Universidad de Alcalá
Escuela Politécnica Superior

Grado en Ingeniería Telemática

Trabajo Fin de Grado

Enrutamiento y Seguridad de Señalización en el núcleo de las Redes Móviles 5G Stand-Alone

Autor: Jesús Andrés González

Tutora: Silvia Jiménez Fernández

2022

UNIVERSIDAD DE ALCALÁ

ESCUELA POLITÉCNICA SUPERIOR

Grado en Ingeniería Telemática

Trabajo Fin de Grado

Enrutamiento y Seguridad de Señalización en el núcleo de las Redes Móviles 5G Stand-Alone

Autor: Jesús Andrés González

Tutora: Silvia Jiménez Fernández

Tribunal:

Presidente: PORTILLA FIGUERAS, José A.

Vocal 1º: UTRILLA MANSO, Manuel

Vocal 2º: JIMÉNEZ FERNÁNDEZ, Silvia

Fecha de depósito: Diciembre 2022

A Sara y Alba.

Resumen

Este TFG ofrece el estudio y análisis de la gestión de la señalización en el núcleo de redes móviles 5G Stand-Alone. Está basado en el estado actual de su estandarización, así como en la implantación real de esta tecnología en las redes móviles existentes a nivel mundial. Mostrando la evolución y los cambios más relevantes con respecto a las tecnologías móviles precedentes, pone foco en las funciones de red específicamente definidas para el manejo de la señalización en el núcleo de las redes móviles 5G Stand-Alone, tanto a nivel de enrutamiento como de seguridad. Ofrece también un análisis específico de los escenarios de roaming, explorando casos de uso reales de señalización extremo a extremo.

En caso de sugerencias o comentarios sobre el mismo, dirigidlas por favor a Jesús Andrés González <jesus.andres.glez@gmail.com>.

Palabras clave: 5G Stand-alone, Señalización del Núcleo de Red, Seguridad en Señalización, Roaming 5G SA.

Abstract

This TFG provides the study and analysis of the architecture and topology required for Core Network Signaling in 5G Stand-Alone mobile networks. It is based on its current standardization status, as well as in the current real implementations of this technology in the existing mobile networks worldwide. Showing the evolution and the most relevant changes compared to the preceding mobile technologies, it focuses on the network functions specifically defined for managing signaling in the core networks of 5G Stand-alone mobile networks, both at routing level and security level. It also offers specific analysis of roaming scenarios, exploring real end to end signaling use cases.

In case of having suggestions or comments, please forward them to Jesús Andrés González <jesus.andres.glez@gmail.com>.

Keywords: 5G Stand-alone, Core Network Signaling, Signaling Security, 5G SA Roaming.

Resumen extendido

A finales del siglo XIX y comienzos del siglo XX, se produjeron dos hitos tecnológicos fundamentales como fueron por un lado la invención del teléfono, y por otro la invención de la radio y las comunicaciones inalámbricas. Durante la primera mitad del siglo XX se extendió el uso de la telefonía fija a nivel mundial, y se mejoraron las capacidades técnicas de los dispositivos de radiotelefonía privados. Desde mediados de siglo hasta la década de los 1970, se desarrollaron las primeras redes públicas de radiotelefonía, que tenían como objetivo conectar unidades móviles (principalmente en vehículos) a la red pública fija *Public Switched Telephone Network (PSTN)*. Utilizaban tecnología de radiodifusión y modulación analógica, y por tanto son comúnmente llamadas tecnologías pre-celulares o «0G». Tenían capacidad muy limitada y no eran escalables. En la década de 1970, la empresa AT&T desarrolló los conceptos fundamentales sobre redes móviles celulares y pudo realizar una demostración funcional junto con Motorola en 1973. Con el desarrollo a nivel mundial de esta tecnología celular, la década de 1980 cambió la manera en la que nos comunicamos, con lanzamientos a nivel mundial de redes móviles celulares analógicas (1G). A finales de los años 1980, existían múltiples estándares de tecnología celular analógica desplegados mundialmente con poca o nula interoperabilidad, y que además padecían las limitaciones intrínsecas de la tecnología analógica. En la década de 1990, aparece la tecnología 2G con redes enteramente digitales y con un mayor nivel de estandarización común entre diferentes operadoras y países, aunque aún con dos principales enfoques a nivel mundial: *GSM* en Europa y *cdmaOne* en Estados Unidos. Con el cambio de siglo, la década de los años 2000 trajo el 3G, principalmente enfocado en mejorar las velocidades de datos ofrecidas por la generación precedente. La década de 2010 supone la llegada de 4G con aun mayores capacidades de velocidad de datos y con un núcleo de red completamente nuevo. Por último, es aproximadamente en el año 2020 cuando la tecnología 5G empieza a dar sus primeros pasos, que exploramos en detalle en este trabajo.

La quinta generación de redes móviles celulares 5G promete habilitar nuevos casos de uso disruptivos, no posibles con las generaciones anteriores. Los nuevos casos de uso prometidos en 5G se engloban en tres áreas de aplicación principales que requieren las capacidades nuevas o mejoradas de 5G: movilidad de banda ancha mejorada o *Enhanced Mobile Broadband (eMBB)*, comunicaciones ultra-confiables de baja latencia o *Ultra Reliable Low Latency Communications (URLLC)* y Comunicaciones tipo máquina masivas o *Massive Machine Type Communications (mMTC)*. El organismo regulador *3rd Generation Partnership Project (3GPP)* es el encargado de publicar las especificaciones de esta tecnología en sus diferentes releases, quedando 5G cubierta con las releases 15, 16 y 17.

Dentro de las redes 5G, se distinguen dos enfoques fundamentales: *5G NSA (Non Stand Alone)* y *5G SA (Stand Alone)*. Con 5G NSA, se mantiene el acceso radio LTE existente junto con el núcleo de red EPC de 4G; y ambos se utilizan como ancla para la gestión de movilidad y cobertura. Sobre ellos se añade la portadora 5G. Esta opción permite a las operadoras ofrecer servicios 5G en un periodo de tiempo más corto y a menor coste. Por otro lado, 5G SA incluye la nueva red radio 5G NR, y un núcleo de red completamente nuevo que incorpora de manera nativa un nuevo conjunto de capacidades como

segmentación de red o *network slicing*, separación de plano de usuario y plano de control o *Control-User Plane Separation (CUPS)*, «cloudificación», soporte multi-Gbps o ultra-baja latencia. La mayoría de las operadoras comienzan con despliegues 5G NSA, puesto que ya disponen de redes 4G. En cualquier caso, se mantiene como objetivo final el desarrollo de redes 5G SA, donde realmente podrán aprovecharse todas las capacidades de 5G.

El ecosistema 5G se divide en 3 áreas fundamentales: los dispositivos móviles de usuario que deben evolucionar para ser compatibles con 5G, la nueva red de acceso 5G (*5G New Radio*) que incluye un mayor conjunto de frecuencias así como múltiples nuevas tecnologías de modulación avanzada; y el nuevo núcleo de red 5G (*5G Core*) que analizamos en detalle en este trabajo.

La arquitectura del núcleo de red 5G está basada en servicios a diferencia de generaciones anteriores y es denominada *Service Based Architecture (SBA)*. Cada función de red del plano de control o *Network Function (NF)* es capaz de exponer los servicios que ofrece al resto de la red, para que éstos puedan ser descubiertos de manera automática por otra NF consumidora que necesite utilizarlos. La comunicación entre ellas se realizará a través de interfaces también basadas en servicios o *Service Based Interfaces (SBI)* ya sea por comunicación directa entre ellas, o mediante comunicación indirecta a través de la entidad de red *Service Communication Proxy (SCP)*. Las diferentes funciones de red realizarán un procedimiento de registro en el *Network Repository Function (NRF)*, encargado de ser un repositorio de los perfiles de NF disponibles en la red. Cuando una función de red consumidora de servicios quiera identificar qué función de red productora puede ofrecer el servicio requerido, iniciará un procedimiento de descubrimiento de servicios contra el NRF. Se utiliza también un procedimiento de autorización de servicios, para garantizar que el consumidor de servicios de NF esté autorizado a acceder al servicio de NF proporcionado por el Productor de servicios de NF. Por último, se utiliza el procedimiento de comunicación entre servicios para las solicitudes enviadas desde los consumidores de servicios de NF a los productores de servicios de NF.

Las interfaces basadas en servicios utilizan el protocolo de señalización HTTP/2, con *JavaScript Object Notation (JSON)* como protocolo de serialización de la capa de aplicación. Para la protección y seguridad en la capa de transporte, todas las NF de 3GPP deben admitir TLS.

Dentro de las múltiples funciones de red definidas por 3GPP para el núcleo de red 5G, en este trabajo nos centramos especialmente en las relacionadas con el enrutamiento y seguridad de señalización del plano de control, como son el NRF, el SCP y el *Security Edge Protection Proxy (SEPP)*. El NRF mantiene el repositorio de funciones de red activas, basándose en los procesos de registro, actualización de registro o cancelación de registro recibidos de las diferentes NFs de la red. Permite el descubrimiento de estos perfiles por parte de otras NFs de la red, y ofrece servicio de autorización de acceso basado en tokens. El SCP es el elemento central de enrutamiento de señalización HTTP2 en el núcleo de red 5G y en la arquitectura SBA. Existen diferentes modelos de comunicación entre NFs en función de que el SCP sea utilizado (comunicación indirecta) o no (comunicación directa). El SEPP es el elemento principal para el roaming tanto a nivel de routing de interconexión como de seguridad en el límite de la red 5G.

La arquitectura de interconexión entre diferentes PLMNs para escenarios de roaming define dos posibles modelos: HR (*Home Routed*) en el que la red doméstica del usuario mantiene el control de las sesiones de usuario en la red visitada; y LBO (*Local Breakout*) en el que se da mayor control a la red visitada para gestionar al roamer recibido evitando transitar internacionalmente hasta la red doméstica. Si bien HR es el modelo utilizado en las generaciones anteriores 2G/3G/4G por la necesidad de los operadores de tener control total sobre sus abonados, es posible que en 5G la adopción de LBO sea inevitable en determinados casos de uso donde la ultra-baja latencia sea requisito indispensable (a pesar de que este modelo también fue propuesto en generaciones anteriores como 4G, pero nunca fue llevado a la práctica).

A nivel de protocolo de señalización, HTTP2 presenta diferencias fundamentales con los protocolos utilizados en generaciones anteriores (SS7 en 2G/3G y Diameter en 4G). El protocolo de señalización SS7 se adoptó para 2G en los años 1990, pero llevaba siendo utilizado desde hacía una década en sistemas de telefonía fija. Pese a tener 50 años, sigue siendo ampliamente utilizado tanto en redes de telefonía fija como en redes móviles 2G/3G. Sin embargo, su principal problema es la seguridad, puesto que los requisitos actuales son diferentes de los de la época en que fue creado. Con 4G se introdujo Diameter, que había sido creado a principios de los años 2000 ofreciendo servicios de AAA (*Authentication, Authorization and Accounting*). Aporta una gran flexibilidad a la hora de añadir nuevas aplicaciones, códigos de comando o parejas de atributo-valor (*Attribute Value Pair*), pero no soluciona los problemas de seguridad existentes en SS7. Con 5G se opta por el protocolo HTTP2, y entre otras medidas de seguridad, se requiere el uso de *Transport Layer Security (TLS)* con autenticación mutua para las comunicaciones entre diferentes NFs.

En el enrutamiento de señalización 5G con el protocolo HTTP2, son varias las cabeceras y pseudo-cabeceras que son utilizadas para identificar la información relevante para tomar las decisiones de enrutamiento. Principalmente, la pseudo-cabecera *«:authority»* contendrá el *Fully Qualified Domain Name (FQDN)* de la NF productora destino final del mensaje, o la FQDN del elemento de red de siguiente salto (que puede ser el SCP en caso de comunicación indirecta o el SEPP en casos de roaming). En los casos en los que la pseudo-cabecera *«:authority»* indica el siguiente salto, se utilizará también la cabecera customizada *«3gpp-sbi-target-apiroot»* para indicar el FQDN del destino final (NF productora) del mensaje. El SCP y SEPP, como proxies de red encargados del enrutamiento, deberán gestionar estas cabeceras, modificando sus valores a medida que procesan las diferentes peticiones de servicio.

En caso de utilizar comunicación indirecta con descubrimiento delegado, las NF consumidoras no llevarán a cabo el descubrimiento de NF productora; simplemente enviarán sus peticiones de servicio al SCP, que antes de enviar la petición a un NF productor, realizará un procedimiento de descubrimiento contra el NRF para identificar la NF productora objetivo. De este modo, el SCP es también el encargado de seleccionar el NF productor adecuado entre la lista de NFs ofrecida por el NRF. Para evitar hacer un nuevo descubrimiento con cada nueva petición de servicio, 5G define el mecanismo de vinculación entre consumidor y productor basado en cabeceras insertadas por ambas partes y que son detectadas por el SCP al tomar sus decisiones de enrutamiento.

