

# Estudio de P4 y su papel en la telemetría de red en banda para las redes 5G y beyond

Edinson Montero Beltre

8 de diciembre de 2025

# Índice general

<b>Dedicatoria</b>	<b>5</b>
<b>Acrónimos</b>	<b>9</b>
<b>1. Introducción</b>	<b>10</b>
1.1. Contexto general de las redes 5G y beyond . . . . .	10
1.2. Motivación del estudio . . . . .	10
1.3. Problemática en la visibilidad y monitoreo de redes 5G y beyond	11
1.4. Hipótesis y preguntas de investigación . . . . .	11
1.5. Objetivos generales y específicos . . . . .	11
1.6. Metodología de trabajo . . . . .	12
1.7. Estructura del documento . . . . .	13
1.8. Resumen de capítulos . . . . .	13
<b>2. Fundamentos de las Redes 5G y Beyond</b>	<b>14</b>
2.1. Evolución de tecnologías móviles . . . . .	14
2.1.1. De 1G a 4G: Hitos clave . . . . .	14
2.1.2. De 4G a 5G SA . . . . .	14
2.1.3. 5G Advanced y Beyond-5G . . . . .	15
2.1.4. Tendencias hacia 6G . . . . .	15
2.2. Arquitectura 5G Non Standalone (NSA) . . . . .	15
2.2.1. Opciones de arquitectura 3GPP para 5G NSA . . . . .	15
2.3. Arquitectura 5G Standalone (SA) . . . . .	17
2.3.1. Visión general de la arquitectura 5G SA . . . . .	17
2.3.2. Service-Based Architecture (SBA) . . . . .	18
2.3.3. Componentes principales: AMF, SMF, UPF, AUSF, UDM, PCF, NRF . . . . .	19
2.3.4. Funciones del plano de control y plano de usuario . . . . .	19
2.3.5. Interfaces clave: N1, N2, N3, N4, N6, N9 . . . . .	19
2.3.6. NGRAN y separación CP/UP . . . . .	19
2.4. Protocolos relevantes en 5G . . . . .	19

## ÍNDICE GENERAL

2.4.1. NGAP . . . . .	19
2.4.2. SCTP en 5G . . . . .	19
2.4.3. GTP-U . . . . .	19
2.4.4. PFCP . . . . .	19
2.5. QoS en 5G . . . . .	19
2.5.1. Conceptos de QFI, 5QI y políticas de flujo . . . . .	19
2.5.2. Programación del tráfico en UPF . . . . .	19
2.5.3. QoS-based routing y desafíos actuales . . . . .	19
2.5.4. Relación entre QoS y telemetría . . . . .	19
<b>3. Telemetría en Redes 5G</b>	<b>20</b>
3.1. Necesidad de telemetría en redes modernas . . . . .	21
3.1.1. Limitaciones de mecanismos tradicionales (SNMP, NetFlow, sFlow) . . . . .	21
3.1.2. Requisitos de visibilidad en URLLC, eMBB y mMTC .	21
3.2. In-Band Network Telemetry (INT) . . . . .	21
3.2.1. Concepto y motivación . . . . .	21
3.2.2. Arquitectura INT: source, transit y sink . . . . .	21
3.2.3. Metadatos INT-MD y su estructura . . . . .	21
3.2.4. INT vs telemetría out-of-band . . . . .	21
3.3. Telemetría aplicada en 5G . . . . .	21
3.3.1. Relevancia de INT en el plano de control N2 . . . . .	21
3.3.2. Relevancia de INT en el plano de usuario N3 . . . . .	21
3.3.3. Métricas clave para tráfico GTP-U . . . . .	21
3.3.4. Desafíos y limitaciones en redes móviles . . . . .	21
<b>4. Fundamentos de P4 y Programación del Plano de Datos</b>	<b>22</b>
4.1. Introducción a P4 . . . . .	22
4.1.1. Historia y evolución . . . . .	22
4.1.2. Modelo de arquitectura PISA . . . . .	22
4.1.3. P4_14 vs P4_16 . . . . .	22
4.2. Herramientas y ecosistema P4 . . . . .	22
4.2.1. BMv2 como switch software . . . . .	22
4.2.2. P4C: compilador . . . . .	22
4.2.3. P4Runtime: control del data plane . . . . .	22
4.3. Implementación de INT en P4 . . . . .	22
4.3.1. Definición de cabeceras INT . . . . .	22
4.3.2. Parsing y deparsing en BMv2 . . . . .	22
4.3.3. Inyección hop-by-hop de metadatos . . . . .	22
4.3.4. Limitaciones del pipeline P4 . . . . .	22

## ÍNDICE GENERAL

<b>5. Entorno de Desarrollo e Implementación</b>	<b>23</b>
5.1. Arquitectura general del sistema . . . . .	24
5.1.1. Descripción general del diseño . . . . .	24
5.1.2. Topología de red . . . . .	24
5.1.3. Interacción entre NGRAN, AMF y switches P4 . . . . .	24
5.2. Despliegue del Core 5G SA . . . . .	24
5.2.1. Componentes utilizados . . . . .	24
5.2.2. Configuración y orquestación . . . . .	24
5.2.3. Integración con UERANSIM . . . . .	24
5.3. Implementación de la red de distribución P4 . . . . .	24
5.3.1. Diseño de los programas P4 . . . . .	24
5.3.2. Tablas y acciones configuradas . . . . .	24
5.3.3. Inserción INT-MD en tráfico N2 (SCTP) . . . . .	24
5.3.4. Inserción INT-MD en tráfico N3 (GTP-U) . . . . .	24
5.4. Servidor de recolección y análisis . . . . .	24
5.4.1. Implementación del sink INT . . . . .	24
5.4.2. Procesamiento y parsing de metadatos . . . . .	24
5.4.3. Modelo de base de datos . . . . .	24
5.4.4. Visualización en Grafana . . . . .	24
5.5. Tecnologías empleadas . . . . .	24
5.5.1. Docker / Docker Compose . . . . .	24
5.5.2. GNS3 / GNS3 VM . . . . .	24
5.5.3. Wireshark y herramientas auxiliares . . . . .	24
<b>6. Resultados Experimentales</b>	<b>25</b>
6.1. Metodología de evaluación . . . . .	26
6.1.1. Escenarios de prueba . . . . .	26
6.1.2. Tráfico analizado . . . . .	26
6.1.3. Métricas recolectadas . . . . .	26
6.2. Validación funcional . . . . .	26
6.2.1. INT en tráfico N2 . . . . .	26
6.2.2. INT en tráfico N3 . . . . .	26
6.2.3. Recepción de metadatos en el servidor . . . . .	26
6.3. Resultados cuantitativos . . . . .	26
6.3.1. Latencia hop-by-hop . . . . .	26
6.3.2. Carga adicional introducida por INT . . . . .	26
6.3.3. Impacto en el plano de usuario . . . . .	26
6.4. Resultados cualitativos . . . . .	26
6.4.1. Dashboards obtenidos . . . . .	26
6.4.2. Interpretación de tendencias . . . . .	26
6.4.3. Problemas encontrados . . . . .	26

