

**Speciale:
Edge Cloud**

3/2019



notiziario tecnico

Indice



Marco Gaspardone, Roberto Procopio

Edge Cloud & Droni Connessi

Droni ed Edge? Quale relazione può legare una rete che consente di elaborare e memorizzare dati critici localmente e il settore in piena espansione degli Aeromobili a Pilotaggio Remoto (APR), comunemente noti come Droni? Le sperimentazioni avviate a Torino da TIM, alla presenza del suo AD Luigi Gubitosi e della sua Chief Innovation & Partnership Office Elisabetta Romano, e della Sindaca di Torino Chiara Appendino.



Federico Bianco Levrin, Ezio Chiocchetti, Anna Maria Fiammengo, Fabrizio Gatti

Edge Computing in ambito Connected Vehicle

Il termine Connected Vehicle rappresenta una realtà molto complessa che vede tutti gli utenti della strada (quindi non solo automobili o camion ma anche pedoni, bici, moto, semafori, infrastrutture, ecc.) comunicare tra loro al fine di acquisire informazioni che consentano di aumentare notevolmente la sicurezza stradale.



Giuseppe Catalano, Daniele Franceschini, Ivano Guardini, Simone Roma

5G cloud native: Business vision e PoC TIM-Ericsson

TIM ha deciso di indirizzare la sua evoluzione tecnologica in modo innovativo, materializzando la propria visione nella 5G Digital Business Platform, che può essere considerata come un ambiente in cui il 5G, il paradigma cloud, le soluzioni per la gestione del ciclo di vita dei servizi ai clienti, uniti all'uso estensivo delle tecniche di intelligenza artificiale, si fondono in una sola piattaforma tecnologica.



Paola Bertotto, Francesco Epifani, Michele Ludovico, Giovanna Zarba

Open Self-Organizing Network: a continuous development for Radio Access Network performance optimization

The evolution of radio access requires a new approach for network creation and management, assuring quality-of-experience to the customer, even with increasing traffic and network complexity, and sustainable costs for the operator.



Andrea Buldorini, Marco Caretti, Gian Michele Dell'Aera, Salvatore Scarpina

Open RAN: dalle specifiche ai trials

Negli ultimi anni si è assistito ad un aumento esponenziale del traffico mobile a cui si è affiancata la domanda crescente di supporto di connessioni di tipo "always connected" e di terminali controllati da remoto (la cosiddetta Internet of Things).



Antonio Manzolini

Edge intelligence: Edge Computing meets Artificial Intelligence

EC (Edge Computing) is about moving part of the service-specific processing and data storage from the Cloud Computing to the edge network nodes. Among the expected benefits of EC deployment in 5G, there are: performance improvements, traffic optimization and new ultra-low-latency services.

EDGE CLOUD & DRONI CONNESSI

Marco Gaspardone; Roberto Procopio

Droni ed Edge?

Quale relazione può legare una rete che consente di elaborare e memorizzare dati critici localmente e il settore in piena espansione degli Aeromobili a Pilotaggio Remoto (APR), comunemente noti come

Droni? Le sperimentazioni avviate a Torino da TIM, alla presenza del suo AD Luigi Gubitosi e della sua Chief Innovation & Partnership Office Elisabetta Romano, e della Sindaca di Torino Chiara Appendino (Figura 1) hanno dimostrato dal vivo i vantag-

gi introdotti al settore droni dall'uso della tecnologia 5G unita al sistema EDGE Cloud : Bassa Latenza, Privacy dei dati, Banda e capacità locale. In questo contesto il sistema di Edge Cloud costituisce un'estensione ed ulteriore declinazione del Cloud di TIM che arriva fino alla periferia della rete per garantire tali vantaggi. L'Edge Cloud consentirà di superare la possibilità per flotte di droni di connettersi attraverso la rete 5G ad una piattaforma Cloud Robotics

Infrastructure di TIM. Tale piattaforma è dotata di opportune Centrali Operative in grado di abilitare nuove soluzioni per la pubblica sicurezza, per il monitoraggio di infrastrutture critiche e, in prospettiva, anche per il trasporto di farmaci di prima necessità, o apparati per il pronto intervento, come mostrato qualche mese fa a Genova - [link](#) nota stampa (Figura 2). Infatti con Edge Cloud sarà possibile pilotare un drone con un radioco-

mando connesso alla rete radiomobile grazie alla bassa latenza e la sicurezza abilitate dalle piattaforme dispiegate all'Edge della rete (Figura 3). Questo permette non solo al pilota di sfruttare un canale radio sicuro (bande licenziate e protette) e reattivo, ma di proteggersi dalle interferenze che in città inficiano pesantemente sulla portata dei radiocomandi ad oggi impostati sulle frequenze libere in cui la presenza di altri segnali sulla stessa banda

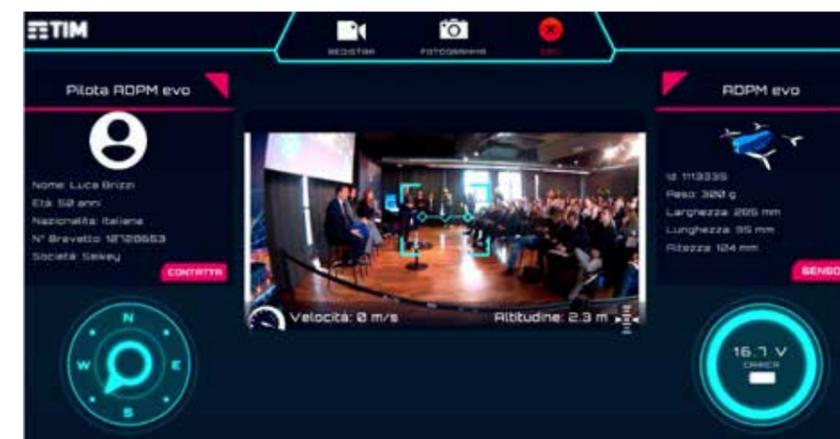
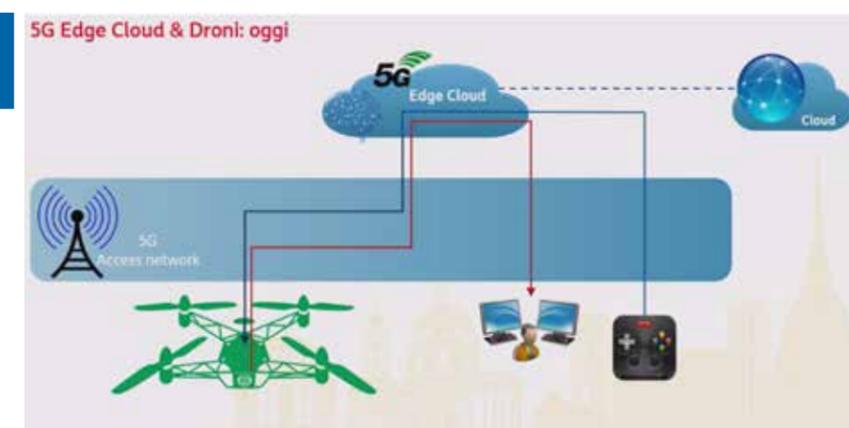


1
Evento 11 Novembre #Torino5G
Edge Cloud & Droni

2
Evento Genova 5G



2
Evento Genova 5G



3
Esempio di schermata di
una Centrale di Controllo

possano disturbare o interrompere la comunicazione con il drone con conseguenze potenzialmente rischiose.

Ma non solo. La nuova normativa Europea relativa all'utilizzo dei Droni ha formalmente inserito le operazioni in BVLOS (Beyond Visual Line of Sight, cioè oltre la linea di vista del pilota) tra le operazioni che potranno essere eseguite in categoria "Specific".

Questo perché molte sono le attività commerciali che potrebbero trarre vantaggio dal BVLOS come

soluzione più efficiente ed economica oltre al fatto che alcune applicazioni sono quasi impossibili da eseguire in modo efficace senza BVLOS (es: ispezioni in ambienti pericolosi, monitoraggio di infrastrutture distribuite sul territorio, ecc).

Ecco che la disponibilità di un nodo di Edge Cloud nelle prossimità del Drone rende possibile anche l'esecuzione di algoritmi intelligenti in rete capaci di supportare il drone stesso e il pilota nella sua missione garantendo livelli di sicurezza mag-

giori che potranno facilitare, coordinati con le normative Europee, l'avvio effettivo di missioni BVLOS anche in contesti "complessi" come quelli urbani (Figura 2) con la possibilità di una costante supervisione da un Centrale di Controllo in rete simile a quella riportata in Figura 3. Queste funzionalità e i voli in modalità BVLOS saranno quindi gli abilitatori di servizi di valore per la PA e per i cittadini per servizi di pubblica sicurezza, e di sorveglianza per protezione di aree/infrastrutture pubbliche e private.

La caratteristica di questi servizi è quella di produrre però grandi quantitativi di dati prodotti dal drone e dal suo Payload costituito da device sotto forma di videocamere 360, fotocamere ad altissima qualità, scansioni infrarosso, sensoristica varia che diventano flussi pesanti di dati da gestire verso una piattaforma in Cloud.

La presenza locale del sistema di Edge Cloud, oltre ad aumentare

complessivamente la capacità del sistema con ovvi vantaggi per tutti i Clienti, permette a questi flussi di dati di rimanere nell'area di interesse, potenzialmente anche "on premises" del Partner/Cliente per garantire i desiderati requisiti di privacy e riservatezza sui dati.

La capacità di segregare i dati dei Clienti localmente è possibile anche in presenza di una miriade di device senza vincoli legati alla connettività verso la cloud, ma

sfruttando la capacità locale di elaborazione dell'Edge Cloud.

Connettività di rete 5G, Edge Cloud ed evoluzione della piattaforma Cloud robotics infrastrutture di TIM potranno essere i cardini su cui far evolvere, in sicurezza e in linea con le nuove normative Italiane ed Europee, gli scenari di servizio in cui i droni troveranno sempre più applicazioni in svariati ambienti e soprattutto nelle nascenti Smart City ■



Marco Gaspardone marco.gaspardone@telecomitalia.it

Ingegnere delle Telecomunicazioni, inizia la sua attività in Telecom Italia nel 2001. Nei primi anni è stato coinvolto in progetti per la definizione di servizi sulla "Intelligent Network". Nel 2008 coordina la messa in esercizio della piattaforma di Energy Management per il monitoraggio delle centrali telefoniche Telecom Italia. Responsabile del centro di ricerca sulla robotica di servizio denominato Joint Open Lab CRAB fino al 2016, è attualmente focalizzato sul tema della Cloud Robotics attraverso la quale la rete radiomobile 5G e il Cloud Computing diventano parte essenziale per l'evoluzione della Robotica di Servizio e Industriale. Negli ultimi anni ha conseguito la graduation per il "Executive Master in Technology and Innovation Management" (Alma Graduate School) e la certificazione PMP® del Project Management Institute. E' anche co-autore di brevetti e pubblicazioni ■



Roberto Procopio roberto.procopio@telecomitalia.it

Ingegnere delle Telecomunicazioni, è entrato in Azienda nel 1998 inizialmente come ricercatore su aspetti di qualificazione degli apparati di commutazione mobile e poi nel campo delle soluzioni di rete per servizi dati su rete mobile (GPRS). Dal 2001 partecipa alle attività di standardizzazione ed in particolare oggi segue il 3GPP CT1. Dal 2001 ha collaborato alle attività di supporto alle partecipate estere su tematiche di UMTS ed IMS. Nel 2004 ha coordinato le attività di innovazione della core network mobile correlate ad UMTS ed IMS. Dal 2008 ha coordinato, come project manager, le attività di innovazione della core network fissa e mobile di Telecom Italia su aspetti relativi ad IMS, data layer, segnalazione Diameter ed interconnessione IP. Oggi lavora come team leader su tematiche di innovazione della Core Network Mobile 5G e dell'Edge Cloud ■

EDGE COMPUTING IN AMBITO CONNECTED VEHICLE

Federico Bianco Levrin, Ezio Chiocchetti,
Anna Maria Fiammengo, Fabrizio Gatti

Il termine **Connected Vehicle** rappresenta una realtà molto complessa che vede tutti gli utenti della strada (quindi non solo automobili o camion ma anche pedoni, bici, moto, semafori, infrastrutture, ecc.) comunicare tra loro al fine di acquisire informazioni che consentano di aumentare notevolmente la sicurezza stradale. Costituisce una delle nuove frontiere della tecnologia mobile, a cui pone importanti sfide in termini di numero di utenti gestiti, latenza e larghezza di banda. Le nuove tecnologie legate al 5G e alla possibilità di trasferire parte dell'intelligenza presente nella rete ai suoi confini (Edge Computing) saranno una componente abilitante fondamentale.

L'Edge Computing nel mondo della mobilità connessa

Le problematiche

La tematica dei veicoli connessi, chiamata così in modo riduttivo in quanto tutti gli elementi della strada possono essere connessi tra loro e non solo i veicoli, pone delle im-

portanti questioni tecnologiche. La sfida è quella di far comunicare tra di loro un insieme di elementi molto eterogenei, alcuni dei quali in condizioni di alta mobilità, con basse latenze ed un elevato livello di qualità e precisione.

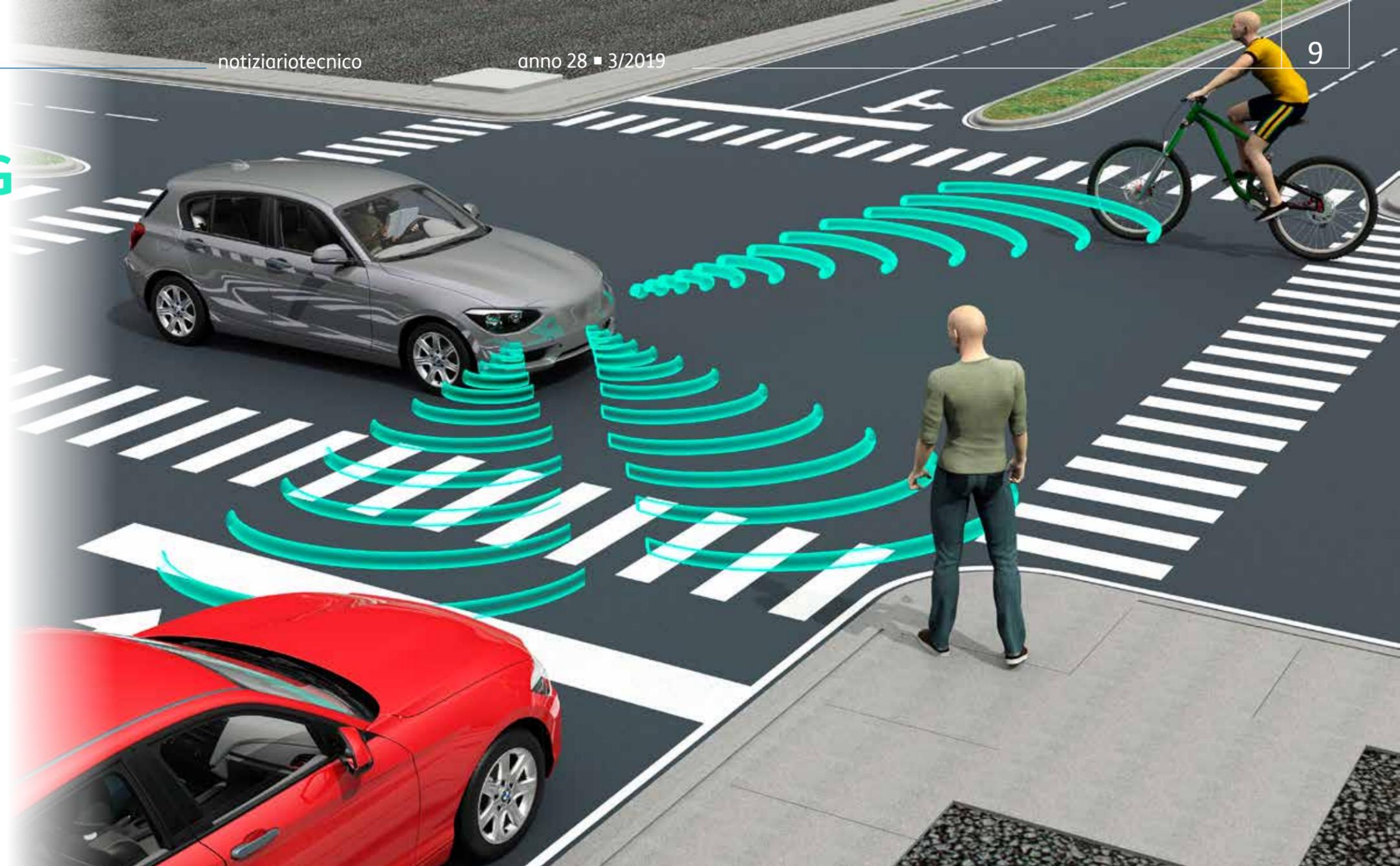
Per capire la complessità di questi scenari possiamo considerare come esempio il caso di due auto che percorrono due differenti tratti di stra-

da che si incrociano. Nell'ambito dello scenario dei connected vehicle tali veicoli invieranno continuamente la loro posizione in modo tale che, tra le altre cose, un "algoritmo anticollisione" possa calcolare le loro traiettorie ed avvertirli in anticipo del rischio di un eventuale scontro in corrispondenza dell'incrocio. Per ottenere questo risultato è necessario che l'algoritmo riceva i dati in tempo

reale (bassa latenza), che li riceva a prescindere dall'operatore mobile che fornisce il servizio di connettività ai veicoli (contesto multi-operatore), che elabori i dati ad una velocità tale che la risposta arrivi in tempo per avvertire il guidatore/veicolo; inoltre deve essere garantita una elevata qualità del servizio (QoS) per tutto il percorso, anche quando ci si trovi in corrispondenza di punti

critici quali i cambiamenti di cella di uno stesso operatore o addirittura tra operatori differenti (seamless coverage communication).

Il quadro si complica ancora di più se consideriamo tutti gli elementi che possono essere connessi. Osservando la *Figura 1* si può notare come siano estremamente numerosi e variabili gli oggetti che appartengono al "field layer", che si distinguo-



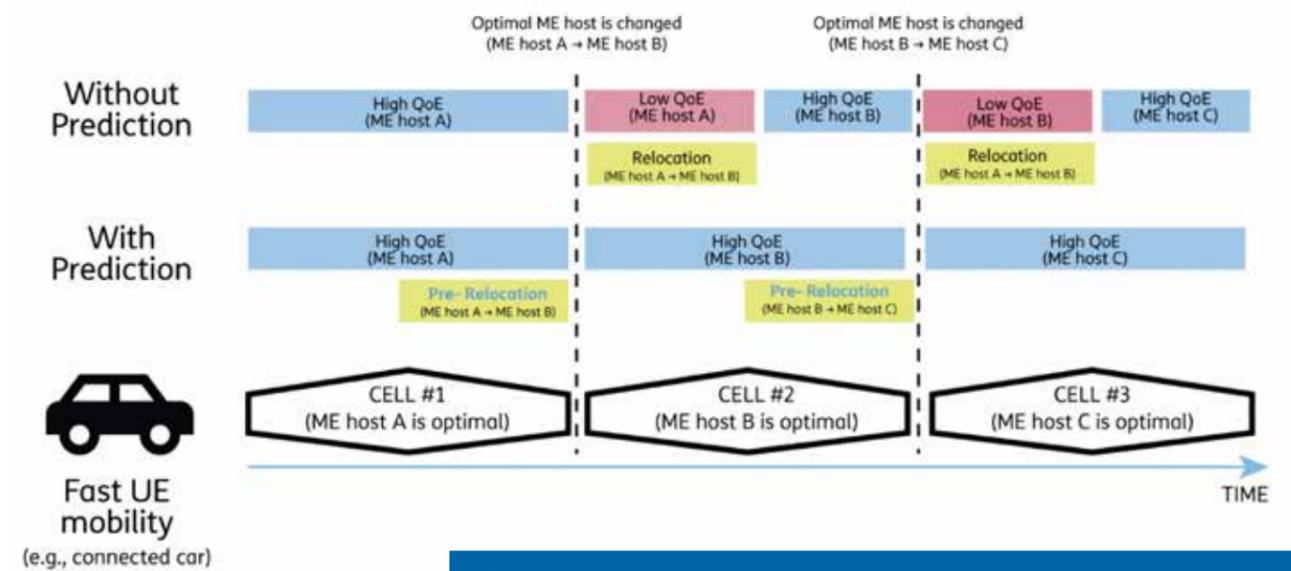
no non solo per la loro eterogeneità (sensori NB-IoT, elementi propri dell'infrastruttura stradale, utenti della strada vulnerabili, veicoli con i loro relativi sensori) ma anche per la tipologia di rete d'accesso utilizzata (contesto multi-access) e la velocità di movimento.

In un contesto così articolato la normale architettura a due livelli di tipo "client-server", con il server posto in cloud, non riesce a fornire i requisiti necessari per il suo corretto funzionamento. L'edge computing è il nuovo paradigma emergente che, portando i building blocks del cloud (potenza computazionale e storage prima di tutto) il più vicino possibile a chi genera e consuma dati ed informazioni può dare una risposta reale al nuovo contesto.

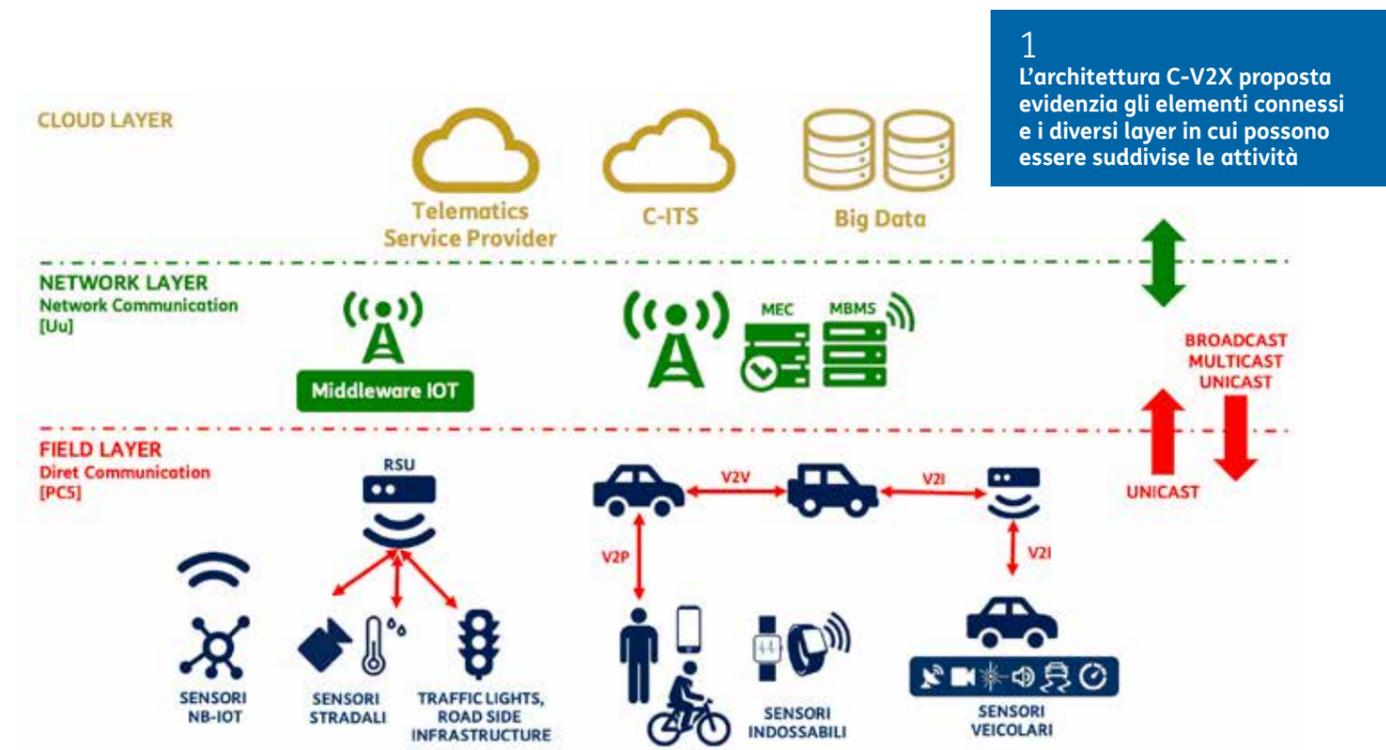
I requisiti

Quali sono le caratteristiche che devono essere offerte da un Edge Computing per poter rispondere alle problematiche evidenziate? Un interessante contributo arriva dal lavoro dell'ETSI, ed in particolare dal documento ETSI "Study on MEC Support for V2X Use Cases" [1], in cui, analizzando i casi d'uso più significativi, si identificano nuovi requisiti che possono essere relativi a servizi, funzioni, interfacce, data model, regole di applicazione ed alcuni particolari punti chiave, indirizzandone possibili soluzioni. Utilizzando tale lavoro e l'esperienza maturata in TIM in questi anni, si evidenziano un insieme di funzionalità che devono essere offerte dall'Edge.