5G incorpora el concepto de seguridad por diseño que, además del uso de TLS ya mencionado, se extiende a otros aspectos como el uso de identificadores de usuarios cifrados como SUPI (*Subscription Concealed Identifier*), o a la autorización en la comunicación entre NFs. Por otro lado, otra nueva capacidad del núcleo de red 5G es el uso de *«slicing»* de red, permitiendo dividir la red en *«slices»* lógicas con diferentes capacidades, recursos, casos de uso y usuarios finales. 5G también supone un salto por parte del sector de las telecomunicaciones móviles a las tecnologías de cloudificación y microservicios, altamente utilizadas ya en el sector TI (Tecnologías de la Información), pero que no han sido utilizadas en las generaciones móviles anteriores.

Particularizando para los escenarios de roaming, el enrutamiento entre diferentes PLMNs presenta algunas características específicas gestionadas por el SEPP. También implica requisitos de seguridad aun más estrictos y la definición por parte de 3GPP de dos modelos diferentes: TLS y *PRotocol for N32 INterconnect Security (PRINS)*, que a día de hoy siguen siendo ampliamente debatidos en foros de *GSM Association (GSMA)* con el objetivo de identificar el modelo final real de implementación.

Todos estos temas son tratados en detalle en este trabajo.

Índice general

Índice de figuras

Lista de acrónimos

256QAM	256 Quadrature Amplitude Modulation.
3GPP	3rd Generation Partnership Project.
3GPP2	3rd Generation Partnership Project 2.
5G-EIR	5G-Equipment Identity Register.
5G-GUTI	5G Global Unique Temporary Identifier.
5G-NSA	5G Non-StandAlone.
5G-SA	5G StandAlone.
5G-VN	5G Virtual Network.
5GC	5G Core.
5GMRR	5G Mobile Roaming Revisited.
5GS	5G System.
5WWC	5G Wireless Wireline Convergence.
AAA	Authentication, Authorization and Accounting.
AF	Application Function.
AI	Artificial Intelligence.
AIaaA	Artificial Intelligence as a Service.
AKA	Authentication and Key Agreement.
ALS	Application Layer Security.
AM	Amplitude Modulation.
AMF	Access and Mobility Management Function.
AMPS	Advanced Mobile Phone System.
AMTS	Advanced Mobile Telephone System.
AN	Access Network.
ANSI	American National Standards Institute.
API	Application Programming Interface.
ARP	Address Resolution Protocol.
ARP	Autoradiopuhelin.
AT&T	American Telephone & Telegraph Corporation.
AUSF	Authentication Server Function.
AVP	Attribute Value Pair.
AWS	Amazon Web Services.
BSC	Base Station Controller.
BSF	Binding Support Function.
BSS	Base Station Subsystem.
BTS	Base Transceiver Station.
CA	Certification Authority.

CA	Carrier Aggregation.
CaaS	Containers as a Service.
CAMEL	Customised Applications for Mobile network Enhanced Logic.
CAP	CAMEL Application Part.
CAS	Channel-Associated Signaling.
CBR	Citizen Band Radio.
CCITT	International Telegraph and Telephone Consultative Committee.
CCS	Common-Channel Signaling.
CD	Continuous Delivery.
CD	Continuous Deployment.
CDG	CDMA Developement Group.
CDMA	Code Division Multiple Access.
CDPD	Cellular Digital Packet Data.
CDR	Call Detailed Record.
CEA	DIAMETER Capabilities Exchange Answer.
CEPT	Conférence européenne des administrations des postes et des télécommunications.
CER	DIAMETER Capabilities Exchange Request.
CGT	Called Global Title.
CHF	Charging Function.
CI	Continuous Integration.
CN	Core Network.
CNCF	Cloud Native Computing Foundation.
CNF	Containerized Network Function.
COTS	Commercial Off-The-Shelf.
CP	Control Plane.
CR	Change Request.
CS	Circuit Switching.
CSD	Circuit Switched Data.
cSEPP	consumer SEPP.
CSFB	Circuit-Switched Fallback.
CT	Continuous Testing.
CTNE	Compañía Telefónica Nacional de España.
CU	gNB Central Unit.
CUPS	Control-User Plane Separation.
DC	Dual Connectivity.
DEA	Diameter Edge Agent.
DevOps	Development and Operations.
DHCP	Dynamic Host Configuration Protocol.
Diameter-FW	Diameter Firewall.
DL	Downlink.
DN	Data Network.
DNN	Data Network Name.
DPA	DIAMETER Disconnect Peer Answer.
DPR	DIAMETER Disconnect Peer Request.

DRA	Diameter Routing Agent.
DRB	Data Radio Bearer.
DS-CDMA	Direct Spread CDMA.
DSL	Digital Subscriber Line.
DSS	Dynamic Spectrum Sharing.
DTMF	Dual-Tone Multi-Frequency Signaling.
DU	gNB Distributed Unit.
DWA	DIAMETER Device Watchdog Answer.
DWR	DIAMETER Device Watchdog Request.
DynaTAC	Motorola DYNamic Adaptive Total Area Coverage.
E-UTRAN	Evolved Universal Terrestrial Radio Access Network.
EDGE	Enhanced Data Rates for GSM Evolution.
eLAA	enhanced Licensed Assisted Access.
eMBB	Enhanced Mobile Broadband.
eMTC	enhanced Machine Type Communication.
EN-DC	E-UTRAN - NR Dual Connectivity.
EPC	Evolved Packet Core.
ePGW	Evolved Packet Data Network Gateway.
EPS	Evolved Packet System.
ETSI	European Telecommunications Standards Institute.
FCC	Federal Communications Commision.
FD-MIMO	Full Dimension MIMO.
FDD	Frequency Division Duplex.
FDMA	Frequency Division Multiple Access.
FM	Frecuencia Modulada.
FOMA	Freedom of Mobile Multimedia Access.
FQDN	Fully Qualified Domain Name.
FRMCS	Future Railway Mobile Communication System.
FSK	Frequency-Shift Keying.
GCI	Global Cable Identifier.
GCP	Google Cloud Platform.
GERAN	GSM EDGE Radio Access Network.
GGSN	Gateway GPRS Support Node.
GLI	Global Line Identifier.
gNB	next Generation Node B.
GPRS	General Packet Radio Services.
GPSI	Generic Public Subscription Identifier.
GRS	General Radio Service.
GSA	Global Mobile Suppliers Association.
GSM	Global Standard for Mobile communications.
GSM	Groupe Speciale Mobile.
GSMA	GSM Association.
GTP	GPRS Tunneling Protocol.
GTT	Global Title Translation.
GUAMI	Global Unique AMF ID.

HLR	Home Location Register.
HNI	Home Network Identifier.
hPLMN	home Public Land Mobile Network.
HSCSD	High-Speed Circuit-Switched Data.
HSDPA	High Speed Downlink Packet Access.
HSPA	High Speed Packet Access.
HSPA+	High Speed Packet Access +.
HSS	Home Subscriber Server.
HSUPA	High Speed Uplink Packet Access.
HTTP	HyperText Transfer Protocol.
HTTP2	HyperText Transfer Protocol version 2.
HTTPS	HyperText Transfer Protocol Secure.
IANA	Internet Assigned Numbers Authority.
IDDD	International Direct Distance Dialing.
iDEN	integrated Digital Enhanced Network.
IETF	Internet Engineering Task Force.
IKE	Internet Key Exchange.
IMPI	IMS Private ID.
IMPU	IMS Public User Identity.
IMS	IP Multimedia Subsystem.
IMSI	International Mobile Subscriber Identity.
IMT-2000	International Mobile Telecommunicacion-2000.
IMT-2020	International Mobile Telecommunicacion-2020.
IMT-2030	International Mobile Telecommunicacion-2030.
IMT-Advanced	International Mobile Telecommunicacion-Advanced.
IMTS	Improved Mobile Telephone Service.
IN	Intelligent Network.
INAP	Intelligent Network Application Part.
IoT	Internet of Things.
IPsec	Internet Protocol Security.
IPUPS	Inter-Plmn User Plane Security.
IPX	IP Exchange Service.
ISAC	Integrated Sensing And Communications.
ISDN	Integrated Services Digital Network.
iSEPP	initiating SEPP.
ISO	International Organization for Standardization.
ISUP	ISDN User Part.
IT	Information Technologies.
ITU	International Telecommunications Unit.
ITU-R	ITU Radiocommunication Standardization Sector.
ITU-T	ITU Telecommunication Standardization Sector.
IWGMSC	Inter Working Mobile Switching Centre.
JOSE	Javascript Object Signing and Encryption.
JSON	JavaScript Object Notation.
JTACS	Japanese TACS.

JWE	JSON Web Encryption.
JWS	JSON Web Signature.
K8S	Kubernetes.
KPI	Key Performance Indicator.
LAA	Licensed Assisted Access.
LCS	LoCation Services.
LEO	Low Earth Orbit.
LI	Lawful Interception.
LMF	Location Management Function.
LMRS	Land Mobile Radio System.
LRF	Location Retrieval Function.
LTE	Long Term Evolution.
LTE-A	Long Term Evolution Advanced.
LTE-A-Pro	Long Term Evolution Advanced Pro.
LTE-MTC	LTE Machine Type Communication.
M2M	Machine to Machine Communications.
M2PA	MTP2 Peer-to-Peer User Adaptation Layer.
M2UA	MTP2 User Adaptation Layer.
M3UA	MTP3 User Adaptation Layer.
MAC	Media Access Control.
MAP	Mobile Application Part.
MC	Mobility Control.
MCC	Mobile Contry Code.
MCC	Mission Critical Communications.
MEC	Mobile Edge Computing.
MF	Multi-Frequency Signaling.
MIMO	Multiple Inputs Multiple Outputs.
MME	Mobility Management Entity.
mMIMO	Massive MIMO.
mMTC	Massive Machine Type Communications.
MNC	Mobile Network Code.
MO-SMS	Mobile Originating - Short Message Service.
MR-DC	Multi-RAT Dual Connectivity.
MSC	Mobile Switching Center.
MSU	Message Signal Unit.
MT-SMS	Mobile Terminating - Short Message Service.
MTA	Mobiltelefonisystem A.
MTB	Mobiltelefonisystem B.
MTD	Mobiltelefonisystem D.
MTP1	Message Transfer Part Level 1.
MTP2	Message Transfer Part Level 2.
MTP3	Message Transfer Part Level 3.
MTS	Mobile Telephone Service.
N32c	N32 control.
N32f	N32 forwarding.

N3IWF	Non-3GPP InterWorking Function.
NAI	Network Access Identifier.
NAI	Nature of Address Indicator.
NAMPS	Narrowband AMPS.
NAS	Non Access Stratum.
NB-IoT	Narrowband Internet of Things.
NE-DC	NR - E-UTRAN Dual Connectivity.
NEF	Network Exposure Function.
NF	Network Function.
NFV	Network Function Virtualization.
NFVI	Network Function Virtualization Infrastructure.
ng-eNB	Next Generation eNodeB.
NGEN-DC	NG-RAN-E-UTRA - NR Dual Connectivity.
NGMN	Next Generation Mobile Networks Alliance.
NMT	Nordic Mobile Telephone.
NPI	Numbering Plan Indicator.
NR-DC	NR - NR Dual Connectivity.
NRF	Network Repository Function.
NSI	Network Slice Instance.
NSSAAF	Network Slice Specific Authentication and Authorization Function.
NSSAI	Network Slice Selection Assistance Information.
NSSF	Network Slice Selection Function.
NTN	Non-Terrestrial Networks.
NTT	Nippon Telephone and Telegraph.
NTT-HiCAP	NTT High CAPacity.
NWDAF	Network Data Analytics Function.
OAM	Operations, Administration and Maintenance.
OBR	Origin-Based Routing.
OFDM	Orthogonal Frequency Division Multiplexing.
OFDMA	Orthogonal frequency-division multiple access.
OLT	Offentlig Landmobil Telefoni.
OSI	Open Systems Interconnection model.
PaaS	Platform as a Service.
PALM	Public Automated Line Mobile.
PC	Point Code.
PCC	Policy and Charging Control.
PCEF	Policy and Charging Enforcement Function.
PCF	Policy Control Function.
PCN	Packet Central Network.
PCRF	Policy and Charging Rules Function.
PCU	Packet Control Unit.
PDC	Personal Digital Cellular.
PDCP	Packet Data Convergence Protocol.
PDN	Packet Data Network.
PDU	Protocol Data Unit.