## ÍNDICE GENERAL

<b>7. Discusión</b>	<b>27</b>
7.1. Análisis crítico de los resultados . . . . .	27
7.1.1. Comparación con el estado del arte . . . . .	27
7.1.2. Validación de la hipótesis planteada . . . . .	27
7.2. Beneficios del uso de P4 en redes 5G . . . . .	27
7.3. Desafíos de escalabilidad . . . . .	27
7.4. Consideraciones de seguridad para INT . . . . .	27
7.5. Limitaciones del trabajo realizado . . . . .	27
<b>8. Conclusiones y Trabajo Futuro</b>	<b>28</b>
8.1. Conclusiones principales . . . . .	28
8.2. Contribuciones del trabajo . . . . .	28
8.3. Limitaciones del estudio . . . . .	28
8.4. Líneas futuras de investigación . . . . .	28
8.4.1. INT en 6G . . . . .	28
8.4.2. UPF programable con P4 . . . . .	28
8.4.3. IA para análisis de telemetría . . . . .	28
<b>A. Apéndices</b>	<b>29</b>
A.1. Código P4 . . . . .	29
A.2. Topologías de red . . . . .	29
A.3. Configuraciones del core 5G . . . . .	29
A.4. Capturas de tráfico . . . . .	29
A.5. Dashboards de Grafana . . . . .	29
A.6. Scripts y herramientas auxiliares . . . . .	29

# Dedicator

A quienes me apoyaron en este proyecto.

# Índice general

# Índice de figuras

2.1. Posibles combinaciones de RAN y Core en una red 5G NSA [2]. 16

# Índice de cuadros

# Acrónimos

---

<b>Acrónimo</b>	<b>Significado</b>
5G	Fifth Generation (Quinta Generación)
5GC	5G Core (Núcleo de Red 5G)
API	Application Programming Interface (Interfaz de Programación de Aplicaciones)
DNS	Domain Name System (Sistema de Nombres de Dominio)
INT	In-band Network Telemetry (Telemetría en Banda)
IP	Internet Protocol (Protocolo de Internet)
MIMO	Multiple Input Multiple Output (Entrada Múltiple Salida Múltiple)
NFV	Network Functions Virtualization (Virtualización de Funciones de Red)
P4	Programming Protocol-Independent Packet Processors
QoS	Quality of Service (Calidad de Servicio)
SDN	Software-Defined Networking (Redes Definidas por Software)
TCP	Transmission Control Protocol (Protocolo de Control de Transmisión)
TLS	Transport Layer Security (Seguridad de la Capa de Transporte)
UDP	User Datagram Protocol (Protocolo de Datagrama de Usuario)
UE	User Equipment (Equipo de Usuario)
UL	Uplink (Enlace Ascendente)
UMTS	Universal Mobile Telecommunications System (Sistema Universal de Telecomunicaciones Móviles)
UPF	User Plane Function (Función del Plano de Usuario)
URLLC	Ultra-Reliable Low-Latency Communications (Comunicaciones de Ultra Fiabilidad y Baja Latencia)

---

# Capítulo 1

## Introducción

### 1.1 Contexto general de las redes 5G y beyond

Las redes de quinta generación (5G) representan un avance significativo en la evolución de las telecomunicaciones móviles, ofreciendo mayores velocidades, menor latencia y una capacidad mejorada para conectar dispositivos masivos. Con la creciente demanda de servicios en tiempo real y aplicaciones críticas, como la realidad aumentada, los vehículos autónomos y la telemedicina, en donde la latencia y la fiabilidad son esenciales debido a que cualquier retraso puede tener consecuencias graves, la necesidad de una visibilidad y monitoreo efectivos de la red se ha vuelto crucial. En esta evolución, las bajas latencias son cruciales para los nuevos servicios en la actualidad y los futuros, incluyendo la cuarta revolución industrial.

### 1.2 Motivación del estudio

La complejidad y dinamismo de las redes 5G y superiores, plantean desafíos significativos para la gestión y el monitoreo de la red. Los métodos tradicionales de telemetría, como SNMP y NetFlow, a menudo no proporcionan la granularidad y la rapidez necesarias para detectar y resolver problemas en tiempo real otorgando una visibilidad de extremo a extremo sobre el estado de la red de forma precisa. La telemetría en banda (In-Band Network Telemetry, INT) emerge como una solución prometedora para abordar estas limitaciones, permitiendo la recopilación de datos detallados directamente desde los paquetes que atraviesan la red. Este estudio se centra en explorar el papel de P4, un lenguaje de programación para definir el comportamiento de los dispositivos de red, en la implementación y optimización de soluciones

## CAPÍTULO 1. INTRODUCCIÓN

de telemetría en banda para redes 5G y beyond.

### 1.3 Problemática en la visibilidad y monitoreo de redes 5G y beyond

A medida que las redes 5G se vuelven más complejas, la visibilidad y el monitoreo efectivos se convierten en desafíos críticos. La diversidad de servicios y la naturaleza dinámica del tráfico generan dificultades para identificar cuellos de botella, latencias elevadas y otros problemas de rendimiento. Además, la necesidad de cumplir con estrictos requisitos de calidad de servicio (QoS) y experiencia del usuario (QoE) exige soluciones de monitoreo que puedan adaptarse rápidamente a las condiciones cambiantes de la red. La problemática radica en cómo implementar mecanismos de telemetría que sean capaces de proporcionar datos precisos y en tiempo real sin introducir una sobrecarga significativa en la red.

