole di inoltro ai dispositivi sottostanti tramite la manipolazione delle tabelle di routing mediante le API SBI dopo aver calcolato, o aggiornato sulla base di eventi, il percorso migliore.

Il controller inoltre è logicamente centralizzato, anche se fisicamente può essere distribui-

to su più dispositivi. In questo modo il piano di gestione (management plane) situato sopra di esso interagisce con un unico punto di accesso [3].

L'application layer comprende le applicazioni e i servizi che sfruttano le capacità della rete SDN per la realizzazione del piano di gestione. Grazie a questo livello si possono definire politiche o intenti da implementare all'interno della rete. Queste regole sono comunicate al controller tramite le API NBI e quest'ultimo si occuperà di farle rispettare mediante il costante monitoraggio delle risorse del data plane.

Questo disaccoppiamento dei vari livelli consente alla rete di diventare direttamente programmabile da un'unica unità centralizzata riuscendo a mantenere una visione globale e permettendo l'astrazione dell'infrastruttura sottostante per affrontare le sfide di gestione incontrate nelle reti moderne.

2.2 Intent-based Networking

L'Intent-Based Networking (IBN) è un paradigma innovativo per la gestione delle reti che permette di separare la complessità di implementazione dal livello di gestione. Esso è nato per rispondere alla crescente ampiezza e dinamicità delle reti moderne, dove la gestione tradizionale basata su comandi manuali e configurazioni dettagliate non è più sostenibile. Negli ultimi anni, grazie a nuove tecnologie come il 5G o l'Internet of Things (IoT), diverse applicazioni stanno emergendo in differenti campi industriali.

In questo contesto, le implementazioni cloud si sono estese ed è diventato essenziale aumentare le capacità di elaborazione, eseguire servizi distribuiti e garantire il comportamento interattivo che queste nuove applicazioni richiedono. Sono state concepite diverse tecnologie emergenti, tra cui l'IBN, per far fronte a queste necessità. Ognuna ha differenti obiettivi e spesso si integrano tra loro [4].

Il Multi-Access Edge Computing[5] (MEC) fornisce funzionalità cloud alla rete per migliorare la qualità dei servizi offerti in tempo reale portando la capacità di calcolo ai punti di accesso.

Il Network Function Virtualization[6] (NFV) permette di distribuire le funzioni di rete (firewall, NAT, DPI) come apparecchi virtuali. Questi vengono forniti in modo flessibile al

cloud, consentendo così modelli innovativi di fornitura di servizi che migliorano la flessibilità e l'agilità della rete.

Il Software Defined Networking (SDN) consente di semplificare la gestione di rete e utilizzare al meglio le risorse facilitando la virtualizzazione all'interno. In particolare riesce a fornire servizi più efficienti per la definizione di percorsi dei dati.

Nonostante queste innovazioni, rimane un divario semantico tra le esigenze delle aziende e gli obiettivi dei service provider che devono continuamente adattare e proteggere la rete in un panorama orientato ai servizi.

L'IBN nasce come un approccio nuovo, concepito dall'IETF [7]. Si occupa della gestione della rete per astrarne la complessità permettendo agli utenti finali di concentrarsi sugli obiettivi di performance senza preoccuparsi dei dettagli tecnici. Gli utenti possono esprimere i propri obiettivi, a livello di applicazione o di servizio, attraverso direttive di alto livello, gli intenti.

Gli intenti sono espressi in un linguaggio naturale che descrive i risultati desiderati, lasciando al sistema IBN il compito di tradurli. Un intento di rete si riferisce infatti a un livello di astrazione in cui la logica dell'applicazione è espressa in termini di cosa deve essere fatto, utilizzando regole di semantica, piuttosto che di come deve essere implementato.

Esistono due tipi di intenti, prescrittivi e descrittivi. Un intento descrittivo utilizza istruzioni di alto livello, come "consentire il traffico tra X e Y", invece un intento prescrittivo contiene un'informazione più specifica, come "da X:10.0.0.2 a Y:10.0.0.1 set rule=allow"[8]. L'idea centrale dell'IBN è di non specificare i dettagli di implementazione della rete; piuttosto, è la rete stessa che deve eseguire le azioni necessarie per soddisfare gli intenti espressi. In questo modo le applicazioni non devono gestire le direttive di rete di basso livello specifiche della tecnologia. Infatti i livelli applicativi possono interagire con l'Intent Layer evitando di apprendere il linguaggio tecnico-specifico del sistema sottostante. L'approccio IBN è reso possibile grazie la mediazione di un Intent Orchestration Layer, che gestisce e regola il ciclo di vita delle richieste di intenti provenienti dalle applicazioni attraverso operazioni di adempimento e garanzia in un flusso di lavoro a ciclo chiuso. Queste operazioni, oltre a includere la traduzione e l'eventuale orchestrazione di configurazione per

la realizzazione dei singoli intenti, mirano a garantire che la rete rispetti effettivamente l'intento desiderato sulla base della raccolta, aggregazione e valutazione in tempo reale dei dati di monitoraggio. L'IBN fa uso di un Intent Repository, un database in grado di interagire con i moduli di gestione e traduzione degli intenti per fornire la mappatura tra l'intento e la sua configurazione. Questo paradigma offre vantaggi anche ai fornitori di rete, infatti permette di migliorare l'agilità, la disponibilità e la gestione delle reti a un livello di astrazione più elevato e verificare continuamente che gli obiettivi siano raggiunti.

2.3 Controller allo stato dell'arte

Prima di analizzare nel dettaglio TeraFlow si introducono due controller che hanno già raggiunto lo stato dell'arte: ONOS e OpenDayLight. Questi controller rappresentano soluzioni già consolidate nel campo del Software Defined Networking e sono ampiamente utilizzati e studiati sia in ambito accademico che industriale. La loro descrizione ci consentirà di mettere in evidenza le differenti caratteristiche per poter analizzare meglio le innovazioni introdotte da TeraFlow.

2.3.1 ONOS

Open Network Operating System (ONOS) [9] è uno dei controller SDN più noti. E' un progetto nato dalla Open Networking Foundation (ONF) [1] al fine di soddisfare le esigenze degli operatori per poter costruire reali soluzioni SDN/NFV. I principali obiettivi sono quelli di introdurre modularità del codice, configurabilità, separazione di interessi e agnosticismo dei protocolli.

Per adattarsi alle esigenze degli utenti è necessario poter sviluppare una piattaforma applicativa modulare ed estendibile. Per questo motivo la base dell'architettura di ONOS è costituita da una piattaforma di applicazioni distribuite, che utilizzano Java come linguaggio di programmazione, collocata sopra OSGi [10] e Apache Karaf [11] così da permettere l'installazione e l'esecuzione dinamicamente. Queste applicazioni offrono delle funzionalità di base e sostegno al livello superiore il quale fornisce una serie di controlli di rete e astrazioni di configurazione necessarie per il corretto funzionamento del controller.

ONOS Distributed Architecture

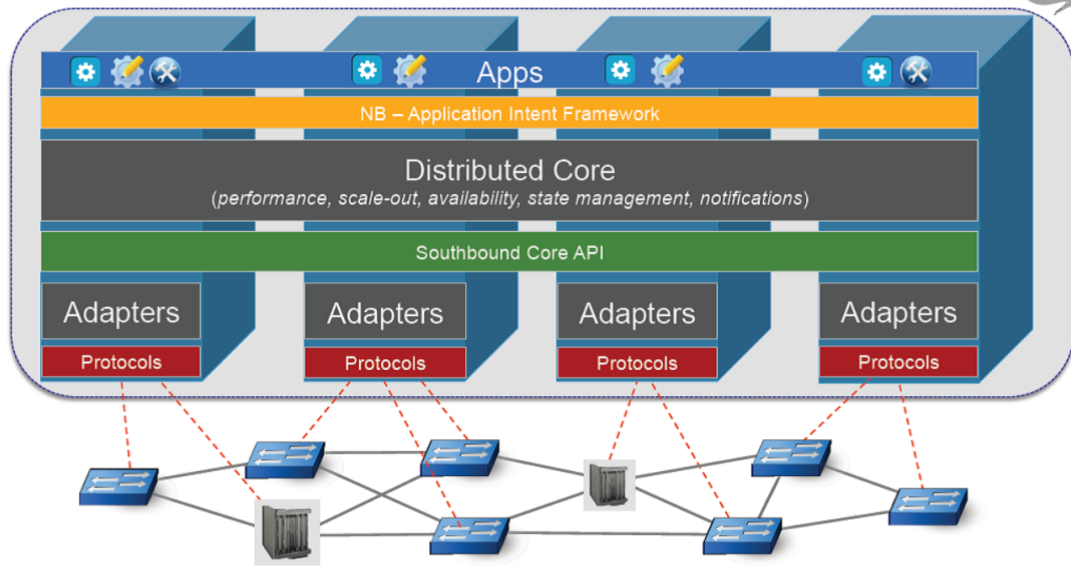


Figura 2.2: Architettura di ONOS

Per estendere le funzionalità, a seconda delle esigenze, sono invece necessarie delle applicazioni ONOS aggiuntive che si comportano come una estensione di quelle già presenti. Ognuna di esse è gestita da un singolo sottosistema che, all'interno del controller, è rappresentato da un modulo. I moduli attualmente installabili che si possono incorporare a quelli inizialmente offerti dal sistema, sono più di 100.

ONOS supporta diverse API northbound tra cui:

- **GUI:** offre un'interfaccia grafica per interagire con l'utente
- **REST API:** facilita l'integrazione con sistemi di orchestrazione e altri controller
- **gRPC:** per un'interazione ad alte prestazioni tra applicazioni e altre entità o protocolli della piattaforma

Per quanto riguarda le API southbound supportate fornisce diversi adattatori che rendono il sistema indipendente dai vari protocolli.