PEI	Permanent Equipment Identifier.
PFD	Packet Flow Detection.
PGW	Packet Data Network Gateway.
PHI	5G NR Physical Layer.
PLMN	Public Land Mobile Network.
PM	Phase Modulation.
PMR	Professional Mobile Radio.
PRINS	PRotocol for N32 INterconnect Security.
PS	Packet Switching.
PSDN	Packet Service Data Node.
pSEPP	producer SEPP.
PSTN	Public Switched Telephone Network.
PTT	Push-To-Talk.
QoS	Quality of Service.
RADIUS	Remote Authentication Dial-In User Service.
RAN	Radio Access Network.
RDSI	Red Digital de Servicios Integrados.
REST	REpresentational State Transfer.
RLC	Radio Link Control.
RNC	Radio Network Controller.
RNS	Radio Network System.
RRC	Radio Resource Control.
RRM	Radio Resource Management.
rSEPP	responding SEPP.
RTMS	Radio Telephone Mobile System.
S-NSSAI	Single - Network Slice Selection Assistance Information.
SAE	System Archicture Evolution.
SBA	Service Based Architecture.
SBI	Service Based Interfaces.
SC-FDMA	Single-carrier FDMA.
SCCP	Signaling Connection Control Part.
SCP	Service Communication Proxy.
SCP	Service Control Point.
SDO	Standards Development Organization.
SDP	Session Description Protocol.
SDR	Software Defined Radio.
SEAF	Security Anchor Function.
SEPP	Security Edge Protection Proxy.
SGSN	Serving GPRS Support Node.
SGW	Serving Gateway.
Sig-FW	Signaling Firewall.
Sigtran	SIGnaling TRANsport.
SIP	Session Initiation Protocol.
SLA	Service Level Agreement.
SM	Session Management.

SM-CP	Short Message - Control Protocol.
SM-RP	Short Message - Relay Protocol.
SMF	Session Management Function.
SMS	Short Message Service.
SMS-GMSC	Short Message Service - Gateway Mobile Switching Centre.
SMSC	Short Message Service Center.
SMSF	Short Message Service Function.
SMTP	Simple Mail Transfer Protocol.
SNI	Server Name Indication.
SP	Signaling Point.
SRB	Signaling Radio Bearer.
SRVCC	Single Radio Voice Call Continuity.
SS5	Signaling System Number 5.
SS6	Signaling System Number 6.
SS7	Signaling System Number 7.
SS7-FW	SS7 Firewall.
SSC	Session and Service Continuity.
SSP	Service Switching Point.
STP	Signaling Transfer Point.
SUA	SCCP User Adaptation Layer.
SUCI	Subscription Concealed Identifier.
SUPI	Subscription Permanent Identifier.
TA	Tracking Area.
TACS	Total Access Communication System.
TAV	Teléfono Automático en Vehículos.
TCAP	Transaction Capabilities Application Part.
TCP	Transmission Control Protocol.
TD-SCDMA	Time Division Synchronous CDMA.
TDD	Time Division Duplex.
TDMA	Time Division Multiple Access.
TIA	Telecommunications Industry Association.
TLS	Transport Layer Security.
TMA-450	Telefonía Móvil Automática 450 Mhz.
TMA-900	Telefonía Móvil Automática 900 MHz.
TR	Technical Report.
TS	Technical Specification.
TSG	Technical Specification Group.
TT	Translation Type.
UCMF	UE radio Capability Management Function.
UDM	Unified Data Management.
UDR	Unified Data Repository.
UDSF	Unstructured Data Storage Function.
UDT	Unit Data.
UE	User Equipment.
UHF	Ultra High Frecuency.

UL	Uplink.
UMB	Ultra Mobile Broadband.
UMTS	Universal Mobile Telecommunications System.
UP	User Plane.
UPF	User Plane Function.
URI	Uniform Resource identifier.
URLLC	Ultra Reliable Low Latency Communications.
UTRAN	UMTS Terrestrial Radio Access Network.
V2X	Vehicle-to-Everything.
VHF	Very High Frecuency.
VLR	Visitor Location Register.
VM	Virtual Machine.
VNF	Virtual Network Function.
VoLTE	Voice over Long Term Evolution.
VoNR	Voice over New Radio.
vPLMN	visited Public Land Mobile Network.
W-CDMA	Wideband-CDMA.
WiMAX	Worldwide Interoperability for Microwave Access.
XR	eXtended Reality.

Capítulo 1

Introducción

Dejen que el futuro cuente la verdad y evalúe a cada uno de acuerdo a sus trabajos y a sus logros. El presente es de ellos; el futuro, por el cual trabajé tanto, es mío.

Let the future tell the truth and evaluate each one according to his work and accomplishments. The present is theirs; the future, for which I really worked, is mine.

Nikola Tesla [1]

Antes de centrarnos en la tecnología 5G como objetivo fundamental de este trabajo, en esta sección ofrecemos una visión generalizada de la historia y evolución de las redes móviles en sus diferentes generaciones desde sus inicios hasta la actualidad. El objetivo no es entrar en los detalles técnicos de estas tecnologías precedentes sino poner en contexto la temática de este trabajo y conocer los hitos tecnológicos más relevantes en la historia de las comunicaciones móviles. 5G es una nueva tecnología que introduce nuevos conceptos tecnológicos disruptivos, pero aún así, se basa y apoya de manera clara en las generaciones tecnológicas precedentes.

Pondremos especial foco en los aspectos relativos al plano de control y la señalización que las diferentes tecnologías han ido utilizando.

1.1 Antecedentes de la telefonía móvil

Estudiaremos los antecedentes y orígenes de las redes móviles, que se encuentran por un lado en la invención del teléfono y posterior desarrollo a nivel mundial de la telefonía fija; y por otro se remontan a la invención de la radio y los sistemas de radio telefonía.

1.1.1 Invención del teléfono

La historia de la invención del teléfono está rodeada de largos procesos judiciales y gran cantidad de litigios y acusaciones sobre la licitud de la patente presentada por el ingeniero de origen escocés *Alexander Graham Bell* (1847-1922) en 1876. *Antonio Meucci* (1808-1889), ingeniero de origen italiano reclamó haber

[1] Nikola Tesla (1856 - 1943), inventor e ingeniero serbio nacionalizado estadounidense. Sobre las controversias en cuanto a patentes relacionadas con la invención de la radio. *Tesla, Master of Lightning (1999). Margaret Cheney* [1].

desarrollado el teléfono con anterioridad. Su «teletrófono» transmitió por primera vez la voz humana en 1849. No pudo comercializar su invento por falta de dinero, pero realizó una demostración pública en 1860 publicada en un periódico italiano de Nueva York. Registró en 1871 un anuncio de invención, más barato que el de patente, pero que requería una renovación anual que no pudo permitirse en 1874. Dos años más tarde, Bell patentó el teléfono con ciertas mejoras, después de haber visto y estudiado el *teletrófono* del italiano, y aplicando experiencias previas del ingeniero eléctrico americano *Elisha Gray* (1835-1901), quien también quiso patentar su prototipo de teléfono y no fue reconocido como inventor del teléfono por una mínima diferencia de tiempo. La patente de Bell abarcaba «el método y el aparato para transmitir los sonidos vocales u otros sonidos telegráficamente, causando ondulaciones eléctricas similares en forma a las vibraciones del aire que acompaña a dicho sonido vocal u otro sonido» [2].

Bell fundó la empresa *Bell Telephone Company* en 1877 y tras perfeccionar el teléfono comprando la patente del micrófono de carbón del inventor americano *Thomas Alva Edison* (1847-1931) en 1879, fue quien lo convirtió en un medio de comunicación de masas a escala internacional, independientemente de que la idea original fuera o no suya. La telefonía se expandió vertiginosamente a nivel mundial así como su empresa *Bell Telephone Company*, que posteriormente evolucionó hacia *American Telephone & Telegraph Corporation (AT&T Corporation)*, durante mucho tiempo la compañía telefónica más grande del mundo.

Meucci falleció en 1889 en la miseria sin ver reconocido su trabajo, aunque más de un siglo después de su muerte, en Junio de 2002, el Boletín Oficial de la Cámara de Representantes de los Estados Unidos publicó la Resolución número 269, por la que se honra la vida y el trabajo del inventor italoestadounidense [3]. En la misma se reconoce que fue más bien Antonio Meucci en vez de Alexander Graham Bell quien inventó el teléfono. Reconoció además que Meucci demostró y publicó su invento en 1860 y concluye con un reconocimiento a su autoría en dicha invención [4].

1.1.2 Invención de la comunicación por radio

Las comunicaciones inalámbricas tienen sus raíces en la invención de la radio en los años 1890, pero abarcó varias décadas desde que se establecieron sus fundamentos teóricos hasta que se logró probar la existencia del fenómeno radioeléctrico y finalmente se desarrollaron las técnicas necesarias para su uso en la transmisión inalámbrica de señales.

Desde principios de 1800, fueron varios los científicos que propusieron que la electricidad y el magnetismo estaban vinculados. En 1831 el científico inglés *Michael Faraday* (1791-1867) comenzó una serie de experimentos en los que descubrió la inducción electromagnética, relación que fue modelada matemáticamente por la *ley de Faraday*. El matemático y científico escocés *James Clerk Maxwell* (1831-1879), basándose en el trabajo experimental anterior de Faraday y otros científicos, describió en 1873 las bases teóricas de la propagación de ondas electromagnéticas; unificando todas las observaciones, experimentos y ecuaciones de la electricidad, el magnetismo y la óptica, que anteriormente no estaban relacionados en una teoría coherente. Su conjunto de ecuaciones demostró que la electricidad, el magnetismo y la luz son manifestaciones del mismo fenómeno: el campo electromagnético. Más adelante, el matemático y físico inglés *Oliver Heaviside* (1850-1925) reformuló las ecuaciones originales de Maxwell y, aunque ellos no transmitieron o recibieron ondas de radio, sus ecuaciones del campo electromagnético establecieron los principios para el diseño de radio y siguen siendo la expresión estándar del electromagnetismo clásico.

El físico alemán *Heinrich Rudolf Hertz* (1857-1894) verificó experimentalmente la teoría de Maxwell entre los años 1886 y 1889, realizando diversas demostraciones prácticas de la teoría sobre el electromagnetismo desarrollada por Maxwell, observando la radiación electromagnética y las ondas de radio. Usó una antena dipolo que constaba de dos cables en línea de un metro con un espacio para la chispa entre sus

extremos internos, así como esferas de zinc unidas a los extremos externos de los cables para la capacitancia, como un radiador. La antena era excitada por pulsos de alto voltaje de unos 30 kilovoltios aplicados entre los dos lados de una bobina de inducción de Ruhmkorff. Recibió las ondas electromagnéticas con una antena resonante de un solo bucle con un espacio de chispa micrométrico entre los extremos. Este experimento emitió y recibió por primera vez lo que ahora se denominan ondas de radio en el rango de frecuencia muy alta.

Tanto Hertz, como varios de los experimentadores que posteriormente exploraron las propiedades físicas del nuevo fenómeno, lo consideraron de poco valor práctico y no describieron ninguna aplicación potencial de la tecnología. Es por ello que el uso de las ondas de radio como medio de comunicación no se desarrolló justo después, aunque se fueron desarrollando distintos componentes electrónicos y métodos para mejorar la transmisión y detección de ondas electromagnéticas.

A principios de la década de 1890, el inventor e ingeniero serbio nacionalizado estadounidense *Nikola Tesla* (1856-1943) comenzó su investigación sobre las corrientes eléctricas de alta frecuencia. El interés principal de Tesla en el fenómeno inalámbrico era como sistema de distribución de energía, pero en 1893 propuso que también podría incorporar señales de comunicación con sus dispositivos y planeó ponerlo en práctica diseñando un proyecto experimental en el que utilizaba una torre como estación de transmisión *(Torre Wardenclyffe)*, que sería un transmisor de potencia y comunicación inalámbrica intercontinental, pero se quedó sin fondos antes de poder completarlo.

En 1895, el físico ruso *Alexander Stepanovich Popov* (1859-1905) presentó un diseño de un radiorreceptor inicialmente enfocado a la detección de rayos en tormentas, que demostró en diferentes experimentos en 1896 en la Universidad de San Petersburgo. Utilizaba un cohesor conectado a la antena, y un circuito separado con un relé y una batería que operaba una campana eléctrica.

El inventor italiano *Guillermo Marconi* (1874-1937) comenzó a trabajar en 1895 en un sistema telegráfico inalámbrico basado en ondas «Hertzianas» (de radio) desarrollando un transmisor de chispa y un receptor basado en un cohesor, que se reiniciaba automáticamente para volver al modo de recepción. Marconi obtuvo una patente británica para la radio descrita como *«Mejoras en la transmisión de impulsos y señales eléctricas»*, cuya especificación completa se presentó en 1897. Basándose en diferentes trabajos previos de otros experimentadores, desarrolló un aparato para la comunicación por radio a larga distancia.