1. Consentire uno switch molto rapido del set di informazioni di contesto del veicolo da una cella alla successiva (seamless coverage). Per indirizzare questo problema, oltre alla definizione di specifiche interfacce, è possibile applicare delle modalità predittive, che anticipando il percorso del veicolo e predicendo a quale cella questo si connetterà, consentano una pre-allocazione delle informazioni di stato sull'Edge target con conseguente aumento della QoE (Figura 2)
2. Risolvere il problema dell'interoperabilità tra operatori (cross operator), collegata sia alla problematica di comunicazione tra veicoli che si trovano in uno



2 Il miglioramento della Quality of Experience in funzione della pre-allocazione delle informazioni di stato presenti sull'Edge (Fonte ETSI [1])



1 L'architettura C-V2X proposta evidenzia gli elementi connessi e i diversi layer in cui possono essere suddivise le attività

stesso luogo ed hanno contratti con differenti MNO sia in particolari situazioni (es. zone di confine) in cui uno stesso veicolo passa da un operatore ad un altro attraverso il roaming. La soluzione di questa problematica non può essere demandata ai livelli alti della rete perché si introdurrebbe troppo ritardo e deve essere mantenuta in local peering points, attraverso una comunicazione orizzontale supportata da una standardizzazione delle API relative

3. Comunicare con le V2X Control Function, un insieme di funzioni in grado di raccogliere e fornire informazioni dalla rete mobile come la lista degli utenti autorizzati, le differenti tipologie di sottoscrizione ai servizi, para-

metri di configurazione dei dispositivi etc.
 4. Raccogliere, utilizzare e notificare ai veicoli informazioni in merito alla QoS. Questo tipo di informazione può essere non solo raccolto dalla rete ma anche ottenuto attraverso uno studio dell'andamento dei veicoli e degli eventi che si verificano sulla strada in quanto mutate condizioni di traffico possono determinare congestioni delle reti. Queste informazioni, fornite di rimando ai veicoli e alla rete stessa in modalità predittiva, potrebbero modificare le informazioni trasmesse dai veicoli, posticipando le comunicazioni non strettamente real time in modo da consegnare in modo prioritario informazioni che ri-

chiedono bassa latenza ed alta affidabilità.
 5. Ospitare funzionalità quali:
 a. Raccolta di informazioni, analisi e memorizzazione per invii successivi. Grazie a questa funzionalità il numero di segnalazioni può essere ridotto e prorogato nel tempo: se N veicoli inviano una segnalazione di lavori in corso, l'Edge potrà re-inviare a quelli che seguono un solo messaggio di avviso, e continuare ad inviarlo nel tempo anche quando non vi saranno più veicoli transitanti in grado di rilevarlo
 b. Elaborazioni di algoritmi specifici sul traffico (anche in coordinamento con il

cloud), che avrebbero il vantaggio di avere delle informazioni relative ad un territorio già partizionato e potrebbero rispondere nei tempi necessari

Sempre in ambito ETSI è in corso un'attività per la standardizzazione di "MEC V2X API", che sarà estremamente importante per consentire un efficace interlavoro tra Edge differenti (per costruttore, rete di appartenenza, tecnologie...) sulle tematiche V2X.

I vantaggi

Un approccio di questo tipo porterebbe ad innumerevoli vantaggi.

1. Bassa latenza, per poter comunicare velocemente informazioni ai veicoli coinvolti

2. Alta densità di connessioni, per consentire a tutti i veicoli presenti in una certa area di comunicare; utilizzando meccanismi di local breakout è possibile concentrare sui nodi periferici il carico, riducendo quello della rete core
3. Gestione di un'alta quantità di dati in zone limitate, che possono risolvere problematiche territorialmente limitate e non appesantiscono l'attività dei sistemi in cloud
4. Segregazione dei dati, funzionale allo use case per ridurre la quantità di informazioni e limitarne la diffusione, con beneficio anche per la sicurezza e la gestione della privacy
5. Invio di informazioni agli oggetti localizzati all'interno di una cer-

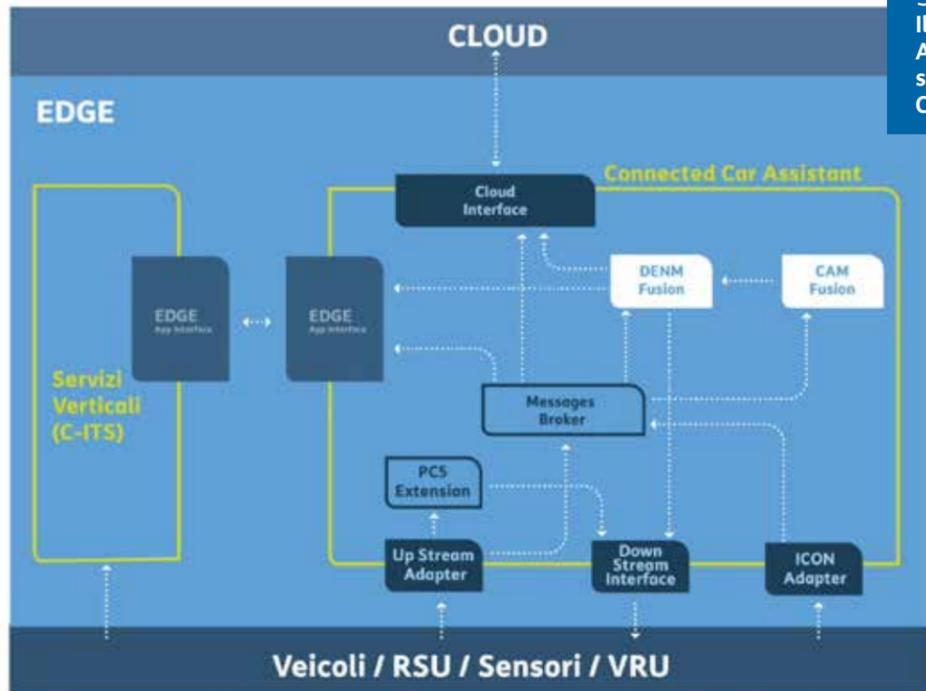
ta area, attraverso il supporto di sistemi che implementano algoritmi di diffusione geografica (es. broker AMQP con opportuna logica)

6. Security, fondamentale per i costruttori di auto

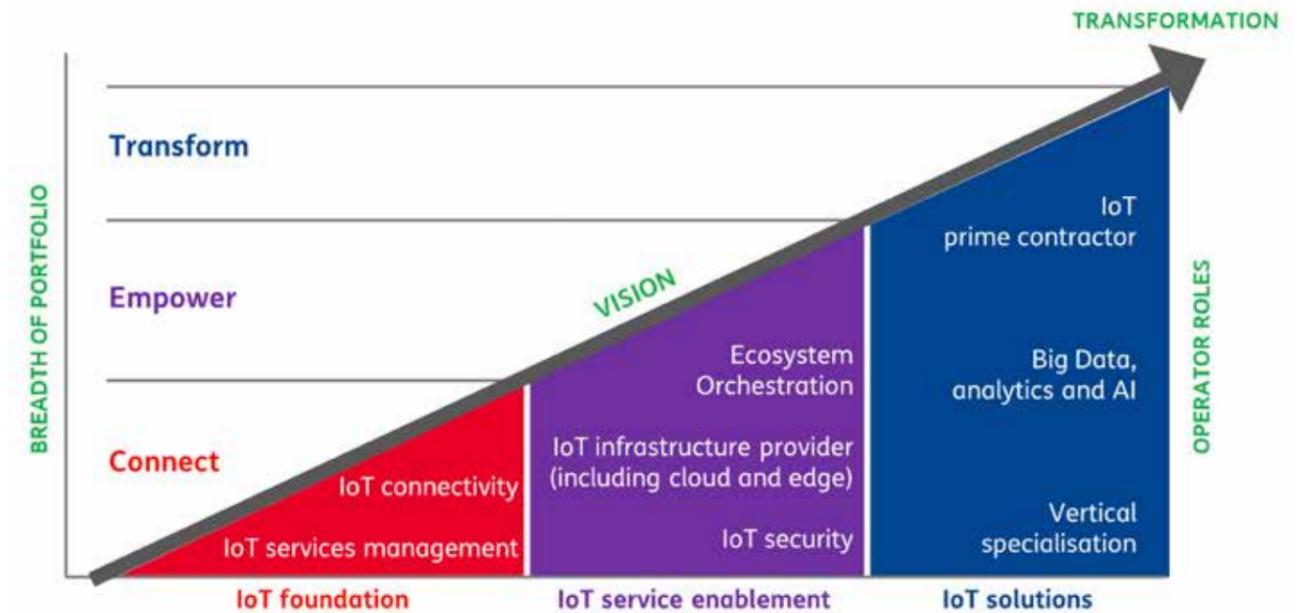
Esempi di casi d'uso applicativi

Progetti di ricerca/ sperimentazioni

Accanto alla parte di normativa sono state avviati diversi progetti e sperimentazioni allo scopo di verificare le potenzialità dell'approccio a tre livelli. Molti operatori TLC stanno



3 Il modulo Connected Car Assistant presente sull'Edge, specifico per le funzionalità CV2X



4 Evoluzione dei ruoli dell'operatore mobile (fonte <https://www.gsma.com/iot/wp-content/uploads/2018/09/New-Roles-for-Operators-in-the-IoT-k.pdf>)

sperimentando soluzioni prototipali per il supporto delle applicazioni di mobilità.

British Telecom e Wind River hanno presentato alcune PoC [4] per dimostrare casi d'uso del 5G Edge Computing per le automobili connesse di prossima generazione analoghi a quelli presi in considerazione da TIM:

- Controllo da remoto del veicolo
- Comunicazione veicolo-veicolo e veicolo-infrastruttura per evitare collisioni

Vodafone sta lavorando con Continental per testare il 5G, con la comunicazione cellulare V2X ed il MEC per migliorare la sicurezza stradale per passeggeri di auto, pedoni e ciclisti, tra cui uno "scudo

digitale per pedoni" e un sistema di avviso di ingorgo.

Inoltre gli operatori, tra cui TIM, sono coinvolti in molti progetti internazionali; se ne riportano alcuni nel seguito.

5G Carmen - 5G for Connected and Automated Road Mobility in the European Union

5G Carmen [2] è un progetto H2020 5G PPP Fase 3 finanziato dalla Commissione Europea per il periodo 1.11.2018 - 31.10.2021 che ha come obiettivo lo studio, l'implementazione e la dimostrazione di soluzioni 5G innovative per la mobilità cooperativa, connessa e automatizzata (CCAM). Gli operatori TLC coinvolti nel progetto sono TIM,

Deutsche Telekom e T-Mobile Austria.

Il progetto ha l'obiettivo di costruire un corridoio stradale abilitato al 5G da Bologna a Monaco per condurre delle prove transfrontaliere delle tecnologie 5G su quattro casi d'uso: Cooperative maneuvering, situation awareness, streaming video streaming e green driving. Per supportare questi servizi che richiedono larga banda e bassi tempi di latenza verrà utilizzata la 5G NR. L'idea che guida il progetto è di definire un "cloud mobile distribuito" che si estende dal veicolo stesso al cloud centralizzato, fornendo una piattaforma finale multi-tenancy e neutral host in grado di abilitare nuovi modelli di business.

Per raggiungere questo obiettivo, si farà leva sui più recenti abilitatori della tecnologia 5G; tra questi l'Edge Computing viene considerato elemento fondamentale dell'architettura, anche in considerazione della necessità di mantenere prestazioni adeguate nelle aree in cui si verificano i cambi di operatore.

5GCAR - 5G Communication Automotive Research and innovation

5GCAR [3] è un progetto H2020 5G PPP Fase 2 finanziato dalla Commissione Europea per il periodo 01.06.2017 - 21.07.2019. L'operatore TLC coinvolto nel progetto è Orange.

Scopo del progetto è la ricerca di tecniche, protocolli e strumenti a supporto della connettività Vehicle-to-Everything (V2X) basata su rete 5G per casi d'uso di advanced driving. Obiettivo principale è migliorare l'affidabilità della comunicazione 5G-V2X, la sua latenza e gli altri indicatori di prestazioni (KPI) correlati alla qualità del servizio mission-critical di advanced driving.

Lo studio condotto su sistema ed architettura V2X ruota attorno a 5 aree di intervento principali, una delle quali è l'elaborazione a bordo rete; tra i miglioramenti proposti per la gestione di veicoli ad alta mobilità vi è il trasferimento preventivo verso l'Edge di destinazione dei processi, allo scopo di ridurre al minimo qualsiasi interruzione del lavoro causata dall'handover.

5GCAR ha anche preso in considerazione approcci in cui le funzionalità

di Edge Computing sono utilizzate congiuntamente alle onde millimetriche (mmWave) per consentire l'ottimizzazione delle risorse di accesso radio e contemporaneamente prendere in considerazione l'offload delle attività di elaborazione.

Progetto Car2MEC

Car2MEC [5] è un progetto durato due anni e concluso nel 2019, finanziato dal Ministero degli Affari Economici della Baviera con lo scopo specifico di ottenere informazioni relativamente al valore dell'Edge Computing in ambito guida connessa; i temi trattati sono stati sia relativi alla tecnologia e all'architettura di rete sia agli aspetti economici. L'operatore telco coinvolto è stato Deutsche Telekom; la sperimentazione è stata condotta sulla pista di prova digitale dell'autostrada A9 in Germania. La conclusione del progetto conferma il ruolo chiave dell'Edge Computing in ambiente V2X: solo se un numero sufficiente di funzioni di calcolo verranno portate a bordo rete, il più possibile vicine alla strada ed ai suoi attori, le comunicazioni V2X veicolo-veicolo e tra veicoli e infrastrutture e servizi stradali saranno in grado di garantire sistemi di guida con gli standard di sicurezza necessari.

Le attività in TIM

All'interno di Chief Technology and Innovation Office - Service Innovation and Ecosystems è in corso da

tempo un'attività il cui scopo è quello di analizzare l'uso dell'Edge Computing nell'ambito dei servizi per il veicolo connesso con l'obiettivo di evidenziare il valore che questa soluzione potrebbe portare alle applicazioni di tale ambito.

L'architettura proposta (Figura 3) non è dissimile da quanto pubblicato nell'ultimo anno da ETSI, e individua all'interno dell'Edge un applicativo specifico denominato "Connected Car Assistant"

Questo elemento è costituito da varie componenti, organizzate logicamente in tre livelli, ognuno dei quali si appoggia su quello inferiore e fornisce le funzionalità di accesso ai propri dati.

- Il primo livello costituisce l'insieme dei punti di accesso con il mondo esterno; sono distinti in Interface, nel caso in cui costituiscano dei semplici banchetti verso l'esterno, oppure Adapter, qualora contengano della logica associata per l'adattamento di dati/protocolli.
- Il secondo ha funzioni di smistamento delle comunicazioni, quali ad esempio la memorizzazione di tutti i messaggi che arrivano all'edge suddivisi secondo le differenti tipologie: messaggi ETSI ITS (CAM, DENM, ecc.), provenienti da sistemi M2M (es. ICON) o da sistemi ITS esterni.
- Il terzo livello presenta le funzioni elaborative e gli algoritmi di analisi in grado di processare tutti i dati disponibili e fornire

meta-informazioni ad applicazioni/servizi per l'utente finale.

Grazie a un'architettura simile a questa e raccogliendo le segnalazioni generate dalle stazioni ITS in movimento sul territorio anche attraverso l'interfaccia di trasmissione Uu su banda licenziata, sarebbe possibile estendere l'area di diffusione e/o la validità temporale dei messaggi al di là dei limiti della comunicazione diretta PC5 in banda libera, dovuti al ridotto raggio di trasmissione o a eventuali interferenze da parte di ostacoli.

Inoltre, le componenti applicative, disegnate ad-hoc per elaborare i messaggi così raccolti con dati provenienti da altre fonti, potrebbero evidenziare in tempi rapidi eventuali problemi alla circolazione e fornire informazioni utili al re-routing dei veicoli.

I frutti di questo tipo di analisi stanno per essere ora portati in ambito sperimentale con la connessione tra l'infrastruttura di Telco Edge messa a punto internamente e le componenti applicative prototipali sviluppate nell'ambito dello specifico verticale. Dopo una prima fase di verifica della corretta interconnessione tra i sistemi e dell'avvenuta comunicazione con i terminali, si procederà a validare in campo l'adeguatezza e la validità di un simile approccio. L'obiettivo è di arrivare in tempi brevi a predisporre un'infrastruttura in grado di permettere l'esecuzione di trial sperimentali e di proporsi come riferimento architeturale per il deployment com-

merciale delle applicazioni per questo settore.

Il valore per TIM di queste soluzioni

Come già evidenziato in un articolo apparso in precedenza sul Notiziario Tecnico (N.3 2018) [6], gli operatori Telco possono cogliere interessanti opportunità di mercato nell'ambito dei servizi per la Smart Mobility. Questo trend rientra nella visione dell'evoluzione del ruolo dell'operatore secondo uno studio condotto da GSMA e IDC [7] in cui si evidenziano i fattori di trasformazione attraverso un percorso che, partendo dagli aspetti di connettività (Connect), aumenta (Empower) e trasforma (Transform) il ruolo degli operatori mobili nell'ambito dei servizi IoT come riportato in Figura 4. Nel caso specifico di questo vertical, l'architettura riportata in Figura 1 evidenzia il ruolo che la componente di Edge Computing può ricoprire. La separazione tra i due nodi (telco e applicativo) facilita la messa a valore di quelle funzionalità/informazioni/dati gestite direttamente dall'operatore che possono essere rese disponibili agli altri player di settore attraverso un'infrastruttura con accesso controllato per garantire quelle latenze dell'ordine di pochi millisecondi necessarie per un'efficace funzionamento dei servizi applicativi. Tale approccio garantisce la necessaria flessibilità per quanto riguarda la componente applicativa che può essere messa in

campo e gestita autonomamente da qualunque altro attore della filiera, favorendo una collaborazione di tipo win-win con l'operatore Telco. Gli esempi riportati nell'articolo dimostrano, inoltre, come questa visione venga perseguita anche dagli altri operatori stranieri.

Conclusioni

Il valore che l'infrastruttura di telecomunicazioni può portare nell'ambito della tematica dei veicoli connessi è fondamentale per il dispiegamento di queste nuove tecnologie. Solo la presenza di nodi intelligenti, posti ai confini della rete e quindi prossimi agli utilizzatori, può soddisfare gli stringenti requisiti che la nuova mobilità richiede. Una rete capillarmente diffusa sul territorio, che garantisce la possibilità di comunicare con un elevato tasso di affidabilità, una copertura pressoché totale della rete stradale, la possibilità di fornire sistemi e funzionalità di rete in grado di portare valore nelle applicazioni specifiche, i PoC e le attività di studio già realizzati fanno di TIM una compagnia telefonica in grado di giocare appieno il suo ruolo in questo ambito ■

Bibliografia

- 1 ETSI GR MEC 022 V2.1.1 (2018-09) Multi-access Edge Computing (MEC); Study on MEC Support for V2X Use Cases https://www.etsi.org/deliver/etsi_gr/MEC/001_099/022/02.01.01_60/gr_MEC022v020101p.pdf
- 2 <https://www.5gcarmen.eu/>
- 3 <https://5gcar.eu/>
- 4 <https://www.windriver.com/news/press/pr.html?ID=21602>
- 5 <https://www.nokia.com/about-us/news/releases/2019/03/21/continental-deutsche-telekom-fraunhofer-esk-mhp-and-nokia-successfully-conclude-tests-of-connected-driving-technology-on-the-a9-digital-test-track/>
- 6 Notiziario Tecnico Telecom Italia -N.3 2018, Smart Mobility: come cambia il ruolo dell'operatore <https://www.telecomitalia.com/content/tiportal/it/notiziariotecnico/edizioni-2018/n-3-2018/N2-Overview-trial-europei.html>
- 7 GSMA, Opportunities in the IoT: evolving roles for mobile operators <https://www.gsma.com/iot/wp-content/uploads/2018/09/New-Roles-for-Operators-in-the-IoT-k.pdf>

Acronimi

5GPPP	5G 5G Public Private Partnership
API	Application Programming Interface
CAM	Cooperative Awareness Message
CCAM	Connected, Cooperative and Automated Mobility
C-V2X	Cellular - Vehicle to everything
DENM	Decentralized Environmental Notification Message
ETSI	European Telecommunications Standards Institute
GSMA	GSM Association
IDC	International Data Corporation
IoT	Internet of Things
ITS	Intelligent Transport Systems
MEC	Multi-access Edge Computing
MNO	Mobile Network Operator
PoC	Proof of Concept
QoE	Quality of Experience
QoS	Quality of Service
V2X	Vehicle To Everything



Federico Bianco Levrin federico.biancolevrin@telecomitalia.it

Laureato in Ingegneria delle Telecomunicazioni entra in azienda nel 2011 a conclusione del master «Innovazione di reti e servizi nel settore ICT» promosso da TIM in collaborazione con il Politecnico di Torino. Inizia l'attività lavorativa nel gruppo «Mobile Application And Services» dove ha seguito lo sviluppo di applicazioni mobili innovative con particolare focus sui device Android. Dal 2015 nel gruppo «Connected Vehicle» ha partecipato a progetti di ricerca sia in ambito italiano che europeo (Scout e AUTOPILOT) ■



Ezio Chiocchetti ezio.chiocchetti@telecomitalia.it

Laureato in Scienze dell'Informazione entra in CSELT nel 1988 occupandosi di Linguaggi Formali di Specifica in progetti europei e presso l'ITU. Nel periodo 1993 - 2006 è responsabile di attività in ambito ITS di Telecom Italia (valutazione prestazionale OSS e collaudo E2E dei processi di delivery dei servizi fonia e dati in Italia e per le controllate europee). A seguire, sempre in ambito ITS, partecipa alle nuove tematiche legate all'introduzione della Software Oriented Architecture. Dal 2009 si occupa della applicazione della metodologia Lean per il miglioramento dei processi in ambito Data center e Top Client. Dal 2011 lavora nell'ambito dei Servizi Innovativi occupandosi in una prima fase di servizi di e-commerce (progetto finanziato Multitaly) e di gestione del telemonitoraggio sulla capillary network, e attualmente nell'ambito dei veicoli connessi, partecipando a progetti di ricerca finanziati in ambito nazionale (URBeLOG) e internazionali (AUTOPILOT) ■



Anna Maria Fiammengo annamaria.fiammengo@telecomitalia.it

Laureata nel 1985 in Scienza dell'Informazione entra in azienda nello stesso anno occupandosi di strumenti CAD per il disegno dei circuiti integrati. In seguito passa ad occuparsi di Sistemi Informativi Geografici (GIS) per la creazione di strumenti innovativi di pianificazione e progettazione della rete fissa, sistemi di raccolta e documentazione dati di traffico per servizi antifrode (progetto RADAR), progettazione di servizi OTP per la telefonia mobile, progetti di innovazione didattica tra cui SMARTINCLUSION, per la partecipazione scolastica da remoto di bambini ospedalizzati. Dal 2015 nel gruppo Connected Vehicle, ha partecipato a progetti di ricerca in ambito italiano (URBeLOG) ed europeo (5GCARMEN, AUTOPILOT) ■



Fabrizio Gatti fabrizio1.gatti@telecomitalia.it

Laureato in Fisica, entra in azienda nel 1992 analizzando le problematiche di qualità e affidabilità dei sistemi tlc. In seguito, passa ad occuparsi della gestione delle tematiche ambientali associate alle attività degli operatori di telecomunicazioni dove ha partecipato a vari progetti EURESCOM e ha rappresentato Telecom Italia in vari WG in ambito ETNO, OMA e ETSI. Dal 2005 lavora nell'ambito dei Servizi Innovativi dove coordina le attività relative alla progettazione e alla prototipazione di nuovi concept per il veicolo connesso/autonomo, l'ITS e la City Logistics, partecipando in rappresentanza dell'azienda a progetti di ricerca finanziati in ambito nazionale come URBeLOG (in cui è Responsabile di Progetto), e internazionali, come "AUTOPILOT", "Team" e "5G-Carmen" ■

5G cloud native: Business vision e PoC TIM-Ericsson

Giuseppe Catalano, Daniele Franceschini,
Ivano Guardini, Simone Roma

TIM ha deciso di indirizzare la sua evoluzione tecnologica in modo innovativo, materializzando la propria visione nella 5G Digital Business Platform, che può essere considerata come un ambiente in cui il 5G, il paradigma cloud, le soluzioni per la gestione del ciclo di vita dei servizi ai clienti, uniti all'uso estensivo delle tecniche di intelligenza artificiale, si fondono in una sola piattaforma tecnologica. Le tecnologie cloud native rappresentano un pilastro di tale approccio, poiché forniscono la flessibilità necessaria ed abilitano nuovi modelli di business più dinamici. TIM ha sperimentato con successo - primo operatore in Italia e tra i primi al mondo grazie alla collaborazione con Ericsson - la core network 5G cloud native presso i propri laboratori di Torino, accelerando così il percorso di evoluzione della rete 5G di TIM verso le tecnologie cloud native.