Abilitando il Transport Layer Security (TLS) per l'interfaccia SBI e l'Hypertext Transfer Protocol Secure (HTTPS) per l'interfaccia NBI, ONOS garantisce una buona sicurezza

monitorando e bloccando l'accesso non autorizzato alle risorse in fase di esecuzione [12]. Per quando riguarda gli intenti ONOS mette a disposizione il framework IMR (Intent Monitor and Reroute). IMR può migliorare la qualità della rete e massimizzare l'uso dei collegamenti su di essa. Il servizio consente una migliore gestione del traffico nel reindirizzamento della rete utilizzando scopi e obiettivi specifici. L'intento viene visto come un sottoinsieme di traffico con valori specifici per un pacchetto ed è collegato a una serie di azioni. Gli utenti possono inoltre richiedere che il percorso risultante passi attraverso un determinato numero di nodi o abbia una certa quantità di banda. Questo permette agli utenti di monitorare e spostare facilmente i percorsi in modo flessibile [13]

ONOS è sviluppato come un sistema simmetrico distribuito in cui ogni entità, dal punto di vista software, è indentica alle altre. In caso di guasto di una componente le altre sono in grado di sostenere mantenere la continuità del servizio, assicurando la disponibilità del sistema. Inoltre, per far fronte ai cambiamenti del carico di lavoro o dell'ambiente, ONOS è dinamicamente scalabile, consentendo una replica virtualmente illimitata della capacità del piano di controllo.

Pur essendo fisicamente disaggregato offre una visione logicamente centralizzata al fine di fornire l'accesso di ogni informazione alle applicazioni in maniera uniforme.

ONOS offre numerosi vantaggi agli operatori di rete, tra cui la capacità di implementare soluzioni altamente modulari e configurabili. Grazie alla sua architettura distribuita e al supporto per diverse API consente di gestire reti in modo efficiente e scalabile. Inoltre è facilmente integrabile con sistemi di orchestrazione esistenti, migliorando la flessibilità e la reattività.

2.3.2 ODL

OpenDaylight[14] è un progetto open source che utilizza protocolli aperti al fine di fornire controlli centralizzati e gestire il monitoring della rete.

Fa parte della fondazione LF Networking [15] che si occupa di stanziare e coordinare il supporto a progetti open source coinvolti nello spostamento o nella comunicazione di dati su una rete. Ciò semplifica il coinvolgimento dei membri e aumenta la collaborazione tra

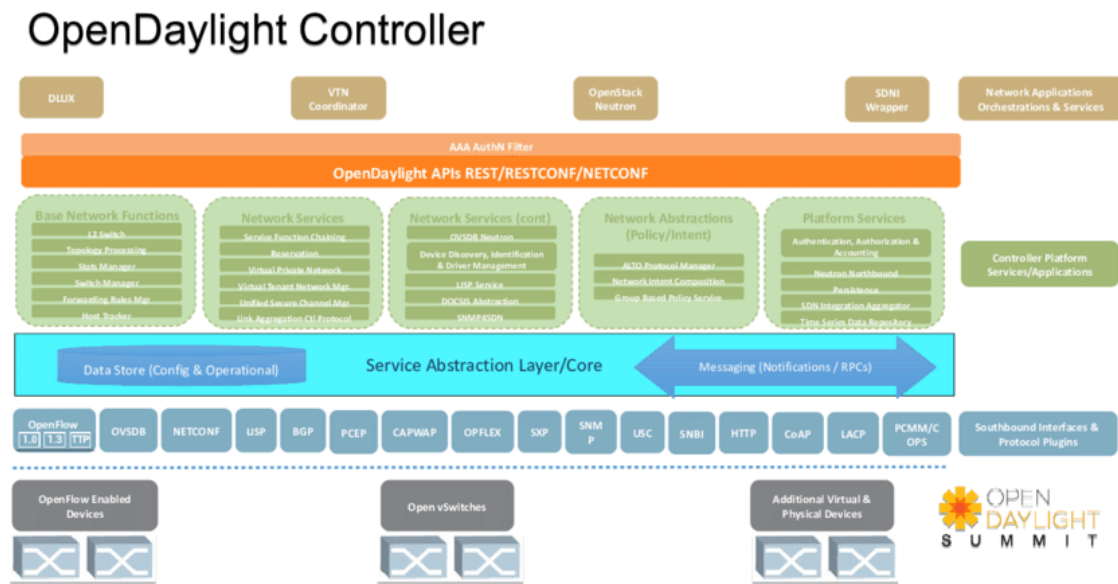


Figura 2.3: Architettura del controller OpenDayLight

i progetti e gli organismi di standardizzazione.

ODL è un framework scritto in Java che consente di soddisfare esigenze specifiche dell'utente al fine di fornire alta flessibilità. Agisce come un software che può essere eseguito su un qualsiasi sistema operativo che supporti java, come una JVM.

L'architettura di OpenDayLight, come mostrato in figura 2.3, è su più livelli. Il livello principale è costituito dal Controller Platform in quanto al suo interno risiede il controller stesso, il quale si occupa di gestire il flusso di traffico andando a modificare le tabelle di inoltro dei dispositivi fisici o virtuali.

Il Service Abstraction Layer (SAL) è il livello inferiore che si occupa di offrire supporto ai vari protocolli SBI come OpenFlow o NETCONF. All'interno di questo livello il collegamento dei moduli tra il controller e i dispositivi avviene dinamicamente al fine di soddisfare il servizio richiesto indipendentemente dal protocollo utilizzato.

Una caratteristica molto importante è l'architettura costituita da microservizi che un utente può decidere se abilitare o meno. Di default sono tutti disabilitati così da permettere una totale personalizzazione.

Questi microservizi sono rappresentati da moduli, contenuti all'interno del controller e collegabili tra loro, che si occupano dell'esecuzione delle varie attività di rete. I moduli sono collegati al SAL dinamicamente.

Per la gestione dei moduli a runtime e l'installazione di funzionalità Karaf da implementare nel software di ODL viene utilizzato Apache Karaf [11]. Attraverso il framework MD-SAL (Model-Driven Service Abstraction Layer) gli sviluppatori sono in grado di creare nuove features sotto forma di servizi e protocolli collegati tra loro. Il controller espone delle API NBI di supporto alle applicazioni. Alcune delle API supportate sono il framework OSGi [10], per le applicazioni in esecuzione all'interno del controller, e REST per comunicare con le applicazioni esterne al controller.

Per far fronte ai problemi di scalabilità, disponibilità e persistenza dei dati si possono avere più istanze di ODL distribuite su macchine differenti che cooperano tra loro attraverso il meccanismo dei cluster.

ODL per gestire gli intenti aveva messo a disposizione una NorthBound Interface che è stata abbandonata nelle release successive a Boron.

Network Intent Composition (NIC) [16] è l'interfaccia che permette all'utente di esprimere uno stato desiderato in una forma neutrale rispetto all'implementazione, detto intento. Quest'ultimo viene applicato tramite la modifica delle risorse disponibili grazie alla gestione dei servizi da parte del controller sulla base delle specifiche. Gli intenti sono descritti al controller tramite l'interfaccia NBI che mette a disposizione la semantica necessaria per la generalizzazione e l'astrazione delle policy invece di specificare i comandi di configurazione dei dispositivi come il resto delle interfacce NBI. È responsabilità dell'implementazione della NIC trasformare l'intento nelle regole di configurazione delle risorse. Questa feature permette di avere a disposizione un modo descrittivo per richiedere il comportamento desiderato della rete. NIC è stato progettato per essere un'interfaccia indipendente dal controller in modo che gli intenti siano trasferibili tra varie implementazioni, in quanto una specifica di intento non dovrebbe contenere specifiche di implementazione e tecnologia.

3. Teraflow

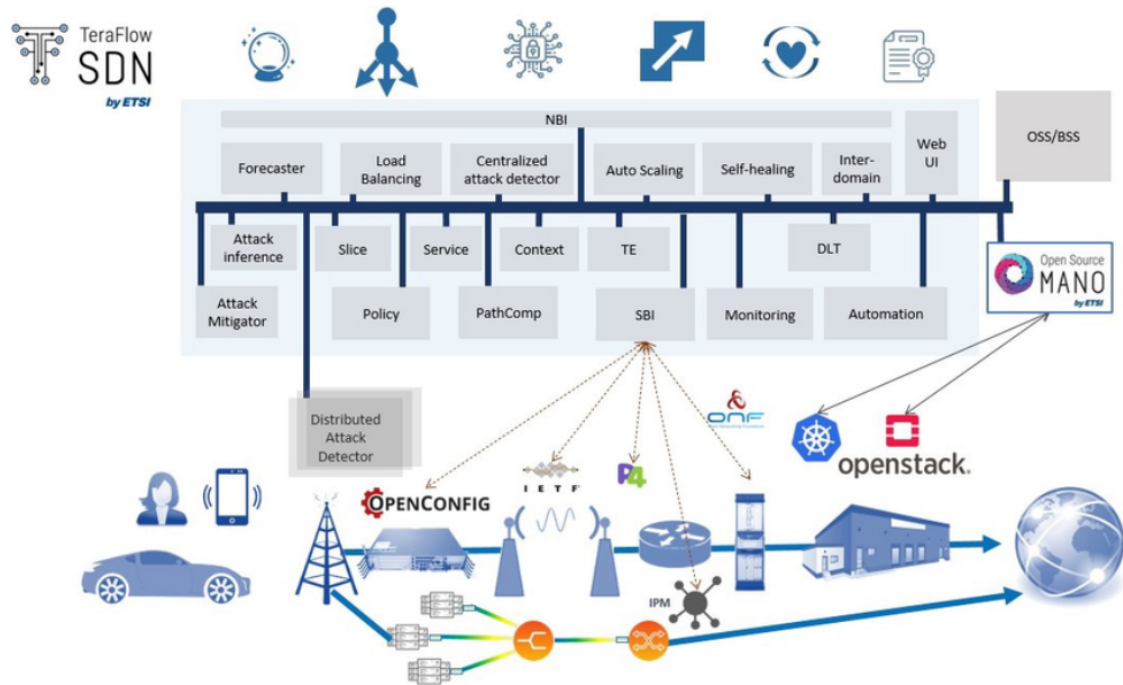


Figura 3.1: Architettura di TeraFlow

Il controller SDN su cui ci focalizzeremo è TeraFlow [17].