Las contribuciones posteriores de múltiples investigadores fueron mejorando la tecnología y rediseñando las capacidades de transmisores y receptores, permitiendo cubrir mayores distancias. El físico alemán *Carl Ferdinand Braun* (1850-1918) introdujo un circuito sintonizador cerrado en la parte generadora del transmisor, y la separación de la parte radiante (la antena) por medio de un acoplamiento inductivo, así como el uso de cristales en los sistemas de recepción.

Para 1901, Marconi consiguió el hito de realizar la primera comunicación trasatlántica por ondas de radio a 3500 Km de distancia. La compañía *Marconi Telegraph Co.* fue instalando estaciones sobre las costas de Inglaterra, Estados Unidos, Italia y otros países.

Marconi y Braun recibieron en 1909 el Premio Nobel de Física *«por sus contribuciones al desarrollo de la telegrafía sin hilos»* y especialmente por las mejoras introducidas en el sistema de transmisión (circuitos resonantes magnéticamente acoplados).

Pese a ser ampliamente atribuida a Marconi, la autoría de la invención de la radio tampoco está exenta de polémica debido a las diferentes aportaciones de otros ingenieros e investigadores como Tesla o Popov.

Sin embargo, es importante resaltar que el sistema de Marconi estaba centrado en la telegrafía sin hilos, pero para transmitir señales, no sonidos. Particularizando en el concepto de telefonía inalámbrica, el

inventor canadiense *Reginald Fessenden* (1866-1932) en 1900 consiguió emitir señales habladas por medio de ondas electromagnéticas, siendo ésta considerada como la primera transmisión de radio de audio. En 1906 realizó la que se considera primera transmisión radiofónica experimental de entretenimiento y música hacia una audiencia pública general, así como una demostración de su nuevo transmisor-alternador, mostrando su utilidad para enlaces punto a punto de telefonía que podía ser utilizada tanto para transmisiones telegráficas como de audio por *amplitud modulada (AM)*.

En los años 1920, la amplificación mediante *válvulas termoiónicas*, o *válvulas de vacío*, revolucionó tanto los transmisores como los receptores. La válvula de vacío es un componente electrónico utilizado para amplificar, conmutar o modificar una señal eléctrica mediante el control del movimiento de los electrones en un espacio vacío a muy baja presión, o en presencia de gases especialmente seleccionados. Aunque el efecto de emisión termoiónica fue originalmente descrito por el físico y químico británico *Frederick Guthrie* (1833-1886) en 1873, es la investigación de *Thomas Alva Edison* el trabajo a menudo más mencionado.

Finalmente en los años 1950, la tecnología radiofónica experimentó un gran número de mejoras tras la generalización del uso del *transistor* y el desarrollo de *componentes de estado sólido*. Esto supuso el ocaso de la tecnología de válvulas de vacío para aplicaciones de comunicaciones radio pues eran mucho más pequeños, baratos y fiables que la válvula. *John Bardeen* (1908-1991), *Walter Brattain* (1902-1987) y *William Shockley* (1910-1989), tres ingenieros de la compañía *Bell Labs*, inventaron los primeros transistores en 1947 [5] revolucionando el mundo de la radio y de la electrónica moderna, recibiendo el premio Nobel de Física en en 1956. El transistor es considerado como uno de los inventos más importantes de todo el siglo XX.

1.1.2.0.1 Radiodifusión Las primeras transmisiones para entretenimiento regulares comenzaron en diferentes países en la década de 1920 impulsadas por el avance tecnológico de las válvulas de vacío, y utilizando modulación AM. En 1933, el ingeniero estadounidense *Edwin Howard Armstrong* (1890-1954) desarrolló la radiodifusión por *modulación en frecuencia (FM)*, aportando un sistema de radio de alta calidad, menos sensibles a interferencias radioeléctricas que la AM, estableciéndose de forma comercial a finales de la década.

1.1.2.0.2 Comunicación radio bidireccional punto a punto En cuanto a los sistemas de radio bidireccional, se mejoraron y extendieron, introduciéndose a principios de la década de 1920 el concepto de *transceptor*, como un dispositivo que cuenta con un transmisor y un receptor. Estos sistemas pueden ser de tipo *símplex* o *dúplex*. En el modo simplex, se dispone de una única señal para transmitir y recibir, por lo que mientras una estación transmite las otras solo pueden recibir. Es el modo más sencillo, pero una estación que empiece a transmitir mientras el canal está en uso, obstruirá las comunicaciones. En el modo dúplex se utilizan dos frecuencias, utilizando un canal para transmisión y otro para recepción. En el caso de sistemas *semidúplex*, ambas partes se turnan para transmitir en ella, estando por defecto en modo receptor, y pasando a modo transmisor al presionar el usuario un interruptor cuando desea hablar; lo que se denomina mecanismo *«push to talk» (PTT)*. Por último, en el modo de operación *full-duplex*, se permite a ambos usuarios hablar y escuchar simultáneamente, requiriendo que el sistema radio transmita y reciba simultáneamente en dos frecuencias separadas.

Los sistemas de radio móvil terrestre *LMRS (Land Mobile Radio System)* son sistemas de comunicación de voz persona a persona que utilizan transceptores radio bidireccionales. Los sistemas públicos de radio móvil terrestre se utilizan para servicios de seguridad pública como policía, bomberos, militar y otros servicios gubernamentales, con frecuencias especialmente reservadas para ello. Los servicios de radio móvil terrestre privado tienen fines comerciales para servicios como taxis, aviación, marítimo, etc. En algunos

países también hay bandas reservadas para uso ciudadano amateur. La asignación de frecuencias de radio amateur se realiza por las autoridades nacionales de telecomunicaciones, con la *ITU (International Telecommunications Unit)* supervisando cuánto espectro se asigna para estos fines. También existen las denominadas bandas de frecuencias ciudadanas, como el *CBR (Citizen Band radio)* en Estados Unidos o el *GRS (General Radio Service)* en Canadá, donde las comunicaciones pueden ser de tipo personal y también empresarial.

En los primeros tiempos de los radioaficionados, los receptores y transmisores eran construidos por separado. Entre 1930 y 1950, no existían equipos comerciales fácilmente disponibles y en la mayoría de los casos los radioaficionados construían sus propios equipos con válvulas de vacío. Estos servicios de radiotelefonía utilizan sistemas de modulación simple como AM o FM.

El primer transceptor de radio portátil apodado *«walkie-talkie»* fue el SCR-300 de Motorola, creado en el año 1940 y utilizado inicialmente para uso militar por parte de Estados Unidos durante la Segunda Guerra Mundial. Era transportado en una mochila. Posteriormente Motorola desarrolló el *«Handie-Talkie»* que podía ser sostenido completamente en la mano. Ambos dispositivos utilizaban tecnología de válvulas de vacío. Tras la Segunda Guerra Mundial, estos dispositivos tuvieron un uso civil, siendo utilizados por operadores de radio amateur.

1.2 Historia de las redes móviles

Tras analizar los antecedentes de la telefonía móvil, estudiaremos aquí los primeros servicios públicos de telefonía móvil con tecnología pre-celular o «0G», para continuar con la primera generación de telefonía móvil celular de tipo analógico (1G); así como las diferentes generaciones de tecnología móvil celular de tipo digital (2G, 3G, 4G) que preceden al desarrollo de 5G.

1.2.1 Inicios de las redes móviles. 0G

Para la década de 1940, teníamos la telefonía fija completamente extendida a nivel mundial, y el avance tecnológico de la radiotelefonía móvil bidireccional desarrollado y disponible. Sin embargo, hasta entonces estos sistemas de radio telefonía no tenían relación o interconexión con las redes públicas de telefonía fija, ni un uso público de propósito general. El desarrollo de las redes móviles de telefonía surge de la combinación de estas dos grandes tecnologías.

Los sistemas de comunicaciones móviles descritos en esta sección tenían como objetivo conectar unidades móviles con la red pública fija *PSTN (Public Switched Telephone Network)* y estaban fundamentalmente pensados para uso en coches o camiones. No utilizaban la tecnología móvil celular que fue introducida en 1G sino radiodifusión y modulación analógica, y por tanto son comúnmente llamadas tecnologías *pre-celulares* o *«0G»*.

Es en los años 1940, la compañía estadounidense *AT&T* junto con su unidad de investigación *Bell Labs*, probó servicios de comunicación públicos de radiotelefonía establecidos en algunas ciudades estadounidenses, con un sistema llamado *MTS (Mobile Telephone Service)* utilizando radio *VHF*. Pensado para ser utilizado en coches, AT&T probó este servicio en 1946 en San Luis (Missouri)[5]. Utilizaba un único transmisor con 6 canales de transmisión y estaba pensado para conectarse a la red fija PSTN. Este servicio de telefonía inalámbrica tenía grandes limitaciones. Un único transmisor en una torre central proveía un reducido número de canales de transmisión para todo un área metropolitana. Los abonados tendrían que escuchar primero a quien estuviera ocupando la línea, esperar a que quedara libre para señalizar al operador y que el operador estableciera la llamada. Una vez en la llamada, el mecanismo era

de push-to-talk. El sistema utilizaba tubos de vacío, ocupaba gran parte del maletero y pesaba más de 35 Kilogramos con un gran consumo de energía y un altísimo coste [6].

En 1958, Alemania lanzó el sistema *A-Netz*, también como red de conmutación manual, con una capacidad máxima de 10.000 usuarios.[7]

La Unión Soviética utilizó el sistema de radiotelefonía móvil *Altai* introducido en 1963 como una red de conmutación automatizada *UHF/VHF* que permitía a nodos móviles conectarse a teléfonos fijos. Se utilizó en las grandes ciudades de la Unión Soviética y fue inicialmente pensado para usos gubernamentales oficiales y para servicios de emergencia, aunque finalmente fue extendido a un uso generalizado.

En 1965, en Estados Unidos, un sistema mejorado de *MTS* llamado *IMTS (Improved Mobile Telephone Service)* que utilizaba radio *VHF* y *UHF* en la banda de 450 MHz., combinado con un pequeño aumento del espectro disponible asignado por el regulador en USA *FCC (Federal Communications Commision)*; permitió la utilización de más canales y más abonados simultáneos, así como la eliminación del mecanismo push-to-talk en favor de una comunicación full-duplex. También permitía marcado directo (conmutación automática) sin necesidad de contactar con un operador. El teléfono originaba una señal de conexión, la radiobase respondía con el tono y el teléfono respondía con su identificación (un código de área y los últimos cuatro dígitos del número telefónico asignado al subscriptor) a veinte pulsos por segundo. Luego el teléfono enviaba el número discado y la radiobase conectaba la llamada con el otro subscriptor. Aún así, la capacidad era muy limitada, hasta el punto de que Bell System limitó el servicio a 40.000 abonados seleccionados en base a acuerdos con las agencias reguladoras estatales. Por ejemplo, 2000 abonados en la ciudad de Nueva York compartían tan solo 12 canales y esperaban típicamente 30 minutos para realizar una llamada. Había una gran lista de espera de potenciales abonados demandando el servicio, demanda que solo podría ser cubierta con una mejora en la tecnología [6].

Japón lanzó en 1965 el sistema de radiotelefonía móvil *AMTS (Advanced Mobile Telephone System)*, utilizado en los sistemas de radio portátiles japoneses y operando en la banda de 900 MHz.[8].

Noruega desarrolló la red móvil *OLT (abreviación en noruego de Offentlig Landmobil Telefoni, o Telefónía Móvil Pública Terrestre)*, introducida en 1966. Utilizaba radio VHF y modulación FM [9].

Un sistema similar se utilizó en Suecia llamado *MTD (abreviación en sueco de Mobiltelefonisystem D, o Sistema de Telefonía Móvil D)* introducido en 1971 como evolución de los sistemas anteriores *MTA* y *MTB* [10].

Finlandia optó por *ARP (abreviatura en finlandés de Autoradiopuhelin, o Radioteléfono de coche)*, que fue lanzada en 1971 y alcanzó un 100 % de cobertura nacional en 1978 con 140 estaciones base[11]. En ocasiones es denominada como tecnología «0.5G» puesto que utilizaba celdas, aunque no se permitía el traspaso entre ellas.

Alemania utilizó el sistema *B-Netz* en 1972. A diferencia de su predecesora *A-Netz*, permitía marcado directo, sin necesidad de un operador humano para conectar llamadas. Al igual que ARP, puede entrar en la categoría de «0.5G» puesto que utilizaba celdas pero no permitía el traspaso entre ellas. También se implementó en los paises vecinos Austria, Holanda y Luxemburgo; aunque sin permitir roaming entre ellos [7].

En España, la *Compañía Telefónica Nacional de España (CTNE)*, (hoy simplemente conocida como *Telefónica*), lanzó el servicio del *«Teléfono Automático en Vehículos» (TAV)* en 1976 en las grandes ciudades de Madrid y Barcelona en la banda de los 160MHz. Era una tecnología pre-celular rudimentaria que utilizaba las técnicas y procedimientos de los equipos de radiotelefonía móvil privada de la época. Utilizaba un transceptor multicanal y una serie de terminales receptores multicanales, que eran muy pesados, necesariamente instalados en vehículos [12].