I nuovi modelli di business abilitati dalle tecnologie cloud native

Il 5G è una keyword della Industry sia Telco sia dei nuovi servizi verticals per l'Industria 4.0 e la Pubblica Amministrazione, come un nuovo paradigma capace di realizzare la trasformazione dei servizi verso il digitale. Se però guardiamo nel fine cosa si intenda per 5G, ci rendiamo conto che esiste una pleora di interpretazioni e aspettative che differiscono molto tra di loro.

In primis il 5G è una nuova tecnologia standardizzata dal 3GPP, così come le precedenti generazioni del Mobile. Molti pensano quindi che il carattere distintivo e l'innovazione del 5G siano legati alla nuova interfaccia radio, che abilita un salto prestazionale. In realtà, sebbene l'interfaccia radio del 5G, la cosiddetta "New radio interface", porterà una serie di innovazioni quali l'uso delle bande millimetriche (come la 26 GHz), delle funzionalità quali le antenne attive, del Multi-User MIMO, e di soluzioni per il supporto di bassissime latenze per servizi verticali "estremi" (il tele controllo di robot, droni, delle auto, etc.), essa non rappresenta l'unico carattere distintivo. Non a caso lo stesso 3GPP qualifica l'evoluzione di LTE dalla release 15 in poi come 5G e prevede il suo collegamento alla nuova core 5G (5GC), rendendo nei fatti LTE la prima interfaccia radio del 5G. Già oggi LTE evoluto per applicazioni eMBB consente di raggiungere velocità di

centina di Mbps in mobilità, soddisfacendo quindi i requisiti dei clienti e della digital agenda europea.

Ma allora quale è il carattere distintivo del 5G?

Una delle più grosse rivoluzioni apportate dal 5G sarà nella rete di back end, andando a creare la prima rete programmabile pensata per gestire in modo autonomo e autonomico una miriade di interazioni non solo human ma soprattutto not human, determinando l'inizio di una nuova era.

Di fatto la tecnologia 5G, unitamente all'impiego delle tecnologie cloud native (le stesse per intenderci già usata dagli hyperscale cloud provider, quali Netflix, Amazon etc), consentirà di portare le reti degli operatori Telco in ambienti cloud, rendendole estremamente più flessibili per la creazione di nuovi servizi digitali (molti dei quali ad oggi non noti).

In questo nuovo paradigma, questi nuovi servizi non saranno più realizzati solo dagli operatori Telco, ma anche da terze parti, ovvero sviluppatori sia del mondo dei vertical sia delle stesse applicazioni che oggi troviamo sugli store e market place dei nostri smartphone.

Le tecnologie pregresse proponevano una rete di telecomunicazioni monolitica, in cui era cablato un set di servizi immutabili, come la voce e gli SMS in mobilità per il 2G e il 3G e i servizi dati per l'HSPA e il 4G. Con il 5G ci troviamo di fronte ad un cambio di paradigma in cui, grazie al cloud e alla programmabilità, la complessità

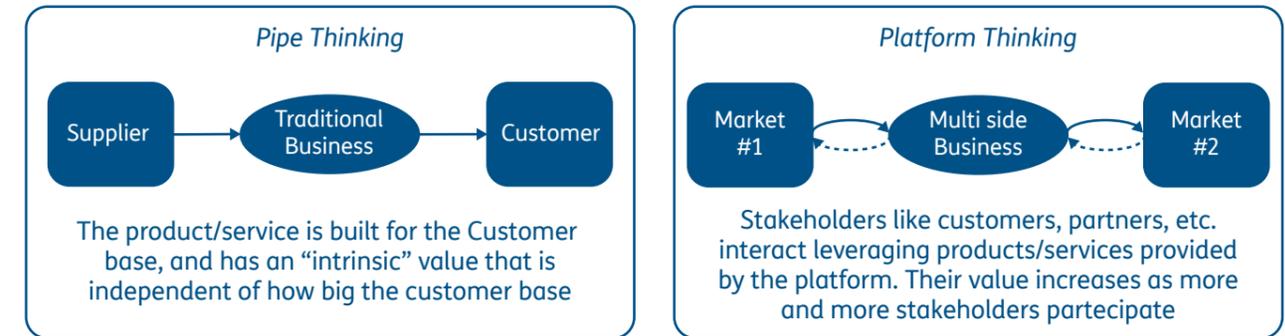
della rete viene nascosta, astratta e messa a disposizione di un ecosistema più ampio di stakeholder e fruitori del mondo consumer e business, che sviluppano i servizi in un ambiente aperto.

La Core Network del 5G è il primo fattore abilitante grazie al Network Slicing e al NEF (Network Exposure Function), che consentono di passare da una rete monolitica, verticale e chiusa in un set finito di servizi, ad una rete aperta, che espone funzionalità e capability per una piena programmabilità sia interna sia esterna da parte di terze parti.

È così quindi che la rete 5G può essere metaforicamente "affettata" per rispondere a esigenze di terze parti, che possono gestire e configurare direttamente le slice a loro dedicate per l'ottimizzazione dei propri servizi. È il caso dei servizi di public safety, dell'automotive, del mondo dell'IoT, etc. solo per fare esempi di alcuni servizi verticals che ad oggi hanno reti dedicate e che domani potranno essere portati sulla rete 5G, incrementando la flessibilità nella creazione, gestione e configurazione dei loro servizi.

Le tecnologie cloud native (descritte nel paragrafo successivo) aggiungono una infrastruttura dinamica, capace di fornire efficienza, scalabilità, esercibilità, resilienza, insieme a livelli elevati di automazione ed intelligenza rendendo l'intervento manuale una eccezione.

In questo senso pensiamo che 5G ed il Cloud sono il compelling event della trasformazione, capaci di abi-



litare un nuovo modello di business che ruota attorno al concetto di "platform thinking" e che crea una nuova modalità di interazione con i clienti, partners, suppliers e l'intera società basata sulla pervasività del digitale.

Il business digitale è la creazione di nuovi servizi, sfocando i confini tra i mondi digitali e fisici grazie alla convergenza di persone, affari e cose in una visione di ecosistema e partnership. Per fare alcuni esempi, il modello di business digitale delle città potrebbe essere rappresentato da città intelligenti con parcheggi, strutture e servizi collegati; quello delle Industrie potrebbe significare fabbriche connesse ad elevata automazione; altri esempi sono quelli delle aziende eco-sistemiche o di piattaforma, come Airbnb e Uber. Indipendentemente dalla varietà di business digitale, la piattaforma digitale del 5G cloud native, grazie alla programmabilità e all'apertura ha le componenti tecnologiche per supportare i nuovi modelli e diventare il centro di questa nuova trasformazione digitale. In TIM siamo

convinti che gli operatori siano in grado di offrire un grosso contributo, creando modelli che facilitino le interazioni tra gli stakeholder/fruitori, attraverso le capacità distintive dei Telco quali:

- L'elaborazione dei dati massivi di rete, ad es. le informazioni anonimizzate di mobilità, con viste consumer e business trasversali
- la presenza capillare ai margini della rete, vicino al cliente e sul territorio
- il controllo della latenza
- il monitoraggio delle prestazioni della rete end2end
- la qualità differenziata e la fornitura di livelli di affidabilità elevati per servizi mission critical.

Questo è un cambiamento radicale nel modo in cui i Telco guardano al mercato e induce un altrettanto radicale cambiamento nel modo in cui la tecnologia viene utilizzata, implementata e gestita.

Gli operatori di telecomunicazioni devono ripensare la loro architettura tecnica fondamentale, le funzionalità da implementare e come promuoverne la diffusione in un ecosistema

più ampio. Una buona piattaforma commerciale richiede supporto per un ambiente aperto ed un ecosistema eterogeneo, per creare valore per tutti i partecipanti. Il business orientato alla piattaforma sposta l'attenzione dal "fare" all'"abilitare", dal concetto di miglioramento di un prodotto, all'aggiungere nuove capacità di interazione e capability alla comunità che ruota attorno all'ecosistema.

In questo senso la 5G Digital Business Platform di TIM è la visione di piattaforma digitale Telco, che consente di creare un insieme interoperabile di servizi (sviluppati in modo aperto), che possano essere combinati per creare applicazioni e flussi di lavoro.

Questa evoluzione è strettamente connessa alla trasformazione "digitale" di tutte le attività e segmenti e non può essere realizzato senza un adeguato supporto tecnologico: l'uso delle tecnologie corrette è uno dei fattori abilitanti chiave della trasformazione della piattaforma di qualsiasi compagnia ed i Telco non fanno eccezione. In questo senso l'articolo

fornisce una descrizione dei primi passi concreti già intrapresi di questo percorso nella nuova direzione descritta.

Core network mobile verso il paradigma cloud-native

TIM ha avviato da tempo un percorso di trasformazione della rete fissa e mobile basato sull'adozione del paradigma NFV (*Network Functions Virtualisation*), il cui scopo è quello di separare il software che implementa una data funzionalità di rete dall'hardware sui cui viene eseguito.

In questo percorso, le appliance fisiche PNF (*Physical Network Function*) diventano funzione di rete virtuali VNF (*Virtual Network Function*), ovvero implementazioni software di funzioni di rete che possono essere rilasciate su un'infrastruttura NFV, con indubbi vantaggi in termini di **risparmi sugli investimenti**, derivanti dall'utilizzo di un'infrastruttura condivisa; **risparmi sui costi operativi** (manutenzione, energia e spazi) abilitati dalla separazione tra hardware e software; **incremento significativo della velocità di realizzazione** di nuovi servizi e prestazioni di rete attraverso l'automazione e l'impiego di soluzioni implementate integralmente in software.

Tuttavia, alcune delle promesse della virtualizzazione sono state

parzialmente disattese, sia per la mancanza di un modello univoco e maturo di gestione del ciclo di vita delle funzioni di rete, sia perché l'ecosistema dei fornitori, per rispondere velocemente alla forte spinta degli operatori sulla virtualizzazione di rete, ha spesso realizzato un semplice porting del proprio software da ambienti fisici basati su hardware specializzato a macchine virtuali in esecuzione su hardware general-purpose.

Oggi, la discontinuità tecnologica introdotta dal 5G offre l'opportunità di un nuovo, più deciso, passo in avanti in questo processo di trasformazione tecnologica, definito *Network Function Cloudification*, in cui il livello di *cloud nativeness* delle nuove funzioni di rete influisce fortemente sul grado di raggiungimento degli obiettivi di automazione e agilità.

La trasformazione non è puramente tecnologica, ma comprende anche un salto culturale e un reskilling delle persone coinvolte, nonché l'adozione di modelli operativi differenti rispetto al passato, come il paradigma Agile/DevOps ed i concetti di responsabilità condivisa.

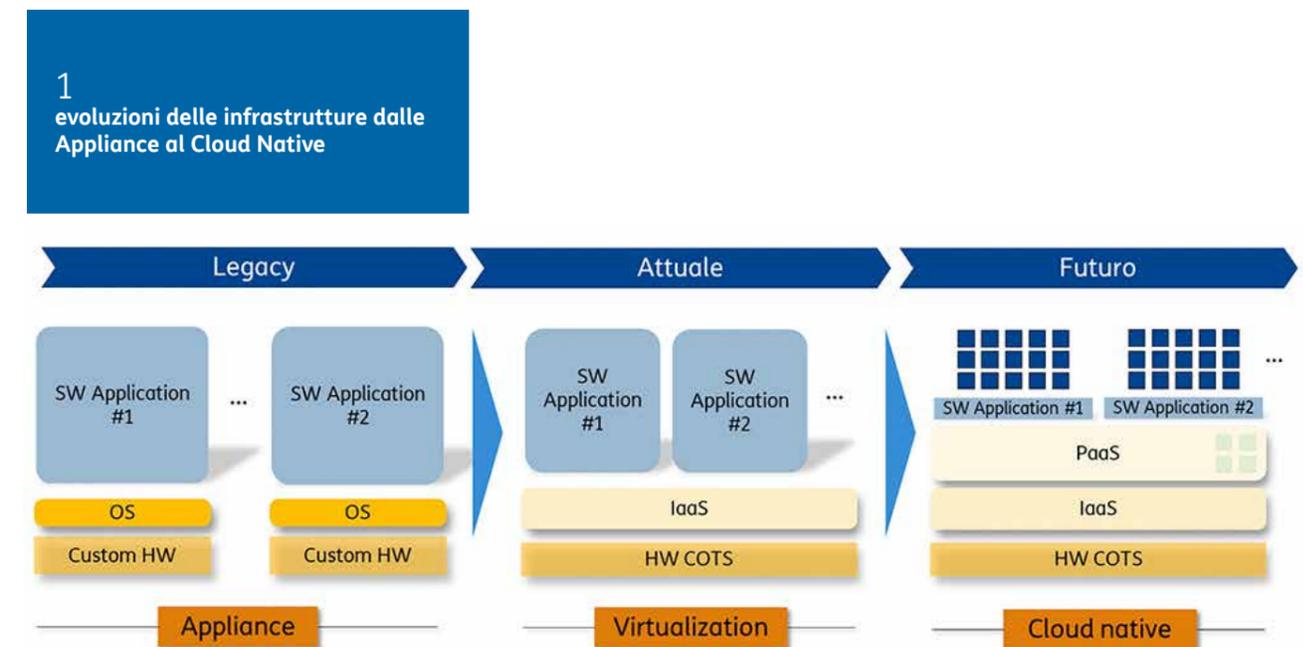
Un'applicazione si definisce Cloud Native se è in grado di trarre vantaggio dalla piattaforma che la ospita. La CNI (*Cloud Native Infrastructure*) è un'infrastruttura nascosta dietro utili astrazioni, controllata da API, gestita da software e che ha lo scopo di eseguire applicazioni.

Una infrastruttura con queste caratteristiche dà origine a un nuovo modello per la gestione delle applicazioni che risiedono su di essa, infatti le CNA (*Cloud Native Application*) devono essere progettate ed ingegnerizzate per dialogare con gli elementi di piattaforma per i fini di resilienza, agilità, operabilità e osservabilità.

TIM nella sua visione individua, tra gli altri, i seguenti principi fondamentali per realizzare una applicazione cloud native:

- **Design per Automazione:** tutte le caratteristiche di una CNA devono essere descritte tramite template, utilizzabili dalle soluzioni di automazione disponibili nella piattaforma, per gestire senza l'ausilio di interventi manuali tutte le fasi del ciclo di vita dell'applicazione ("zero-touch configuration"). Un corretto design abilita:

1. **Continuous Integration/Continuous Delivery:** per automatizzare la build, il testing e il deployment di pacchetti software attraverso tool come Jenkins e Spinnaker. Non solo: una catena CI/CD consente di automatizzare anche processi come il canary testing e il rollback
2. **Elasticità:** al variare del carico smaltito deve essere possibile ridurre o ampliare automaticamente, sulla base di metriche di tipo infrastrutturale e/o



applicativo, la capacità erogata da una CNA decrementando o incrementando il numero di moduli che implementano specifiche componenti dell'applicazione (scale in/out)

3. **Osservabilità e recovery automatico:** una CNA deve essere osservabile, ovvero deve offrire meccanismi di monitoraggio e notifica che permettano di verificarne lo stato di salute, rilevare prontamente eventuali problemi ed eventualmente innescare automaticamente meccanismi di healing mirati a ripristinare la piena operatività dell'applicazione. La CNA gestisce i fallimenti piuttosto che cercare di

prevenirli. Logging e monitoraggio possono fornire preziose informazioni sull'uso del sistema e sul comportamento e possono essere utilizzati in forma aggregata per fornire una misura della salute generale del sistema. È la corretta manipolazione di queste informazioni che può innescare degli automatismi.

- **Architettura Modulare e Stateless:** i moduli che compongono una CNA devono essere il più possibile disaccoppiati tra loro in termini di sviluppo, dispiegamento, testing e gestione del ciclo di vita. Le comunicazioni tra moduli devono essere basate su protocolli standard noti alla CNI (e.g. utilizzo di API

REST e HTTP invece di Diameter o altri protocolli "telco"), in quanto questo è un fattore abilitante per poter sfruttare i servizi a catalogo messi a disposizione dalla piattaforma. Esiste ovviamente un trade-off tra i vantaggi sopra citati e un overhead crescente che deriva dalle molteplici interazioni tra i vari moduli software. La memorizzazione dello stato del sistema è l'aspetto più difficile da implementare per un'architettura cloud native e distribuita. Pertanto, è necessario progettare il sistema in modo che sia intenzionale quando e come archiviare lo stato. Componenti stateless consentono di:

1. **Scalare** rapidamente;
2. **Riparare** una componente fallita spegnendola e

istanziandola nuovamente;

3. Eseguire in modo semplice il **Roll-back** di un intero ambiente ad una versione precedente;

4. **Load-Balancing** su più istanze, dato che ogni componente è in grado di gestire ogni richiesta.

- **Riuso dei servizi offerti dalla piattaforma.** Per svolgere le proprie funzioni una CNA deve potersi avvalere dei servizi a catalogo messi a disposizione dalla CNI (e.g. Service Discovery, Identity and Access Management, “DB as a Service”, “Load Balancer as a service”, “Firewall as a Service”), in modo da consentire di delegare alcune funzioni “comuni” alla CNI, lasciando quindi lo sviluppatore della CNA libero di concentrarsi sullo sviluppo delle funzionalità con valore di business.

Seppur non obbligatorio, l'utilizzo di **Microservizi e Container** da una parte e l'approccio **Dichiarativo** dall'altro, si stanno affermando come standard de facto per implementare le CNA.

La core network 5G è stata progettata fin dall'inizio privilegiando scelte tecniche a supporto di funzionalità cardine come il network slicing, che richiede la capacità di creare rapidamente istanze di rete dedicate a specifici clienti e/o scenari di servizio, o la network exposure che presenta la core 5G come

una piattaforma aperta agli utilizzatori dei servizi.

La Service Based Architecture (SBA) introduce un nuovo approccio al design dell'architettura della core mobile, più allineato ai paradigmi delle architetture cloud. In questo contesto l'adozione di tecnologie cloud native appare particolarmente adatta, poiché fornisce ulteriori elementi di flessibilità nel dispiegamento di un sistema concepito per essere altamente dinamico e supportare nuovi modelli di servizio con cicli di vita radicalmente diversi dal passato.

Trial in lab della core network 5G cloud native

TIM sta conducendo delle attività di trial sulla core network 5G realizzata secondo i principi cloud native esposti.

Una prima fase è consistita nella sperimentazione della 5GC fornita da Ericsson.

In particolare, Ericsson ha reso disponibile una 5GC composta dai building block essenziali della architettura 3GPP [1] descritti in Figura 2.

Le Network Function componenti la core network sono quelle relative ai domini Data Layer (UDM, AUSF, UDM, NRF) e Packet Core (AMF, SMF, UPF), mentre l'accesso in tecnologia NR SA (New Radio Stand Alone) ed il corrispondente device

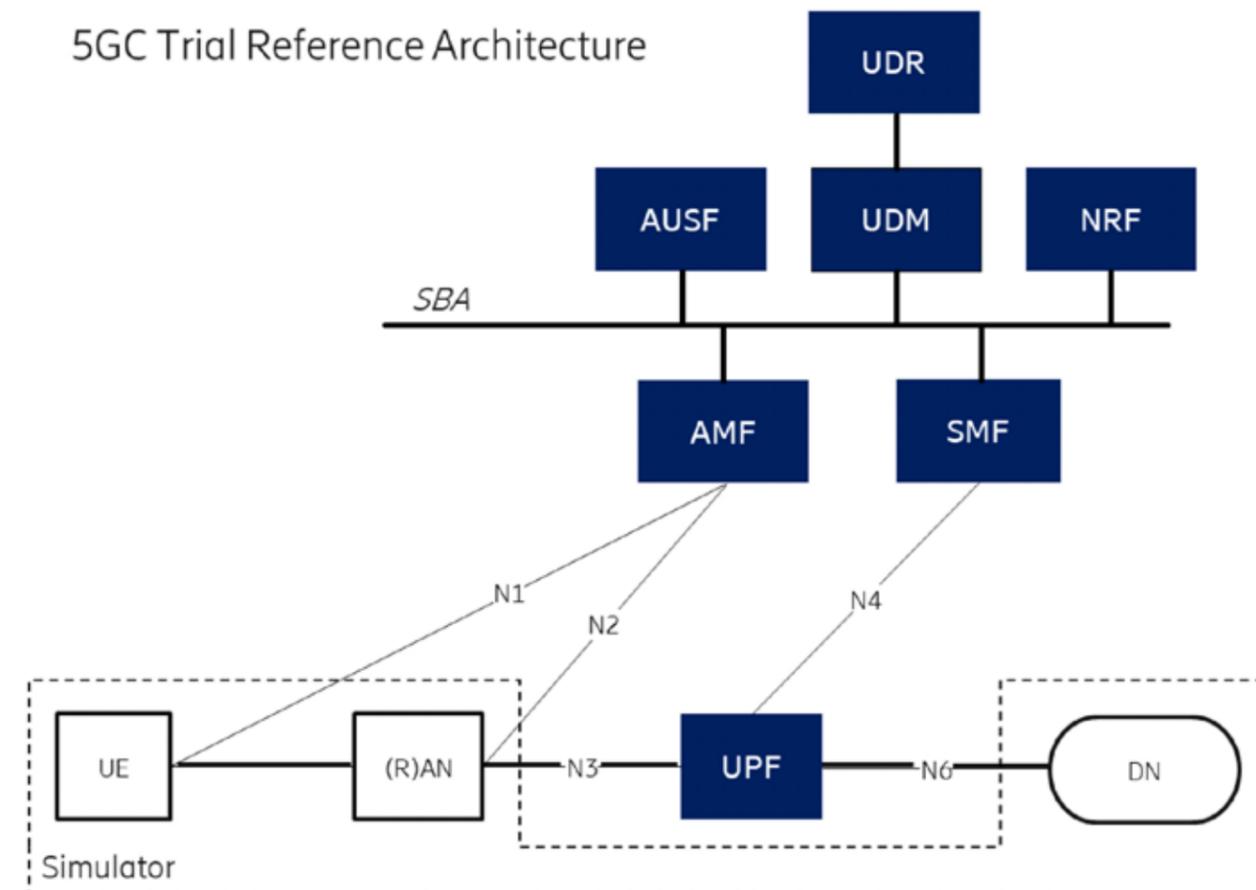
UE (User Equipment) sono stati implementati con un simulatore.

La soluzione di trial ha un piano di evoluzione che introduce nuove Network Function 5GC e quelle EPC, realizzando così una dual mode core, che sarà ugualmente oggetto di trial nel corso dell'anno. La 5GC è dispiegata sulla piattaforma Ericsson Cloud Container Distribution (CCD), a sua volta dispiegata su un ambiente RHOSP 10 (Red Hat OpenStack 10). La CCD agisce come layer Container as a Service (CaaS) per la gestione e l'orchestrazione dei container (deployment, scaling, management), ed è realizzata utilizzando Kubernetes come elemento centrale, la cui composizione (ad esempio in termini di plugin per networking e storage) può essere gestita in maniera flessibile [2].

Il trial ha coperto tanto l'aspetto funzionale, ovvero le procedure relative alla parte applicativa delle Network Functions, quanto quello non funzionale, ovvero le procedure relative Cloud Native Infrastructure.

Più in dettaglio, i test funzionali hanno coperto le aree dei Network Service (NF registration/deregistration, update,...), delle NF selection, del Provisioning, del Mobility e del Session Management, della Security, andando a provare tanto le procedure più convenzionali di messa in piedi delle sessioni e di mobilità, quanto i nuovi aspetti legati alla SBA, che richiedono procedure di registrazione e discovery delle Network Functions.

5GC Trial Reference Architecture



2
5GC trial architecture

I test non funzionali sono stati relativi al dispiegamento e configurazione delle NF utilizzando template Helm; alle procedure di scale in e scale out e self-healing, sfruttando le capability di Kubernetes; alla raccolta di metriche di performance utilizzando tool open source (Prometheus e Grafana); alle procedure di rolling update delle NF, ovvero l'upgrade delle Network Function senza il fuori servizio delle stesse. Sebbene si tratti di soluzioni nelle loro release iniziali, le potenzialità dell'ambiente cloud native sono risultate già evidenti, in particolare per quanto riguarda il dispiega-

mento della intera core network 5GC, con una corretta configurazione applicativa, realizzato in pochi minuti.

Allo stesso modo è di grande interesse la flessibilità della gestione operativa, che permette di installare, aggiornare e riconfigurare la rete con procedure automatizzate, controllate da remoto.

Conclusioni

L'adozione delle tecnologie cloud native rappresenta un passo fon-

damentale per lo sviluppo dei nuovi servizi 5G, consentendo una gestione operativa avanzata della rete, che ottimizza le risorse di rete, che permette di dispiegare e gestire funzioni di rete e servizi in modo dinamico, immediato e automatico, e garantisce l'alta qualità del servizio richiesto, grazie all'elasticità intrinseca a tali tecnologie.

Oltre a questo, il cloud native rappresenta l'abilitatore di nuovi modelli di servizio, dove il service

provider non offre ai clienti servizi statici ed immutabili, ma, al contrario, servizi aperti e flessibili, che si sviluppano continuamente grazie al contributo della community dei clienti e degli sviluppatori.