TeraFlow è stato finanziato dall'unione europea nell'ambito del programma di ricerca e innovazione Horizon 2020[18] e supportato dal 5G PPP[19], un'iniziativa congiunta tra la Commissione Europea e l'industria europea delle telecomunicazioni. Nonostante la quantità di controller SDN i fondatori di TeraFlow hanno riscontrato come problema comune il declino delle contribuzioni negli ultimi anni. Di conseguenza, i nuovi bisogni e requisiti delle reti moderne che si stanno sviluppando in questo momento non saranno più supportati.

L'obiettivo di TeraFlow è implementare un controller che soddisfi i requisiti attuali ed eventualmente futuri, sia architetturali che infrastrutturali, per le reti. Il controller mira a migliorare le capacità di elaborazione dei flussi consentendo la gestione di oltre un Tera di

servizi di connettività, questo rappresenta una capacità cruciale per supportare le esigenze delle reti B5G. Per raggiungere questo scopo è necessaria una contribuzione attiva da parte degli utenti. A tal fine TeraFlow è un progetto OpenSource al quale tutti i membri della comunità ETSI (European Telecommunications Standard Institute) [20] possono contribuire.

Un ulteriore proposito è di riuscire a diminuire il gap tra ciò che le industrie richiedono e ciò che si può ricavare dagli standard SDN. Questo controller, che è ancora in fase di sviluppo, sarà in grado di integrarsi con gli attuali framework NFV e MEC e fornire l'integrazione delle apparecchiature di rete ottica e a microonde. E' inoltre in grado di interfacciarsi con altri controller come ONOS, ma anche altre istanze di TeraFlow che gestiscono diversi domini, per poter sfruttare funzionalità non ancora implementate e comunicare con le altre reti.

A differenza dei controller presentati precedentemente con un'architettura monolitica, TeraFlow ha un'architettura nativa cloud basata su container, tecniche di virtualizzazione leggera che consentono di isolare le componenti utilizzando una minore quantità di risorse rispetto alle VM (Virtual Machine). L'ambiente creato da un contenitore integra sia il codice di esecuzione del microservizio che le sue dipendenze. Questi container sono sviluppati sopra Kubernetes, un orchestratore responsabile dell'allocazione delle risorse in termini di capacità di calcolo, memoria e archiviazione[21]. Kubernetes presenta diverse funzionalità che garantiscono al controller dinamicità, autoriparazione, integrità e bilanciamento del carico [22] [23]. L'ambiente è sviluppato presso la sede del CTTC a Barcellona.

Ogni container, o componente, ha delle responsabilità specifiche e definisce un microservizio che interagisce con gli altri attraverso la connessione di rete rendendo il controller disaggregato. Le componenti principali sono implementate in Java (solo quelle di Automation e Policy) e Python.

Per ottimizzare l'accesso concorrente e gestire la configurazione di rete, TeraFlow utilizza una componente chiamata Context. Questa componente memorizza la configurazione di rete, inclusi topologie, dispositivi, collegamenti e servizi, in un database No-SQL[24]. La componente garantisce la coerenza dei dati gestiti dai vari componenti del controller

SDN.

La componente di Monitoring gestisce invece le diverse metriche configurate per le apparecchiature e i servizi di rete memorizzando i dati di monitoraggio relativi alle Key Performance Indicators (KPI) selezionate. I servizi sono semplici e dettagliati rendendo possibile l'uso di protocolli leggeri. Dal punto di vista della sicurezza utilizza un sistema di Machine Learning per la prevenzione e la mitigazione di attacchi.

Altri elementi fondamentali del controller sono rappresentati da[25]:

- **usabilità:** realizzata grazie l'interfaccia utente webui che consente la configurazione di servizi predefiniti e la possibilità di visualizzare le metriche in modo personalizzabile in base alle esigenze.
- **scalabilità:** è intrinseca nel design del controller con la replicazione automatica di microservizi per gestire elevati volumi di richieste in ingresso.
- **affidabilità:** garantita attraverso robusti meccanismi di monitoraggio che supervisionano lo stato dei microservizi e dei flussi attivando automaticamente dei processi di ripristino se necessari.

TeraFlow si propone di affrontare le sfide delle reti moderne attraverso un'architettura innovativa e la collaborazione attiva della comunità. Il suo focus è sviluppare un controller SDN Carrier Grade (reti o infrastrutture ben testate con livelli estremamente elevati di affidabilità, ridondanza e sicurezza) per le reti B5G che automatizzi la gestione della rete e sia in grado di scalare per gestire miliardi di dispositivi.

Inoltre, TeraFlow mira a ottimizzare l'uso delle risorse di rete per migliorare l'efficienza energetica e ridurre i costi operativi.

Questo progetto riveste un ruolo chiave nel panorama delle tecnologie 5G, contribuendo a unire diverse università e istituti di ricerca per sviluppare soluzioni all'avanguardia lavorando con organismi di standardizzazione per garantire l'adozione su scala globale.

3.1 Componenti

Le componenti TeraFlow sono classificate in due categorie; le componenti principali del sistema operativo e le netapp sovrapposte [26]. Implementano tutte delle interfacce di comunicazione all'avanguardia indipendenti dal linguaggio di programmazione per formare un potente sistema operativo di rete (Network Operating Sistem 'NOS') che fornisce servizi di connettività sulle infrastrutture di rete programmabili più avanzate tra cui P4, OpenConfig, TAPI...

TeraFlow utilizza un bus gRPC (Google Remote Procedure Call) come protocollo interno per la comunicazione tra di esse. Questo protocollo è stato preferito a REST per la sua capacità più concreta di descrivere le interazioni tra entità in modo efficiente e preciso. gRPC sfrutta i Protocol Buffers per definire in modo preciso gli schemi dei messaggi condivisi e scambiati tra le componenti, consentendo ai vari servizi di implementare le adeguate funzioni di comunicazione indipendentemente dalla piattaforma e dal linguaggio. Ogni componente di TeraFlow implementa i propri Protocol Buffers. Questi si possono trovare nella pagina dedicata del sito [27].

TeraFlow fornisce, inoltre, funzionalità di automazione SDN avanzate basate su policy e parti interessate. Nello specifico, TeraFlow Automation sfrutta importanti eventi di sistema per realizzare la (ri)configurazione di servizi e dispositivi in modo zero-touch, contribuendo così notevolmente alla riduzione delle spese operative per gli operatori di rete[26].

3.1.1 Device Level Abstraction

Il Device Level Abstraction permette l'interazione con i dispositivi presenti all'interno della rete e corrisponde all'Infrastructure layer descritto nell'architettura SDN 2.1. Una componente fondamentale per questo livello di astrazione è la componente SBI. Per far comunicare più tipi di device possibili con il controller la componente presenta eterogeneità offrendo supporto a differenti protocolli. Il suo compito principale è stabilire una connessione con i dispositivi per integrarli nell'ecosistema del controller e essere in grado di configurarli dinamicamente a tempo di esecuzione. Dopo aver stabilito la connes-