Años más tarde, en 1983, pero aún como tecnología pre-celular, Canadá desarrolló el sistema *Autotel*, también llamado *PALM (Public Automated Line Mobile)* como un punto intermedio entre las tecnologías pre-celulares y las celulares. El sistema no era celular, utilizaba canales VHF de alta potencia y usaba canal de voz analógico. Sin embargo sí incorporaba la utilización de señalización digital para mensajes de supervisión (establecimiento de llamada, tono, asignación de canal, etc), como varios de los sistemas 1G que estaban desarrollándose en paralelo; y es por lo que puede ser considerada como tecnología «0.5G». Se utilizó en las zonas rurales de Columbia Británica ofrecido por *British Columbia Telephone Company* (posteriormente *Telus*), donde construir una red con terminales celulares de baja potencia para cubrir áreas rurales y forestales habría sido excesivamente costoso. Terminó su servicio en 2009. [13] [14].

Todos estos sistemas predecesores de la tecnología móvil celular, también denominados «0G», tenían como principal problema que la tecnología no era escalable. Rápidamente los sistemas quedaban saturados debido a su limitada capacidad.

1.2.2 1G

1.2.2.1 Comienzos de las redes móviles celulares

AT&T había estado solicitando más espectro a FCC durante las décadas de 1950 y 1960, pero fue finalmente en 1968 cuando *FCC* solicitó a *AT&T* una propuesta para usar una franja significativa de espectro que se había asignado a los canales de televisión UHF previamente, pero que no estaba siendo utilizada.

Un equipo de ingenieros de investigación en Bell Labs, había comenzado a trabajar en telefonía móvil avanzada dos años antes. La idea incluía el uso de múltiples transmisores y receptores de baja potencia repartidos en una región o a lo largo de una autopista en series de celdas, con diferentes frecuencias usadas en celdas adyacentes pero reutilizadas en la ciudad o a lo largo de la autopista; y una manera de conmutar las llamadas a celdas adyacentes de manera automática según el vehículo se movía por la carretera. Desarrollando una red de celdas hexagonales y con la disponibilidad de computadoras y electrónica de estado sólido, pudieron proponer un sistema funcional. En 1971, AT&T presentó su informe a la FCC para que AT&T operara un sistema de telefonía celular analógica en el espectro asignado por la FCC[6].

Curiosamente, la idea original de un sistema de este tipo se remontaba muchos años atrás, cuando en 1947, el ingeniero de Bell Labs *Douglas H. Ring* (1907-2000), escribió un memorando interno en el que proponía el desarrollo de un sistema telefónico celular para AT&T [15]. Se necesitó de un largo periodo de tiempo hasta que el desarrollo de la tecnología permitió llevar a cabo esta idea, pero los fundamentos estaban aquí presentados.

AT&T consideró el sistema como una mejor manera de brindar servicio telefónico a los vehículos en movimiento. Dado que Motorola había fabricado para AT&T con anterioridad todo el equipamiento de vehículo del sistema existente, AT&T compartió su trabajo en el nuevo sistema con Motorola. Motorola por un lado argumentó ante la FCC que el nuevo espectro y la tecnología no deberían ser competencia exclusiva de la compañía telefónica, y por otro puso a trabajar a un equipo en el desarrollo de un teléfono celular, bajo la dirección del vicepresidente de la compañía, *Martin Cooper* (1928 - ...). El equipo de Cooper tuvo éxito, y consiguió realizar una demostración del prototipo funcional *Motorola DynaTAC (DYNamic Adaptive Total Area Coverage)* en abril de 1973. Pesaba 1,28 Kg. y su batería recargable solo duraba unos 30 minutos de llamada; pero funcionó. Su primera llamada, realizada desde una calle de la ciudad de Nueva York minutos antes de la manifestación pública programada, fue a su homólogo *Joel S. Engel* (1936 - ...) en Bell Labs[5].

La demostración atrajo la atención de los medios y de la FCC. No obstante, la FCC tardó ocho años en decidir qué hacer. Mientras los comisionados de la FCC deliberaban, autorizaron a AT&T a demostrar un prototipo de sistema comercial en funcionamiento en Chicago en 1978, y a Motorola un sistema en Washington el año siguiente.

Del mismo modo, otros países a nivel mundial también comenzaron sus desarrollos de redes móviles celulares analógicas.

1.2.2.2 Estándares 1G

La década de los 1980 cambió la manera en la que nos comunicamos, contemplando el lanzamiento de las redes móviles celulares analógicas a nivel mundial. 1G hace referencia a esta primera generación de redes móviles celulares de tipo analógico. La tecnología celular utiliza una red de celdas a lo largo de un área geográfica utilizando transmisores radio de baja potencia y, en esta primera generación analógica, tecnología *FDMA (Acceso Múltiple por División de Frecuencia)* o *FDD (Duplexación por División de Frecuencia)*, dividiendo la banda de frecuencias disponibles en múltiples canales, cada uno pudiendo transportar una conversación de voz, que no estaba encriptada. Por lo general, para la señalización de control, utilizan modulaciones digitales sencillas *FSK (Frequency-Shift keying o Modulación por desplazamiento de Frecuencia)*, en la que se utilizan 2 o más frecuencias diferentes para cada símbolo.

Nippon Telephone y Telegraph (NTT) lanzó en Japón la primera red móvil celular a nivel mundial en 1979 como un estándar privado llamado *NTT System* operando en la banda de 800 MHz en la ciudad de Tokio, con *modulación PM* (modulación en fase) para los canales de voz y utilizando modulación digital FSK para la señalización del plano de control. Posteriormente evolucionó a *NTT-HiCap (NTT High Capacity, alta capacidad)* en 1988. Estas redes fueron apagadas en 1999 [16].

Fue seguido en 1981 por el lanzamiento de *Nordic Mobile Telephone (NMT)* como estándar abierto en los países escandinavos Suecia y Noruega, así como en Arabia Saudí; y posteriormente en otros países como Dinamarca, Finlandia, España, Austria, Bélgica, Luxemburgo y Holanda, entre otros muchos. Inicialmente utilizaba la banda de 450 MHz (*NMT-450*), y en 1986 introdujo la variante *NMP-900* en la banda de 900 MHz para aportar un mayor número de canales disponibles. Los canales de voz utilizaban *modulación en frecuencia (FM)*. Las diferentes redes NMT a nivel mundial fueron discontinuadas durante los primeros años de la década de los 2000 [17].

En España, basándose en el estándar NMT, la CTNE (Telefónica) inició en 1982 la prestación del servicio de *Telefonía Móvil Automática (TMA-450)* en la banda de 450 MHx. Estaba pensada para coches y precisaba de grandes equipos, teniendo escasa aceptación [18].

En Estados Unidos, tras 8 años de espera, la FCC finalmente asignó frecuencias para la telefonía móvil celular en 1981. AT&T y su subsidiaria Bell Labs lanzaron el sistema celular analógico *AMPS (Advanced Mobile Phone System)* en 1983, también como estándar abierto. Fue adoptado en 1984 por Canadá y Corea del Sur, y por muchos otros países en los siguientes años. Utilizaba la banda de 800-900 MHz y modulación analógica FM.

En AMPS, la señalización entre el móvil y la estación base es de 10 kbps, con la modulación FSK y *codificación Manchester* que lleva la velocidad de bits a 20 kbps. El control de errores se logra mediante la repetición múltiple de cada palabra de señalización (5 u 11 veces dependiendo del canal), con una votación mayoritaria aplicada en el receptor para corregir los errores. También se aplica un código de bloque de tipo BCH para detectar cualquier error no corregido [19].

Países como Alemania, Italia y Francia optaron por estándares privados (*Net-C*, *Radio Telephone Mobile System RTMS* y *Radiocom 2000* respectivamente) lanzados en 1985 que también usaban la banda

de 450 MHz. Al igual que en el caso de Japón, se trataron de estándares desarrollados por las administraciones nacionales de Correos, Telegrafía y Telecomunicaciones en colaboración con proveedores de electrónica locales potentes (*Siemens* en Alemania, *Italtel* en Italia, *Matra* en Francia y una combinación de *NEC, Fujitsu y Mitsubishi* en Japón), que eran los únicos proveedores de infraestructura y dispositivos del estándar (al menos durante sus primeros años). Estos estándares no fueron adoptados en ningún otro país más allá del propio [20].

Reino Unido e Irlanda en 1985, adoptaron una variante del sistema estadounidense AMPS denominado *TACS (Total Access Communication System)*, que tuvo también lanzamiento posterior en muchos otros países incluyendo a Hong-kong, China, España, Italia y Austria. ETACS (Extended TACS) fue una versión extendida que utilizaba más canales [21]. Los últimos servicios ETACS en Reino Unido e Irlanda fueron discontinuados en 2001 [22]. Japón, que utilizó múltiples estándares 1G, utilizó también una versión de TACS llamada *JTACS (Japanese TACS)* [16].

En España, fue en 1990 cuando Telefónica puso en funcionamiento el sistema *TMA-900 (Telefonía Móvil Automática)* en la banda de 900 MHz, basado en ETACS [12]. Desde 1993 fue comercializado bajo la marca *Moviline* y desapareció por completo en 2004 [23], al decidir Telefónica centrarse en el servicio de telefonía móvil digital (bajo la marca Movistar) [24].

1.2.2.3 Necesidad de salto tecnológico

A finales de los años 1980 y principios de los años 1990, existían múltiples estándares de tecnología celular analógica desplegados mundialmente. Entre ellos existía poca o nula interoperabilidad, y además padecían las limitaciones intrínsecas de la tecnología analógica, a nivel de capacidad limitada o seguridad en la transmisión de voz. Era clara la necesidad de desarrollar estándares comunes de tipo digital que permitieran mayores capacidades y la interoperabilidad en comunicaciones internacionales.

Las limitaciones de la tecnología existente analógica también dieron lugar a que varias empresas investigaran sobre un uso más eficiente del espectro a diferencia de la tecnología FDMA utilizada en los sistemas analógicos. Esto resultó en dos nuevos sistemas de transmisión digital distintos, *TDMA (Acceso Múltiple por División en el Tiempo)*, introducido en 1989 por la *TIA (Telecommunications Industry Association)* y *CDMA (Acceso Múltiple por División de Código)*, introducido en 1995 por *Qualcomm, Inc.*. Estas tecnologías combinaban compresión digital de voz y modulación digital. Multiplicaban por 10 o 20 las capacidades de las modulaciones analógicas, aunque eran incompatibles entre ellas. Sobre ellas se construirían los primeras estándares de tecnología celular digital.

1.2.2.3.1 Estados Unidos En esta situación, Estados Unidos intentó sacar el máximo provecho de la tecnología celular analógica desplegada, mientras se preparaba para adentrarse en la telefonía celular digital.

De este modo, Motorola propuso en 1991 un sistema todavía analógico, como mejora de AMPS, llamado *NAMPS (Narrowband Advanced Mobile Phone Service)* o *IS-88*. Utilizaba canales con solo 10KHz de ancho de banda (en lugar de los 30KHz de AMPS), permitiendo así transportar el triple de llamadas que el sistema original. Es una solución intermedia entre la tecnología analógica AMPS y la tecnología digital. Su uso principal fue proporcionar mayor capacidad en las zonas urbanas donde la capacidad de AMPS era insuficiente [25].

De modo similar, a principios de los años 1990 se desarolló *CDPD (Cellular Digital Packet Data)* como una mejora sobre AMPS desplegado por primera vez en 1994. CDPD permitía la transferencia de paquetes de datos sobre canales analógicos utilizando canales vacíos para transmisiones cortas de datos con velocidades de 19.2 Kbps [26].

1.2.2.3.2 Europa La situación en Europa, con diferentes países implementando diferentes estándares de red analógicos con poca o nula interoperabilidad, así como las limitaciones intrínsecas de la tecnología analógica, dejaba clara la necesidad de desarrollar un estándar común de tipo digital que permitiera la interoperabilidad en comunicaciones internacionales.

En 1982, un consorcio llamado *Groupe Speciale Mobile (GSM)* fue formado por la *Conferencia Europea de Administraciones de Correos Telecomunicaciones (CEPT)* y comenzó a definir la nueva generación de redes móviles. El estándar resultante, denominado *GSM (Global Standard for Mobile communications)* en referencia a sus creadores, permitiría a las redes integrarse entre ellas de manera transparente, estableciendo las bases de las comunicaciones móviles digitales tal cual las entendemos hoy día [21].

1.2.3 2G

En la tecnología 1G de los años 1980, el audio era codificado como señales de radio analógicas, aunque en algunos casos el establecimiento de llamada y otras comunicaciones de red en el plano de control fueran digitales. La gran diferencia incorporada en 2G es que las redes eran enteramente digitales. Aparecieron a finales de la década de 1980 y comienzos de la década de 1990.