TIM ha sperimentato con successo - primo operatore in Italia e tra i primi al mondo grazie alla collaborazione con Ericsson - la core network 5G cloud native presso i propri laboratori di Torino, accelerando così il percorso di evolu-

zione della rete 5G di TIM verso le tecnologie cloud native. Il trial ha incluso le procedure di controllo e trasmissione del traffico definite dagli standard internazionali 3GPP, le procedure di creazione e configurazione delle funzioni di rete, di scale-in e scale-out delle stesse, insieme alla capacità della piattaforma di reagire automaticamente ai guasti e di eseguire un upgrade della Core Network mentre è operativa, secondo le tecno-

logie innovative dell'ecosistema cloud native.

In questa prospettiva, le tecnologie cloud native saranno l'ambiente di sviluppo dei sistemi 5G futuri e rappresentano la base su cui realizzare una piattaforma digitale di business del 5G (5G Digital Business Platform) su cui poter sviluppare nuovi modelli di business e servizi, valorizzando al massimo le potenzialità del 5G ■

References

[1] TS 23.501 - System Architecture for the 5G System; Stage 2

[2] <https://www.ericsson.com/en/portfolio/digital-services/cloud-infrastructure/cloud-container-distribution>



Giuseppe Catalano giuseppe.catalano@telecomitalia.it

Si occupa di pianificazione di rete ed architetture con particolare esperienza sui temi del mobile, sviluppato come elemento ricorrente nei diversi gruppi di cui ha fatto parte, a partire dai modelli e strumenti di misura della QoE, passando per l'innovazione dell'accesso wireless e della core mobile, fino alle Strategie e, oggi, nel Planning TIM. Ha sviluppato un'esperienza trasversale occupandosi di progetti di ricerca, enti di standard (delegato al RAN WG3 del 3GPP dal 2006 al 2016, GSMA 2013-2014), brevetti (3) e docenze sia interne al gruppo sia presso Università (Politecnico di Torino, Scuola Superiore Sant'Anna di Pisa), ed allo stesso tempo di attività di testing on field e nei Testing Lab di Torino, di dispiegamento in rete di tecnologie innovative, di redazione di RFI/RFQ, attività svolte sia in ambito Domestico sia presso le consociate del Gruppo (TIM Brasil, TIM Hellas, Entel Chile, Telecom Personal, Amena). Ingegnere delle Telecomunicazioni (Università di Pisa), è in Telecom Italia dal 1998, vive e lavora a Torino ■



Daniele Franceschini daniele.franceschini@telecomitalia.it

ingegnere delle Telecomunicazioni, è in Telecom Italia dal 1998, attualmente responsabile del gruppo di Technology, Planning and Architecture per la pianificazione tecnico-economica pluriennale di TIM e della definizione delle architetture di riferimento per tutti i segmenti di rete. In tale ambito si occupa della definizione del Piano Tecnologico di Gruppo e del Piano di Sviluppo di Technology nonché il governo dei piani di efficienza tecnica e di ammodernamento delle reti ■



Ivano Guardini ivano.guardini@telecomitalia.it

ingegnere elettronico in Telecomunicazioni, nel 1995 entra in Telecom Italia, dove attualmente ricopre il ruolo di responsabile della funzione Network Function Cloudification and Automation. Dal suo ingresso in Telecom Italia si è occupato di networking IP ed innovazione della core network mobile a pacchetto, attività che lo ha portato ad approfondire temi quali IPv6, l'interlavoro con accessi Wi-Fi, le reti mobili ad-hoc e l'evoluzione verso i sistemi mobili di quarta e quinta generazione. Negli anni ha accumulato un'ampia esperienza negli enti di standardizzazione, in particolare IETF e 3GPP. Dal 2011 al 2015 è stato Vice Chair del gruppo 3GPP SA WG2, che standardizza l'architettura di sistema della rete mobile. Più recentemente si è occupato di ingegneria e sviluppo di soluzioni per la virtualizzazione, l'automazione e l'evoluzione verso il paradigma cloud-native delle funzioni di rete. Questi temi costituiscono ad oggi il suo principale ambito di attività ■



Simone Roma simone.roma@telecomitalia.it

Ingegnere delle Telecomunicazioni, è in Telecom Italia dal 2015, dove attualmente ricopre il ruolo di Network Evolution Engineer nella funzione Network Function Cloudification and Automation. Nel 2015 ha conseguito il titolo di Dottore di Ricerca presso l'Università di Pisa, in collaborazione con Telecom Italia, nell'ambito della progettazione e ottimizzazione della rete di accesso radiomobile 4G. Dal suo ingresso in TIM si è occupato di design di architetture cloud, dello scouting di tecnologie innovative per piattaforme di virtualizzazione e di attività di testing in laboratorio e on field. Recentemente ha preso parte al gruppo di lavoro per la definizione dei requisiti architetturali per la progettazione e dispiegamento della 5G Digital Business Platform di TIM ■

OPEN SELF-ORGANIZING NETWORK: A CONTINUOUS DEVELOPMENT FOR RADIO ACCESS NETWORK PERFORMANCE OPTIMIZATION

Paola Bertotto, Francesco Epifani, Michele Ludovico, Giovanna Zarba



TIM RAN digitalization Path

The evolution of radio access requires a new approach for network creation and management, assuring quality-of-experience to the customer, even with increasing traffic and network complexity, and sustainable costs for the operator. **Digitalization and automation** are key elements to address this challenge.

TIM is engaged in a **transformation process** which, applied in the radio access domain, involves the introduction of open and cloud software architectures that enable the automation of network functions.

Two main initiatives were exploited in the last few years:

- Starting in 2015, TIM defined an “Open SON” reference architecture based on API (*Application Programming Interface*) availability for the exchange of information among the involved functional blocks [1];
- Starting in 2016, TIM, as a worldwide pioneer, tested and deployed Virtual RAN (vRAN) with several partners [2] [3].

The “digitalization” (i.e. automation and softwarization) path of TIM is represented in *Figure 1*.

It is worth noticing how digitalization proceeds “from the center to the edge”, i.e. from the centralized NMS (*Network Management Systems*) going down to the radio access sites, leveraging on cloud infrastructure and operators’ software

development and integration capabilities.

Open SON represents a **first automation step**: starting from already existing design and optimization tools, an automation framework was built by:

- Introduction of centralized SON (cSON) software modules, integrated with existing RAN architectures (and existing distributed SON features - dSON) and providing API for “closed loop” interaction with TIM design and optimization tools;
- Usage of TIM cloud infrastructure for onboarding both centralized SON and design/optimization tools. In particular, NFV-I (*Network Function Virtualization Infrastructure*) was exploited.

With RAN virtualization (**second step**), part of the baseband components of the BBUs (*Base Band Units*) nodes become virtualized functions integrated on NFV as well (indicated as “Centralized Units”), while lowest protocol layers (e.g. physical level) still exploit physical dedicated HW, together with the RRUs (*Remote Radio Units*), installed next to the antennas. Network Management Systems, on the other hand, evolve to enable orchestration and workflow management within RAN domain. The digitalization path completes with the **third step**, aligned with O-RAN (Open RAN) architecture [4]:

- Virtualization of the whole base-band protocols on NFV (*Network Function Virtualiza-*

tion), including physical layer. Baseband is split into Centralized units (CU) and Distributed units (DU). The only physical component is the remote Unit implementing Radio frequency (RF) transmission and reception;

- Architecture with open interfaces, from the fronthauling (i.e. interconnection between RU and DU) and the mid-hauling, up to the management interfaces, thus enabling the growth of “an ecosystem of innovative new products that will form the underpinnings of the multi-vendor, interoperable, autonomous RAN” [4];
- RAN intelligent controller introduction, that can be considered as an evolution of Centralized SON module in previous TIM architectures, providing Open APIs towards higher level systems (Management & Orchestration according to ONAP architectures) [5].

The “cloudification” of RAN protocols down to the physical layer implies the availability of distributed data centers, designed taking into account band/latency requirements of fronthauling interconnections. This requirement can be managed in synergy with Edge Computing, also leading to a reduction of the distance between the network computational capability and the customer. For such a reason, RAN virtualization and Edge Computing can be seen as main

drivers for distributed data centers realization.

Open RAN introduction is the lever to transform Network Management systems and processes, by fully exploiting open environment and HW/SW decoupling. Automation based on the evolution of SON paradigm is even more crucial in a multi-vendor environment that will require enhanced control by the Operator. In this context the capability of “in house development” of SW modules by the operator can be a crucial asset, enabling also introduction of

advanced “machine learning” capabilities, exploiting an “augmented intelligence” approach (see [6] for details)

The Open SON architecture

The SON (*Self-Organizing Network*) concept was introduced by 3GPP since the initial releases of LTE standardization, in order to help mobile access network operators

to cope with the increasing complexity of configuration, optimization and assurance processes.

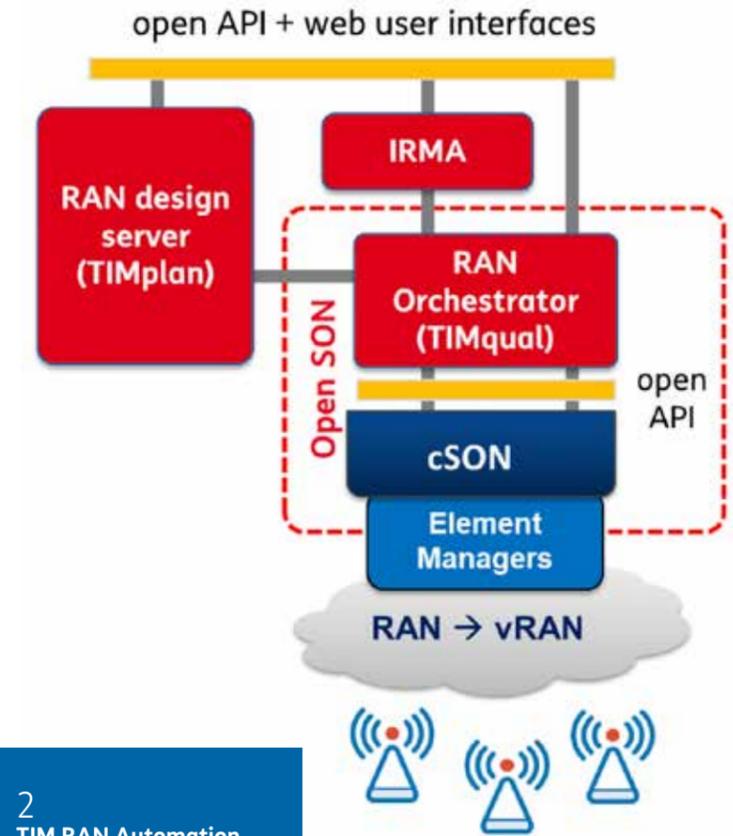
TIM has played an active role in 3GPP SON standardization activities since their beginning, combining the innovation vision with the “in field” experience in 2G/3G/LTE radio access design and optimization [7]. The work of TIM has been mainly focused in promoting open, interoperable and flexible solutions.

Based on this experience, in the last years TIM defined the “Open

SON” approach, including “in house” software development, as a key initiative in the path toward virtualization and automation of the radio access network.

The “Open SON” architecture can be applied to both the legacy 2G/3G/LTE radio access networks and the evolution towards the so called “Virtual RAN”, based on the NFV paradigm within the radio access domain.

The “Open SON” approach enables two different “closed loop” optimization processes:

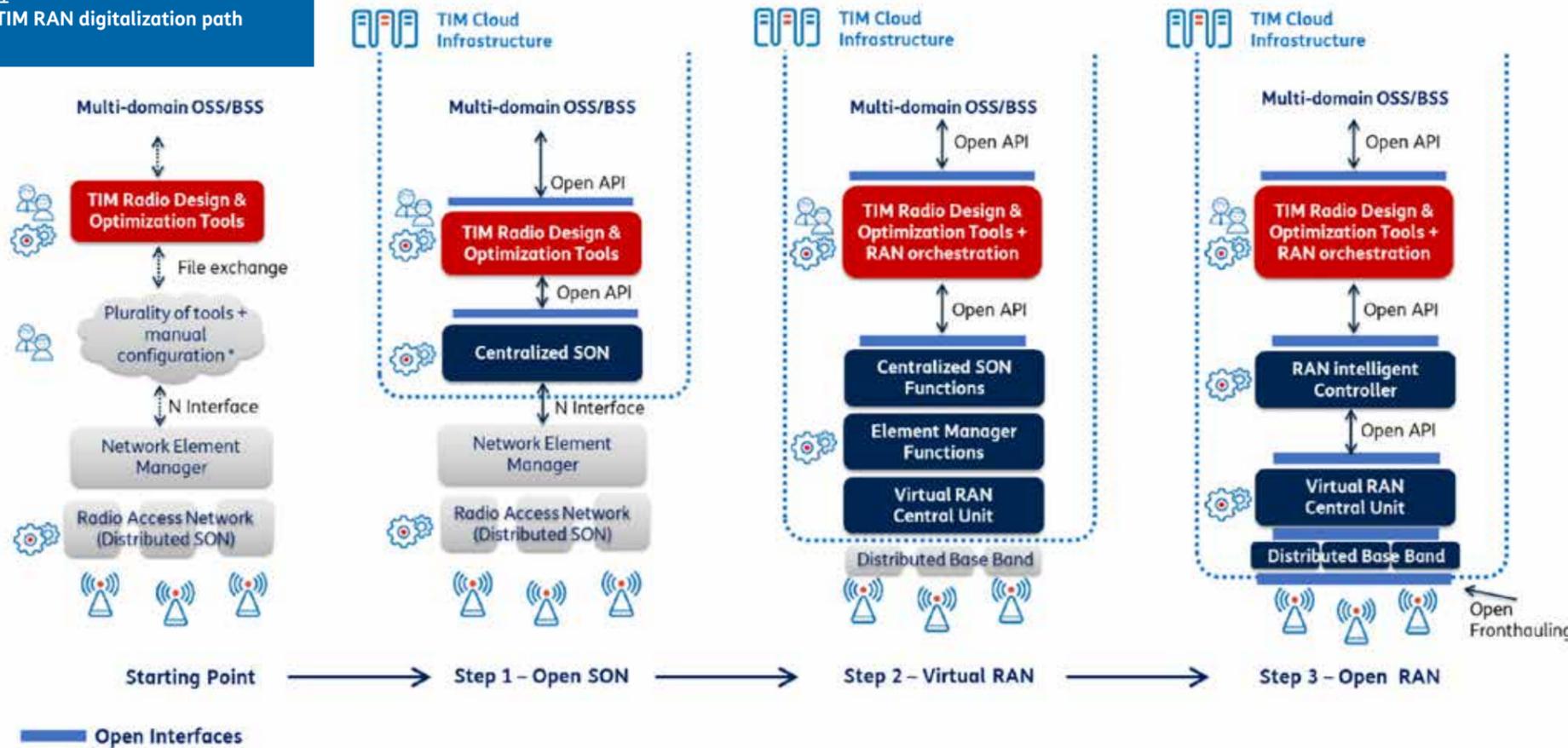


2 TIM RAN Automation Architecture

- “Automatic Closed Loop” addressing basic optimization and configuration activities, for which a complete automation can be envisaged;
- “Human Closed Loop” addressing more complex activities, where TIM radio access specialists are supported by software applications (mainly developed “in house”), enabling also an effective control of automatic functions (through API).

The architecture defined by TIM is summarized *Figure 2*. Centralized SON Solution was specified as an open platform: modular, flexible, programmable (through API). Baseline platform was integrated (through NFV) in 2017, with the first use cases. Since then, new use cases can be easily introduced exploiting Open API and leveraging on development of radio access design and optimization tool in TIM domain. In particular:

1 TIM RAN digitalization path



USER THROUGHPUT IMPROVEMENT EXPLOITING MLB ALGORITHMS

giuseppe.barillaro@telecomitalia.it, gianpiero.galati@telecomitalia.it, davide.missan@telecomitalia.it, isabella.vario@telecomitalia.it

MLB (*Mobility Load Balancing*) is an example of implementation of SON use case through the CI / CD process (see *Figure 4*).

In this context, the patented algorithm exploited by TIM [note 1] selects as target cells those characterized by a carried traffic greater than the admissible maximum value (i.e. defined as “loaded cells”); these cells may suffer of performance degradation due to “radio resource sharing” among the multiple concurrent users. To mitigate such an effect, the adopted algorithm identifies the neighbouring cells that can act as “support cells” (i.e. able to “help” loaded cells), (being characterized by a limited carried traffic and by significant overlapping area with the addressed loaded cells).

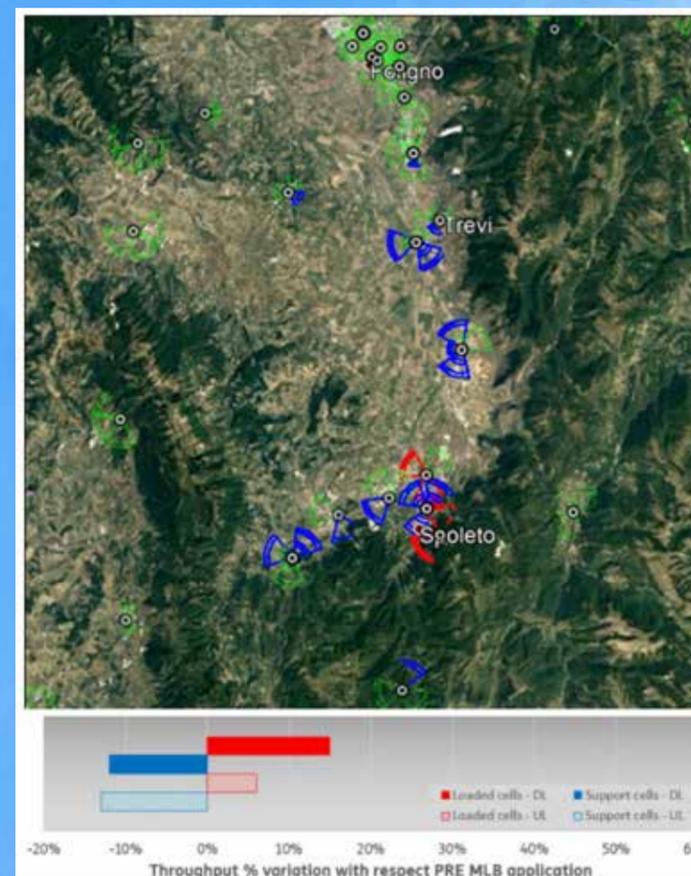
In the MLB feasibility study, both idle and connected mode mobility parameters were analyzed, taking into account already active radio resource management algorithms (depending on RAN vendor implementation). Already available SON APIs were exploited for all the three RAN vendors operating in TIM radio access network, without need for further implementation. During the testing phase, the involved mobility parameters were iteratively tuned for both “loaded cells” and “support cells” to optimize automatic selection algorithm in order to achieve the expected benefits.

Results of MLB application in Rome (dense urban area) for seven loaded cells (colored in red) with the “support” cells identified by the load balancing algorithm (colored in blue); improvements (%) for downlink (filled curves) and uplink (dotted curves) user throughput for the loaded cells (red) and support



A
MLB Cluster tested in Rome

cells (blue) are also reported. It is worth noticing that performance improvements for the loaded cells are obtained without penalizing the performance of the support cells chosen for balancing the traffic. In *Figure B* analogous results are reported for four loaded cells in Spoleto area (small town in rural environment). Performance improvements depend essentially on the load of the other cells in the area; in some cases, additional algorithms can be applied in cooperation with MLB in order to further enhance performance by optimizing cell footprints (e.g. CCO that acts on electrical tilts). After the field testing phase, the MLB algorithm was integrated in TIM tools exploiting already available SON APIs for each of the RAN vendors implementations ■

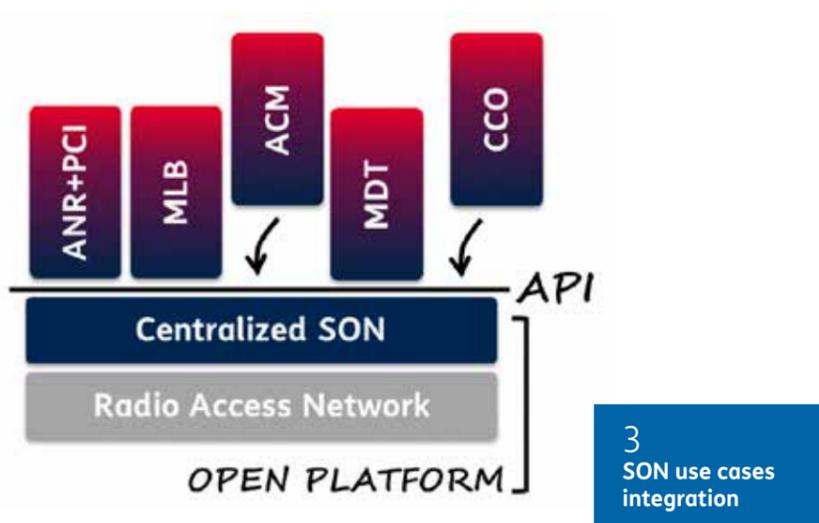


B
MLB Cluster tested in Spoleto

- TIMplan server, described in [6] providing radio design data;
- IRMA: implementing configuration policies (macro-optimization);
- TIMqual: providing micro-optimization (i.e. cell-by-cell optimization) and overall process orchestration.

The interworking between design, optimization and configuration tools is a crucial aspect of the proposed architecture, due to the intrinsic characteristics of Radio Access management: each new site leads to changes in mobility management, traffic distribution and interference conditions also for the neighbouring sites. In this context, the exploitation of accurate coverage/interference simulations [6], specific peculiarity of TIM approach, is fundamental. The following use cases have been considered so far:

- **ANR+PCI** (*Automatic Neighbours Relations*): neighbouring lists (defined by distributed SON) are controlled by centralized SON and PCI conflicts (caused by new neighbours) are automatically managed;
- **ACM** (*Automatic Configuration Management*): automatic check and enforcement of configuration parameters (based on configuration guidelines);
- **MLB** (*Mobility Load Balancing*): Inter-cell load balancing to redistribute traffic to avoid/delay sites congestion;



- **CCO** (*Coverage/Capacity Optimization*): based on tilt adjustment and interference control;
- **Automatic lock/unlock**: automatic execution of lock/unlock of cells/nodes.
- **Auto-provisioning** of new nodes (also known as “zero touch” provisioning).

As a further use case, **MDT** (*Minimization of Drive Test*) was introduced, also exploiting the same reference architecture.

Optimization use cases

TIM methodology for Open SON use cases implementation is depicted in Figure 4, where three interconnected loops are represented:

- “Inner” loop, describing core “DevOps” activities carried out “in house” by TIM for the Con-

tinuous Development (CD) of design, configuration and optimization tools (i.e. TIMplan, TIMqual, IRMA). SW modules by external providers can be also integrated. SW solutions are developed starting by a “Minimum Viable Solution” that is tested in field, followed by an iterative CI/CD process, tuning both algorithms performance and tools usability;

- “Outer” loop, describing the exploitation of SON use cases, leading to process evolution (through automation): new use cases introduction is typically triggered by field activities feedbacks and/or technology innovation opportunities;
- “External” loop, representing overall market and ecosystem evolution. TIM contributes also to this global evolutionary path through participation in standardization bodies (e.g. 3GPP,

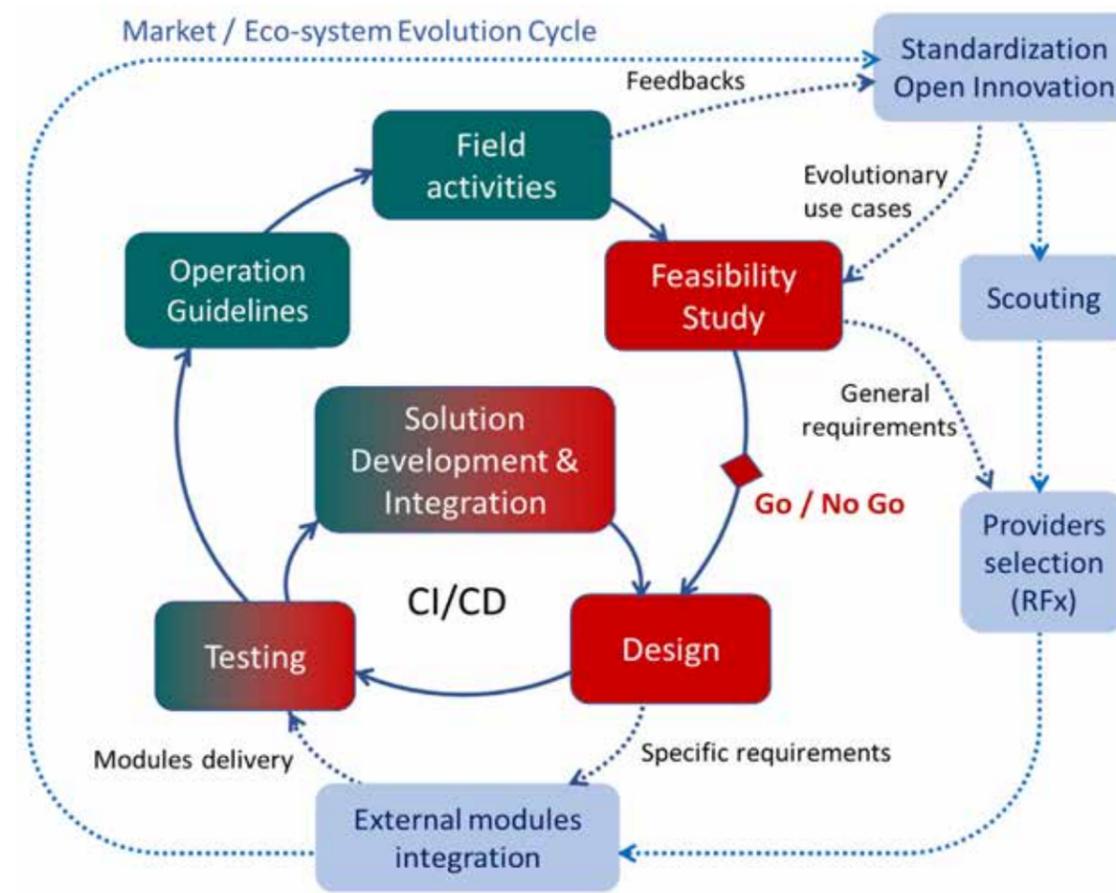
O-RAN...) and other innovation initiatives (European projects, trials...), but also through continuous co-operation with partners (selected through scouting and evaluation activities). To better clarify the process described in Figure 4, specific application examples can be considered. The workflow for setting up a use case can start from a requirement coming from field activities, such

as in the case of MLB, which addresses over-loaded cells (identified through KPI analysis), or when new opportunities are made available by technology evolution, as in the case of MDT feature, which provides a new methodology for “in-field” performance monitoring. The second step is the feasibility study, which analyses the matching between field requirements and evolutionary “know-how” in order to

decide whether and how to proceed with the implementation (GO / No GO decision).