### 1.4 Hipótesis y preguntas de investigación

La hipótesis central de este estudio es que la implementación de telemetría en banda utilizando P4 puede mejorar significativamente la visibilidad y el monitoreo de las redes 5G y beyond, permitiendo una gestión más eficiente y una mejor calidad de servicio. Las preguntas de investigación que guían este estudio incluyen:

- ¿Cómo puede P4 facilitar la implementación de soluciones de telemetría en banda en redes 5G y beyond?
- ¿Qué beneficios específicos ofrece la telemetría en banda en comparación con los métodos tradicionales de monitoreo?
- ¿Cuáles son los desafíos y limitaciones asociados con el uso de P4 para telemetría en banda en entornos 5G y beyond?
- ¿Cómo afecta la telemetría en banda al rendimiento general de la red y a la experiencia del usuario?

### 1.5 Objetivos generales y específicos

El objetivo general de este estudio es evaluar el papel de P4 en la implementación de telemetría en banda para mejorar la visibilidad y el monitoreo

## CAPÍTULO 1. INTRODUCCIÓN

de las redes 5G y beyond. Los objetivos específicos incluyen:

- Analizar las capacidades de P4 para definir y programar el comportamiento de los dispositivos de red en el contexto de la telemetría en banda, específicamente INT-MD.
- Diseñar e implementar un prototipo de telemetría en banda (INT-MD) utilizando P4 en un entorno de red 5G SA simulado.
- Evaluar el rendimiento y la efectividad de la solución propuesta en términos de visibilidad, latencia y sobrecarga de la red.
- Identificar los desafíos y limitaciones encontrados durante la implementación y proponer posibles soluciones o mejoras.

### 1.6 Metodología de trabajo

Este estudio adoptará un enfoque experimental y analítico para investigar el papel de P4 en la telemetría en banda para redes 5G y beyond. La metodología incluirá las siguientes etapas:

- Revisión bibliográfica: Se realizará una revisión exhaustiva de la literatura existente sobre telemetría en banda, P4 y redes 5G y beyond para establecer un marco teórico sólido.
- Diseño del prototipo: Se diseñará un prototipo de telemetría en banda utilizando P4, definiendo las tablas y acciones necesarias para insertar metadatos INT-MD en el tráfico N2 (SCTP) y N3 (GTP-U).
- Implementación: El prototipo se implementará en un entorno emulado usando herramientas como GNS3 y VMware workstation que contará con Routers Cisco y switches P4 (BMv2), los cuales serán programados para soportar INT-MD y por último, un servidor sink para la recolección y análisis de los datos de telemetría usando influxDB y Grafana para representar los datos. Este prototipo integrará un core 5G SA básico y una red de distribución programable con P4.
- Evaluación: Se llevarán a cabo pruebas para evaluar el rendimiento del prototipo, midiendo métricas como la latencia, la sobrecarga de la red y la precisión de los datos recopilados, así como el impacto de la telemetría en banda en la calidad del servicio.
- Análisis de resultados: Los resultados obtenidos se analizarán críticamente para identificar beneficios, desafíos y áreas de mejora.

## CAPÍTULO 1. INTRODUCCIÓN

### 1.7 Estructura del documento

El documento se estructura en varios capítulos que abordan diferentes aspectos del estudio:

- Capítulo 1: **Introducción** - Presenta el contexto, la motivación, los objetivos y la metodología del estudio.
- Capítulo 2: **Redes 5G** - Proporciona una visión general de las redes 5G, sus características y desafíos.
- Capítulo 3: **Telemetría en Redes 5G** - Explora la necesidad de telemetría, los métodos tradicionales y la telemetría en banda (INT).
- Capítulo 4: **Fundamentos de P4** - Describe el lenguaje P4, su arquitectura y capacidades relevantes para la telemetría.
- Capítulo 5: **Entorno de Desarrollo** - Detalla el entorno utilizado para implementar y probar el prototipo de telemetría en banda.
- Capítulo 6: **Resultados** - Presenta los resultados obtenidos durante las pruebas del prototipo.
- Capítulo 7: **Discusión** - Analiza críticamente los resultados, comparándolos con el estado del arte y discutiendo beneficios y desafíos.
- Capítulo 8: **Conclusiones y Trabajo Futuro** - Resume las conclusiones principales, contribuciones, limitaciones y propone líneas futuras de investigación.

### 1.8 Resumen de capítulos

Cada capítulo del documento se enfoca en aspectos específicos del estudio, proporcionando una comprensión integral del papel de P4 en la telemetría en banda para redes 5G. A lo largo del documento, se integran conceptos teóricos con aplicaciones prácticas, culminando en un análisis crítico de los resultados obtenidos y su relevancia para el campo de las telecomunicaciones móviles.

## Capítulo 2

# Fundamentos de las Redes 5G y Beyond

### 2.1 Evolución de tecnologías móviles

Las redes móviles han evolucionado significativamente desde su inicio, pasando por varias generaciones que han mejorado la capacidad, velocidad y latencia de las comunicaciones inalámbricas, pasando a ser una pieza clave en la infraestructura de telecomunicaciones global, impactando positivamente en diversos sectores como la industria, la salud y el entretenimiento, entre otros.

#### 2.1.1. De 1G a 4G: Hitos clave

**1G:** Introducción de la comunicación analógica.

**2G:** Digitalización de la voz y servicios básicos de datos (SMS).

**3G:** Introducción de datos móviles y servicios multimedia.

**4G:** Redes IP y mayor velocidad de datos.

**LTE y LTE-Advanced:** Mejoras en eficiencia espectral y latencia.

#### 2.1.2. De 4G a 5G SA

**5G Non-Standalone (NSA):** Uso combinado de 4G y 5G para una transición suave. Aumentando la capacidad y velocidad de la red.

## CAPÍTULO 2. FUNDAMENTOS DE LAS REDES 5G Y BEYOND

**5G Standalone (SA):** Arquitectura completamente nueva basada en servicios y virtualización, permitiendo nuevas funcionalidades y mejoras en latencia y confiabilidad.