1.2.3.1 GSM, GPRS, EDGE/EGPRS

El estándar *GSM (Global Standard for Mobile Communications)* fue desarrollándose durante la década de 1980. En 1986, la Unión Europea solicitó a los países miembros liberar las bandas de frecuencia de los 900 MHz para esta futura red móvil celular paneuropea. En 1987, 13 países miembros se comprometieron a implementar un sistema móvil celular digital en sus redes, incluyendo el roaming entre ellos para el 1 de Julio de 1991; fecha que posteriormente se retrasaría un año. En 1989, la responsabilidad de las especificaciones GSM pasó de CEPT al recién creado *ETSI (European Telecommunications Standards Institute)*. En 1990 se completaron las especificaciones y en Diciembre de 1991 la operadora *OY Radiolinja AB* de Noruega se convirtió en la primera red comercial 2G GSM [21]. En España, el servicio GSM se lanzó en 1995 por parte de *Movistar* y de *Airtel* (ahora Vodafone), tras liberalizarse el mercado de telefonía móvil el año anterior.

GSM es un sistema radio de tipo TDMA y bandas portadoras con un ancho de banda de 200 KHz. Cada banda se compone de 8 ranuras portadoras. Es un sistema de *Conmutación de Circuitos, Circuit Switching (CS)* en el que cada ranura portadora (un circuito) se asigna a cada canal de comunicación de voz por lo que hasta ocho abonados pueden compartir una banda portadora. Una única celda soporta múltiples bandas portadoras. En GSM las frecuencias radio usadas para las bandas portadoras pueden ser reutilizadas entre celdas mientras que los radiotransmisores que utilicen esas bandas no estén en celdas adyacentes [27].

GSM soporta servicios básicos de datos *CSD (Circuit Switched Data)* con capacidad limitada ya que se utiliza un solo canal de voz para la transmisión de datos. Solo se asigna una ranura de portadora inalámbrica de una banda portadora GSM a la transferencia de datos, de modo que la tasa de transferencia está limitada a 9,6 kbps. Además, la tarificación es por tiempo de conexión, como en el caso de las llamadas de voz. Una mejora en el mecanismo de transmisión de datos llamada *HSCSD (High-Speed Circuit-Switched Data)* fue aprobada por ETSI en 1997 y desplegada comercialmente por primera vez en 1999. Permitía asignar al usuario hasta 4 canales temporales y transmisión con menor detección de errores, aumentando la velocidad hasta 57,6 Kbps [28].

La red móvil de un operador, también denominada *PLMN (Public Land Mobile Network, o Red Móvil Terrestre Pública)* consta de estos elementos en una red GSM según los estándares ETSI:

- **BSS**: *Subsistema de Estación Base* que incluye BTS y BSC.
  - **BTS**: *Estación Transceptora Base* que incluye las antenas y maneja la transmisión de radio a los terminales móviles.
  - **BSC**: *Controlador de la Estación Base* que gestiona varios BTS y transmite llamadas de voz al MSC. Al incorporar GPRS, también contiene la *Unidad de Control de Paquetes (PCU)* para manejar el tráfico de datos a la red GPRS.
- **MSC**: *Centro de Conmutación Móvil*. Conmuta las llamadas de voz entre los terminales móviles y la red telefónica pública conmutada (*PSTN*). Maneja el establecimiento de llamadas y asignación de circuitos entre terminales móviles y la PSTN, o entre terminales móviles. El *VLR* es el *Registro de Ubicación de Visitantes*, a menudo ubicado junto al MSC como base de datos que almacena información temporal sobre los terminales móviles en su área.
- **HLR** *Registro de Ubicación de Inicio*. Esta base de datos contiene la información del perfil de los suscriptores móviles que incluye la lista de servicios suscritos. El HLR autentica los terminales móviles que desean acceder a la red móvil y también registra las ubicaciones de los terminales móviles en la red.

La figura 1.1 muestra la arquitectura de red de una PLMN utilizando GSM [27].

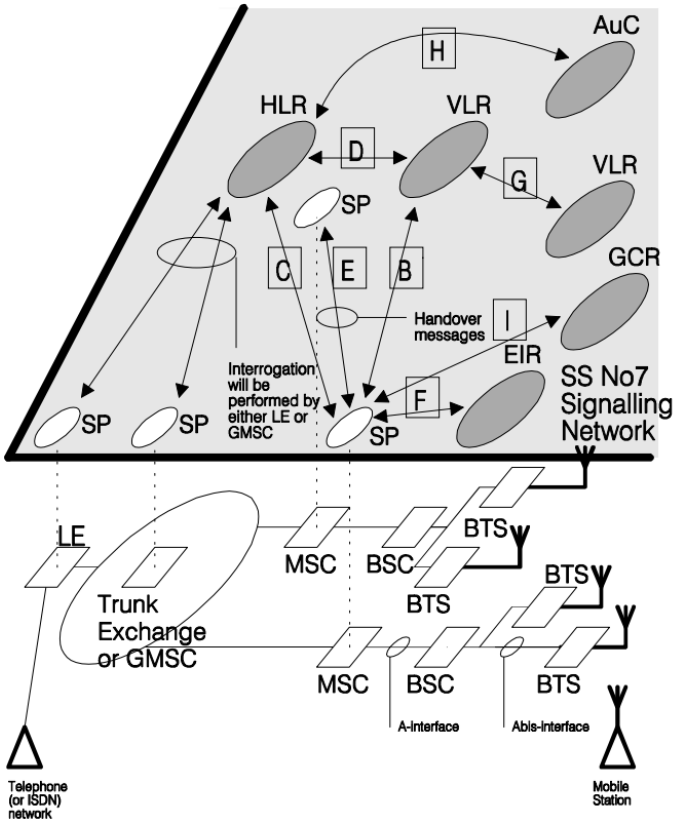

Figura 1.1: Arquitectura de red GSM en una PLMN

Para ofrecer un mejor soporte a servicios de datos, ETSI desarrolló entre 1997 y 1999 *GPRS (General Packet Radio Services)*, un sistema de *Conmutación de Paquetes o Packet Switching (PS)* que permite agregar varias portadoras, se superpone a GSM y es compatible con redes externas de paquetes de datos como Internet. Permite tarificación por cantidad de datos consumidos y alcanza velocidades de hasta 115

Kbps. GPRS es considerado un sistema de comunicación móvil de generación 2.5 o 2.5G y sus primeros despliegues comerciales se introdujeron a principios de los años 2000 [29].

En un sistema GPRS, a cada terminal móvil se le asigna una dirección IP. La asignación puede ser estática (según lo determine el operador celular), o dinámica (por conexión). Cuando el terminal móvil está encendido, siempre está conectado a GPRS. La red GPRS incluye dos nuevos nodos:

- **SGSN**: *Nodo de Soporte de GPRS de Servicio* es responsable de rastrear los terminales móviles GPRS en su área y de enrutar los paquetes de datos a los terminales móviles. Lleva un registro del BSC al que está asignado cada móvil de su área.
- **GGSN**: *Nodo de Soporte de GPRS de Puerta de Enlace* sirve como un enrutador que hace de interfaz entre Internet (u otra red de paquetes de datos) y la red GPRS basada en IP. El GGSN asigna direcciones IP a los terminales móviles cuando éstos se asignan dinámicamente y enruta los paquetes destinados al móvil al SGSN apropiado.

La figura 1.2 muestra la arquitectura de red de una PLMN utilizando GSM y GPRS [27].

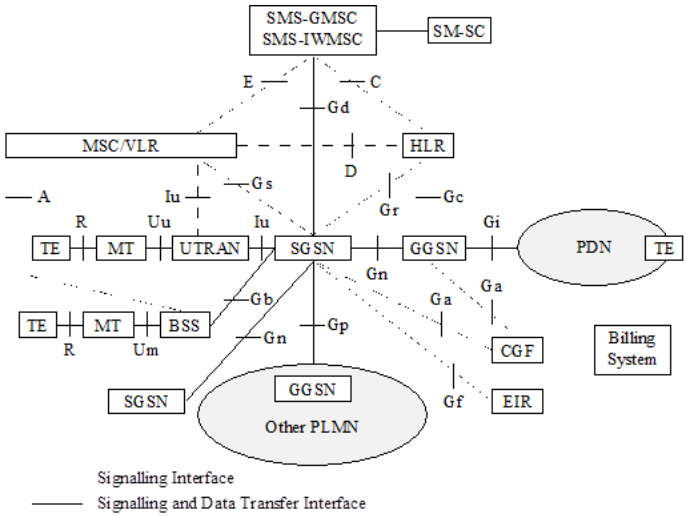

Figura 1.2: Arquitectura de red GSM-GPRS en una PLMN

EDGE (Enhanced Data Rates for GSM Evolution o tasas de Datos Mejoradas para la Evolución del GSM), es una evolución de GPRS que ofrece mayores velocidades de datos basándose en mejores esquemas de modulación y codificación. Utiliza 8PSK a diferencia de GMSK utilizado en GSM/GPRS y aumenta de 4 a 9 los diferentes esquemas de codificación utilizados con respecto a GPRS. EDGE puede llegar hasta 384 Kbps, que es comparable a las primeras implementaciones de W-CDMA para la tercera generación de UMTS. Es por esto que EDGE se considera como el puente entre la segunda y tercera generación de los sistemas de comunicaciones móviles, a veces denominado como 2.75G [30].

El sistema de acceso para redes GSM/EDGE fue posteriormente denominado *GERAN (GSM EDGE Radio Access Network)*.

1.2.3.2 NA-TDMA

El desarrollo de la tecnología celular digital radio en los Estados Unidos comenzó más tarde que en Europa por el hecho principal de que para la mayoría de los clientes AMPS, el roaming no era problema y pudo ser implementado con Canadá, México y el resto de sistemas analógicos AMPS como Hong-Kong

o Australia. En cualquier caso, a finales de los años 1980, la preocupación era la capacidad máxima que un sistema analógico como AMPS podría soportar.

La *TIA (Telecommunications Industry Association)* se crea en 1988 como *Organización de Desarrollo de Estándares (SDO, Standards Development Organization)* acreditada por el *Instituto Americano de Estándares nacionales (ANSI, American National Standards Institute)*, y asigna al subcomité TR-45 la tarea de definir un nuevo estándar celular digital. Se debatió sobre la conveniencia de utilizar tecnología TDMA o CDMA.

En 1990 se decidió por TDMA, o también *NA-TDMA (North American TDMA)*, y se creó el estándar *IS-54*, especificando la transición de AMPS a *D-AMPS* aumentando por tres la capacidad pero todavía permitiendo un modo dual de canales analógicos y digitales para compatibilidad con AMPS [31].

En 1994, con *IS-136*, una especificación completamente digital D-AMPS-TDMA estaba disponible, incoporando un TDMA mejorado y servicios adicionales como mensajes de texto. Una versión revisada en 1996 especificaba su operación en la banda norteamericana de 1900 MHz.

IS-54 fue el primer estándar 2G y el primero en utilizar TDMA superando las capacidades de la tecnología 1G avanzada *N-AMPS* creada por Motorola. Las operadoras estadounidenses y canadienses con redes TDMA (D-AMPS) fueron apagándolas en beneficio de GSM/UMTS para finales de la década de 2010.

1.2.3.3 cdmaOne

Un segundo estándar celular digital llamado *IS-95* o *cdmaOne* fue creado en 1993 en Estados Unidos utilizando tecnología CDMA e impulsado por *Qualcomm. TIA/EIA IS-95 (Telecommunications Industry Association/Electronic Industries Association Interim Standard - 95)* utiliza una tecnología de radio diferente denominada espectro ensanchado donde el espectro de radio se divide en portadoras que tienen aproximadamente 1.23 MHz de ancho. En cdmaOne, a cada canal de voz se le asigna un código único dentro del operador y la señal de voz se propaga a una velocidad de transmisión de aproximadamente 1.23 Mbps. Dado que todas las llamadas de usuario en una celda dada comparten la misma banda de canales, la única forma de distinguir entre las llamadas es a través del código único asignado a cada canal de voz. El código único se usa para difundir la señal original y luego para decodificar la señal en el extremo del receptor. La red cdmaOne implementa la reutilización de frecuencia universal donde se puede reutilizar la misma frecuencia en cada celda ya que lo que distingue a los canales de voz son los códigos únicos. Esto permite una mayor capacidad de red en comparación con los sistemas basados en TDMA.

La revisión *IS-95A* se publicó e implementó en 1995 y es la base de múltiples sistemas comerciales 2G CDMA alrededor del mundo. IS-95A describe la estructura de los canales CDMA de 1.25 MHz, el control de potencia, el procesamiento de llamada, el traspaso entre celdas y las técnicas de registro extremo a extremo. Ofrece conexiones de datos basadas en conmutación de circuitos de hasta 14.4 kbps.