Several aspects are evaluated during this stage:

- **Stakeholders**: involved subjects for each phase of the work are individuated, analyzing whether they belong to internal structures of the organization, beneficiaries of the solution or development structures, or ex-



4 SON use cases implementation through CI/CD (Continuous development, Continuous integration) process

MINIMIZATION OF DRIVE TESTS

EXPLOITATION IN SITE CERTIFICATION ACTIVITY

antonio.bernini@telecomitalia.it, dario.boccia@telecomitalia.it

TIM adopts a “Quality by design” approach to provide best-in-class performance to the customer: using proprietary tools (TIMplan) coverage and interference are controlled for each new installed node. Advanced simulation models are used in the radio design phase, providing accurate high-resolution coverage/interference predictions (e.g. on 10x10 m bins resolution).

For this reason, the last step in the network design process is the sites certification activity, verifying the coherence between the radio project and the “in field” implementation and assessing coverage targets defined during the radioelectric design phase.

Examples of issues analyzed during certification:

- Crossed cables (between cells and/or MIMO ports);
- Mechanical/electrical tilt misalignment;
- Interference with neighbouring sites.

Site certification is typically carried out using specific measurements collected with drive tests, compared with the simulations performed by TIM tool TIMplan.

Certification activity can confirm the correct implementation of the design or alternatively lead to further correction and optimization, e.g. exploiting SON uses cases such as CCO (*coverage and Capacity Optimization*).

Thanks to the MDT functionality (3GPP MDT - Minimization Drive Test), available today in the radio access

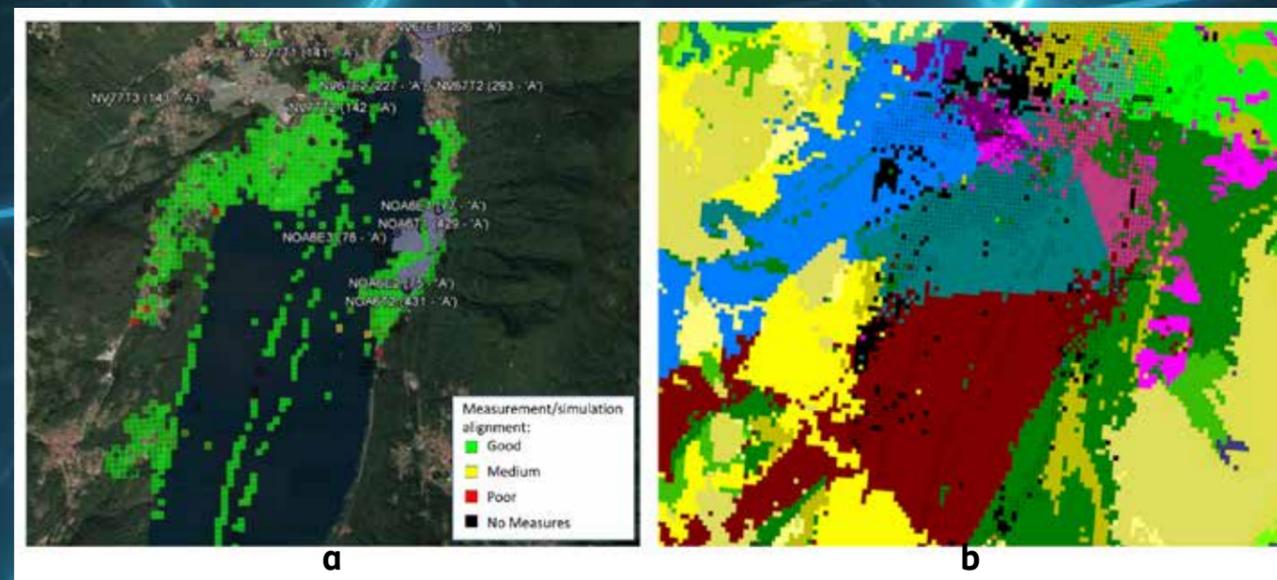
network, the certification activity can be carried out from the office, through measurements derived from tracing events collected from user equipments.

3GPP (in TS 37.320, 32.422, 32.423) standardized the MDT feature, adding the GPS position to the set of radio measurements that user equipments send to the network thus enabling accurate traces localization.

Certification activity target is the entire pool of newly activated sites, including Macro Sites, Small Cells, temporary installations (T-MAST [note 2]).

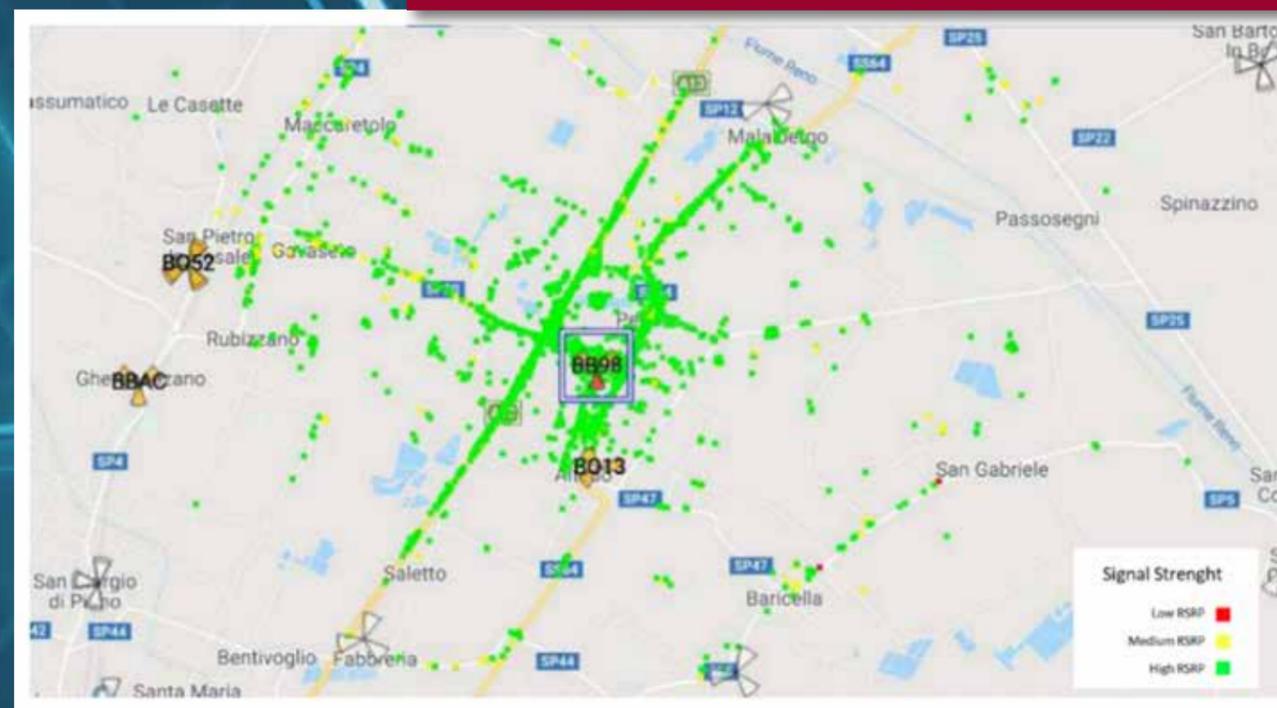
The exploitation of MDT allowed a significant increase of certification activities with the same effort (the recovery of efficiency exceeds 20% in the regions where the use of MDT is already up and running), thanks to the reduction of time-consuming activities such as drive tests planning (routes individuation), execution (including driving from the TIM premises to the target areas) and measurements collection. The observed efficiency in site certification activity represents a common trend for SON use cases.

Automation provides process improvements and reduces time-consuming activities, enhancing operational efficiency and re-focusing specialists from routine to “know-how intensive” activities ■



C Comparison between TIMplan simulated data and MDT data, binned on 10x10 m pixels (Orta lake area): a) received power; b) best server

D Site certification geo-localized KPIs (Bologna area): signal strength indicator based on RSRP reported by UEs

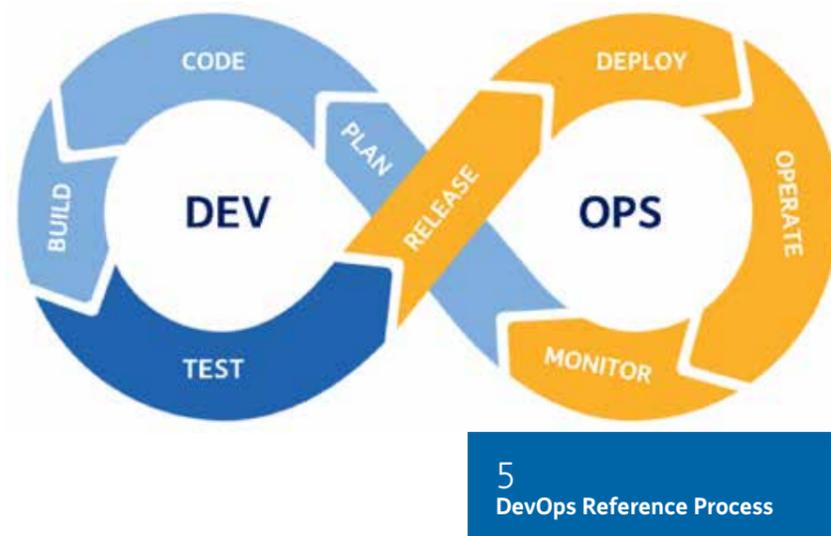


ternal subjects that provide tools and / or re-usable knowledge;

- **Technical feasibility:** this point focuses specifically on the technical aspects. The technical viability of proposed use case is assessed by evaluating the architecture, the availability of “actionable” network parameters, the KPIs to be considered for evaluating the benefits of the use cases;
- **High level solution definition:** the high-level design of the software solution is carried out. Needed SW modules may be already available or may need to be added to the existing framework: new modules are typically implemented leveraging on “in house” development capabilities. External SW modules may be integrated as well, selected in cooperation with procurement departments;
- **Cost sustainability:** it assesses the economic viability of a proposed use case by evaluating the set-up costs, operating expenses and required efforts with respect to the benefits (Balanced Capex /Opex /Efforts and Benefits) The “cost of not doing” is also taken into account in this phase.

For example, in the case of MDT the feasibility was based on:

- the availability of the MDT feature at the radio level (included in RAN vendors solutions);



- the development of a platform for the collection and processing of measures;
- the definition of a methodology, leveraging internal SW tools, for managing the processed data.

Once the feasibility study is completed (i.e. with a “go” decision), the “inner” loop is activated, exploiting the paradigm of “continuous development”.

As a matter of fact, fast evolving technologies lead to a major shift in the design and delivery of SW solutions. The waterfall paradigm cannot be applied anymore, because the context itself is continuously evolving, leading to the need of a SW production flowchart based on continuous development, which means continuous integration, continuous testing, continuous delivery. For the same reason, a strict cooperation is needed, even in the SW design/development phase, among developers and final users (typically

belonging to the Operations departments).

For such a reason the “DevOps” paradigm, already adopted for the development of planning and optimization tools, quite perfectly matches the methodology adopted by TIM for SON use cases definition (Figure 5).

During the Design phase, a “Minimum Viable Solution” is defined to start as soon as possible with “in field” activities and then to learn and to continuously improve the solution. This way of working conflicts with the linear, iterative design process that works towards a perfect problem-solution fit. It is a remarkable cultural change, which requires fast implementation rather than risk reduction, because the real use of the solution and the evaluation of its benefits are visible only in field, since it is impossible to capture “a priori” all the possible behaviors.

In the “digital world”, sub-optimal solutions may be released and then adjusted in continuous way, considering feedbacks from the users, and thus making the solutions more and more adherent to the operational needs.

“Learning in field” is the key to assess the real “added value” for which the solution was initially designed. This working method involves:

- a fast development methodology, making available a preliminary release of the use case;
- a fast testing carried out together with operation departments (does the solution work?);
- an efficient “knowledge sharing” between the development team and the operations team for the functional testing (does the solution provide the expected benefits?).

As a final step, “Operation guidelines” phase includes the final solution deployment and the drafting of the guidelines for recurrent use, encompassing both the operational steps and the related responsibilities (e.g. through the definition of a proper Responsibility assignment matrix). This phase leads to a continuous cycle too, since in-field deployment, in an evolving scenario, continuously requires process adjustments and SW solutions upgrades.

The methodology application results are reported in the “insights” sections, providing network performance improvements (e.g for MLB: Mobile Load Balancing) and opera-

tional efficiency improvements (e.g. for MDT exploitation in Certification Activities). A third application example related to VoLTE performance improvements derived from PCI optimization can be found in [1]

Conclusions and acknowledgments

TIM Open SON initiative was described, by focusing on three main aspects:

- Automation is a crucial factor in TIM RAN digitalization path and “in house” development of design, optimization, configuration and orchestration tools is a key asset for “taking the control” in the Operator domain, both in the present and in the evolutionary “Open RAN” scenarios.
- Continuous development, based on DevOps approach, has been adopted also for Open SON to fully exploits TIM’s know how on Radio Access and to evolve automation tools based on real in-field “experience”
- For each use case introduced, network performance and operational efficiency benefits have to be monitored in order to enable a continuous tuning/improvement of deployed solutions.

The described methodology and the reported use cases are the results of an extensive cross-functional ac-

tivity (still on-going) which involved many TIM colleagues, belonging to several TIM areas (innovation, planning, procurement, engineering, SW development, NFV implementation, network development, network maintenance, in field activity coordination...). The authors gratefully acknowledge the valuable contribution of each of the involved colleagues ■

References

1. “DigiRAN: il valore dell’automazione nell’accesso radio” Notiziario Tecnico TIM. N. 1 2018. <https://www.telecomitalia.com/tit/it/notiziariotecnico/edizioni-2018/n-1-2018/N6-DigiRAN-valore-automazione-accesso-radio.html>
2. <https://www.telecomitalia.com/tit/en/archivio/media/note-stampa/market/2016/TIM-AltioStar-VNAN.html>
3. <https://www.telecomitalia.com/tit/en/archivio/media/note-stampa/corporate/2018/Nota-Stampa-TIM-Ericsson.html>
4. <https://www.o-ran.org/> [o link ad articolo nello stesso notiziario tecnico]
5. <https://www.onap.org/architecture>
6. “Managing complexity: augmented intelligence for 5G radio access design and optimization” Notiziario Tecnico TIM. N. 2 2019 <https://www.telecomitalia.com/content/tiportal/it/notiziariotecnico/edizioni-2019/n-2-2019/N5-Managing-complexity-augmented-intelligence-5G-radio-access-design-optimization.html>
7. “Mobile trend: Self-Organizing Networks” Notiziario Tecnico TIM, N.2 2014. <http://www.telecomitalia.com/tit/it/notiziariotecnico/archivio/2014-2/capitolo-06.html>

Notes

1. Automatic Method For Mobility Load Balancing In Mobile Telecommunications Networks – Patent No US 10,433,209,B2 – Oct. 1, 2019 and Patent No EP 3 318 082 B1
2. Example of T-MAST installed at Tirrenia for Italian acrobatic team exhibition: https://www.telecomitalia.com/content/dam/telecomitalia/it/archivio/documenti/media/note_stampa/mercato/2019/Carrato_Tirrenia1.jpg



Paola Bertotto paola.bertotto@telecomitalia.it

Laureata in Ingegneria Elettronica con indirizzo Telecomunicazioni presso l’Università di Trieste nel 1996, con una tesi sviluppata presso l’allora CSELT di Torino. Entra in azienda nel 1997 per occuparsi di tematiche legate all’esposizione umana alle radio frequenze, nello specifico di analisi del SAR – Specific absorption rate. Dal 2001 si è occupata, nell’ambito dell’Accesso Radio, di aspetti legati alle stime di traffico e al dimensionamento della rete, sviluppando applicativi SW per l’ottimizzazione delle reti 2G/3G/4G/5G. Dal 2014 lavora nel gruppo “Radio Propagation and Self Organizing Network”, dove segue le attività di dispiegamento di nuovi use cases SON, del testing e del rilascio di linee guida. Nel 2015, grazie all’opportunità offerta dall’azienda, ottiene la laurea di primo livello in Comunicazione, Media e Pubblicità ■



Francesco Epifani francesco.epifani@telecomitalia.it

Ingegnere elettronico, ha iniziato nel 2000 l’esperienza lavorativa nel gruppo TIM come tesista presso l’unità di Microonde dell’allora CSELT di Torino. Entra in azienda a marzo 2001 per occuparsi dello sviluppo di modelli di propagazione di previsione di campo elettromagnetico e algoritmi di progettazione, dimensionamento e ottimizzazione delle risorse radio, tematiche per le quali è co-inventore di diversi brevetti. Dal 2007 si occupa delle evoluzioni degli strumenti SW, che TIM sviluppa “in house”, a supporto del network design e dell’ottimizzazione della rete di accesso dei sistemi 2G/3G/LTE/5G ■



Michele Ludovico michele.ludovico@telecomitalia.it

Ingegnere elettronico, ha iniziato ad operare nel gruppo TIM progettando sistemi a micro-onde per comunicazioni via satellite. Dal 2001 si occupa di strumenti e metodologie di progettazione ed ottimizzazione dell’accesso radio, che TIM sviluppa “in house” a supporto dell’evoluzione della rete mobile. Dal 2014 è responsabile della funzione di TIM che assicura l’ingegneria e lo sviluppo delle soluzioni di automazione per la rete di accesso radio, secondo il paradigma “Self Organizing Network”. Ha svolto, inoltre, attività di formazione e consulenza in Italia ed all’estero ed è co-inventore di diversi brevetti nel campo della progettazione wireless e della gestione delle risorse radio ■



Giovanna Zarba giovanna.zarba@telecomitalia.it

Laureata in Ingegneria Elettronica al Politecnico di Torino, inizia la sua esperienza lavorativa nell’unità Microonde dello CSELT dove si occupa della progettazione di componenti in guida d’onda per satellite (diplexers, triplexers). Dal 2002 si occupa di metodologie di ottimizzazione e gestione dell’accesso radio focalizzando la sua attività sulla definizione e sperimentazione di nuovi use case SON da implementare nei tool sviluppati in TIM. È co-autrice di articoli e brevetti nel campo delle comunicazioni mobili e della gestione delle risorse radio ■

OPEN RAN: DALLE SPECIFICHE AI TRIALS

Andrea Buldorini, Marco Caretti, Gian Michele Dell'Aera, Salvatore Scarpina

Verso l'Open RAN

Negli ultimi anni si è assistito ad un aumento esponenziale del traffico mobile a cui si è affiancata la domanda crescente di supporto di connessioni di tipo "always connected" e di terminali controllati da remoto (la cosiddetta *Internet of Things*). Questa rapida evoluzione

spinge gli operatori a cospicui investimenti al fine di aggiornare la rete per far fronte alle sempre crescenti richieste di capacità, a cui non corrisponde un analogo aumento dei profitti anche a causa della forte competizione in atto tra i diversi operatori. Queste dinamiche stanno spingendo gli operatori mobili a cercare nuove soluzioni architetture

rali che permettano da una parte di introdurre più velocemente nuovi servizi e nuove funzionalità per aumentare la capacità offerta dalla rete, e che dall'altra parte aiutino a rendere i costi di ampliamento e gestione della rete compatibili con lo scenario fortemente competitivo. Nel caso del segmento dell'accesso radio, visti i maggiori investimenti richiesti, si fa più pressante l'esigenza di ricercare nuove soluzioni ed approcci che permettano di realizzare un'architettura più flessibile, anche a supporto dell'introduzione delle nuove funzionalità legate al 5G. In quest'ottica si stanno affermando diverse iniziative che propongono, con diverse sfumature, il concetto di 'Open RAN', con l'obiettivo di introdurre una maggiore competizione nel segmento radio e favorire

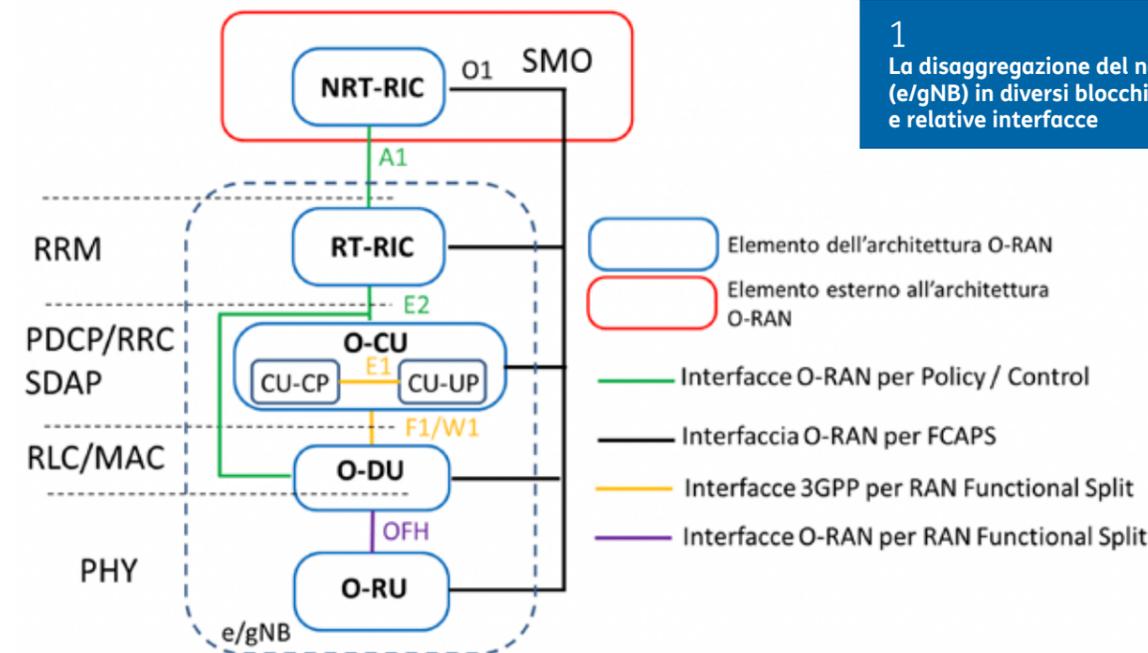
anche l'ingresso di nuovi attori. In particolare, è possibile individuare alcuni obiettivi comuni alle diverse iniziative:

- La possibilità di utilizzare dispositivi, apparati e applicativi da diversi fornitori, garantendo quindi l'interoperabilità tramite la definizione di interfacce standard aperte;
- L'utilizzo di hardware non specializzato su cui poter installare i moduli software (realizzati anche da fornitori differenti) che implementano le diverse funzionalità di rete;
- L'introduzione di maggiore intelligenza nel segmento radio, sfruttando anche tecniche di AI/ML (Artificial Intelligence/Machine Learning) al fine di automatizzare molte operazioni volte

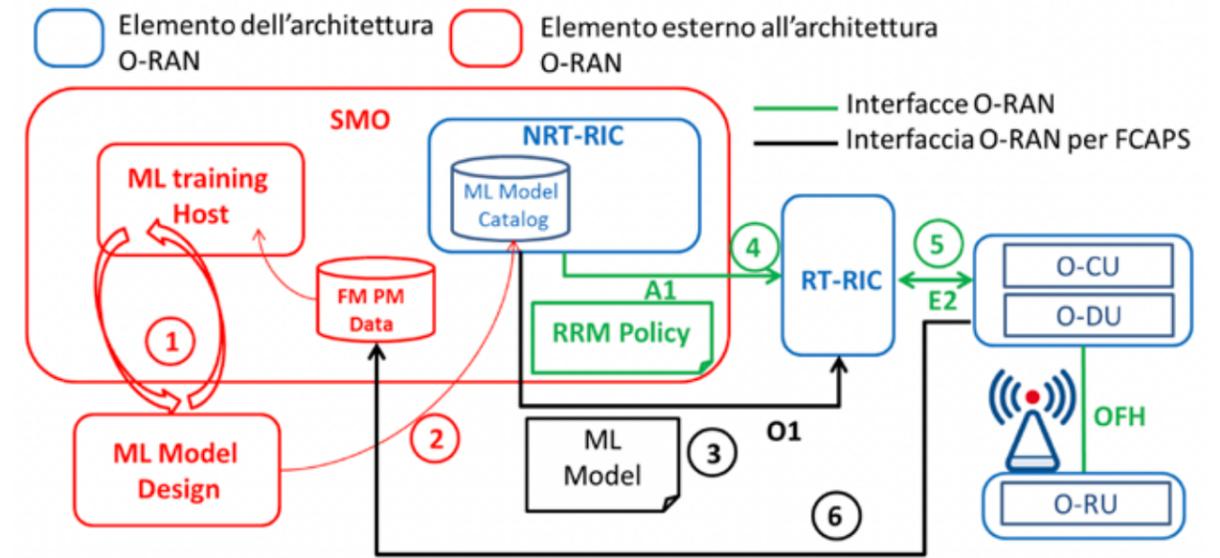
alla gestione, configurazione e ottimizzazione radio.