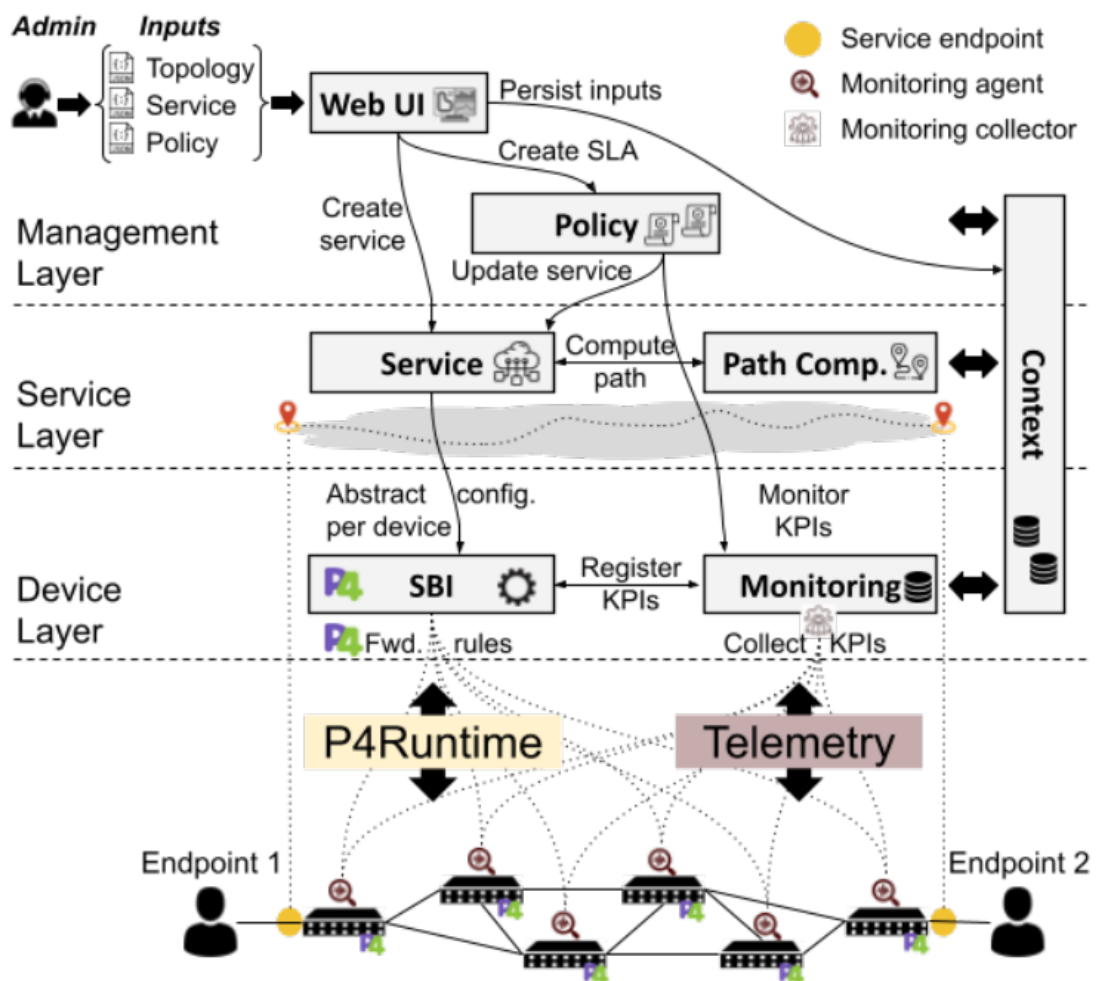


Figura 3.2: Interazione tra alcune componenti in TeraFlow

ne con un determinato dispositivo e aver verificato la disponibilità, la componente offre una API che permette di mandare in input le configurazioni scelte dal controller TeraFlow tramite un file JSON. Il controller provvederà a inserire le giuste entry nelle tabelle degli switch o router pertinenti.

Un'altra componente che fa parte di questo livello è quella di monitoring il cui compito è offrire supporto al Management-level. E' essenziale per l'automazione dei servizi e per prendere le decisioni in tempo reale sulla base di eventi. Questa componente interagisce con i dispositivi per catturare lo stato della rete attraverso delle KPI (Key Performance Indicators) persistendo le informazioni all'interno di un database. Quando un valore KPI

registrato supera determinate soglie predefinite, la componente di Monitoring utilizza il canale API degli eventi del controller per notificare le componenti interessate. Ogni evento è composto da una KPI che identifica la regola a cui si riferisce, un timestamp e un KPIValue, che rappresenta il valore della metrica monitorata in tempo reale richiesta dalla KPI al momento specificato dal timestamp.

Per il corretto funzionamento la componente di Monitoring deve essere in grado di recuperare le metriche da tutti i diversi dispositivi monitorati. Questi implementano spesso protocolli diversi per notificare le KPI, per questo motivo sono stati inclusi una serie di sottomoduli che si connettono agli elementi monitorati utilizzando i protocolli SBI necessari. I dati estratti vengono memorizzati nel Metrics Database così da poter fornire dati dimensionali con serie temporali visualizzabili su Grafana [28]. Tuttavia quanto si tratta di gestire topologie più complesse che coinvolgono molti dispositivi è necessario un livello di astrazione superiore per specificare la connessione tra vari end-points. A tal fine si introduce il Service-level.

3.1.2 Service Level Abstraction

Il Service Level è responsabile della creazione e dell'aggiornamento dei servizi di rete. Questo livello di astrazione permette agli utenti di definire intenti specifici per la connessione tra gli end-points attraverso la componente di Service. La componente di Service supporta diversi tipi di servizi ed è in grado di utilizzare vari protocolli per configurare i dispositivi di rete. Implementa inoltre una Service Handler API che consente agli operatori di rete di definire i comportamenti necessari per ciascun tipo di servizio.

- **L2-VPN**: servizio per dispositivi OpenConfig
- **L3-VPN**: servizio per dispositivi emulati o OpenConfig con supporto per ACLs
- **Connectivity**: servizio per dispositivi TAPI
- **L2 service Handler**: servizio per dispositivi P4
- **Microwave service Handler**

La componente ha al suo interno un blocco gRPC che espone una NBI al resto delle componenti del controller. Gestisce i vari servizi inviando richieste di calcolo del percorso alla PathComp e successivamente, grazie a uno scheduler, si occupa di configurare i dispositivi in base alle connessioni restituite. Questo permette di astrarre la complessità del livello sottostante all'utente in quanto la componente è in grado di tradurre l'intento in un insieme di regole che vengono propagate all'SBI attraverso dei file JSON.

La componente PathComp (path computation) si occupa di gestire la selezione del percorso tra gli end-points per i servizi di connettività di rete. Riceve richieste dalla componente di Service e, interagendo con la componente di Context, recupera le informazioni sulle topologie sottostanti al fine di creare percorsi che soddisfino i requisiti del servizio di rete richiesto. La PathComp rappresenta un'entità singola e specializzata dove possono essere ospitati diversi algoritmi. Questo permette che qualsiasi nuovo algoritmo utilizzato non impatti su altre componenti del controller. Per confrontare i percorsi viene utilizzato inizialmente un algoritmo regolare K-SP dove i k percorsi sono ordinati per numero di passi (hop), ritardo end-to-end e larghezza di banda disponibile sul link più congestionato. Per automatizzare e rendere reattivo questo processo agli eventi che si verificano nello stato della rete, è essenziale introdurre un ulteriore livello di astrazione.

3.1.3 Management Level Abstraction

Questo livello di astrazione è stato introdotto per consentire l'interazione dinamica con la componente di Service per la creazione, l'aggiornamento o l'eliminazione di un servizio in risposta agli eventi notificati. Una delle componenti fondamentali che ne fanno parte è la componente di Policy. Si occupa di definire condizioni di politica che possono essere applicate sia a livello di singoli dispositivi che a livello di dominio della rete. Le politiche possono includere più regole collegate tra loro tramite condizioni AND/OR.

Ogni regola è composta da una KPI che la identifica, un operatore numerico di paragone e un KPIValue, che rappresenta, insieme all'operatore numerico, l'insieme dei valori accettabili o non per quella determinata metrica. Quando le condizioni definite dalla politica sono verificate (ad esempio, il KPIValue identificato per la latenza supera il valore massimo indicato), la componente di Monitoring genera degli allarmi. Questi allarmi vengono

inviati al componente di origine per poi eseguire l'azione predefinita specificata nella politica (ad esempio ricalcolo del percorso). Questo meccanismo consente al controller di reagire agli eventi e di ripristinare un set di dispositivi ad uno stato desiderato.

Quando nella policy è specificato l'id del servizio, la componente interroga il database di contesto per recuperare l'insieme di dispositivi attraverso i quali transita il servizio. In caso contrario la componente itera su una lista di dispositivi per identificare quali devono rispettare la politica di input. Una regola di politica può avere vari stati:

- **inserted** (inserita)
- **validated** (convalidata)
- **provisioned** (provvista)
- **actively enforced** (attivamente applicata)
- **failed** (fallita)

L'integrazione della componente di Policy con quella di Monitoring ha reso quindi possibile l'associazione delle condizioni di politica con gli allarmi del sistema di Monitoring.

3.2 gRPC

RPC (Remote Procedure Call) era un popolare protocollo inter-processo[29]. Esso utilizza uno scambio di messaggi con le caratteristiche di basso sovraccarico, semplicità e trasparenza. Per questi motivi è stato ampiamente usato per diverso tempo in sistemi applicativi distribuiti. Con RPC un client può invocare da remoto una funzione di un metodo. Un utente di RPC non può distinguere dove viene eseguita la procedura chiamata, infatti viene chiamata come una procedura locale [30]. Tuttavia, la maggior parte delle convenzionali implementazioni sono complesse dato che sono costruite sopra protocolli di comunicazione come TCP che ostacolano l'interoperabilità e richiedono specifiche eccessive.

Google Remote Procedure Calls (gRPC[31]) è un framework OpenSource sviluppato da Google per facilitare la comunicazione tra applicazioni distribuite. Permette di connettere, invocare, operare e fare debug di programmi eterogenei in modo semplice.

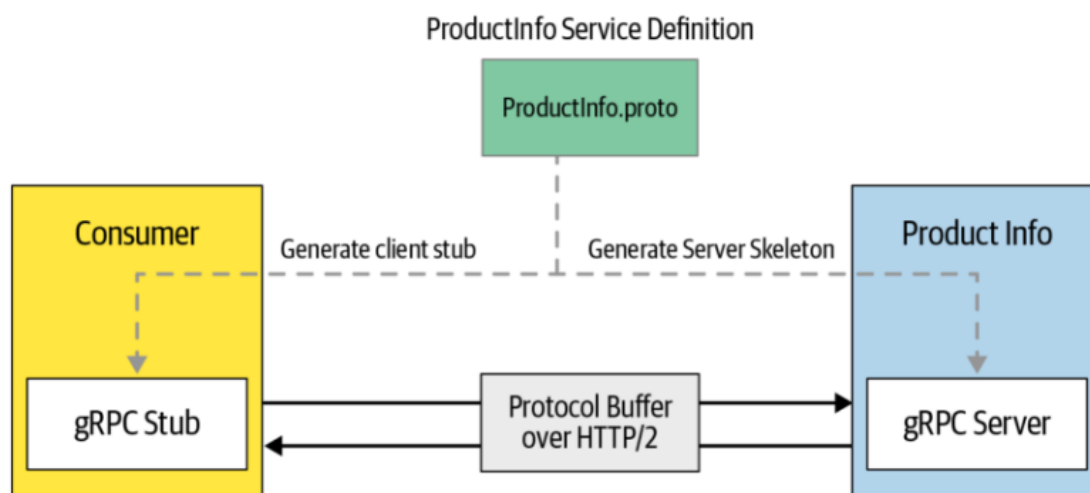


Figura 3.3: Funzionamento del protocollo gRPC

gRPC è basato sul protocollo di trasporto HTTP/2 che supporta la comunicazione bidirezionale. Include supporto per il bilanciamento del carico, il tracciamento, il controllo dello stato e l'autenticazione[32] [33]. Inoltre, ha la capacità di supportare alte performance di streaming in modalità push da parte del server, consentendo l'indipendenza basata sul set di dati e definizioni, grazie ai metodi YANG, e la compressione di dati binari.