La revisión *IS-95B* es un sistema 2.5G. Ofrece servicios de datos mediante conmutación de circuitos de 64 Kbps y se desplegó por primera vez en 1999 en Corea del Sur, extendiéndose posteriomente a otros países [32].

1.2.3.4 iDEN

Motorola desarrolló en 1994 un estándar digital no compatible llamado *iDEN (integrated Digital Enhanced Network)* ofreciendo voz y mensajes de texto. Utiliza tecnología TDMA y modulación digital M16-QAM como formato propietario de Motorola. Se desplegó en múltiples países además de en Estados Unidos [33].

1.2.3.5 PDC

A menor escala existió también el sistema 2G japonés *PDC (Personal Digital Cellular)* que comenzó a ser definido en 1990 y fue implementado en 1992. Utiliza TDMA y opera en las bandas de 800 MHz (*PDC-800*) y 1,5GHz (*PDC-1500*).

Para la transmisión de datos se introdujo el PDC-P (PDC Mobile Packet Data Communication System) basado en conmutación de paquetes y ofreciendo una tasa de transferencia de 28,8 kbit/s [34].

1.2.3.6 SMS

En 1993 se introdujo el servicio de los mensajes de texto *SMS* sobre la red GSM.

La mensajería SMS es asíncrona, sin correlación entre la respuesta de un suscriptor móvil y un mensaje recibido previamente. Los mensajes SMS se envían y reciben desde un *Centro de servicio de SMS (SMSC)* a través del Protocolo de transporte de mensajes cortos (SMTP). El SMSC está conectado a la red del operador celular a través de un SMS Gateway/MSC interfuncional utilizando el sistema de señalización de ITU número 7 (*SS7*) para transportar el mensaje SMS al MSC.

1.2.4 3G

Hasta finales de la década de 1990, los servicios de voz habían sido considerados como la fuente fundamental de tráfico, pero era obvio que un ancho de banda mayor sería necesario para los servicios multimedia emergentes; como Internet, e-mail o vídeo.

En 1996, un grupo de estudio de la *ITU (Unión Internacional de Telecomunicaciones)*, comenzó a considerar las especificaciones para crear un conjunto de estándares celulares de «tercera generación» (3G), conocidos colectivamente como *IMT-2000 (International Mobile Telecommunicacion -2000)*. El objetivo claro era evitar la fragmentación ocurrida en la primera y segunda generación (principalmente entre GSM y cdmaOne), para facilitar, por ejemplo, la itinerancia móvil entre los distintos sistemas. También era necesario ofrecer mayor capacidad, velocidades de datos más rápidas de hasta 2 Mbps y mejor calidad de servicio (QoS). El desarrollo de este estándar se produjo al mismo tiempo que Internet crecía enormemente en popularidad [35].

El esfuerzo de IMT-2000 no pudo llegar a un acuerdo sobre un único estándar común y finalmente consistió en una familia de estándares para manejar la evolución de GSM y cdmaOne. Las especificaciones de acceso radio soportan diferente variantes como FDD (Frequency Division Duplex) o TDD (Time Division Duplex) y múltiples bandas de frecuencias.

1.2.4.1 3GPP. UMTS W-CDMA. HSPA

El sistema *Universal Mobile Telecommunications System (UMTS)*, o *Sistema Universal de Comunicaciones Móviles* se basa en *CDMA de banda ancha (W-CDMA, Wideband-CDMA)*. Fue desarrollado por los creadores originales de GSM y gestionado por el *Proyecto de Asociación de Tercera Generación (3GPP, 3rd Generation Partnership Project)* establecido en 1998. Los objetivos declarados por 3GPP son desarrollar un sistema móvil 3G basado en redes centrales GSM evolucionadas y las tecnologías de acceso de radio que soportan. Además, 3GPP también gestiona la evolución y desarrollo de GSM y EDGE.

UMTS emplea un ancho de portadora de canal de 5 MHz para brindar estas velocidades de datos más altas y una mayor capacidad. Ofrece una portadora mucho más amplia en comparación con la portadora

de 1,23 MHz de ancho de redes 2G. Además, a diferencia de las generaciones anteriores, en UMTS se incluye el aseguramiento de la calidad de servicio (QoS).

El sistema W-CDMA utiliza tecnología CDMA de *secuencia directa de banda ancha (DS-CDMA, Direct Spread CDMA)* en un ancho de banda de 5 MHz para admitir los requisitos de velocidad de datos IMT-2000 de cobertura de área amplia de 384 kbps y cobertura local de 2 Mbps.

En la parte de acceso de la red UMTS, llamada *UTRAN (UMTS Terrestrial Radio Access Network)*, el BSS pasa a denominarse *sistema de red de radio (RNS)*, y el BTS se denomina *Nodo B*. La funcionalidad BSC se reemplaza en UMTS por el *controlador de red de radio (RNC)*. La red central incluye los mismos elementos de la red 2.5G (GPRS) y admite servicios de circuitos (voz) a través de MSC y servicios de paquetes (datos) a través de SGSN y GGSN.

La Figura 1.3 muestra los elementos de radio UMTS y sus interfaces a la red central [36].

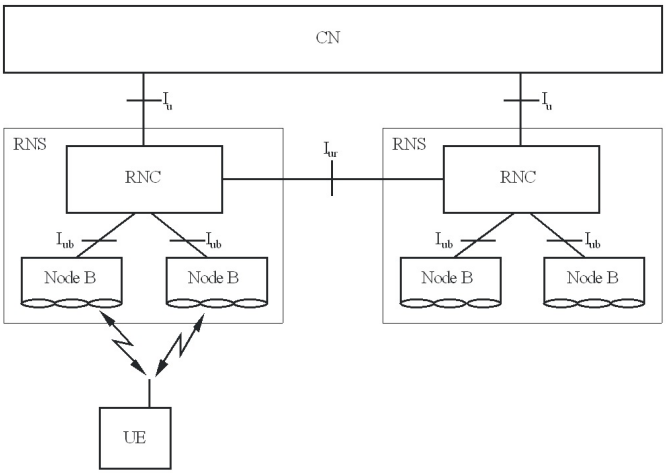

Figura 1.3: Red de acceso UTRAN en la red UMTS

Con el objetivo de mejorar las velocidades de datos, se incorporaron diversas mejoras sobre la red 3G basadas en la optimización de la tecnología espectral. La tecnología *HSDPA (High Speed Downlink Packet Access)* fue introducida en la Release 5 de 3GPP y conseguía tasas de bajada de hasta 14 Mbps. *HSUPA (High Speed Uplink Packet Access)* fue introducido en la Release 6 de 3GPP en 2007 y mejoraba la tasa de datos de subida llegando hasta 7,2 Mbps. Se nombra *HSPA (High Speed Packet Access)* a la fusión de los dos protocolos móviles y finalmente en 2008 se lanzó el estándar 3GPP *HSPA+ (High Speed Packet Access +)*, desplegado a partir de 2010 con velocidades pico de 337 Mbps en el enlace descendente y 34 Mbps en el enlace ascendente [37]. Las variantes de HSPA son consideradas en muchos casos como tecnología 3.5G o 3.75G.

1.2.4.2 3GPP2. CDMA2000

La evolución del estándar 2G *cdmaOne*, denominado *cdma2000*, está a cargo de otro organismo de estándares: *3GPP2 (3rd Generation Partnership Project 2)*, establecido en 1999. 3GPP2 es un proyecto colaborativo de especificaciones de telecomunicaciones 3G que incluye intereses norteamericanos y asiáticos que desarrollan especificaciones globales para redes celulares ANSI/TIA/EIA-41. Otra organización que participa activamente en la progresión de las redes CDMA es *CDMA Development Group (CDG)*, un consorcio internacional de empresas unidas para liderar la adopción y evolución de los sistemas inalámbricos CDMA en todo el mundo.

El concepto CDMA2000 engloba la familia de estándares que representan la sucesiva evolución en fases de las capacidades tecnologícas de CDMA2000 sobre su antecesor cdmaOne. También es denominado *CDMA Multi-Carrier*, a diferencia de la implementación UMTS que utiliza *CDMA Direct Spread*.

El sistema inicial *CDMA2000 1X (IS-2000) o 1xRTT*, fue introducido en 1999 y desplegado en 2000. Ofrece voz por conmutación de circuitos y servicios de datos de alta velocidad de hasta 153 Kbps. Este sistema para la parte de voz (CC) sería posteriormente mejorado con *1X-Advanced* mejorando eficiencia de voz y optimizado para soluciones *M2M (Machine to Machine Communications)*.

La mejora para la parte de datos, denominada *CDMA2000 EV-DO (Evolution-Data Optimized) (IS-856)* o *1xEV-DO Release 0*; fue lanzada comercialmente en 2002 e introduce nuevas técnicas de conmutación de paquetes asignando una portadora inalámbrica de 1,25 MHz separada para datos. De este modo, las tasas de datos de bajada pueden alcanzar hasta 2.4 Mbps con 153 Kbps de subida. Mejoras posteriores realizadas sobre EV-DO aportan un progresivo aumento en las velocidades de datos y disminución en los tiempos de latencia. *1xEV-DO Revision A (IS-856A)* aumenta la tasa de bajada hasta los 3.1 Mbps y la de subida hasta 1.8 Mbps y fue lanzada en 2006. *EV-DO Revisión B (IS-856B)* incluye capacidad multi-portadora y ofrece una tasa de datos en canal descendente de 14.7 Mbps, y de 5.4 Mbps en canal ascendente. Se desplegaron en 2010.

Las versiones EV-DO incluyen diferentes mejoras tecnológicas como la modulación *OFDM (Orthogonal Frequency Division Multiplexing)* o nuevas técnicas de antena como *MIMO (Multiple Inputs Multiple Outputs)* y son consideradas de tipo 3.5G.

El núcleo de la red CDMA2000 y las interfaces definidas también mejora la red cdmaOne e incluye una red de datos separada, similar a la red GPRS en GSM. Esta red de datos, la *Red Central de Paquetes (PCN)*, incluye un nuevo elemento de red, el *Nodo de Servicio de Datos en Paquetes (PDSN)*, un BSC mejorado que enruta las llamadas de voz al MSC y los paquetes de datos al PDSN. Otra novedad es el *servidor de autenticación, autorización y contabilidad (AAA)* utilizado en el establecimiento de sesiones IP [32].

Por último, Qualcomm intentó trabajar en una evolución de EV-DO hacia la cuarta generación, que competiría con LTE y WiMAX. Incialmente fue llamado *EV-DV (Evolution Data and Voice)* y finalmente *EV-DO Revision C*. Fue definido por 3GPP2 y TIA (TIA-1121) en 2007 y 2008 respectivamente. El nombre *UMB (Ultra Mobile Broadband)* fue introducido en 2006 como sinónimo para este estándar. En Noviembre de 2008 Qualcomm anunció que terminaba el desarrollo de este tecnología, en favor de LTE [38].

1.2.4.3 Otros

En Japón, la implementación de NTT de WCDMA se llamó *FOMA (Freedom of Mobile Multimedia Access)* siendo el primer servicio 3G. Por su parte, China utilizó *TD-SCDMA (Time Division Synchronous CDMA)*

1.2.5 4G

Las crecientes exigencias impuestas a los teléfonos móviles para manejar incluso más datos de los que podría manejar 3G llevaron al desarrollo de la tecnología 4G. En 2008, la ITU presentó una lista de requisitos para lo que denominó *IMT-Advanced*, o 4G. Estos requisitos incluían tasas de datos de 1 Gbps para un usuario estacionario y 100 Mbps para un usuario en movimiento, así como suportar múltiples redes de acceso. Las tecnologías 4G utilizan alto ancho de banda, baja latencia y una red IP subyacente con

servicios de alto nivel (como la voz) sobre ella. El despliegue masivo de redes 4G permite la utilización de aplicaciones no posibles con las velocidades de 3G, como puede ser el streaming de video de alta definición o los juegos móviles.

En 2010, la ITU decidió que dos tecnologías; *LTE-Advanced (Long Term Evolution Advanced)* y *WirelessMan-Advanced* (también llamada *WiMAX 2*), cumplían con los requisitos. Las tecnologías *Long Term Evolution (LTE)* y *WiMAX* no cumplían todos los criterios establecidos en IMT-Advanced, aunque en este mismo comunicado, ITU consideró que el termino 4G podría ser también aplicado a ellas [39].