Un approccio cloud, mediante l'impiego di opportune soluzioni di IaaS/PaaS per l'hosting anche delle funzioni radio è visto come un importante abilitatore per il raggiungimento degli obiettivi sopra elencati. Al fine di perseguire questi obiettivi TIM ha aderito a due consorzi internazionali, l'O-RAN Alliance [1] e il Telecom Infra Project (TIP) [2], rispettivamente sin dal 2018 e dal 2017.

L'O-RAN Alliance è un'iniziativa guidata da operatori e risultante dalla fusione tra xRAN Forum e C-RAN Alliance al fine di spingere verso una maggiore apertura delle interfacce nella rete di accesso radio dei sistemi mobili di futura generazione. Supportata da ope-



1 La disaggregazione del nodo radio (e/gNB) in diversi blocchi funzionali e relative interfacce



2 Possibile sequenza (da step 1 a step 5) per l'utilizzo di Modelli di ML in rete

ratori quali AT&T, China Mobile, Deutsche Telekom, NTT DOCOMO, Orange, Verizon e TIM, persegue l'obiettivo di fare evolvere l'accesso radio secondo i paradigmi della disaggregazione delle funzionalità radio, virtualizzazione e disaccoppiamento hardware/software con utilizzo di hardware generico (cosiddetto whitebox), automazione e standardizzazione di interfacce aperte per abilitare una vera interoperabilità.

Il TIP è un'iniziativa patrocinata da Facebook, a cui partecipano operatori e costruttori, con un obiettivo più generale che può riassumersi con la frase 'connect the unconnected': supportare l'innovazione tramite lo sviluppo di Proof of Concept e trial al fine di sviluppare soluzioni cost-effective che per-

mettano di portare internet ovunque. Per il segmento radio questo obiettivo ha portato a considerare elementi comuni all'O-RAN: disaggregazione delle funzionalità, virtualizzazione e disaccoppiamento hardware/software con utilizzo di whitebox.

O-RAN Alliance

Uno degli obiettivi dell'O-RAN Alliance è favorire la realizzazione di una RAN aperta, basata sulla disaggregazione del nodo di accesso radio secondo quanto previsto dall'architettura 3GPP, alla quale sono stati aggiunti ulteriori elementi funzionali e le relative interfacce. A tal fine è stata definita un'architettura

di riferimento (riportata in Figura 1) dove sono rappresentati diversi componenti associati alle varie funzionalità radio:

O-RU (O-RAN Radio Unit):

La parte del nodo radio connessa alle antenne (o che le integra nel caso di sistemi di antenne attive) e responsabile della conversione del segnale da digitale ad analogico e a radiofrequenza in trasmissione e viceversa in ricezione. Il blocco contiene la parte bassa del livello protocollare fisico (Low-PHY).

O-DU (O-RAN Distributed Unit):

La componente del nodo radio che contiene la parte alta del livello fisico (High-PHY) insieme ai livelli protocollari MAC e RLC. Questo blocco, contenendo sia lo scheduler che la codifica di canale, è quello che richiede il maggior carico computazionale. Al

Open Fronthaul (FH)

bruno1.melis@telecomitalia.it

Le specifiche dell'interfaccia di fronthaul recentemente rilasciate dal WG4 dell'O-RAN nella versione 2.0 definiscono un'interfaccia standard fra gli apparati di banda base (O-DU) e le unità radio (O-RU) per LTE e NR nel caso di split funzionale 7-2 (secondo la classificazione 3GPP, [4]) all'interno del livello fisico (Low Layer Split, LLS), permettendo l'interoperabilità fra apparati di vendor diversi (a.k.a. multivendor RAN).

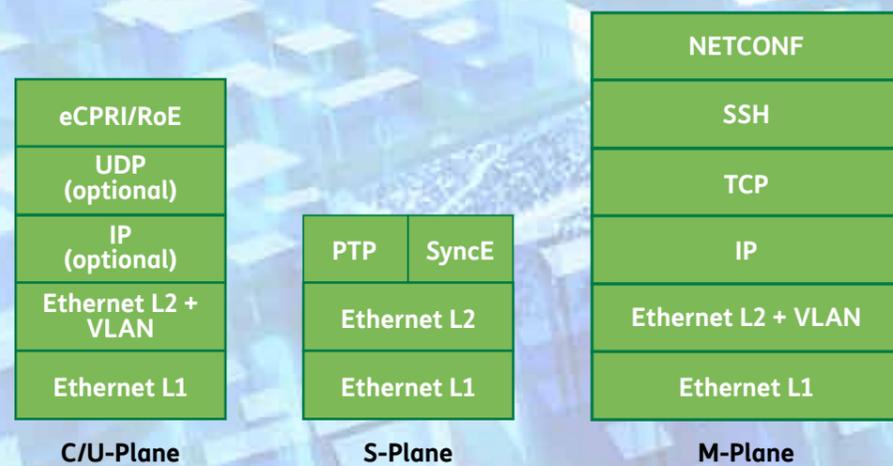
Il protocollo standardizzato da O-RAN, definisce il piano utente, controllo, sincronizzazione e management e si appoggia sul protocollo eCPRI (o in alternativa sul protocollo IEEE 1914.3, anche noto come Radio over Ethernet, RoE).

A livello 3 è previsto l'utilizzo opzionale del protocollo UDP/IP (Control/User plane) o TCP/IP (Management plane) che a sua volta si appoggia sul protocollo Ethernet per il trasporto a livello 1 e 2. La struttura dei protocolli su cui si appoggia il protocollo Open Fronthaul

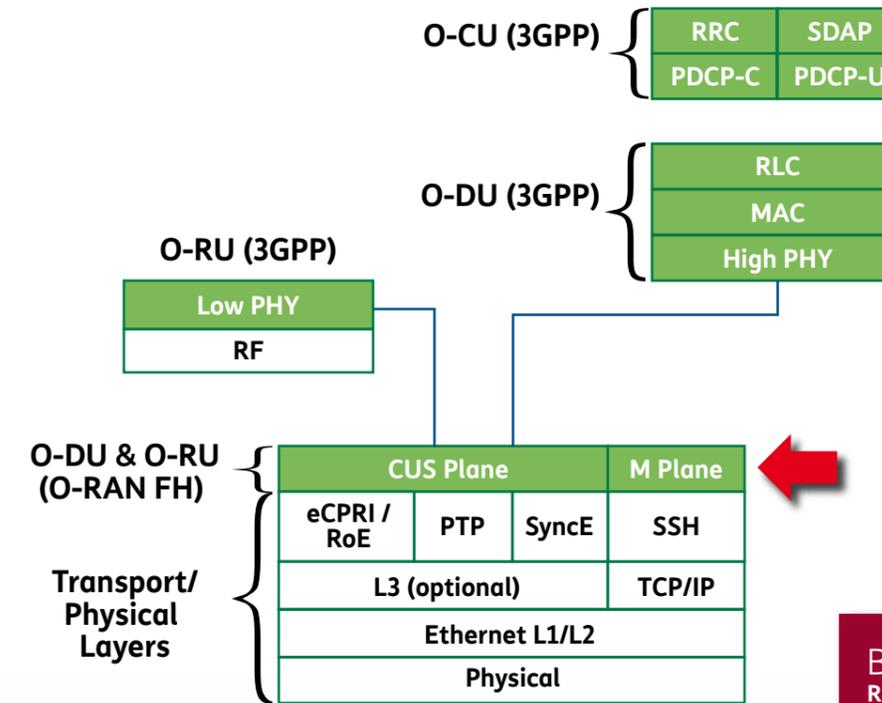
è illustrata in *Figura A* mentre la ripartizione dei protocolli fra i vari nodi che costituiscono l'architettura Open RAN è mostrata in *Figura B*.

I messaggi del piano utente dell'O-FH trasportano i campioni del segnale OFDM (Orthogonal Frequency Division Multiplexing) nel dominio della frequenza, quantizzati ed eventualmente compressi mediante specifici algoritmi standardizzati dalla specifica O-RAN.

I messaggi del C-Plane inviati dalla O-DU alla O-RU trasportano invece informazioni di controllo associate ai dati trasportati sul piano utente, quali ad esempio le risorse tempo/frequenza (i.e. scheduling) su cui si mappano i dati e le informazioni di supporto al beamforming (es. l'indice del beam i cui pesi sono pre-caricati nella O-RU o alternativamente i pesi di beamforming calcolati nella O-DU da applicare nella O-RU ai dati). La specifica O-FH standardizza diversi metodi per lo scambio delle informazioni di controllo legate al be-



A Struttura dei protocolli su cui si appoggia Open Fronthaul



B Ripartizione dei protocolli fra i nodi che costituiscono l'Open RAN

amforming rendendo quindi possibile l'interoperabilità fra O-DU ed O-RU di vendor diversi.

Il throughput dell'interfaccia O-FH è un elemento importante da cui dipendono le scelte di dispiegamento e l'opportuno dimensionamento della rete di trasporto. Caratteristica dello split funzionale 7-2 è che il throughput sul fronthaul, a differenza dell'interfaccia CPRI (Common Public Radio Interface), scala in proporzione al carico sull'interfaccia radio rappresentato essenzialmente dal numero di Resource Block (RB) utilizzati. Questa caratteristica permette la moltiplicazione statistica nel caso di aggregazione del traffico generato/ricevuto da molteplici O-RU, aumentando l'efficienza della rete di trasporto e permettendo di effettuare un dimensionamento di tipo statistico non vincolato al valore di picco come nel caso dell'interfaccia CPRI. Per quanto riguarda il MIMO, poiché il beamforming digitale e/o analogico con l'applicazione dei relativi pesi è effettuato nella O-RU, il throughput del piano utente scala proporzionalmente al numero di layer/stream

MIMO trasmessi e non in funzione del numero di antenne (come nel caso della CPRI), rendendo possibile l'utilizzo di antenne con un numero elevato di elementi radianti (a.k.a. massive MIMO) mantenendo requisiti sul trasporto ragionevoli. A titolo di esempio, la *Figura C* mostra una stima ottenuta per via analitica del throughput di picco generato in downlink da una O-DU sull'interfaccia O-FH nel caso di una portante NR, in funzione della larghezza di banda della portante e del numero di layers trasmessi (ottenibili in NR anche con meccanismi di trasmissione MU-MIMO). Il calcolo è effettuato supponendo un utilizzo completo di tutte le risorse disponibili in termini di Resource Blocks (carico massimo sull'interfaccia radio), spaziatura delle sottoportanti OFDM di 30 kHz ed una quantizzazione su 8 bit dei segnali I/Q trasmessi. Il throughput sul piano dati dipende oltre che dal numero di risorse utilizzate (RB) e dal numero di layer trasmessi (quindi indirettamente dal numero di utenti serviti) anche dal livello di quantizzazione che a sua volta è legato allo schema di com-

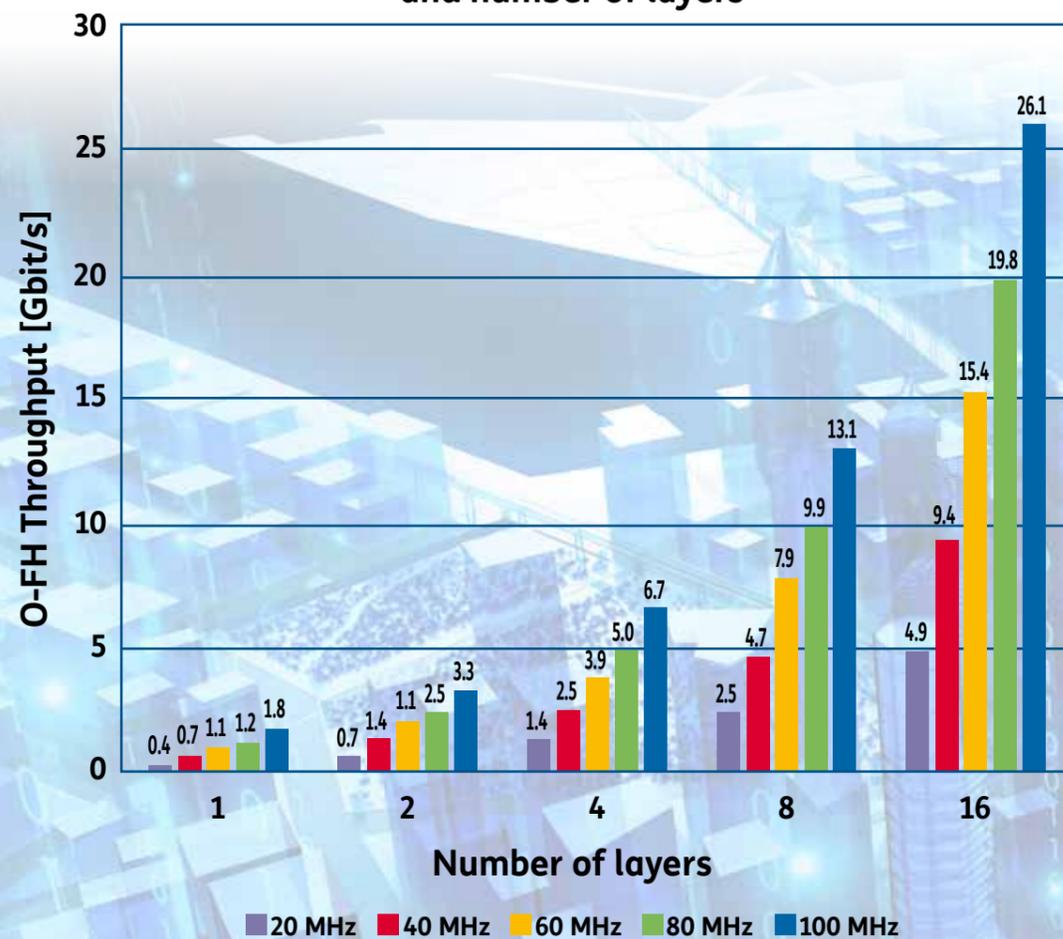
pressione utilizzato. Al throughput generato sul piano utente è stato sommato il throughput generato sul piano di controllo che in una situazione media è stato stimato pari al 5% del throughput sul piano utente. Il throughput sul piano di controllo dipende principalmente dal numero di utenti serviti simultaneamente e dallo schema di beamforming utilizzato (ad es. dal periodo con cui sono ricalcolati e trasmessi i coefficienti di beamforming).

Altro elemento importante da considerare per le opportune scelte di dispiegamento di una soluzione O-RAN è la massima latenza supportata da quest'interfaccia. I valori considerati nella prima versione della specifica sono dell'ordine delle centinaia di microsecondi. Nella

seconda versione è stato introdotto il supporto anche di valori superiori, sulla scorta di quanto svolto nell'ambito del progetto vRAN Fronthaul del TIP (vedere il box di approfondimento "Non ideal Fronthaul e attività TIM nell'ambito del TIP") ■

C
Throughput sull'interfaccia O-FH sul piano utente nel caso di portante NR

O-RAN Fronthaul Throughput vs. Channel Bandwidth and number of layers



fine di perseguire un approccio cloud anche per la componente RAN l'idea è virtualizzare anche questa componente e centralizzarla in un opportuno data centre consentendo l'abilitazione di algoritmi di coordinamento quali Carrier Aggregation e CoMP. In tal caso, visti gli elevati requisiti in termini di processing e esigenze real time, al fine di avere una soluzione più efficiente, è prevedibile l'impiego di soluzioni di accelerazione hardware (quali FPGA o GPU) in aggiunta a hardware non specializzato (tipicamente basato su architetture x86).

O-CU (O-RAN Central Unit):

questo elemento è standardizzato anche in 3GPP e viene ulteriormente composto da una parte che gestisce lo User Plane (contenente il protocollo PDCP e il protocollo SDAP) e una parte di Control Plane (contenente il protocollo RRC). Facendo riferimento agli split funzionali analizzati in 3GPP (si veda [4]), l'interfaccia tra la O-CU e la O-DU (indicata con F1 per NR e W1 per LTE) è basata pertanto sull'opzione 2. Anche tenendo conto delle prime soluzioni di virtual RAN, la virtualizzazione di questo elemento è fattibile impiegando unicamente hardware non specializzato.

RT-RIC (Real Time Radio Intelligence Controller):

è l'elemento funzionale in cui vengono centralizzate alcune procedure di RRM (Radio Resource Management) quali ad esempio l'Admission Control, l'Handover e il Measurement Report, permettendo anche l'utilizzo di implementazioni fornite da terze parti e l'impiego di tecniche di AI/ML.

NRT-RIC (Non Real Time Radio Intelligence Controller):

è l'elemento funzionale presente all'interno del layer di Service Management and Orchestration (SMO) che permette il controllo del nodo RAN attraverso l'invio di Policy e monitorando le informazioni FCAPS (Fault, Configurations, Alarms, Performance and Security) provenienti dai vari blocchi dell'architettura.

Per favorire l'apertura dell'ecosistema dell'accesso radio a diversi player, l'O-RAN si è posto l'obiettivo di assicurare l'effettiva interoperabilità tra i diversi elementi di rete. Per questo, a partire dalle specifiche standard già definite in 3GPP, vengono definiti opportuni profili ottenuti selezionando sottoinsiemi delle opzioni previste dallo standard. Essendo tuttavia presenti nuovi elementi nell'architettura, l'O-RAN specifica anche nuove interfacce:

- L'interfaccia **O-FH (Open Fronthaul)** è l'interfaccia tra la O-DU e la O-RU, che abilita l'interoperabilità tra O-DU (che può essere installata presso il sito radio o presso un data centre) e O-RU (installata vicino alle antenne) e che possono essere fornite da costruttori differenti. Facendo riferimento agli split funzionali analizzati in 3GPP (si veda [4]), per tale interfaccia è stata selezionata l'opzione 7-2.
- Le interfacce A1 e E2, che consentono il monitoraggio e la riconfigurazione della rete in modo dinamico e automatico in funzione delle esigenze di servi-

zio/copertura, applicando anche tecniche di AI/ML. Come si vede nell'esempio in *Figura 2* l'interfaccia O1 permette di passare al RT-RIC un ML Model (step 3). Il Modello passato verrà usato dal RT-RIC per ottimizzare le decisioni da applicare in rete attraverso input e output sulle interfacce A1 e E2 (step 4 e 5). I contatori di rete vengono poi collezionati nel SMO (step 6) e vengono usati per fare il training di nuovi modelli ML (step 1), i modelli ritenuti validi vengono caricati come quelli utilizzabili dal NRT-RIC (step 2) pronti per aggiornare quelli esistenti nei RT-RIC.

L'attività O-RAN è organizzata in *Working Group* che hanno lo scopo di definire le interfacce descritte in *Figura 1*.

WG1: è il gruppo che si occupa di standardizzare l'interfaccia O1 che corrisponde all'interfaccia di OAM (Operation & Maintenance). Le informazioni FCAPS provenienti dai vari blocchi vengono catturate da questa interfaccia, che allo stesso tempo può essere usata per la configurazione del nodo.

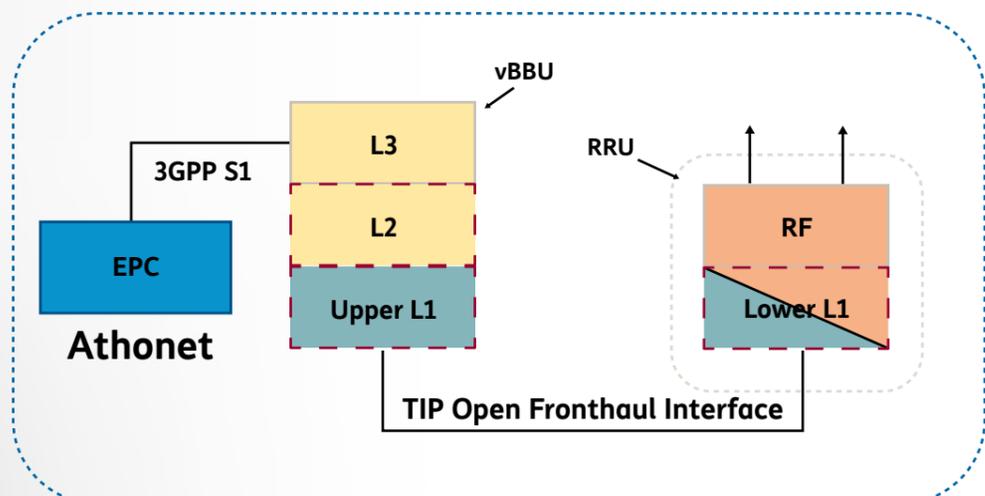
WG2: è il gruppo che si occupa di standardizzare l'interfaccia A1, il cui scopo è inviare informazioni relative alle policy con cui la RAN deve gestire gli utenti o i servizi legati ai diversi utenti. La stessa interfaccia può essere usata per trasmettere alla RAN modelli di Machine Learning (ML) e/o Enrichment Information (EI) al fine di migliorare la gestione delle procedure Radio per

Non ideal Fronthaul e attività TIM nell'ambito del TIP

antonio.orlando@telecomitalia.it;
alessandro.torielli@telecomitalia.it

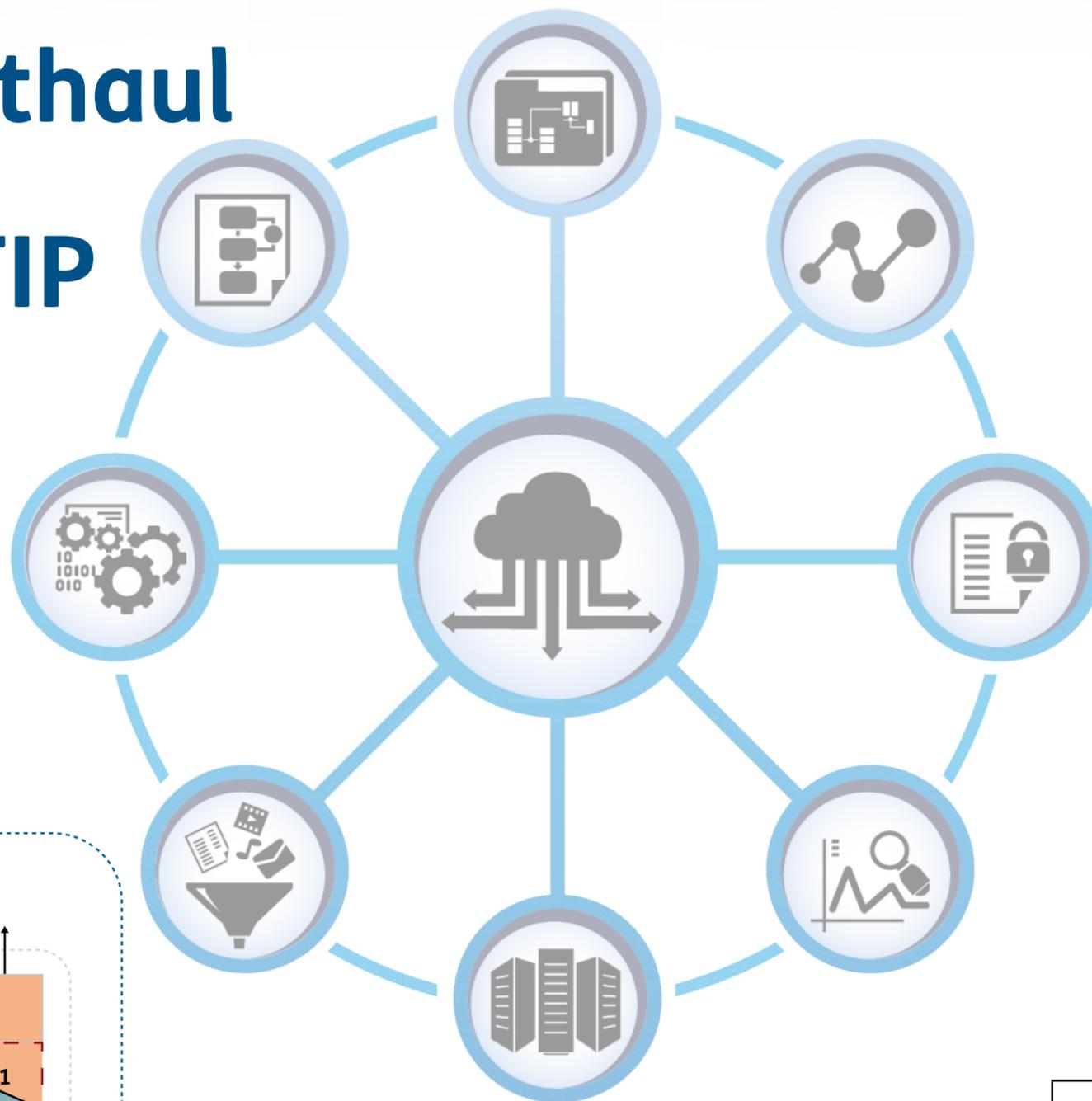
La soluzione di accesso mobile sperimentata nel TIP Community Lab di TIM a Torino è una soluzione multi vendor volta a verificare il comportamento del sistema nel caso di un FH non ideale trasportato su una rete a pacchetto. Differenti fornitori hanno contribuito alla sua realizzazione, in particolare:

- Phluido ha fornito l'Upper Layer 1 della BBU (funzionalmente corrispondente all'insieme O-CU + O-DU definiti in O-RAN) e il Lower Layer 1 installato sulla FPGA della RRU (funzionalmente corrispondente alla O-RU definita in O-RAN);



Phluido
Radisys
Baicells
Requires split-specific enhancements

D
Eco-sistema multi vendor realizzato nel TIP CL di TIM



- Radisys ha fornito Layer 2 e Layer 3 della BBU;
- Baicells ha fornito la RRU, costituita da una small cell indoor;
- Athonet ha messo a disposizione una core LTE carrier grade;
- Tech Mahindra insieme a TIM ha svolto il compito di System Integrator ed ha avuto la responsabilità dell'esecuzione dei test previsti, con analisi dei risultati ottenuti e generazione del report finale dell'attività svolta.