### 2.1.3. 5G Advanced y Beyond-5G

**5G Advanced:** Mejoras en IA, eficiencia energética y soporte para nuevas aplicaciones, como XR y comunicaciones vehiculares.

**Beyond-5G:** Investigación en tecnologías emergentes como comunicaciones cuánticas y redes holográficas, con miras a la futura generación 6G.

### 2.1.4. Tendencias hacia 6G

**Visión de 6G:** Redes ultra confiables, latencia casi nula y capacidades de inteligencia artificial integradas de manera nativa.

**Tecnologías emergentes:** Comunicaciones terahertz, redes auto-organizadas y computación en el borde (Edge Computing).

**Aplicaciones futuras:** Realidad extendida (XR), ciudades inteligentes (Smart Cities) y redes de sensores masivos (IoT masivo).

**Desafíos:** Seguridad, privacidad y sostenibilidad ambiental.

## 2.2 Características de RAN para 5G y beyond

En 5G (Release 15) el UE puede conectarse a la red mediante dos tipos de acceso: LTE (E-UTRA) y NR (New Radio). La red Core puede ser EPC (Evolved Packet Core) o 5GC (5G Core), dependiendo de si la red es NSA o SA, respectivamente. Posteriormente, esto se amplió para que ambas celdas puedan pertenecer a 5G NR, en cuyo caso la CN es exclusivamente 5G Core. Estas diversas opciones se agrupan en el término Multi-Radio Dual Connectivity (MR-DC). MR-DC es una generalización de Intra-E-UTRA Dual Connectivity (EN-DC) y E-UTRA-NR Dual Connectivity (NE-DC) que permite que un UE se conecte simultáneamente a dos nodos de acceso, que pueden ser E-UTRA o NR. En MR-DC, un nodo de acceso actúa como el nodo maestro o master node (MN) y el otro como el nodo secundario o secondary node (SN). El MN es responsable de la señalización y el control, mientras que el SN se utiliza principalmente para la transferencia de datos.

## 2.3 Arquitectura 5G Non Standalone (NSA)

La arquitectura 5G Non-Standalone (NSA) es una configuración de red que permite la coexistencia y colaboración entre las tecnologías 4G LTE y 5G NR. En esta arquitectura, la red 4G LTE actúa como la red principal para la señalización y el control, mientras que la red 5G NR se utiliza principalmente para el transporte de datos de alta velocidad. Esta configuración permite a los operadores de red aprovechar la infraestructura existente de LTE mientras implementan nuevas capacidades de NR, facilitando una transición más suave hacia las redes 5G completas.

Desde una perspectiva operacional, la arquitectura NSA permite a los proveedores de servicios móviles activar capacidades 5G de forma gradual sin necesidad de reemplazar completamente la infraestructura de núcleo de red existente. El despliegue de NSA comienza típicamente con la instalación de nuevas estaciones base NR que coexisten con las estaciones base LTE existentes. Los dispositivos del usuario (UE) que soportan tanto LTE como NR pueden conectarse simultáneamente a ambas redes, aprovechando la mejor señal disponible para la señalización de control a través de LTE y utilizando NR para el tráfico de datos cuando está disponible. Este enfoque dual proporciona beneficios inmediatos de rendimiento sin requerir una arquitectura de núcleo completamente nueva, lo que reduce significativamente los costos de inversión durante la transición.

La arquitectura NSA introduce el concepto de "Dual Connectivity", donde un dispositivo móvil mantiene conexiones simultáneas a dos estaciones base: una primaria (master) y una secundaria (secundaria). En el caso más común de NSA, la estación base LTE actúa como el nodo maestro, proporcionando la conexión primaria y controlando el plano de control, mientras que la estación base NR actúa como un nodo secundario dedicado al tráfico de datos de alta velocidad. Esta separación de funciones permite que los operadores optimicen el uso del espectro radioeléctrico, permitiendo que LTE continúe manejando eficientemente la señalización de control y la cobertura amplia, mientras que NR proporciona capacidad adicional de datos donde la densidad de usuarios es más alta.

Desde el punto de vista del usuario final, la arquitectura NSA ofrece una experiencia mejorada en comparación con las redes LTE puras. Los usuarios pueden experimentar velocidades de descarga más altas (potencialmente en el rango de gigabits por segundo), menor latencia en aplicaciones específicas de datos, y mejor eficiencia general de la red. Sin embargo, las ventajas están limitadas principalmente al tráfico de datos, ya que los servicios de señalización y control siguen estando limitados por las especificaciones de

## CAPÍTULO 2. FUNDAMENTOS DE LAS REDES 5G Y BEYOND

LTE. Para casos de uso que requieren baja latencia extrema o requisitos ultra confiables (como comunicaciones críticas), la arquitectura NSA puede no ser suficiente, lo que subraya la necesidad eventual de migrar a 5G SA para alcanzar todas las capacidades de 5G.

### 2.3.1. Opciones de arquitectura 3GPP 5G NSA

El trabajo realizado por el 3GPP sobre la arquitectura de la red 5G dio como resultado un conjunto de alternativas arquitectónicas, fundamentadas en tres decisiones clave del propio 3GPP. Este estudio esta documentado en el reporte técnico 3GPP TR 23.799 [1], donde se describen las diferentes opciones arquitectónicas para la implementación de redes 5G NSA. A continuación, se presentan las principales opciones:

- **Opción 1:** Arquitectura basada en LTE/EPC, donde el Core de la red sigue siendo el EPC y se utiliza NR como acceso adicional.
- **Opción 2:** Arquitectura basada en 5G Core (5GC), donde tanto el acceso como el Core de la red son 5G (SA).
- **Opción 3:** Arquitectura híbrida que combina elementos de LTE/EPC y 5GC, permitiendo una transición gradual hacia 5G.

El hecho de que la red de acceso LTE y NR puedan coexistir en una misma red 5G NSA, permite a los operadores de red aprovechar la infraestructura existente de LTE mientras implementan nuevas capacidades de NR. Por lo que la red LTE (RAN) tiene dos formas de conectarse con el Core (Core):

- **Conexión mediante S1 al EPC:** En esta configuración, la red LTE se conecta al EPC a través de la interfaz S1, mientras que la red NR se conecta al EPC a través de una interfaz adaptada. Esta opción permite una integración más sencilla con la infraestructura LTE existente.
- **Conexión mediante N2/N3 al Core 5GC:** En esta configuración, la red LTE se conecta al Core 5GC a través de las interfaces N2 y N3, mientras que la red NR también se conecta al 5GC. Esto permite una integración más estrecha entre las redes LTE y NR, facilitando la gestión y el control de la red.