Consente di definire i servizi, i loro metodi di comunicazione e trasportare messaggi attraverso dei file di descrizione dell'interfaccia detti Protocol Buffer, o più semplicemente file proto.

I file proto sono un meccanismo indipendente dal linguaggio e dalla piattaforma per la serializzazione delle strutture dati. Questo sistema di serializzazione è più efficiente rispetto a formati come JSON o XML in termini sia di dimensione dei messaggi che di velocità di serializzazione e deserializzazione.

Per sviluppare un'applicazione gRPC è necessario definire un'interfaccia dei servizi. Questa contiene informazioni su come il servizio deve essere usato e quali sono i metodi, i parametri e il formato da utilizzare per i messaggi. Essa viene definita in un file proto, che a prima vista può sembrare un file di testo ordinario, ma in realtà specifica i metodi con i parametri di input e i valori di ritorno. Tutti i metodi nella definizione di questa interfaccia possono essere invocati dal client da remoto. Una volta definita si può generare il codice

lato server (server skeleton) e il codice lato client (client stub) nel linguaggio desiderato usando il compilatore Protobuf *protoc*[29].

Quando il client invoca un servizio gRPC, il lato client utilizza i protocol buffer per serializzare la chiamata di procedura remota nel formato appropriato. Successivamente, la richiesta viene mandata tramite HTTP/2. Sul lato server, viene deserializzata e viene invocata la relativa procedura usando i protocol buffers. La risposta segue il flusso inverso da server a client[29].

Il framework gRPC sottostante gestisce tutta le complessità che normalmente sono associate all'imposizione di vincoli di servizio, serializzazione dati, comunicazioni di rete, autenticazione e molto altro. gRPC è progettato per trasportare messaggi peer-to-peer in modo distribuito e non durevole, consentendo a più servizi di scambiarsi informazioni attraverso un bus condiviso. Grazie all'uso di HTTP/2 e alla codifica orientata ai byte, gRPC riesce a introdurre bassa latenza, rendendolo adatto a sistemi altamente distribuiti e scalabili. Un'altra caratteristica importante è la sicurezza, infatti supporta nativamente il protocollo TLS (Transport Layer Security), garantendo che le comunicazioni tra i servizi siano cifrate e sicure.

3.3 P4

P4 [34], che sta per "Programming Protocol-independent Packet Processor", è un linguaggio di programmazione ad alto livello che consente di descrivere il comportamento di un elemento di rete e del piano dati in modo flessibile. Il linguaggio permette di personalizzare il modo in cui i dispositivi di rete elaborano i pacchetti, senza essere vincolati a un set predefinito di protocolli.

P4 è stato introdotto per l'esigenza di superare le limitazioni con cui si stava interfacciando OpenFlow. Per quest'ultimo protocollo l'hardware e il software non erano più sincronizzati. In alcuni casi gli switch non riuscivano a supportare tutte le funzionalità che OpenFlow poteva offrire a causa delle limitazioni hardware in quanto progettati per supportare un insieme fisso di protocolli e funzionalità. Per introdurre un cambiamento dell'hardware sono richiesti vari sforzi per lo sviluppo e il finanziamento, inoltre richiederebbe mesi se

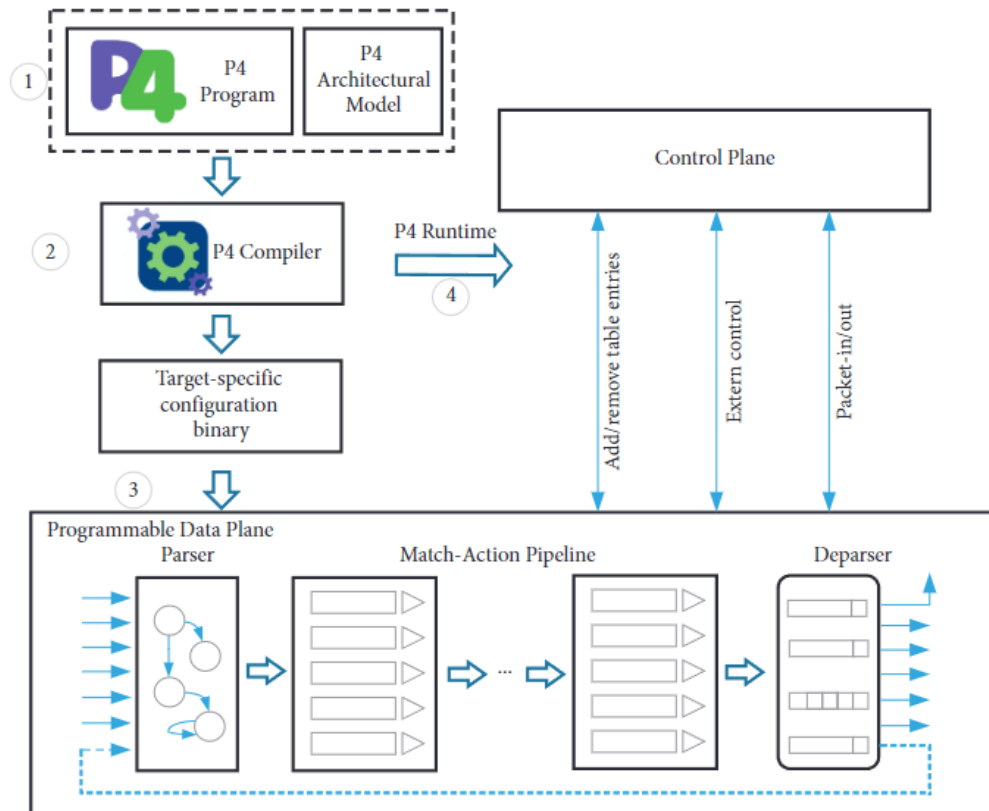


Figura 3.4: Workflow di P4 sul piano dati

non anni per essere portato a termine.

P4 è nato quindi con l'obiettivo di definire nuove astrazioni per programmare il piano dati senza dover andare incontro alle limitazioni riscontrate. Offre un linguaggio di programmazione che permette di definire regole per cui i pacchetti vengono processati direttamente nei dispositivi di rete. Di seguito vengono indicati i vari passi che si eseguono durante un flusso di lavoro di un programma P4 che si possono ritrovare anche nell'immagine 3.4.

Un programma P4 è costituito da diverse sezioni [35], ognuna delle quali descrive un aspetto specifico del trattamento dei pacchetti. La dichiarazione degli header permette di specificare i protocolli che si vogliono analizzare e riconoscere o inventarne di nuovi per scopi di ricerca o esperimenti. Questo include campi come indirizzi IP, numeri di porta e tutti i dati di protocollo. Il Parser specifica come estrarre e interpretare i vari header dai pacchetti. Esso definisce uno stato macchina che determina, attraverso le informazioni in

arrivo, come passare da uno stato all'altro.

La Pipeline di Elaborazione comprende tabelle e azioni che definiscono come i pacchetti vengono processati dopo aver effettuato il parsing. I Controlli del flusso coordinano il parser, le tabelle di elaborazione e le azioni. Definiscono quindi i controlli necessari per il flusso corretto di un pacchetto.

Dopo aver definito il programma esso viene mandato al compilatore P4 che genera due tipi di output. Il primo è un file binario P4 ed è ciò che viene installato all'interno del dispositivo target. Questo file binario è specifico per un target e quindi per uno specifico hardware (ad esempio Asics, fpga..). Il secondo output è indipendente dal target ed è diretto verso la parte NorthBound del controller. Questo è chiamato P4 info e genera i metadati necessari per consentire al piano di controllo e al piano dati di comunicare attraverso P4 Runtime.

P4Runtime è un API messa a disposizione da P4 che permette al controller di connettersi ai dispositivi, vedere cosa c'è attualmente nel pipeline e poter mandare le configurazioni rilevanti nella relativa tabella. Permette inoltre di definire il piano dati in modo dinamico collegandolo al piano di controllo. Per il piano di controllo, P4Runtime protegge i dettagli hardware del piano dati ed è indipendente dalle funzionalità e dal protocollo supportati. P4Runtime riesce quindi a raggiungere l'indipendenza dal target, dal pipeline e dal protocollo.

I nodi programmabili che possono essere ottenuti tramite software o hardware sono definiti P4 target. Essi hanno una pipeline di elaborazione dei pacchetti la cui struttura è specifica per il target ed è descritta da un determinato modello di architettura.

I principali obiettivi quindi includono:

- **Protocollo-indipendenza:** P4 non è vincolato a nessun protocollo specifico consentendo addirittura di definire nuovi protocolli o modificare quelli esistenti
- **Target-indipendenza:** Il codice può essere compilato per una varietà di target sia hardware che software
- **Riprogrammabilità:** permette di aggiornare e modificare il comportamento del dataplane in tempo reale rispondendo ai cambiamenti nei requisiti di rete.

P4 rappresenta un passo significativo verso reti più flessibili e programmabili, consentendo agli sviluppatori di adattare e innovare rapidamente in risposta ai cambiamenti nei requisiti di rete. Si propone come una soluzione innovativa e versatile per superare le limitazioni degli attuali protocolli e dispositivi di rete offrendo un linguaggio dinamico e indipendente dall'hardware.