Nella Figura D si riporta uno schema a blocchi con le connessioni logiche

È importante segnalare che, sebbene uno degli obiettivi iniziali del progetto fosse quello di specificare un'interfaccia open per il fronthaul, le attività del gruppo si siano limitate poi alla realizzazione del PoC.

La soluzione implementata sul FH è pertanto proprietaria e non allineata alle specifiche O-RAN, anche se è stata considerata la stessa opzione di functional split 7-2. Le attività sperimentali condotte nel TIP Community Lab di TIM sono state focalizzate sulla caratterizzazione del Sistema in configurazione RRU singola e multipla (abilitando anche algoritmi CoMP) secondo quanto riportato nella figura precedente (Figura D).

Nel seguito si riportano alcuni risultati ottenuti in configurazione singolo vendor (Phase 0 del progetto, in cui tutto lo stack protocollare è stato fornito da Phluido).

Il sistema è stato testato considerando la configurazione LTE MIMO 2x2. Un primo risultato interessante è la banda misurata sul FH in corrispondenza di un singolo utente connesso (single user peak throughput) in grado di generare quindi 150Mbps in Downlink e 50Mbps in Uplink sulla

	Downlink	Uplink
Option 2	150Mbps	75 Mbps
Option 7-2	1.08 Gbps	1.08 Gbps
Option 8 (CPRI)	2.46 Gbps	2.46 Gbps
TIP solution measured in lab	175 Mbps	210 Mbps

T1
Fronthaul peak throughput per differenti functional splits

tratta radio, che mette in evidenza l'efficienza dell'algoritmo di compressione implementato.

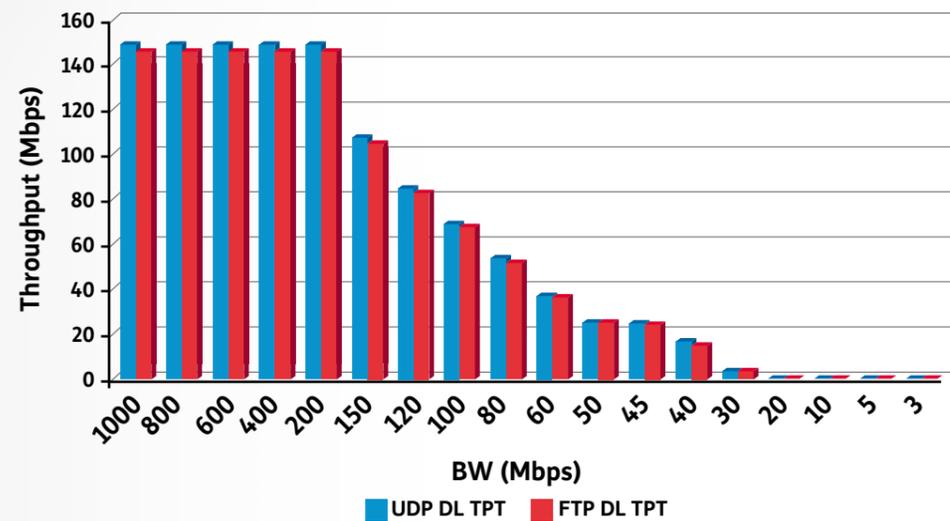
I valori misurati sul fronthaul sono confrontati nella tabella (Tabella T1) con i valori teorici di alcuni degli split analizzati in 3GPP (si veda [4]), calcolati in assenza di algoritmi di compressione.

Un altro risultato interessante è ottenuto confrontando il throughput singolo utente sulla tratta radio con la banda corrispondente misurata sul FH nel caso di limitazione

della banda disponibile sul trasporto (si veda Figura E). Dai dati misurati è possibile non solo stimare l'overhead del FH ma soprattutto evidenziare come il sistema sia in grado di lavorare in caso di reti di trasporto congestionate anche se con una degradazione nelle prestazioni massime raggiungibili.

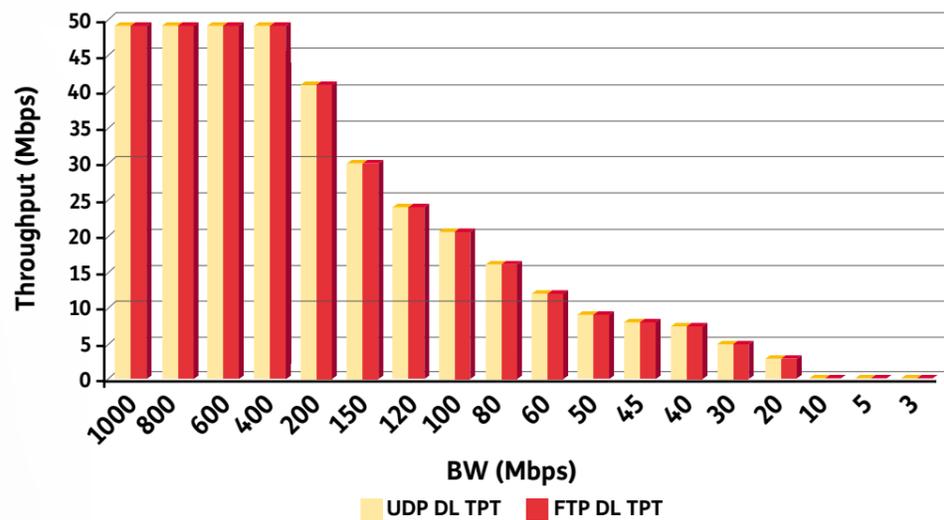
Al fine di valutare la robustezza del sistema alla presenza di possibili delay in rete di trasporto è stata introdotta della latenza aggiuntiva sulla connessione di FH (si veda

BW Reduction - DL Throughput

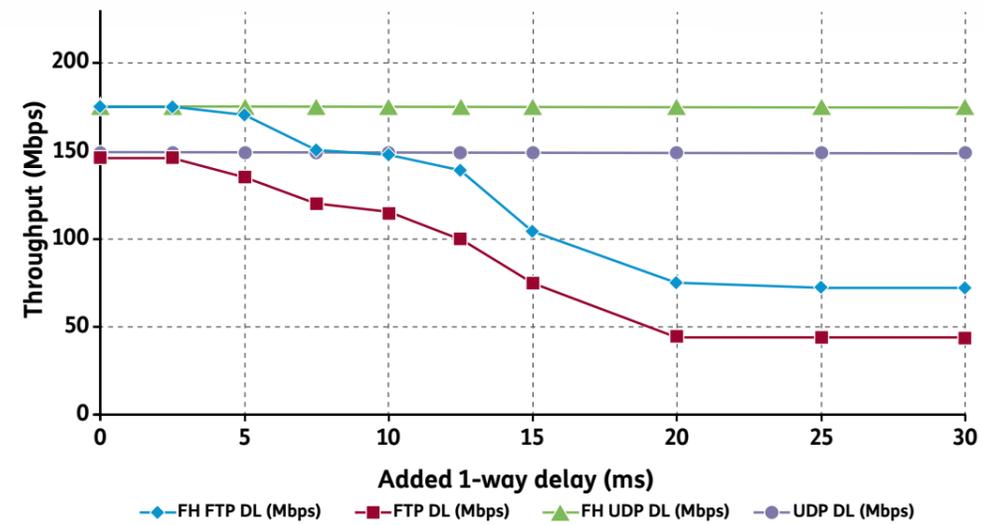


E Throughput Downlink e Uplink vs banda limitata sul fronthaul

BW Reduction - UL Throughput



DL Throughput - 20 MHz BW - Std Latency



F Throughput in Downlink e Uplink vs Latenza sul fronthaul

UL Throughput - 20 MHz BW - Std Latency

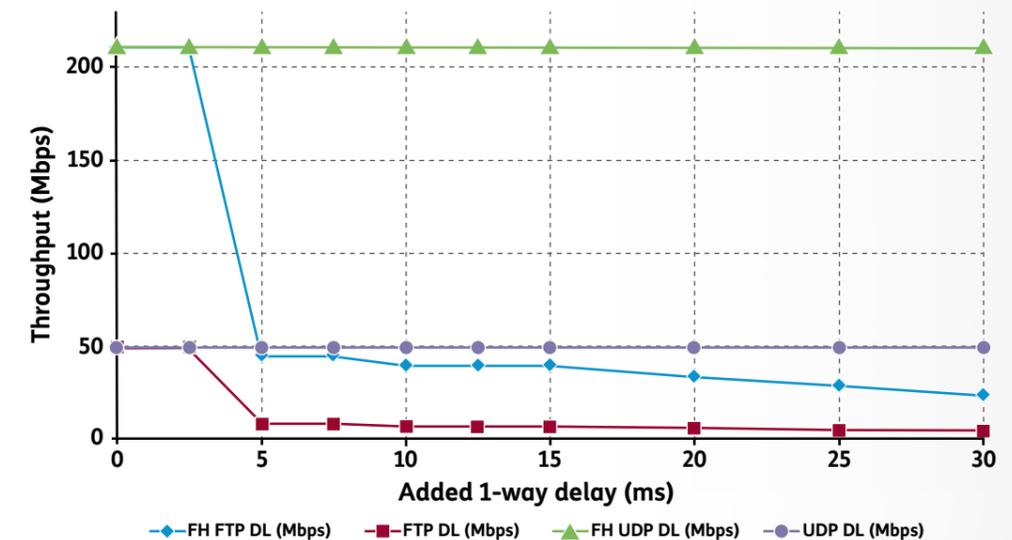


Figura F). Il throughput TCP mostra una degradazione delle prestazioni per valori di latenza superiori ai 5 ms (dovuti alla mancanza di ottimizzazione degli algoritmi di controllo del TCP) mentre il traffico UDP mostra un buon comportamento anche per valori di latenza estremamente alti. I risultati ottenuti sono tanto più interessanti se confrontati con quanto previsto in letteratura e nella specifica di O-FH O-RAN (almeno nella versione 1.0), dove si parla di valori massimi dell'ordine delle centinaia di microsecondi. Le soluzioni adottate per gestire questi elevati valori di

latenza possono comportare una riduzione di prestazione in alcuni scenari; in altre situazioni, tipicamente caratterizzate da bassa mobilità e da una bassa densità di utenti, ci si aspetta una degradazione delle prestazioni trascurabile se non nulla (si veda l'annex L in [5] per approfondimenti). L'implementazione di soluzioni di questo tipo consentirebbe pertanto una maggiore flessibilità per l'operatore nell'implementazione di una soluzione RAN in ottica cloud e il supporto di alcuni scenari quali small cell e FWA (Fixed Wireless Access) con impatti limitati sul trasporto ■

raggiungere l'obiettivo prefissato a livello di SMO.

WG3: è il gruppo che si occupa di definire l'interfaccia E2, il cui compito è quello gestire le procedure radio del nodo in real-time (RT). Per fare questo deve comunicare con i moduli che gestiscono il C-Plane nel nodo radio ovvero: RRC della O-CU e lo scheduler a livello MAC che risiede nella O-DU.

WG4: è il gruppo che si occupa di definire l'interfaccia O-FH che collega la O-DU alla O-RU. Il gruppo definisce anche le specifiche dell'interfaccia di gestione della O-RU e di conformance test e di interoperabilità dell'interfaccia O-FH

WG5: è il gruppo che si occupa delle interfacce tra O-DU e O-CU (F1/W1), tra C-Plane e U-Plane (E1) e anche delle interfacce tra altri nodi radio (X2/Xn). Essendo queste interfacce oggetto di standardizzazione in 3GPP, il compito di questo gruppo è assicurarsene l'effettiva interoperabilità andando a ridurre le opzioni e gli aspetti lasciati all'implementazione.

Il compito di O-RAN non si limita solo alla definizione di interfacce logiche tra funzionalità radio: nel consorzio sono presenti ulteriori working group con l'obiettivo di promuovere l'implementazione di soluzioni O-RAN in ottica cloud, al fine di disaccoppiare l'hardware dal software e permettere l'utilizzo di hardware non specializzato:

WG6: è il gruppo che si occupa di definire l'architettura cloud per il nodo RAN, in funzione delle di-

verse ipotesi di dispiegamento. L'attività comprende la definizione di modalità per l'utilizzo di piattaforme di accelerazione hardware (come FPGA/DSP/ASIC/GPU) necessarie per alcune delle funzioni radio più onerose dal punto di vista computazionale.

WG7: è il gruppo che si occupa della definizione di un reference design per l'hardware di una "white box" in grado di implementare il nodo radio (in particolare per la componente O-DU ed O-RU)

WG8: è il gruppo che si occupa della definizione di un reference design del software che implementa il nodo radio.

Un passo ulteriore verso la realizzazione di RAN multivendor consiste nella possibilità di certificare il livello di conformance allo standard e l'interoperabilità delle soluzioni sviluppate. Per questo è stato istituito il **TIFG** (*Test and Interoperability Focus Group*) il cui compito è definire un processo di test e di certificazione delle diverse interfacce al fine di validare la bontà delle soluzioni proposte dai diversi vendor. A tal fine il gruppo prevede la costituzione di diversi laboratori **OTIC** (*O-RAN test and integration certification*) responsabili della certificazione di apparati O-RAN compliant. Il gruppo si occupa anche dell'organizzazione di *plugfest*, eventi nei quali i vari fornitori possono verificare il livello di interoperabilità raggiunto delle proprie soluzioni con quelle sviluppate da altri fornitori. Infine, è importante segnalare l'intento di promuovere lo

sviluppo open source dei moduli che compongono l'architettura O-RAN. Per questo è stata stabilita una relazione con la Linux Foundation dove è stata istituita la **OSC** (*O-RAN Software Community*) il cui compito è fornire un'implementazione open source dei vari moduli, partendo in particolare dalle piattaforme di NRT-RIC e RT-RIC e da altri moduli del Service Management and Orchestration basati sulla piattaforma ONAP e opportunamente specializzati per le applicazioni in ambito RAN.

Telecom Infra Project

Nell'ambito del TIP, TIM ha partecipato sin dalla sua costituzione al progetto vRAN Fronthaul [3]. Il progetto, nato nel 2017, ha l'obiettivo di sviluppare una soluzione di RAN virtualizzata multi vendor in grado di poter essere dispiegata in presenza anche di soluzioni di trasporto per il fronthaul cosiddette non ideali, ovvero in scenari dove non è sempre disponibile la fibra e le tecnologie di trasporto sono caratterizzate da una banda limitata e latenze anche superiori al millisecondo. All'interno del progetto sono stati selezionati 4 laboratori (*Community Labs*) sponsorizzati da altrettanti operatori, dove diverse soluzioni prototipali multi-vendor sono state sviluppate al fine di supportare diversi use case:

- CableLabs: La sperimentazione in questo caso è focalizzata a misurare l'efficienza di compressio-

ne della banda e di robustezza a elevati valori di latenza sul fronthaul tali da permettere il dispiegamento della soluzione vRAN sulla rete DOCSIS esistente. Gli scenari di dispiegamento considerano quindi femto celle indoor e small cell in ambito urbano.

- BT: nel caso di BT l'interesse è per soluzioni di vRAN con fronthaul efficiente in grado di essere trasportato su reti ethernet (quali quelle presenti in campus universitari o grandi aziende), G.Fast e ponti radio. Gli scenari di interesse in tal caso considerano anche funzionalità multi-operatore/neutral host.
- Airtel: anche nel caso dello use case considerato da Airtel l'efficienza e robustezza alla latenza sono importanti in quanto si

prevede l'utilizzo di ponti radio multi-hop per il trasporto del fronthauling.

- TIM: Lo use case considerato da TIM indirizza un utilizzo efficiente della fibra già dispiegata, dove le capacità di compressione della banda è la prima priorità per assicurare lo sfruttamento adeguato degli asset in fibra e loro evoluzione. In tal caso lo scenario considerato è quello a supporto del dispiegamento di small cell con utilizzo di tecnologia PON (*Passive Optical Network*) e sue evoluzioni per il trasporto del fronthauling. Pensando ai meccanismi di allocazione della banda in questi sistemi punto-multipunto, anche la robustezza nei confronti della latenza è un fattore importante.

Un dettaglio sulle attività svolte nell'ambito del Community Lab di TIM è illustrato nel box di approfondimento "Non ideal Fronthaul e attività TIM nell'ambito del TIP". È importante segnalare come le varie aziende che hanno partecipato al progetto vRAN Fronthaul hanno poi proposto un work item in O-RAN per estendere le specifiche dell'**O-FH**, al fine di considerare anche latenze tipiche di questi scenari "non ideal". Questa estensione è stata approvata nella release 2.0. In aggiunta, al fine di permettere il supporto da parte dell'**O-FH** di tecnologie di trasporto quali DOCSIS e PON è stato avviato un altro work item in O-RAN (*CTI, Cooperative Transport Interface*) tutt'ora in corso ■

References

- [1] <https://www.o-ran.org/>
- [2] <https://telecominfraproject.com/>
- [3] <https://vran.telecominfraproject.com/>
- [4] 3GPP TR 38.801, 3rd Generation Partnership Project, Technical Specification Group Radio Access Network, Study on new radio access technology: Radio access architecture and interfaces (Release 14).
- [5] ORAN-WG4.CUS.0-v02.00, O-RAN Fronthaul Working Group, Control User and Synchronization Plane Specification.
- [6] O-RAN A1 interface: General Aspects and Principles Version 1.0 - October 2019 (ORAN-WG2.A1.GA&P-v01.00)
- [7] O-RAN NR U-plane profile for EN-DC Version 1.0 - June, 2019 (ORAN-WG5.U.0-v1.00)
- [8] O-RAN NR C-plane profile for EN-DC Version 1.0 - June, 2019 (ORAN-WG5.C.1-v1.00)
- [9] O-RAN Cloud Architecture and Deployment Scenarios for O-RAN Virtualized RAN Version 1.0 - October 2019 (O-RAN-WG6.CAD-V01.00.00)
- [10] O-RAN Operations and Maintenance Architecture Version 1.0 - July 2019 (O-RAN-WG1.OAM Architecture -v01.00)
- [11] O-RAN Operations and Maintenance Interface Version 1.0 - July 2019 (O-RAN-WG1.OAM Interface Specification-v1.0)

Acronimi

ASIC	Application Specific Integrated Circuit
BBU	Baseband Unit
CoMP	Coordinated Multi Point
DOCSIS	Data Over Cable Service Interface Specification
DSP	Digital Signal Processor
eCPRI	enhanced CPRI
FPGA	Field Programmable Gate Array
GPU	Graphics Processing Unit
LTE	Long Term Evolution
MAC	Medium Access Control
MIMO	Multiple Input Multiple Output
MU-MIMO	Multi User MIMO
NR	New Radio
PDCP	Packet Data Convergence Protocol
RAN	Radio Access Network
RLC	Radio Link Control
RRC	Radio Resource Control
RRM	Radio resource Management
RRU	Radio Remote Unit
SDAP	Service Data Adaptation Protocol
TCP	Transmission Control Protocol
UDP	User Datagram Protocol



Andrea Buldorini andrea.buldorini@telecomitalia.it

Ingegnere elettronico, indirizzo telecomunicazioni, entra in Azienda nel 1997 e partecipa ad attività di ricerca sui sistemi radiomobili. Attualmente è nella funzione Technology Innovation. Si è occupato di tematiche relative alle tecnologie radio, rappresentando Telecom Italia in vari enti di normativa internazionale (3GPP, O-RAN, NGMN, ETSI) - E' delegato TIM in O-RAN Alliance e in 3GPP RAN WG3, il gruppo tecnico responsabile della standardizzazione dei protocolli di rete di accesso radio 4G e 5G. Si occupa di aspetti di ottimizzazione di rete, Self-Organizing Networks, RAN Management and Orchestration ■



Marco Caretti marco.caretti@telecomitalia.it

Ingegnere elettronico, è entrato nel 2000 in Azienda per occuparsi di analisi delle prestazioni tramite simulazione e di tematiche di dimensionamento e pianificazione di sistemi radiomobili 2G/3G/4G. In ambito standard ha seguito i lavori del gruppo TSG RAN WG1 e successivamente del WG4 del 3GPP. Attualmente lavora all'analisi di prestazioni del sistema 4G/5G e segue le attività di sperimentazione e prototipazione di sistemi di accesso radio in architettura Cloud e Virtual RAN. In quest'ambito segue le attività in O-RAN Alliance (principalmente WG4 e WG6) e il progetto vRAN Fronthaul nel Telecom Infra Project (TIP) ■



Gian Michele Dell'Aera gianmichele.dellaera@telecomitalia.it

Laureato al Politecnico di Torino in ingegneria delle Telecomunicazioni nel 2006 con una tesi sui ricevitori multi antenna. Entra in azienda nel 2007 per occuparsi di simulazioni di link per la rete radio mobile 4G. Ha seguito lo standard 3GPP WG1 dal 2012 al 2016. Dal 2016 ad oggi segue gli sviluppi dall'accesso Radio verso il 5G. Attualmente è delegato per TIM della O-RAN Alliance ■



Salvatore Scarpina salvatore.scarpina@telecomitalia.it

Laureato in ingegneria elettronica presso Polito, lavora in TIM dal 2005 nell'ambito dell'innovazione e della ricerca. Inizialmente focalizzato sullo sviluppo di tecnologie sui terminali mobili, successivamente si è orientato allo sviluppo di servizi web ottimizzati per diverse tipologie di terminali (mobili e non); in questo contesto ha sviluppato competenze in ambito di project management e di ambienti cloud riapplicate successivamente in ambito Accesso Mobile (nel quale lavora dalla fine del 2017), insieme a nuove skill in ambito Management di reti mobili, Machine Learning applicato al SON, accelerazione hardware, Edge Computing. Fin dal 2007 partecipa ad attività di standardizzazione presso diversi enti, tra cui OMA DM (chair), ETSI MEC e O-RAN ■

EDGE INTELLIGENCE: EDGE COMPUTING MEETS ARTIFICIAL INTELLIGENCE

Antonio Manzalini

EC (*Edge Computing*) is about moving part of the service-specific processing and data storage from the Cloud Computing to the edge network nodes. Among the expected benefits of EC deployment in 5G, there are: performance improvements, traffic optimization and new ultra-low-latency services. If today EC is getting momentum, we're witnessing, at the same time, a growing development of AI (*Artificial Intelligence*) for a wide spectrum of applications, such as: intelligent personal assistants, video/audio surveillance, smart cities' applications, self-driving, Industry 4.0. The requirements of these applications seem calling an AI's resources-hungry model, whose cloud-centric execution appears in the opposite direction with a migration of computing, storage and networking resources at the edge. In reality, the two technology trends are crossing in the EI (*Edge intelligence*): an emerging paradigm meeting the challenging requirements of future pervasive services

scenarios where optical-radio networks requires automatic real-time joint optimization of heterogeneous computation, communication, and memory/cache resources and high dimensional fast configurations (e.g., selecting and combining optimum network functions and inference techniques). Moreover, the nexus of EI with distributed ledger technologies will enable new collaborative ecosystems which can include, but are not limited to: network operators, platform providers, AI technology/software providers and Users.