Teniendo en cuenta que RAN de LTE y NR pueden coexistir en una misma red 5G NSA, existen cuatro formas de implementar LTE y/o NR. [1]:

- **Opción 1:** Solo LTE para todo el tráfico de datos y señalización.

## CAPÍTULO 2. FUNDAMENTOS DE LAS REDES 5G Y BEYOND

- **Opción 2:** Solo NR para todo el tráfico de datos y señalización.
- **Opción 3:** Una combinación de LTE y NR donde LTE tiene la mayor cobertura y se utiliza para señalización, mientras que LTE y NR se utilizan para tráfico de datos.
- **Opción 4:** Una combinación de LTE y NR donde NR tiene la mayor cobertura y se utiliza para señalización, mientras que LTE y NR se utilizan para tráfico de datos.

Considerando la integración de 2 Core diferentes (EPC y 5GC) en una red 5G NSA, existen 8 formas de implementar la arquitectura 5G NSA. [2].

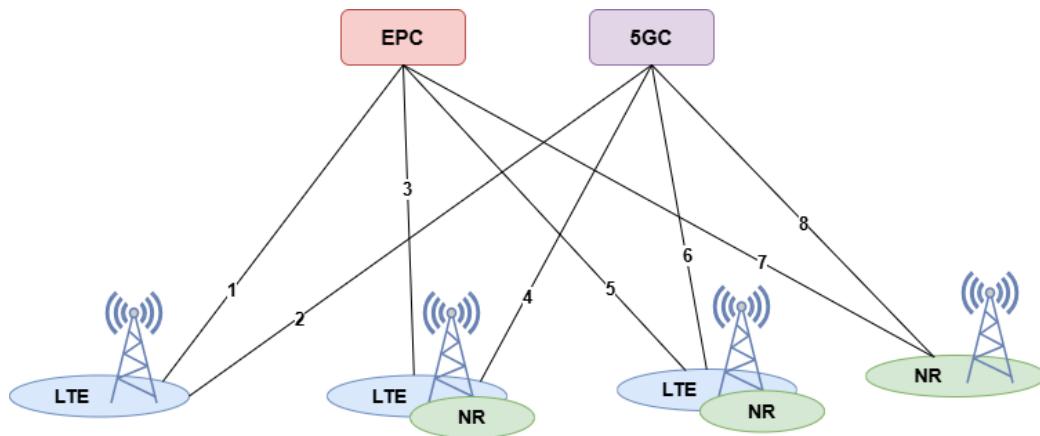


Figura 2.1: Posibles combinaciones de RAN y Core en una red 5G NSA [2].

En la figura 2.1 se muestran las diferentes combinaciones posibles de RAN y Core en una red 5G NSA  $4 \times 2 = 8$ . Las cuales son:

- **Opción 1:** RAN LTE conectada al EPC. (Solo LTE)
- **Opción 2:** RAN LTE conectada al 5GC. (Solo LTE)
- **Opción 3:** RAN LTE y NR conectadas al EPC. (LTE para señalización y datos, NR para datos)
- **Opción 4:** RAN LTE y NR conectadas al 5GC. (LTE para señalización y datos, NR para datos)
- **Opción 5:** RAN NR y LTE conectadas al EPC. (NR para señalización y datos, LTE para datos)

## CAPÍTULO 2. FUNDAMENTOS DE LAS REDES 5G Y BEYOND

- **Opción 6:** RAN NR y LTE conectadas al 5GC. (NR para señalización y datos, LTE para datos)
- **Opción 7:** RAN NR conectada al EPC. (Solo NR)
- **Opción 8:** RAN NR conectada al 5GC. (Solo NR)

Las opciones 3, 4, 5 y 6 son las que permiten la coexistencia de LTE y NR en la misma red 5G NSA, aprovechando las capacidades de ambas tecnologías para mejorar la cobertura y el rendimiento de la red. Estas opciones representan diferentes estrategias de integración que los operadores pueden adoptar según sus objetivos de despliegue y requisitos de infraestructura.

En las opciones 3 y 4, LTE actúa como la tecnología principal para la señalización de control y la cobertura amplia, mientras que NR se despliega como una tecnología complementaria para proporcionar capacidad adicional de datos de alta velocidad en áreas densamente pobladas. Esta configuración permite a los operadores mantener una cobertura LTE extendida (que puede alcanzar kilómetros de distancia) mientras aprovechan las bandas de frecuencia más altas de NR (donde la propagación es más limitada) para aumentar la capacidad de datos. Las opciones 3 y 4 difieren en el núcleo de red utilizado: la opción 3 utiliza el EPC tradicional de 4G, lo que requiere menos cambios en la infraestructura existente, mientras que la opción 4 utiliza el 5GC, permitiendo una integración más profunda con las funciones de red 5G.

Las opciones 5 y 6 invierten el rol de las tecnologías, donde NR se convierte en la tecnología primaria para la señalización y el control, mientras que LTE actúa como una tecnología secundaria para datos adicionales. Este enfoque es útil en escenarios donde los operadores han desplegado ampliamente NR y desean optimizar la utilización del espectro LTE existente o en regiones donde LTE proporciona una cobertura superior. Similar a las opciones anteriores, la opción 5 utiliza EPC como núcleo de red, mientras que la opción 6 utiliza 5GC, proporcionando mayor flexibilidad en la arquitectura de red.

La selección entre estas opciones (3, 4, 5 o 6) depende de varios factores estratégicos: la cobertura existente de LTE en la región, la disponibilidad de espectro NR, la madurez de la infraestructura 5GC disponible, los objetivos de calidad de servicio requeridos para diferentes tipos de aplicaciones, y las inversiones ya realizadas en infraestructura de red anterior. Los operadores típicamente comienzan con las opciones 3 o 4 (donde LTE es primario) para aprovechar su cobertura establecida, y gradualmente evolucionan hacia las opciones 5 o 6 (donde NR es primario) conforme la cobertura de NR se expande y se vuelve más dominante en la red.