3.4 Mininet

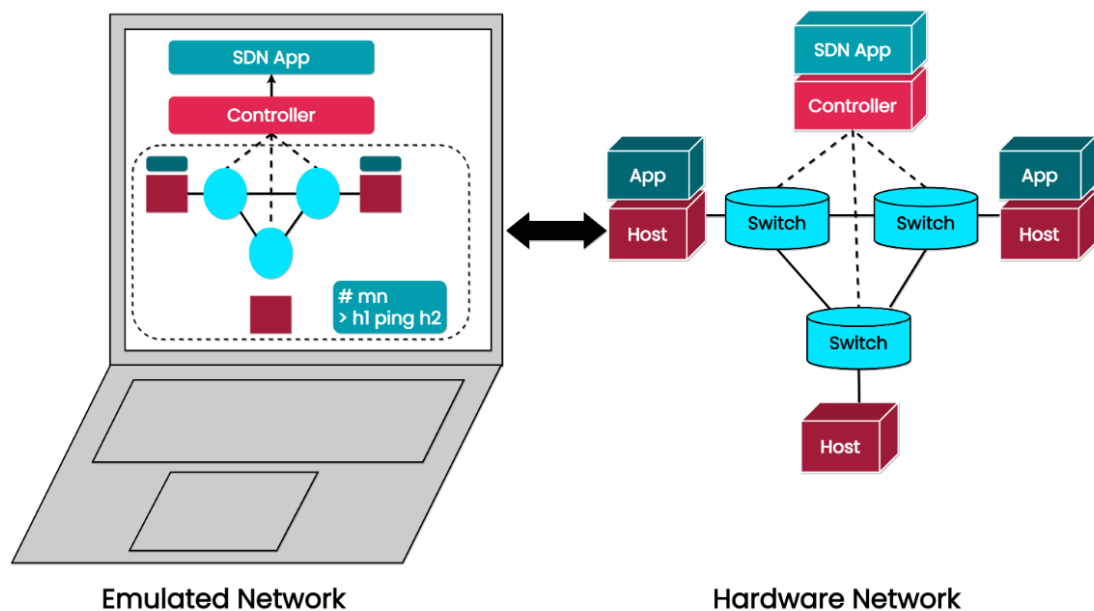


Figura 3.5: Rete mininet

Mininet [36] è un sistema open source di orchestrazione per l'emulazione di reti su un unico ambiente Linux che permette di simulare un'intera rete su un singolo computer. E' ampiamente utilizzato per la creazione, il test e la sperimentazione di reti virtuali realistiche. Rispetto ad altri emulatori presenti in circolazione che emulano ogni dispositivo su una macchina virtuale, Mininet offre una serie di vantaggi. Inanzitutto, permette l'avviamento rapido di una rete, la capacità di eseguire test e programmi con tipologie ampie e personalizzate. Inoltre consente di personalizzare l'inoltro dei pacchetti per testare diverse funzionalità con la possibilità di condividere e replicare il codice. A tal proposito Mininet

presenta delle API e un interprete Python che consentono di definire e gestire facilmente delle topologie di rete. Ciò è possibile anche tramite interfaccia a riga di comando (CLI). In entrambi i casi possono essere sia predefinite che personalizzate con la possibilità di aggiungere e rimuovere switch, router, host, controller e link, tutti eseguiti su un unico computer. Mininet utilizza infatti dei container di virtualizzazione leggera per creare nodi di rete ognuno con la propria pila di rete in modo tale che possano comunicare con gli altri nodi come farebbero in una rete fisica. Ciò consente di testare nuove applicazioni, protocolli e algoritmi in un ambiente controllato e modificabile prima di implementarli su reti reali.

Mininet mette a disposizione tre livelli differenti di API [37]:

- **Low-level:** consiste nelle classi dei nodi e dei link istanziati individualmente e usati per creare una rete.
- **Mid-level:** aggiunge un container per nodi e link, l'oggetto Mininet, e fornisce metodi per la configurazione di rete.
- **High-level:** aggiunge l'astrazione della topologia di rete, la classe Topo. Offre la possibilità di creare modelli di topologia riusabili passandoli al comando mn da linea di comando.

Si possono configurare i link come up o down e inserire metriche specifiche come quelle di banda, ritardo, perdita o massima lunghezza della coda di ricezione per rendere la rete più realistica e adatta a esperimenti di test.

Gli host su Mininet condividono il filesystem root del server sottostante. Ciò significa che non è necessario trasferire file tra gli host virtuali perché tutti accedono agli stessi file direttamente. Tuttavia, questa condivisione del filesystem può creare problemi se un programma ha bisogno di file di configurazione specifici per ogni host. In tal caso, è necessario creare un file di configurazione separato per ogni host e specificare quale file utilizzare quando si avvia il programma. Un'altra limitazione riguarda la condivisione delle risorse del sistema su cui è in esecuzione che dovranno essere bilanciate tra tutti gli host della rete.

Mininet è stato progettato per essere facilmente integrabile con altri software e sistemi di

rete. Consente anche di connettere un controller SDN remoto, quindi esterno alla rete, agli switch, indipendentemente dal PC su cui è installato, in modo da fornire un ambiente adatto allo sviluppo e al test.

3.4.1 Alcuni comandi fondamentali

Linea di comando

Inanzitutto è fondamentale creare una topologia di rete con il seguente comando[38]:

```
$ sudo mn
```

Di default viene inizializzata la topologia minimale (`-topo=minimal`) che consiste in uno switch connesso a due host e un controller OpenFlow. All'interno di Mininet si possono trovare altre topologie disponibili e visualizzabili con il comando

```
$ sudo mn -h
```

che si possono specificare tramite l'opzione `--topo`.

Per avviare la topologia esistono diverse opzioni da poter applicare. Ad esempio, l'opzione `--controller` seguito dall'indirizzo IP specifica il controller al quale gli switch dovranno collegarsi al posto del predefinito offerto da Mininet.

Una volta creata la topologia per avere informazioni su di essa esistono diversi comandi:

- *nodes*: per visualizzare i nodi presenti.
- *net*: per visualizzare i nodi e i link presenti.
- *dump*: per visualizzare tutte le informazioni di dump dei nodi.
- *h1 ifconfig*: per visualizzare le interfacce del nodo h1.

Alcuni comandi per interagire con la rete e fare dei test minimali sono:

- *h1 ping -c 1 h2* : verifica il corretto funzionamento del percorso tra h1 e h2.
- *pingall*: esegue il ping tra tutti gli host connessi alla rete.
- *iperf*: esegue un test di banda tra 2 degli host della rete.
- *xterm h1*: permette di avviare il terminale relativo al nodo h1.

- *exit*: esce dalla rete.

Per manipolare le metriche relative ai link invece vengono messi a disposizione i seguenti comandi:

- *link s1 h1 down*: disabilita un link, in questo caso quello tra lo switch s1 e l'host h1.
- *link s1 h1 up*: attiva un link, in questo caso quello tra lo switch s1 e l'host h1.
- *s2 tc qdisc add dev s2-eth2 root netem loss 50%* : aggiunge una packet loss del 50% sulla porta eth2 dello switch s2.
- *s2 tc qdisc add dev s2-eth2 root netem delay 200ms*: aggiunge un ritardo di 200ms sulla porta eth2 dello switch s2.
- *s2 tc qdisc del dev s2-eth2 root netem loss 50%* : elimina una packet loss del 50% sulla porta eth2 dello switch s2.
- *s2 tc qdisc del dev s2-eth2 root netem delay 200ms*: elimina un ritardo di 200ms sulla porta eth2 dello switch s2.

API Python

Le API Python di Mininet permettono di creare e gestire topologie di rete in modo più flessibile e programmabile. Di seguito esponiamo alcune classi e comandi della Mid-level API:

- *Mininet*: classe per creare e gestire la rete. Il costruttore prende in input diversi parametri la topologia, gli host, gli switch, i controller, i link e ritorna un oggetto di rete.
- *addSwitch()*: aggiunge uno switch alla topologia.
- *addHost()*: aggiunge un host alla topologia.
- *addLink()*: aggiunge un link alla topologia. Si possono specificare parametri come la banda espressa in Mbit (bw=10), il ritardo (delay='5ms'), massima dimensione

della coda espressa in numero di pacchetti (`max_queue_size=1000`), la loss espressa in percentuale (`loss=10`)

- *start*: avvia la rete
- *stop*: esce dalla rete
- *pingall*: esegue il ping tra tutti gli host connessi alla rete
- *h1.cmd('comando da eseguire')*: esegue un comando su h1 da CLI e prende l'output

Con le API in Python si può anche estendere il comando *mn* usando l'opzione *-custom* per invocare la topologia ricreata nello script.

```
sudo mn -your_script.py -topo your_topo
```

4. Studio e sperimentazione della gestione di policy in Teraflow

Lo scopo della tesi è quello di stabilire una connessione tra due end points nella rete che rispetti determinate caratteristiche a tempo di esecuzione utilizzando il controller SDN Teraflow. Inizialmente, attraverso il Service Level, è stato introdotto un servizio nella rete sotto forma di intento per stabilire la connessione e il percorso. Successivamente, tramite il Management Level, è stata inserita una politica basata sugli eventi. Questa politica permette agli operatori di associare a un servizio scelto un service level agreement (SLA) con specifiche condizioni che devono essere rispettate durante l'esecuzione.

Per iniziare il lavoro si è partiti da una demo già esistente per poi apportare delle modifiche successivamente.

Servizio end-to-end

Come già detto in precedenza il Device Level sfrutta una South-Bound Interface (SBI) per interagire con i device tramite l'API P4Runtime. Inizialmente, il codice p4 compilato, ossia i vari artefatti, viene copiato nel pod SBI per poter inserire le giuste configurazioni nelle tabelle dei dispositivi. Il passo successivo è registrare i dispositivi e i link al controller SDN per permettere una corretta comunicazione tra di essi. A questo punto, siamo in grado di richiedere una connessione tra due end points specificando solamente i dispositivi finali tramite un servizio.