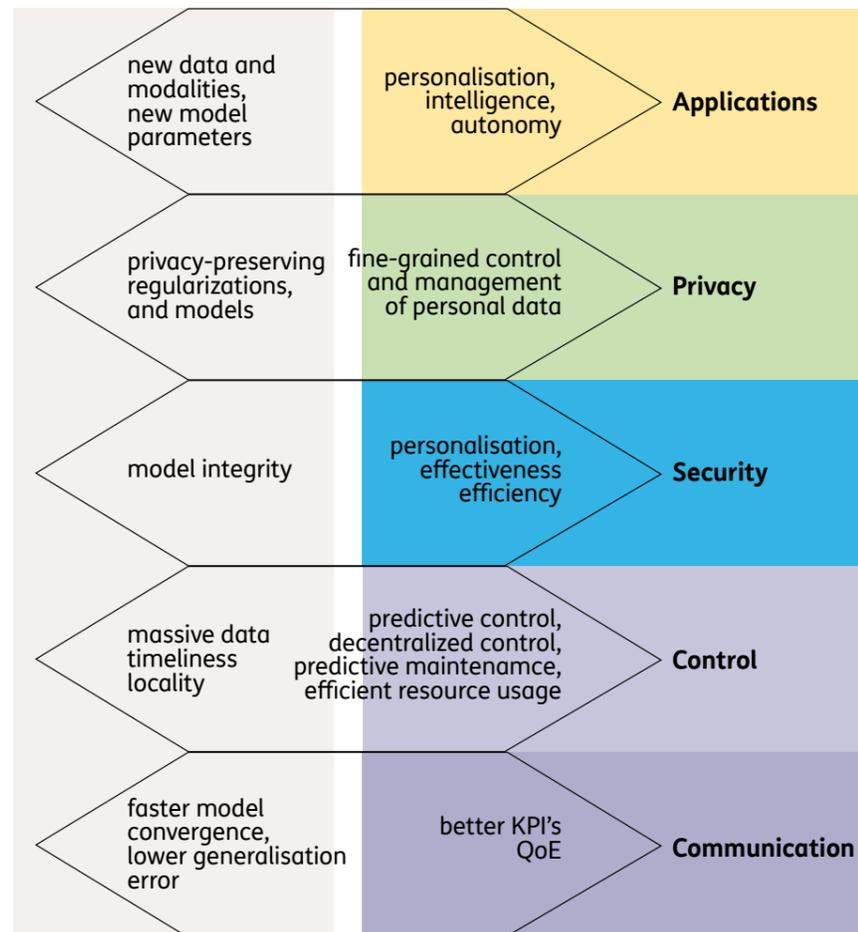
If AI is booming and Edge Computing is coming soon

In the last few years, we have witnessed a growing market interest and exploitation of AI solutions in a wide spectrum of ICT applications. AI services are becoming more and more popular, such as intelligent personal assistants, video/audio surveillance, smart cities' operations, self-driving vehicles, Industry 4.0 [1]. AI flourished again (after the AI winter) due to the current availability of powerful and low-cost processing and storage resources in the Cloud Computing: an abundance meeting the resources hungry requirements of AI methods and sys-

tems, called to elaborate an enormous quantity of Big Data. On the other hand, it is well-known that most of these big data elaborated by AI are generated at the edge, nearer the Users. The wide diffusion of smart terminals, devices and mobile computing, the IoT (*Internet-of-Things*) and the proliferation of sensors and video cam are generating several hundreds of ZB at the network edge. For example, Cisco estimates that nearly 850 ZB will be generated by people, machines, and things at the network edge by 2021 [2].

Indeed, the edge is where the Cloud Computing model is also being extended. EC, in fact, is about moving part of the service-specific processing and data storage from the Cloud Computing to the edge network nodes. Among the expected benefits of EC deployment in 5G, there are: performance improvements, traffic

1
Benefits provided for both AI and edge computing (Picture credits [3])



optimization and new ultra-low-latency services.

These trends are, obviously, encouraging to consider how extending AI frontier towards the edge ecosystems, in the last mile(s): in fact, the interplay of AI and EC will bring clear advantages in several dimensions. Figure 1 provides a brief summary of some of the cross benefits provided by entangling AI with EC [3].

Because of these trends, a new emerging innovation avenue is emerging rapidly, called "Edge Intelligence" (EI). This avenue brings several technical challenges: to make it simple, while EC aims at distributing computational and storage resources across the network architecture, real decentralized AI solutions are still in infancy.

The diffusion of EI doesn't mean, obviously, that there won't be a future for a centralized CI (*Cloud Intelligence*). The orchestrated use of Edge and Cloud virtual resources, in fact, is expected to create a continuum of intelligent capabilities and functions across all the Cloudified infrastructures. This is one of the major challenges for a successful deployment of an effective and future-proof 5G.

How a pervasive AI will bring to new network models

Software Defined Network (SDN) and Network Function Virtualization (NFV) technologies and models are

offering the unprecedented opportunity of designing and operating networks infrastructures (e.g., 5G), with two decoupled architectural levels, hardware and software. This is bringing higher flexibility and programmability, but, at the end of the day, it doesn't change the foundations about the network layering and stack protocols, as historically designed with the OSI models and the Internet. It's layers all the way down, where each layer controls a subset of the decision variables, and observes a subset of constant parameters and the variables from other layers. When transmitting data from one terminal to another, the data flow starts at the Application layer proceeding down to the Physical layer; then over the channel to the receiving terminal and back up the hierarchy.

Future services scenarios may take benefits from enhancing this three-decades old model by exploiting a pervasive intelligence. For example, network performance would be radically improved if instead network layering and stack protocols would be systematically designed as distributed solutions to decision making and optimization context-specific problems. In some cases, it might not be necessary entirely crossing up and down the OSI hierarchy.

The idea of a new mathematical theory of networks has been nicely described in [4], where it is argued that if network dynamics is modeled by a generalized network utility maximization problem, each layer corresponds to a decomposed sub-

problem, and the interfaces among layers are quantified as functions of the optimization variables coordinating the subproblems. The reverse-engineering of network layering and stack protocols - originally designed with ad hoc heuristics - discovers the original underlying mathematical (decision making and optimization) problems [4] to be solved. This consequent flexibility would bring to several alternative decompositions, leading to a wider choice of different layering and software architectures. In turn, this would mean greater modularity, flexibility, adaptability and cost effectiveness of the overall network infrastructure.

This is food for an extended AI, which is already dealing with very complex problems requiring classification, regression, optimization and decision making. Embedding pervasive DL (*Deep Learning*) methods and systems into the network domain it likely to bring to radically new paradigms in network design and management as well as new service models.

As a matter of fact, recent advances in AI might have already considerable impact in different sub-fields of ongoing networking. The paper [5] provide an interesting overview of the state-of-art in this direction, together with a number of significative examples. For instance, traffic prediction and classification are two of the most studied applications of AI in the networking field. DL is also offering promising solution for efficient resource management and network adaption thus improving,

EDGE INTELLIGENCE AND DISTRIBUTED LEDGER TECHNOLOGIES FOR FUTURE INTERNET

Today Internet has lost much of its end-to-end, peer-to-peer original characteristics, which was the ability of any host connected to it to act both as a server and a client. For example, intra- and inter-domain routing, DNS, and so on were designed to operate in a distributed manner. However, over time the current deployment model for applications and some infrastructure services evolved to become more centralized and hierarchical.

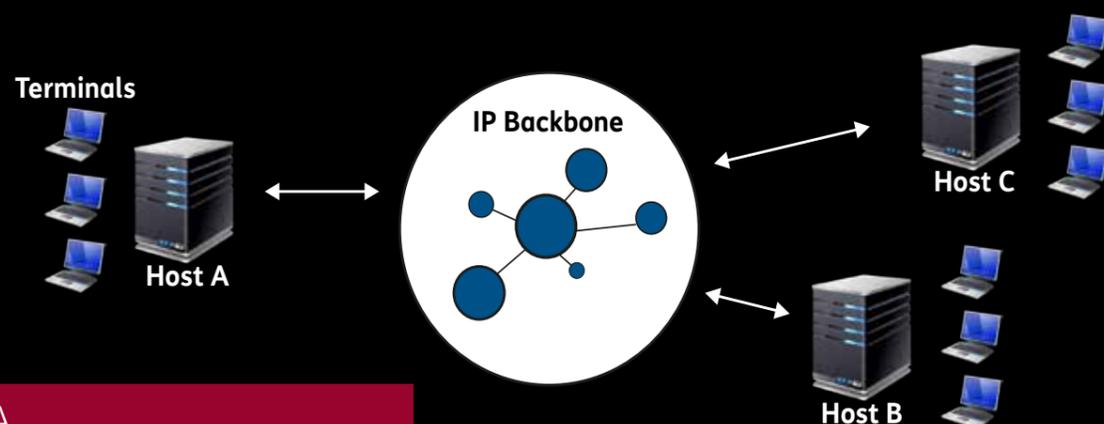
Internet is still there but is has been buried under the weight of several middleboxes, performing functions such as NAT, firewalls, port filters, and others. Nothing wrong in this, as it was a natural response to the fact that Internet protocols did not provide clear definitions of trust, security, and authentication.

A pervasive intelligence (Cloud Computing integrated with Edge Intelligence in the network nodes and smarter-and-smarter terminals) will pave the way to a future-proof Internet. Consider, for instance, also leveraging

the achievements of the emerging distributed ledger technologies and the platforms. Some examples include name resolution (Namecoin, Ethereum Name Service), identity management (OneName), distributed storage (IPFS, MaidSafe), distributed applications, or DApps (Blockstack), and IP address allocation and delegation. These topics are currently investigated in [A] connect the domain expertise in the IRTF and IETF with the distributed systems and decentralized ledgers community ■

Reference

[A] IRTF Decentralized Internet Infrastructure Research Group - <https://trac.ietf.org/trac/irtf/wiki/blockchain-federation>



A
Original Internet connection topology

even today, network system performance (e.g., traffic scheduling, routing and TCP congestion control). Another area where EI could bring performance advantages is efficient resource management and network adaptation. Example issues to address include traffic scheduling, routing, and TCP congestion control.

On the other hand, today it is rather challenging to design a real-time system with heavy computation loads and big data. This where EC enters the scene. An orchestrated execution of AI methods in the computing resources not only in the cloud but also at the edge, where most data are generated, will help in this direction. Moreover, collecting and filtering a large amount of data that contain both network profiles and performance metrics is still very critical, and that the question becomes even more expensive when considering the need of data labelling. Even these bottlenecks could be faced enabling EI ecosystems capable of engaging win-win collaborations between Network/Service Providers, OTTs, Technology Providers, Integrators and Users.

A further dimension is that a network embedded pervasive intelligence (Cloud Computing integrated with Edge Intelligence in the network nodes and smarter-and-smarter terminals) could also pave the way to leveraging the achievements of the emerging distributed ledger technologies and platforms (read more on the side).

Towards a network that design and operate itself

The idea that future networks should be able to mitigate a growing complexity by self-organizing is well known: think about the past advances in Autonomic Networking, SON (*Self-Organizing Networks*), Zero-touch Networks and the adoption of AI methods and systems in OSS/BSS and orchestrations processes.

This trend, today, is accelerating with the idea that AI could (almost) autonomously design, shape and operate the architectures of future networks, and it could do that optimally meeting the techno-economic and business real-time requirements.

It is not new, in fact, that one of the most challenging open question in AI is finding the way to automatically write code from kinds of specifications that humans can easily provide, such as via natural language instruction. A key idea is that a flexible combination of pattern recognition and explicit reasoning can be used to approach that open question. For instance, SketchAdapt [6], trained on tens of thousands of programs examples, learns how to compose short, high-level programs, while letting a second set of algorithms find the right sub-programs to fill in the details. This is just a very simple example, but it is showing an important trend, growing exponentially. AI might have the

potentiality to design and create new network modules and components with limited or even without human involvement. As a matter of fact, Generative Adversarial Networks (GANs) have also shown that the DL models have these abilities in generating new elements, modules and developing strategies that humans fail to discover with the same rapidity (e.g., AlphaGo).

A reverse side of the coin: AI is still power hungry and time consuming

The technical reason why DNNs (*Deep Neural Networks*) are so performing is not fully understood yet. One possible explanation, described in literature, is that being DNNs based on an iterative coarse-graining scheme, their functioning is somehow rooted to some fundamental theoretical physics laws or tools (e.g., Renormalization Group). It might not be a surprise that DNNs functioning resembles some profound law of Nature (imagine the complex neurons networks in the brain); but in one thing we are still very far away from emulating Nature: energy consumption.

Current AI solutions are quite resources/energy-hungry and still time-consuming. In fact, today DNNs (as other AI models) still rely on Boolean algebra transistors to do an enormous amount of digital

computations over huge data sets. The roadblock is that chipsets technologies aren't getting faster at the same pace as AI software solutions are progressing in serving markets' needs. Today, for example, Cloud servers and data centers currently account for around 2% of power consumption in the U.S. According to some forecasts, data centers will consume one fifth of the world's electricity by 2025 [7].

Will this energy consuming trend be really sustainable in the long term? We remind that in basic functioning of a DNN, each high-level layer learns increasingly abstract higher-level features, providing a useful, and at times reduced, representation of the features to a lower-level layer. This similarity is suggesting [8], the intriguing possibility that DNNs principles are deeply rooted in Quantum Field Theory and Quantum Electromagnetics [9]. This aspect is, perhaps, offering a way to bypass above roadblocks: developing AI technologies based on photonic/optical computing systems which are much faster and much less energy-consuming than current ones.

As a matter of fact, while, in line with the Moore's law, electronics starts facing physically fundamental bottlenecks, nanophotonics technologies [10] are considered promising candidates to overcome electronics future limitations. Consider that DNNs operations are mostly matrix multiplications, and nanophotonic circuits can make such operations almost at the speed of light and

very efficiently due to the nature of photons. In simple words, photonic/optical computing uses electromagnetic signals (e.g., via laser beams) to store, transfer and process information. Optics has been around for decades, but until now, it has been mostly limited to laser transmission over optical fiber. Today technologies, using optical signals to do computations and store data, would accelerate AI computing by orders of magnitude in latency, throughput and power efficiency.

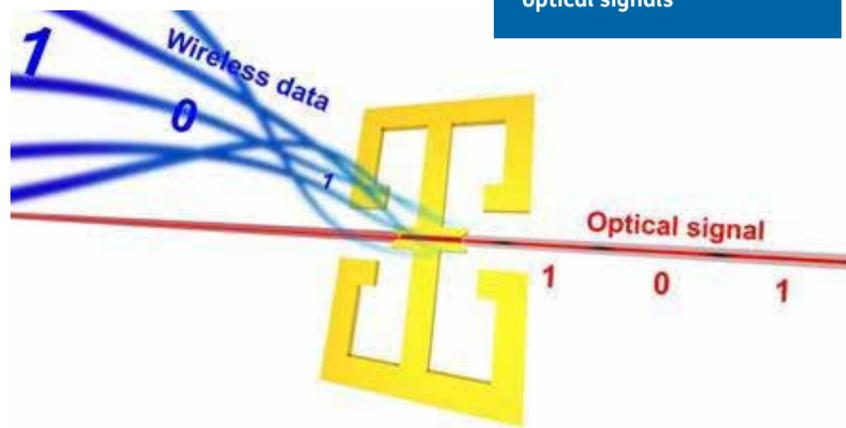
Matrix multiplication, for example, is one of the most expensive and time-consuming calculations involved in DNN, and being able to speed it up will help create much faster DNN models. Optical computing allows making any matrix multiplication, regardless of the size, in one CPU clock, while electronic chips take at least a few hundred clocks to perform the same operations. Research and innovation activities are already producing concrete results and prototypes (Box 2).

Markets movements and start-ups

are confirming this trend. Lightelligence [11], for example, is a start-up proposing a new architecture for a fully optical neural network that, in principle, could offer an enhancement in computational speed and power efficiency over state-of-the-art electronics for conventional inference tasks. LightOn [12] is another example of start-up building Optical Processing Units (OPUs) which uses light to perform computations matrix-vector multiplications with nonlinearities for AI. The OPU can just be plugged onto a standard server or workstation, and accessed through a toolbox that can be integrated within familiar programming environments.

Conclusions

A next exploitation of Edge Computing in 5G will allow migrating part of the service-specific processing and



2
Metamaterials coupling THz radio signals with optical signals

AI MATRIX MULTIPLICATION WITH "LIGHT"

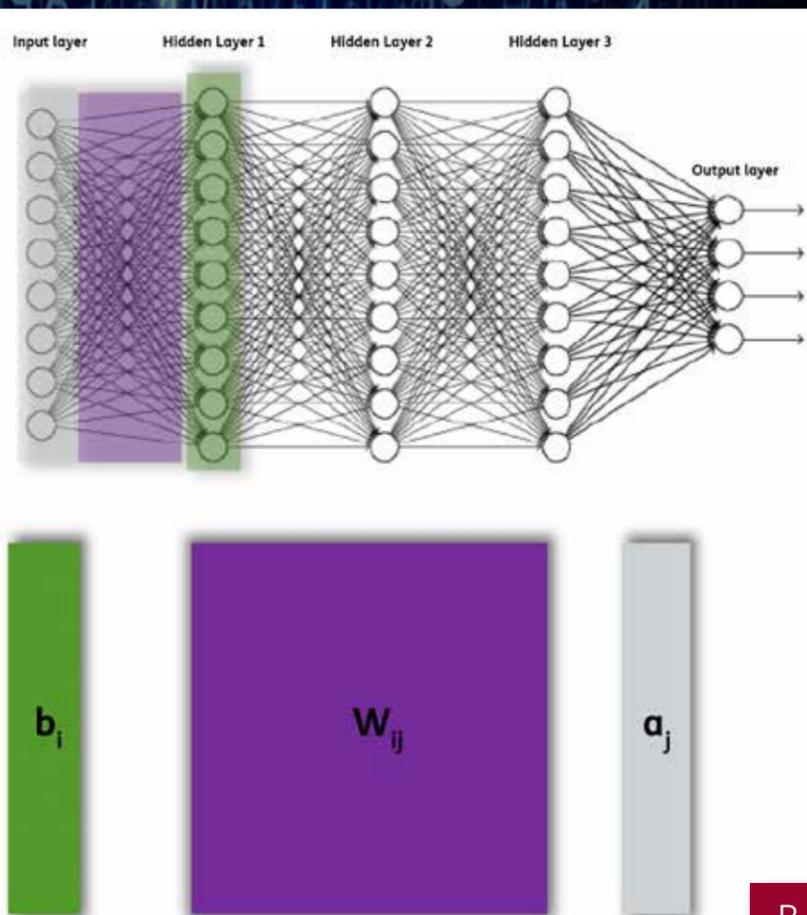
Matrix multiplication is the most power hungry and time-consuming operations in AI algorithms. Extrapolating current trends, speed of electronics components performing matrix calculations is likely to be insufficient for supporting future AI applications, at least in the long term. Reducing the electric energy consumption is also another strict requirement for sustainability.

Neuroscience may hold the solution. Our brain is not digital, it's analogue, and it makes calculations all the time using electromagnetic signals, consuming just 30 W.

The advantage of using light processing to do matrix multiplication plays significantly in speeding up calculations and power savings. In fact, instead of using streams of electrons, the calculations are performed by beams of photons that interact with one another, in a medium, and with optical resonators and guiding components. To make it simple: unlike electrons, photons have no mass, travel at light-speed and draw no additional power once generated.

Interesting prototypes of an all-optical DNN are already available. For example [B] shows the feasibility study of an all-optical diffractive DNN (D²NN). The prototype is made of a set of diffractive layers, where each point (equivalent to a neuron) acts as a secondary source of an electromagnetic wave directed to the following layer.

The amplitude and phase of the secondary wave are determined by the product of the input wave and the complex-valued transmission or reflection coefficient at that point (following the laws of transformation optics). The transmission/reflection coefficient of each point of a layer is a learnable network parameter, which is iteratively adjusted during the



B
Matrix multiplication in a DNN

training process (e.g., performed in a computer) using a classical error back-propagation method. After the training, the design of the layer is fixed as the transmission/reflection coefficients of all the neurons of all layers are determined. Other examples of prototypes are based on the metamaterials, information metasurface, and optical field-programmable gate array based on Mach-Zehnder Interferometers ■

Reference

- [B] Lin, Xing, et al. "All-optical machine learning using diffractive deep neural networks." *Science* 361.6406 (2018): 1004-1008.

data storage from the Cloud Computing to the edge network nodes. Expected benefits includes: performance improvements, traffic optimization and new ultra-low-latency services.

Today, in this scenario, we are also witnessing a growing development of Artificial Intelligence solutions for a wide spectrum of applications and services. The evolving requirements are demanding an AI more and more resources-hungry model. Therefore cloud-centric execution appears in

the opposite direction with a migration of computing, storage and networking resources at the edge. In reality, the intertwining of Edge Computing and Artificial Intelligence trajectories has already given rise to a new emerging innovation trend, called Edge Intelligence.

The orchestrated use of Edge and Cloud virtual resources, in fact, is expected to create a continuum of intelligent capabilities and functions across all the Cloudified infrastructures. This is one of the major chal-

lenges for a successful deployment of an effective and future-proof 5G. A major roadblock to this vision is the long-term extrapolations of the energy consumption needs of a pervasive Artificial Intelligence embedded into future network infrastructures. Low-latency and low-energy neural network computations can be a game changer. In this direction, fully optical neural network could offer impressive enhancements in computational speed and reduced power consumptions ■

References

- [1] Zhou, Zhi, et al. "Edge Intelligence: Paving the Last Mile of Artificial Intelligence with Edge Computing." arXiv preprint arXiv:1905.10083 (2019)
- [2] Cisco Global Cloud Index: Forecast and Methodology, 20162021 White Paper. [Online]. Available: <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/white-paper-c11-738085.html>
- [3] Lovén, Lauri, et al. "EdgeAI: A Vision for Distributed, Edge-native Artificial Intelligence in Future 6G Networks." *The 1st 6G Wireless Summit* (2019): 1-2
- [4] Chiang, Mung, et al. "Layering as optimization decomposition: A mathematical theory of network architectures." *Proceedings of the IEEE* 95.1 (2007): 255-312
- [5] Wang, Mowei, et al. "Machine learning for networking: Workflow, advances and opportunities." *IEEE Network* 32.2 (2017): 92-99.
- [6] Nye, Maxwell, et al. "Learning to infer program sketches." arXiv preprint arXiv:1902.06349 (2019).
- [7] <https://data-economy.com/data-centres-world-will-consume-1-5-earths-power-2025/>
- [8] Lee, J.W. Quantum fields as deep learning. arXiv preprint arXiv:1708.07408 2017
- [9] Manzalini, Antonio. "Complex Deep Learning with Quantum Optics." *Quantum Reports* 1.1 (2019): 107-118
- [10] Yao, K.; Unni, R.; Zheng, Y. Intelligent nanophotonics: merging photonics and artificial intelligence at the nanoscale. arXiv preprint arXiv:1810.11709 2018
- [11] <https://www.lightelligence.ai/technology>
- [12] <http://www.lighton.io>



Antonio Manzalini manzalini.antonio@telecomitalia.it

ingegnere elettronico, Ph.D è entrato in Telecom Italia nel 1990. Ha partecipato a diversi progetti di ricerca internazionali riguardanti reti di trasporto SDH ed ottico (WDM), occupando varie posizioni di responsabilità. Ha inoltre contribuito a molte attività di standardizzazione, guidando alcuni gruppi di lavoro in ITU-T, IEEE e GSMA.

Attualmente si occupa di tecnologie ed architetture di reti 5G, basate sull'integrazione di SDN, NFV con Cloud-Edge Computing e sistemi di Intelligenza Artificiale. Nel 2019, IEEE gli ha assegnato il premio Industrial Distinguished Lecturer Award. È autore di oltre un centinaio di pubblicazioni internazionali e di sette brevetti ■



Notiziario Tecnico

Anno 28 - Numero 1, Aprile 2019

www.telecomitalia.com/notiziariotecnico

ISSN 2038-1921

Registrazione

Periodico iscritto al n. 00322/92 del Registro della Stampa
Presso il Tribunale di Roma, in data 20 maggio 1992

*Gli articoli possono essere pubblicati solo se autorizzati
dalla Redazione del Notiziario Tecnico.*

*Gli autori sono responsabili del rispetto dei diritti di
riproduzione relativi alle fonti utilizzate.*

*Le foto utilizzate sul Notiziario Tecnico sono concesse
solo per essere pubblicate su questo numero;
nessuna foto può essere riprodotta o pubblicata senza
previa autorizzazione della Redazione della rivista.*