## 2.4 Arquitectura 5G Standalone (SA)

### 2.4.1. Visión general de la arquitectura 5G SA

La arquitectura 5G Standalone (SA) representa una evolución significativa en comparación con las generaciones anteriores de redes móviles. A diferencia de las implementaciones Non-Standalone (NSA), que dependen de la infraestructura existente de 4G LTE, la arquitectura SA está diseñada desde cero para aprovechar al máximo las capacidades de la tecnología 5G. Esta arquitectura se basa en una serie de principios clave que permiten una mayor flexibilidad, eficiencia y capacidad para soportar una amplia gama de servicios y aplicaciones. Entre los aspectos más destacados de la arquitectura 5G SA se encuentran:

- **Red basada en servicios:** La arquitectura 5G SA adopta un enfoque basado en servicios, donde las funciones de red se implementan como servicios independientes que pueden ser orquestados y gestionados de manera flexible.
- **Virtualización y desagregación:** La arquitectura permite la virtualización de funciones de red (NFV) y la desagregación de hardware y software, lo que facilita la implementación en entornos de nube y mejora la escalabilidad.
- **Separación del plano de control y plano de usuario:** Esta separación permite una gestión más eficiente del tráfico y una mejor calidad de servicio (QoS) para diferentes tipos de aplicaciones.
- **Soporte para nuevas tecnologías:** La arquitectura 5G SA está diseñada para integrar tecnologías emergentes como la inteligencia artificial (IA), el edge computing y la Internet de las cosas (IoT).

### 2.4.2. Modos de implementación de 5G SA

Existen tres modos principales de implementación de redes 5G NSA, que permiten la coexistencia de tecnologías LTE y NR en una misma red. Estos modos son:

- **NGEN-DC (NG-RAN E-UTRA–NR Dual Connectivity):** el UE (User Equipment) se conecta simultáneamente a una estación base LTE (eNodeB) como master node (MN) y a una estación base NR (gNodeB) como secondary node (SN), utilizando el EPC como Core

## CAPÍTULO 2. FUNDAMENTOS DE LAS REDES 5G Y BEYOND

de la red. El gNodeB (en-gNB) se puede conectar al EPC mediante la interface S1-U o x2-U para el plano de usuario y la interfaz S1-MME o X2-C para el plano de control.

- **NE-DC (NR–E-UTRA Dual Connectivity)**: el UE se conecta simultáneamente a una estación base NR como master node (MN) y a una estación base LTE como secondary node (SN), utilizando el 5GC como Core de la red.
- **NR-DC (NR–NR Dual Connectivity)**: el UE se conecta simultáneamente a dos estaciones base NR, una como master node (MN) y la otra como secondary node (SN), utilizando el 5GC como Core de la red.

La descripción de DC se encuentra en el documento 3GPP TS 37.340 [?].

### 2.4.3. Perspectivas en 5GC

La arquitectura del núcleo de red 5G (5GC) introduce una serie de innovaciones y mejoras en comparación con las arquitecturas de núcleo de red anteriores, como el EPC utilizado en 4G. Estas mejoras están diseñadas para soportar las demandas crecientes de conectividad, velocidad y baja latencia que caracterizan a las redes 5G.

## 2.5 Arquitectura basada en servicios (SBA)

La diferencia más destacada en 5G frente a las arquitecturas 3GPP anteriores es la adopción del concepto de interfaces basadas en servicios (**SBA**) . Esto implica que las funciones de red que contienen la lógica y los mecanismos para procesar los flujos de señalización ya no se conectan mediante interfaces punto a punto, sino que **exponen sus capacidades como servicios** accesibles para otras funciones de red. En cada intercambio, una función actúa como **consumidora de servicios** y la otra como **proveedora de dichos servicios**. 2.2 muestra la arquitectura 5GC basada en SBA.

## CAPÍTULO 2. FUNDAMENTOS DE LAS REDES 5G Y BEYOND

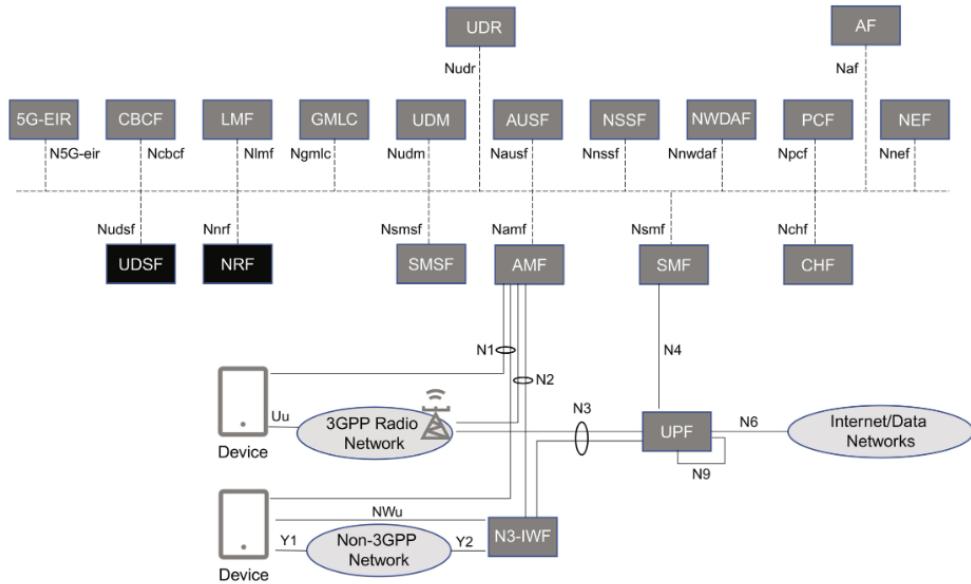


Figura 2.2: Arquitectura 5GC basada en interfaces basadas en servicios. [?].