La creazione del servizio è realizzata in due passi. Inizialmente, tramite la funzione *CreateService*, si crea un servizio di connettività vuoto nel quale viene specificato solo il tipo associato per poi ritornare l'identificativo a cui è stato correlato. Successivamente viene aggiornato il servizio popolandolo i campi richiesti come endpoints, vincoli e configurazioni di servizio. Di questo si occupa la funzione *UpdateService* [39]. La componente di Service a questo punto si rivolge prima alla componente di Context per recuperare la versione più aggiornata del servizio e settare lo stato a Planned (pianificato), poi alla PathComp per calcolare un percorso.

Per eseguire questa operazione la PathComp utilizza delle informazioni della rete che risiedono nel database logicamente centralizzato della componente di Context. La PathComp è in grado di soddisfare anche richieste di servizio che attraversano più livelli, in questo caso, la risposta includerà uno o più sottoservizi con le relative sottoconnessioni che li supportano.

Una volta ricevute le connessioni ed eventuali sottoservizi la componente Service istanzia un *Task Scheduler* con le relative funzioni, quest'ultimo è responsabile dell'esecuzione delle attività di installazione e smantellamento dei servizi e collegamenti nell'ordine appropriato. Infine viene eseguito il metodo *Execute* del *Task Scheduler* per realizzare tutte le operazioni di configurazione richieste per i dispositivi lungo il percorso tramite l'SBI; inoltre viene modificato lo stato del servizio. Al termine del processo, la componente Context viene aggiornata con le nuove informazioni e l'identificatore del servizio viene restituito all'entità chiamante.

Per mantenere questo processo agnostico rispetto ai dettagli della tecnologia, la componente di Service sfrutta una definizione minima permettendo agli utenti di esprimere cosa vogliono connettere, lasciando che sia il sistema sottostante a decidere come realizzare la connessione reale. La componente traduce automaticamente questa definizione minimale del servizio in modelli di configurazioni astratte dei dispositivi. Queste vengono a loro volta tradotte in regole P4 dal driver del dispositivo P4 della SBI.

Vincoli e configurazioni di servizio

È possibile richiedere azioni supplementari, come l'aggiunta di vincoli o configurazioni di servizio specifiche. Queste vengono specificate inizialmente alla creazione di un servizio e devono essere rispettate finché quest'ultimo non verrà eliminato. Ciò semplifica la gestione dei servizi, in quanto la dichiarazione di queste informazioni aggiuntive può sostituire l'associazione di una politica.

Nella demo presa in considerazione i vincoli non erano specificati ma sono stati aggiunti per quanto riguarda la latenza e la capacità del percorso 4.1. Questa aggiunta non ha prodotto cambiamenti nella definizione del percorso poiché le funzioni relative all'effettivo funzionamento di queste funzionalità non sono ancora state implementate nel controller ma sono solo pianificate per release future. Altri esempi di vincoli possono essere rappre-

sentati dalla posizione di un dispositivo, che può essere sia terminale che non, dal tempo o dal numero di passi massimo di un determinato percorso.

Service svc:SW1/4==SW8/4 (d5261206-1047-51c0-8ef2-89b4c601afe7)

[Back to service list](#)
[Delete service](#)

Context: 43813baf-195e-5da6-af20-b3d0922e71a7
UUID: d5261206-1047-51c0-8ef2-89b4c601afe7
Name: svc:SW1/4==SW8/4
Type: L2NM
Status: ACTIVE

Endpoint UUID	Name	Device	Endpoint Type
82b62384-23c7-5794-9d41-63b7ad775258	4	SW1	
8dd86fc0-c0c7-5540-a384-a4d70141b43c	4	SW8	

Constraints:

Kind	Key/Type	Value
SLA Capacity	-	10.0 Gbps
SLA E2E Latency	-	15.2 ms

Configurations:

Key	Value
Connection Id	Sub-Service Path
4a3db8a1-adf4-4e9b-a98d-b2dcd32ba627	SW1 / 4 SW1 / 2 SW4 / 1 SW4 / 2 SW5 / 1 SW5 / 2 SW8 / 2 SW8 / 4

Figura 4.1: Servizio creato in fase di sperimentazione

Politica

Una parte fondamentale nei sistemi moderni è la gestione a run-time del servizio stabilito[40]. A tale scopo si sfrutta la componente di Monitoring che permette di associare il monitoraggio delle metriche nel proprio database con condizioni che devono essere rispettate. Quando questi requisiti non vengono soddisfatti, la componente di Monitoring solleva un allarme che fa scattare l'azione prestabilita.

Nella sperimentazione abbiamo introdotto una politica per il servizio stabilito con due condizioni: una per la latenza (che non deve superare i 100ms) e una per la loss (che non deve superare il 5%) 4.2. Queste due condizioni sono legate tra loro tramite un "OR", quindi appena una delle due non è più rispettata viene invocata l'azione. In questo caso l'azione consiste nel ricalcolo del percorso.

Verifica

Per verificare il comportamento del controller e assicurarsi che le condizioni di latenza e di loss vengano rispettate, non avendo a disposizione una rete reale, abbiamo utilizzato una topologia di rete su Mininet di switch P4 basati su bmv2[41]. La topologia iniziale era composta da 4 switch ed è stata poi ampliata a 8 per renderla più complessa con l'ag-

eraFlow Home Device Link Service Slice Policy Rules Grafana Debug Load Generator BGPLS About									
Selected Context(admin)/Topology(admin)									
Policy Rules									
1 policy rules found in context admin									
UUID	Kind	Priority	Condition	Operator	Action	Service	Devices	State	Message
uuid: "c4b5e66e-fa99-5075-9b6e-760476791fc1"	service	0	[kpid { kpi_id { uuid: "1" } } numericalOperator: POLICYRULE_CONDITION_NUMERICAL_GREATER_THAN kpiValue { floatVal: 5.0 } , kpid { kpi_id { uuid: "2" } } numericalOperator: POLICYRULE_CONDITION_NUMERICAL_GREATER_THAN kpiValue { floatVal: 100.0 }]	2	[action: POLICY_RULE_ACTION_RECALCULATE_PATH action_config { }]	context_id { context_uuid { uuid: "43813baf-195e-5da6-af20-b3d0922e71a7" } } service_uuid { uuid: "d5261206-1047-51c0-8ef2-89b4c601afe7" } }	[]	ACTIVE	Successfully transitioned from PROVISIONED to ACTIVE state

Figura 4.2: Politica creata in fase di sperimentazione

giunta di più percorsi possibili tra i due end-points. I dispositivi della topologia sono stati connessi alla componente SBI, come per i dispositivi reali, in modo tale che potessero comunicare con il controller. Successivamente per dimostrare la politica basata sul servizio e il ricalcolo delle configurazioni, abbiamo aggiunto artificialmente delle condizioni di ritardo e packet loss all'interno di uno switch presente nel percorso. Ad esempio, abbiamo utilizzato il comando *switch2 tc qdisc add dev switch2-eth2 root netem delay 200ms* per aggiungere un ritardo di 200ms oppure il comando *switch2 tc qdisc add dev switch2-eth2 root netem loss 10%* per simulare una perdita di pacchetti del 10%. Inoltre, abbiamo disattivato un link tra due switch con il comando *link switch6 switch7 down*.

Questi cambiamenti sono stati monitorati grazie a un probe in Python, una funzionalità che rileva lo stato di integrità delle istanze dell'applicazione. Inizialmente, si avvia l'agent che, ascoltando gli eventi della componente di Context, crea due Kpi Id differenti per loss e latenza ogni volta che si viene creato un servizio 4.3. Successivamente si attiva un secondo script che riesce a monitorare le due differenti metriche attraverso il comando ping del terminale e le invia tramite una socket all'agent.

L'agent, ottenuti i relativi valori, li inoltra alla componente di monitoring tramite una KPI composta dal valore, il kpi id corretto e un timestamp.

La componente, ricevute le metriche, riconosce che è avvenuto un cambiamento nella rete e solleva un allarme per una potenziale violazione della politica che sarà poi validata dal-

```
stream: New CREATE event:
context_id {
  context_uuid {
    uuid: "43813baf-195e-5da6-af20-b3d0922e71a7"
  }
}
service_uuid {
  uuid: "d5261206-1047-51c0-8ef2-89b4c601afe7"
}

loss: kpi_id {
  uuid: "1"
}

latency: kpi_id {
  uuid: "2"
}
```

Figura 4.3: Creazione delle due KPI id

la componente di Policy. Questo evento causa l'esecuzione dell'azione predefinita nella politica, ovvero l'aggiornamento del percorso. Quando il nuovo tragitto è calcolato dalla Path Comp, la componente di Service compila una lista di configurazioni di dispositivi per validare gli aggiornamenti seguita da un'altra lista per la cancellazione delle vecchie configurazioni. Queste istruzioni sono tradotte in regole P4 dalla componente SBI prima di essere imposte al piano dati tramite il P4Runtime.

In tutte e tre le tipologie di alterazione della rete, si è riscontrato un cambiamento del percorso volto a rispettare le condizioni desiderate.

5. Conclusioni

All'interno di questo elaborato sono state inizialmente descritte le motivazioni che stanno spingendo la ricerca di nuove tecnologie per superare le limitazioni dell'infrastruttura di rete tradizionale. Abbiamo presentato il paradigma SDN come uno dei più promettenti per conseguire tale obiettivo e sono state discusse le implementazioni più popolari per i controller SDN. Ci siamo concentrati sull'architettura e sulle funzionalità di TeraFlow, un controller SDN emergente progettato per le reti B5G, e abbiamo analizzato i suoi meccanismi di creazione e gestione dei servizi di rete. La sperimentazione si è focalizzata sulla creazione di servizi di connettività tra due end points, l'introduzione di politiche di rete specifiche, e la verifica della risposta del sistema a condizioni introdotte artificialmente. I test sono stati resi possibili dall'emulatore di rete Mininet. La valutazione è stata fatta su prove di raggiungibilità e di performance che, pur essendo abbastanza semplici, rappresentano un punto di partenza per analisi più complesse ed eseguibili su reti più ampie. In una rete reale, tuttavia, andrebbero svolti ulteriori accertamenti per garantire le stesse performance.