Las funciones de red en 5GC se comunican entre sí a través de una interfaz común llamada **Service-Based Interface (SBI)**. Esta interfaz utiliza protocolos estándar como HTTP/2 y RESTful APIs para facilitar la comunicación entre las funciones de red. La adopción de SBA permite una mayor flexibilidad y escalabilidad en la arquitectura de la red, ya que las funciones pueden ser desarrolladas, desplegadas y actualizadas de manera independiente. Además, esta arquitectura facilita la integración con tecnologías emergentes como la virtualización de funciones de red (NFV) y la computación en la nube, permitiendo a los operadores de red adaptarse rápidamente a las demandas cambiantes del mercado y ofrecer nuevos servicios de manera más eficiente.

La arquitectura SBA también se puede representar como punto a punto, donde cada función de red se conecta directamente con las demás funciones que requieren sus servicios, utilizando la interfaz SBI para la comunicación. Como se muestra en la figura 2.3.

## CAPÍTULO 2. FUNDAMENTOS DE LAS REDES 5G Y BEYOND

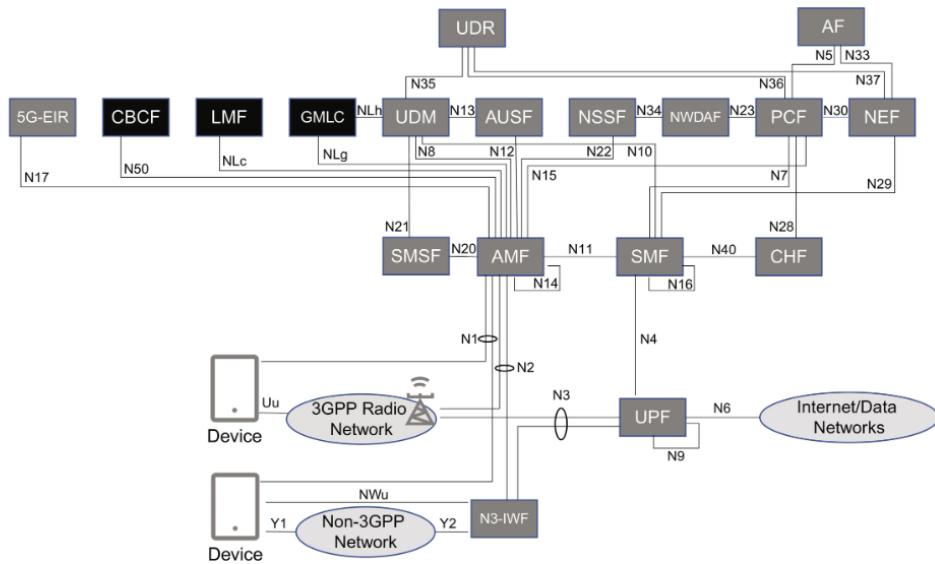


Figura 2.3: Arquitectura 5GC con interfaces punto a punto. [?].

### 2.5.1. Interfaces HTTP REST

En la arquitectura 5G Core (5GC), las funciones de red se comunican entre sí utilizando una interfaz basada en servicios conocida como Service-Based Interface (SBI). Esta interfaz utiliza el protocolo HTTP/2 junto con RESTful APIs para facilitar la comunicación y el intercambio de información entre las diferentes funciones de red. A continuación, se describen los aspectos clave de las interfaces HTTP REST en 5GC:

- **Protocolo HTTP/2:** 5GC utiliza HTTP/2 como el protocolo de transporte para las comunicaciones entre funciones de red. HTTP/2 ofrece varias ventajas sobre su predecesor, HTTP/1.1, incluyendo una mayor eficiencia en la multiplexación de solicitudes, compresión de encabezados y reducción de la latencia, lo que es crucial para las aplicaciones de baja latencia en 5G.
  - **RESTful APIs:** Las funciones de red en 5GC exponen sus capacidades a través de RESTful APIs, que son interfaces basadas en principios REST (Representational State Transfer). Estas APIs permiten a las funciones de red interactuar de manera sencilla y estandarizada, utilizando métodos HTTP como GET, POST, PUT y DELETE para realizar operaciones sobre los recursos.

## CAPÍTULO 2. FUNDAMENTOS DE LAS REDES 5G Y BEYOND

- **Formato de datos JSON:** La comunicación entre funciones de red a través de las RESTful APIs generalmente utiliza JSON (JavaScript Object Notation) como formato de datos para el intercambio de información. JSON es ligero y fácil de leer, lo que facilita la integración entre diferentes sistemas y tecnologías.
- **Seguridad:** La seguridad es un aspecto crítico en las comunicaciones entre funciones de red. En 5GC, se implementan mecanismos de autenticación y autorización para garantizar que solo las funciones autorizadas puedan acceder a los servicios expuestos a través de las RESTful APIs. Además, se utilizan protocolos seguros como TLS (Transport Layer Security) para proteger la integridad y confidencialidad de los datos transmitidos.
- **Escalabilidad y flexibilidad:** La adopción de interfaces HTTP REST permite una mayor escalabilidad y flexibilidad en la arquitectura 5GC. Las funciones de red pueden ser desarrolladas, desplegadas y actualizadas de manera independiente, lo que facilita la adaptación a las demandas cambiantes del mercado y la incorporación de nuevas tecnologías.

## 2.6 Componentes de 5G Core (5GC)

En 5G Core (5GC), los componentes tienen funciones específicas a diferencia de las arquitecturas anteriores, las cuales combinaban algunas funciones como por ejemplo, el MME y el SGW en una sola entidad llamada AMF.

### 2.6.1. Componentes principales de 5GC

#### NRF

El Network Repository Function (NRF) es un componente clave en la arquitectura 5G Core (5GC) que actúa como un repositorio centralizado para la gestión y descubrimiento de servicios de red. Su función principal es mantener un registro actualizado de todas las funciones de red disponibles en la red 5G, permitiendo que otras funciones de red puedan descubrir y comunicarse con ellas de manera eficiente. acorde a [?]. Algunas de las funciones principales del NRF incluyen:

## CAPÍTULO 2. FUNDAMENTOS DE LAS REDES 5G Y BEYOND

### 2.6.2. AMF

El Access and Mobility Management Function (AMF) es responsable de la gestión de la movilidad y el acceso de los dispositivos de usuario (UE) en la red 5G. Sus funciones principales incluyen:

Gestión de la conexión y autenticación del UE.

Manejo de la movilidad del UE, incluyendo el seguimiento y la gestión de la ubicación.