Durante la sperimentazione si sono riscontrate delle problematiche dovute alla mancanza di implementazione di alcune funzioni che hanno limitato lo sviluppo completo delle funzionalità desiderate.

Nonostante queste limitazioni, l'approccio sperimentale adottato ha permesso di ottenere risultati significativi, evidenziando la capacità del sistema di rispondere a condizioni di rete variabili e di applicare politiche di gestione del traffico in modo efficace.

Per lavori futuri si potrebbe pensare di implementare il servizio tramite interfaccia Web. Attualmente la creazione dei servizi tramite WebUI non supporta il tipo L2MV utilizzato in questa sperimentazione.

Si potrebbe inoltre sperimentare con topologie più grandi e politiche più complesse per approfondire ulteriormente le capacità e i limiti del sistema. Infine, si potrebbero verificare le stesse prove con altri protocolli e tipi di switch differenti da P4, ampliando così le possibilità di applicazione e la robustezza delle soluzioni proposte.

6. Appendice

6.1 Installazione

Il controller viene eseguito all'interno di una Java Virtual Machine (JVM), quindi è necessario verificare quali versioni di Java supporta la distribuzione che si decide di installare. La versione stabile più recente di ODL è compatibile con le versioni di Java superiori alla 17, mentre per le versioni più datate quest'ultime non vanno bene. Per l'installazione si deve scaricare la distribuzione desiderata che si trova sul loro sito ufficiale nella pagina di download del software [42]. Successivamente è necessario fare l'unzip del file, navigare nella cartella e eseguire il seguente comando da terminale per avviare il controller.

```
./bin/karaf
```

La versione di ODL dell'immagine 6.1 è Potassium.



```
PS C:\Users\ferra\Downloads\karaf-0.19.2\karaf-0.19.2> ./bin/karaf
Apache Karaf starting up. Press Enter to open the shell now...
100% [=====]
Karaf started in 23s. Bundle stats: 349 active, 350 total

ODL

Hit '<tab>' for a list of available commands
and '[cmd] --help' for help on a specific command.
Hit '<ctrl-d>' or type 'system:shutdown' or 'logout' to shutdown OpenDaylight.
```

Figura 6.1: Installazione ODL

Successivamente, si può trovare una lista completa delle feature disponibili eseguendo il seguente comando.

```
feature:list
```

Per installarle invece

```
feature:install <feature1> <feature2>..
```

```
opendaylight-user@root>feature:list
```

Name	Description	Version	Required	State	Repository
odl-bgpcep-concepts	OpenDaylight :: BGPCEP :: Concepts	0.20.6		Uninstalled	odl-bgpcep-concepts-0.20.6
odl-bgpcep-routing-policy-config-loader	OpenDaylight :: BGPCEP :: BGP Routing Policy Conf	0.20.6		Uninstalled	odl-bgpcep-routing-policy-config-loader
odl-mdsal-uint24-netty		0.0.0		Uninstalled	odl-mdsal-uint24-netty
odl-mdsal-rfc8294-netty	OpenDaylight :: MD-SAL :: RFC8294 :: Netty	12.0.4		Uninstalled	odl-mdsal-uint24-netty
odl-mdsal-binding-runtime	OpenDaylight :: MD-SAL :: Binding Runtime	12.0.4		Started	odl-mdsal-binding-runtime
odl-ws-rs-api	OpenDaylight :: Javax WS RS API	13.0.10		Started	odl-ws-rs-api

Figura 6.2: Alcune features disponibili

Bibliografia

- [1] Open networking foundation. <https://opennetworking.org/>.
- [2] Nicholas Brasini. *Analisi e sviluppo di un'interfaccia web per gestire controller SDN*. PhD thesis, Alma mater studiorum - universita' di Bologna, 2017.
- [3] Daniel Barattini. Supporto a micro-servizi per controller ad alta scalabilità e affidabilità. Master's thesis, Alma mater studiorum - universita' di Bologna, 2020.
- [4] B. Martini, M. Gharbaoui, and P. Castoldi. Intent-based zero-touch service chaining layer for software-defined edge cloud networks. *Computer Networks*, 212:109034, 2022.
- [5] MEC. url: <https://www.etsi.org/technologies/multi-access-edge-computing>.
- [6] Nfv. <https://www.etsi.org/technologies/nfv>.
- [7] Ietf. <https://www.ietf.org/>.
- [8] Nathan Sousa, Nazrul Islam, Danny Perez, and Christian Esteve Rothenberg. Policy-driven network traffic rerouting through intent-based control loops. 07 2019.
- [9] Onos. <https://opennetworking.org/onos/>.
- [10] Osgi. <https://www.osgi.org/>.
- [11] Apache karaf. <https://kafka.apache.org/>.
- [12] Ashwin Rajaratnam, Raturaj Kadikar, Shanthi Prince, and M. Valarmathi. Software defined networks: Comparative analysis of topologies with onos. In *2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, pages 1377–1381, 2017.

- [13] Putri Monika, Ridha Negara, and Danu Sanjoyo. Performance analysis of software defined network using intent monitor and reroute method on onos controller. *Bulletin of Electrical Engineering and Informatics*, 9(5):2065–2073, 2020.
- [14] Opendaylight. <https://www.opendaylight.org/>.
- [15] Lfnetworking. <https://lfnetworking.org/>.
- [16] Modulo nic di odl. [https://test-odl-docs.readthedocs.io/en/stable-boron/user-guide/network-intent-composition-\(nic\)-user-guide.html#](https://test-odl-docs.readthedocs.io/en/stable-boron/user-guide/network-intent-composition-(nic)-user-guide.html#).
- [17] Teraflow. <https://www.teraflow-h2020.eu/%>.
- [18] Progetto horizon 2020. https://research-and-innovation.ec.europa.eu/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-2020_en.
- [19] TeraFlow. *5GPPP*, 2021. url: <https://5g-ppp.eu/teraflow/>.
- [20] Etsi. <https://www.etsi.org/>.
- [21] Daniel Adanza, Lluís Gifre, Pol Alemany, Juan-Pedro Fernández-Palacios, Oscar González de Dios, Raul Muñoz, and Ricard Vilalta. Enabling traffic forecasting with cloud-native sdn controller in transport networks. *Computer Networks*, 250:110565, 2024.
- [22] D5.3 final demonstrators and evaluation report. *teraflow-h2020.eu*, 2023.
- [23] D1.4: Final project periodic report. *teraflow-h2020.eu*, 2023.
- [24] D3.1: Preliminary evaluation of life-cycle automation and high performance sdn components. *teraflow-h2020.eu*, 2022.
- [25] D2.2 - preliminary requirements, architecture design, techno-economic studies and data models. *teraflow-h2020.eu*, 2023.

- [26] Componenti principali di teraflow. <https://www.teraflow-h2020.eu/blog/teraflowsdn-core-components>.
- [27] Protocol buffer delle componenti in teraflow. <https://gitlab.com/teraflow-h2020/controller/-/tree/develop/proto>.
- [28] Grafana. <https://grafana.com/>.
- [29] K. Indrasiri and D. Kuruppu. *gRPC: Up and Running: Building Cloud Native Applications with Go and Java for Docker and Kubernetes*. O'Reilly Media, 2020.
- [30] Jong-Kun Lee. A group management system analysis of grpc protocol for distributed network management systems. In *SMC'98 Conference Proceedings. 1998 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No.98CH36218)*, volume 3, pages 2507–2512 vol.3, 1998.
- [31] grpc. <https://grpc.io/>.
- [32] F. Paolucci, A. Sgambelluri, M. Dallaglio, F. Cugini, and P. Castoldi. Demonstration of grpc telemetry for soft failure detection in elastic optical networks. In *2017 European Conference on Optical Communication (ECOC)*, pages 1–3, 2017.
- [33] Ricard Vilalta, Noboru Yoshikane, Ramon Casellas, Ricardo Martínez, Shohei Bep-pu, Daiki Soma, Seiya Sumita, Takehiro Tsuritani, Itsuro Morita, and Raul Muñoz. Grpc-based sdn control and telemetry for soft-failure detection of spectral/spacial superchannels. In *45th European Conference on Optical Communication (ECOC 2019)*, pages 1–4, 2019.
- [34] P4. <https://p4.org/>.
- [35] Ya Gao and Zhenling Wang. A review of p4 programmable data planes for network security. *Mobile Information Systems*, 2021(1):1257046, 2021.
- [36] Mininet. <https://mininet.org/>.

- [37] Nikhil Handigol Bob Lantz and Vimal Jeyakumar Brandon Heller. *Introduction to Mininet*. url: <https://github.com/mininet/mininet/wiki/Introduction-to-Mininet>.
- [38] Mininet.org. *Mininet Walkthrough*. url: <https://mininet.org/walkthrough/>.
- [39] D3.2: Final evaluation of life-cycle automation and high performance sdn components. *teraflow-h2020.eu*, 2023.
- [40] Panagiotis Famelis, Georgios P. Katsikas, Vasilios Katopodis, Carlos Natalino, Lluís Gifre Renom, Ricardo Martinez, Ricard Vilalta, Dimitrios Klonidis, Paolo Monti, Daniel King, and Adrian Farrel. P5: Event-driven policy framework for p4-based traffic engineering. In *2023 IEEE 24th International Conference on High Performance Switching and Routing (HPSR)*, pages 1–3, 2023.
- [41] P4 Language. *Behavioural model (bmv2) reference P4 software switch*, March 2023. Available: <https://github.com/p4lang/behavioral-model>.
- [42] Installazione opendaylight. <https://docs.opendaylight.org/en/latest/downloads.html>.