Coordinación con otras funciones de red para garantizar una experiencia de usuario fluida durante los cambios de celda y las transiciones entre diferentes tecnologías de acceso.

Gestión de la seguridad y la integridad de las comunicaciones entre el UE y la red.

Coordinación con el Session Management Function (SMF) para la gestión de sesiones de datos.

### 2.6.3. PCF

El Policy Control Function (PCF) es responsable de la gestión y aplicación de políticas en la red 5G. Sus funciones principales incluyen:

Definición y aplicación de políticas de calidad de servicio (QoS) para diferentes tipos de tráfico.

Gestión de políticas de acceso y control de recursos en función de las condiciones de la red y las necesidades del usuario.

Coordinación con otras funciones de red, como el AMF y el SMF, para garantizar que las políticas se apliquen de manera coherente en toda la red.

## CAPÍTULO 2. FUNDAMENTOS DE LAS REDES 5G Y BEYOND

- 2.6.4. SMF
- 2.6.5. UPF
- 2.6.6. AUSF
- 2.6.7. UDM
- 2.6.8. PCF
- 2.6.9. NSSF
- 2.6.10. NEF
- 2.6.11. AF
- 2.6.12. Funciones del plano de control y plano de usuario
- 2.6.13. Interfaces clave: N1, N2, N3, N4, N6, N9
- 2.6.14. NGAP y N2
- 2.6.15. NGRAN y separación CP/UP

## 2.7 Protocolos relevantes en 5G

- 2.7.1. NGAP
- 2.7.2. SCTP en 5G
- 2.7.3. GTP-U
- 2.7.4. PFCP

## 2.8 QoS en 5G

- 2.8.1. Conceptos de QFI, 5QI y políticas de flujo
- 2.8.2. Programación del tráfico en UPF
- 2.8.3. QoS-based routing y desafíos actuales
- 2.8.4. Relación entre QoS y telemetría



# Capítulo 3

## Telemetría en Redes 5G

### 3.1 Necesidad de telemetría en redes modernas

- 3.1.1. Limitaciones de mecanismos tradicionales (SNMP, NetFlow, sFlow)
- 3.1.2. Requisitos de visibilidad en URLLC, eMBB y mMTC

### 3.2 In-Band Network Telemetry (INT)

- 3.2.1. Concepto y motivación
- 3.2.2. Arquitectura INT: source, transit y sink
- 3.2.3. Metadatos INT-MD y su estructura
- 3.2.4. INT vs telemetría out-of-band

### 3.3 Telemetría aplicada en 5G

- 3.3.1. Relevancia de INT en el plano de control N2
- 3.3.2. Relevancia de INT en el plano de usuario N3
- 3.3.3. Métricas clave para tráfico GTP-U
- 3.3.4. Desafíos y limitaciones en redes móviles

## Capítulo 4

# Fundamentos de P4 y Programación del Plano de Datos

### 4.1 Introducción a P4

4.1.1. Historia y evolución

4.1.2. Modelo de arquitectura PISA

4.1.3. P4\_14 vs P4\_16

### 4.2 Herramientas y ecosistema P4

4.2.1. BMv2 como switch software

4.2.2. P4C: compilador

4.2.3. P4Runtime: control del data plane

### 4.3 Implementación de INT en P4

4.3.1. Definición de cabeceras INT

4.3.2. Parsing y deparsing en BMv2

4.3.3. Inyección hop-by-hop de metadatos

4.3.4. Limitaciones del pipeline P4



## Capítulo 5

# Entorno de Desarrollo e Implementación

### 5.1 Arquitectura general del sistema

5.1.1. Descripción general del diseño

5.1.2. Topología de red

5.1.3. Interacción entre NGRAN, AMF y switches P4

### 5.2 Despliegue del Core 5G SA

5.2.1. Componentes utilizados

5.2.2. Configuración y orquestación

5.2.3. Integración con UERANSIM

### 5.3 Implementación de la red de distribución P4

5.3.1. Diseño de los programas P4

5.3.2. Tablas y acciones configuradas

5.3.3. Inserción INT-MD en tráfico N2 (SCTP)

5.3.4. Inserción INT-MD en tráfico N3 (GTP-U)

### 5.4 Servidor de recolección y análisis

5.4.1. Implementación del sink INT

5.4.2. Procesamiento y parsing de metadatos

5.4.3. Modelo de base de datos



## CAPÍTULO 6. RESULTADOS EXPERIMENTALES

# Capítulo 6

# Resultados Experimentales

## 6.1 Metodología de evaluación

### 6.1.1. Escenarios de prueba

### 6.1.2. Tráfico analizado

### 6.1.3. Métricas recolectadas

## 6.2 Validación funcional

### 6.2.1. INT en tráfico N2

### 6.2.2. INT en tráfico N3

### 6.2.3. Recepción de metadatos en el servidor

## 6.3 Resultados cuantitativos

### 6.3.1. Latencia hop-by-hop

### 6.3.2. Carga adicional introducida por INT

### 6.3.3. Impacto en el plano de usuario

## 6.4 Resultados cualitativos

### 6.4.1. Dashboards obtenidos

### 6.4.2. Interpretación de tendencias

### 6.4.3. Problemas encontrados

# Capítulo 7

## Discusión

### 7.1 Análisis crítico de los resultados

#### 7.1.1. Comparación con el estado del arte

#### 7.1.2. Validación de la hipótesis planteada

### 7.2 Beneficios del uso de P4 en redes 5G

### 7.3 Desafíos de escalabilidad

### 7.4 Consideraciones de seguridad para INT

### 7.5 Limitaciones del trabajo realizado

## Capítulo 8

# Conclusiones y Trabajo Futuro

- 8.1 Conclusiones principales
- 8.2 Contribuciones del trabajo
- 8.3 Limitaciones del estudio
- 8.4 Líneas futuras de investigación
  - 8.4.1. INT en 6G
  - 8.4.2. UPF programable con P4
  - 8.4.3. IA para análisis de telemetría

# Capítulo A

## Apéndices

A.1 Código P4

A.2 Topologías de red

A.3 Configuraciones del core 5G

A.4 Capturas de tráfico

A.5 Dashboards de Grafana

A.6 Scripts y herramientas auxiliares