1.2.5.1 LTE, EPC

LTE es un estándar ETSI 3GPP de comunicaciones móviles de banda ancha como evolución para sistemas GSM/UMTS o CDMA2000. Estuvo disponible en 2008 y la compañía telefónica sueca TeliaSonera introdujo la primera red 4G LTE en Estocolmo en Diciembre de 2009 (aunque los primeros dispositivos móviles LTE y despliegues comerciales aparecieron durante 2010 y 2011). Las operadoras con despliegues 3G de tipo CDMA2000 planearon inicialmente la posibilidad de utilizar Ultra Mobile Broadband (UMB) o WiMAX, pero finalmente en la mayoría de los casos escogieron LTE. En España, LTE fue lanzado comercialmente en 2013.

La red de acceso es denominada *Evolved Universal Terrestrial Radio Access Network (E-UTRAN)*, incompatible con W-CDMA de 3G y utiliza *Orthogonal frequency-division multiple access (OFDMA)* para el enlace de bajada y *Single-carrier FDMA (SC-FDMA)* para el enlace de subida. LTE ofrece tasas pico de 300 Mbps en bajada y 75 Mbps en subida soportando tanto *Frequency Division Duplex (FDD)* como *Time Division Duplex (TDD)* y latencia de 5 milisegundos. La Red de Acceso Radio de LTE está formada por un único nodo, el *eNodo-B* [40].

El núcleo de red LTE está basado en IP y es descrito como *SAE (System Architecture Evolution)*. SAE está formado por *EPC (Evolved Packet Core)*, como una red pura de conmutación de paquetes, sin incluir conmutación de circuitos usada para voz y sms en las tecnologías precedentes. EPC consta de los siguientes elementos de red clave:

- **MME**: *Mobility Management Entity* es el nodo de control clave para la red de acceso. Termina la señalización cifrada *NAS (Non Access Stratum)* gestionando las llaves de seguridad. Es responsable del paging, de la activación/desactivación de portadoras y de escoger Serving Gateway durante el registro inicial del dispositivo *UE (User Equipment)* o durante el traspaso intra-LTE. También se encarga de la autenticación del usuario (interactuando con HSS), entre otras funciones.

- **HSS**: *Home Subscriber Server* es la base de datos central con toda la información relativa a usuarios y suscripciones. Se encarga de la gestión de movilidad, soporte en establecimiento de llamadas y sesiones, autenticación y autorización de usuarios.

- **SGW**: *Serving Gateway* enruta paquetes de datos de usuario actuando como anclaje de movilidad para el plano de usuario durante los traspasos inter-eNodeB o la movilidad entre LTE y las tecnologías 2G/3G.

- **PGW**: *Packet Data Network Gateway* ofrece conectividad entre el UE y las redes de paquetes de datos *PDNs* externas como punto de salida y entrada. También se utiliza como punto de anclaje de movilidad con tecnologías no-3GPP como *WiMAX* o *3GPP2* (CDMA 1X y EV-DO).

- **ePGW**: *Evolved Packet Data Network Gateway* requerido para el acceso no confiable de tipo no-3GPP [41].

- **PCRF**: *Policy and Charging Rules Function* se encarga de la gestión y aplicación de las políticas de control y provee las interfaces con los sistemas de tarificación.

La Figura 1.4 muestra los elementos de radio E-UTRAN y sus interfaces a la red EPC [42].

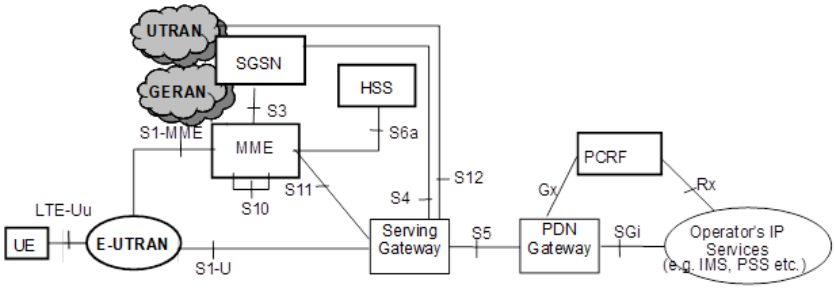

Figura 1.4: Arquitectura no-roaming de accessos 3GPP a LTE EPC

Para el soporte de servicios de voz en EPC se especifican 2 opciones por 3GPP:

- **VoLTE**: Voz sobre LTE, utilizando Voz sobre IP en la red IMS. Además, 3GPP define *SRVCC (Single Radio Voice Call Continuity)* para el traspaso de llamadas VoLTE a redes 2G/3G. Las llamadas de conmutación de paquetes de LTE utilizando el sistema IMS pueden ser traspasadas a las redes de conmutación de circuitos GSM/UMTS.
- **CSFB**: (Circuit-Switched Fallback) en la que el UE cambia su tecnología de acceso radio de LTE a 2G/3G cuando envía o recibe llamadas, y requiere una nueva interfaz (SGs) ente MME y MSC.

Para el soporte de SMS en EPC, 3GPP especifica 2 opciones:

- **SMS sobre IP**: utilizando la red IMS.
- **SMS sobre SGs**: utilizando la interfaz introducida entre MME y MSC para soportar CSFB.

Tanto para voz como para SMS, 3GPP considera las soluciones sobre la red IMS como las soluciones a largo plazo. Sin embargo, 3GPP definió estas alternativas sobre la red de conmutación de circuitos consciente de que los despliegues LTE e IMS requerirían cierto tiempo de adopción.

1.2.5.1.1 IMS *IP Multimedia Subsystem* se desarrolló por primera vez como parte de 3GPP Release 5 en 2002. Se basa en protocolos *IETF (Internet Engineering Task Force)* ampliamente utilizados como *SIP (Session Initiation Protocol)* y *SDP (Session Description Protocol)*. Estas tecnologías han sido adoptadas por la industria como mecanismo de señalización para aplicaciones multimedia.

LTE se apoya en la red IMS para ofrecer los servicio de voz y SMS, al no incluir una red de conmutación como en las tecnologías precedentes. En Release 7 se optimizó IMS y los protocolos de soporte para garantizar que la voz y otros medios fueran compatibles con la misma eficiencia que en las redes de conmutación de circuitos [43].

En la realidad, muchas operadoras continúan utilizando LTE solo para datos, utilizando CSFB y SMS over SGs para voz y SMS. Las dificultades para implementar IMS no se deben a los protocolos ni a las especificaciones. La problemática no esta solo relacionada con los aspectos técnicos, sino también con el completo cambio de paradigma de la industria de los servicios CS a un entorno verdaderamente basado en IP. Esto aplica a aspectos como la migración del servicio, las políticas, la interoperabilidad y el plan de implementación. Sin embargo, estas complejidades deben abordarse para poder realmente proporcionar un entorno de servicio avanzado. Esto afecta no solo a 4G LTE, sino también a 5G que tampoco incluye conmutación de circuitos y debe ofrecer servicios de voz sobre la red IMS.

1.2.5.2 LTE Advanced

LTE Advanced (LTE+ o LTA-A) fue estandarizado en 2011 por 3GPP como mejora de LTE, y desplegado a partir de 2014 [44]. Tiene como objetivo cumplir (e incluso superar) el requisito inicial de 4G de llegar a velocidades de 1Gbps, para lo que utiliza 3 nuevas tecnologías de modulación combinadas:

- **Carrier Aggregation**: o Agregación de portadora en la que se combinan múltiples portadoras LTE para un mayor ancho de banda.
- **MIMO 4x4**: *Multiple Inputs Multiple Outputs* que utiliza más antenas para una mejor eficiencia espectral.
- **256QAM**: *256 Quadrature Amplitude Modulation* que transmite más bits por símbolo (8) mejorando también la eficiencia espectral.

1.2.5.3 LTE Advanced Pro

LTE Advanced Pro (LTE-A Pro, 4.5G o Gigabit LTE) es introducido por 3GPP en las releases 13 y 14 en 2015 [45] como evolución de LTE-A describiendo tasas de datos por encima de 3Gbps como camino intermedio hacia la generación 5G. Utiliza mejoras en la tecnología de CA (Carrier Aggregation), permitiendo utilizar 32 portadoras (5 en el caso de LTE-A). Esto se consigue introduciendo el concepto de *LAA (Licensed Assisted Access)* por el que se utiliza espectro no licenciado disponible en la banda de 5 GHz actualmente utilizado por las redes WiFi. LAA permite el uso simultáneo del espectro licenciado de LTE y el no licenciado para un uso más eficiente del espectro total disponible. *eLAA (enhanced LAA)* se introduce en Release 14 para ofrecer la misma funcionalidad en el canal de subida. LTE-A Pro también explora las capacidades del uso de celdas pequeñas (Small cells) que son ideales para el uso en el espectro no licenciado, así como la agregación de macro celdas y celdas pequeñas con conectividad dual para unas mayores tasas de datos por usuario. LTE-A Pro también utiliza *FD-MIMO (Full Dimension MIMO)* a la par que la tecnología de antenas evoluciona hacia *Massive MIMO*, un habilitador clave para 5G.

NB-IoT (Narrowband Internet of Things) y *LTE-MTC (LTE Machine Type Communication)* o LTE-Cat-M1 que incluye *eMTC (enhanced Machine Type Communication)* son también estándares de acceso radio desarrollados por 3GPP en la Release 13 (LTE Advanced Pro) en 2016 con aplicación específica en los servicios *IoT (Internet of Things)*. Por prestaciones, pueden ser consideradas como tecnologías 5G.

Los despliegues de LTE-A Pro comenzaron en 2018 [46], aunque en muchos casos las operadoras empezaron a centrarse en 5G-NSA con el objetivo de declarar comercialmente el lanzamiento de 5G. En ambos casos, las mejoras se centran en la tecnología de antenas.

1.2.5.4 WiMAX

WiMAX (Worldwide Interoperability for Microwave Access), representa una familia de protocolos de comunicaciones inalámbricas de banda ancha basados en el conjunto de estándares IEE 802.16 [47].

El *WiMAX Forum* se formó en 2001 y describe WiMAX como una tecnología que permite ofrecer acceso de banda ancha inalámbrico de última milla como alternativa al cable y DSL. Este WiMAX original introducido en 2001 ahora es conocido como *WiMAX Fijo* (Fixed WiMAX) y fue inicialmente diseñado para ofrecer velocidades de 30 a 40 Mbps, mejorando en 2011 hasta 1 Gbps para estaciones fijas. Fue pionero en el uso de tecnología MIMO.

El *WiMAX Móvil* (Mobile WiMAX) basado en 802.16e-2005 fue introducido en 2005 y es la revisión que fue desplegada en diferentes países permitiendo traspaso entre estaciones base. Pretendía cubrir las

necesidades de las nuevas redes 4G y estuvo disponible antes que LTE. La versión evolucionada fue el estándar IEEE 802.16m-2011, que fue la base de *WirelessMan-Advanced o WiMAX 2*, propuesto con anterioridad al *ITU* como solución a la estandarización *IMT-Advanced* 4G. Tenía planes de despliegue para 2011-2012.

Sin embargo mientras que LTE era una evolución de los sistemas existentes 2G/3G; WiMAX (o WirelessMan-Advanced) se trataba de una tecnología completamente nueva. Finalmente LTE se impuso como el estándar 4G puesto que las grandes operadoras pudieron reutilizar y extender sus inversiones en conocimiento, equipamiento y espectro de 3G a LTE.

Finalmente, la versión 2.1, conocida como *WiMAX 2+*, fue lanzada a principios de 2010 y rompía la compatibilidad con versiones WiMAX anteriores. Varias operadoras migraron a este nuevo estándar que es compatible con TD-LTE.

1.3 Conclusiones

Partiendo del siglo XIX con la invención del teléfono y las radiocomunicaciones, y llegando hasta los despliegues actuales de redes 4G; en este capítulo hemos realizado un análisis cronológico de la evolución y acontecimientos relativos al sector tecnológico de las comunicaciones de telefonía móvil.

No es posible analizar correctamente la tecnología 5G sin comprender el largo camino de avances tecnológicos experimentados por las diferentes generaciones tecnológicas precedentes. El estudio realizado en este capítulo aporta esta base de conocimiento y un mejor entendimiento de las motivaciones detrás de cada salto tecnológico.

Es importante resaltar la corta duración de las distintas generaciones. Cada una de las cuatro generaciones anteriores a 5G han tenido una duración aproximada de 10 años, antes de la necesaria llegada de la siguiente generación para cubrir sus carencias y límites tecnológicos intrínsecos. Aproximadamente, se puede considerar a 1G como la tecnología móvil de la década de los 1980, 2G para la década de los 1990, 3G en la de los 2000 y 4G en la de los 2010. Del mismo modo, es también alrededor de 2020 cuando 5G empieza a ser desplegado. Sin embargo, dentro de la duración de cada generación también se producen múltiples cambios y mejoras intermedias que dan lugar a diferentes terminologías no estandarizadas del tipo «2.5G», «3.5G» o «4.5G».

Todo esto se debe a la rápida demanda de nuevos servicios y capacidades, y probablemente ocurra del mismo modo con la tecnología 5G. A modo de curiosidad, la tecnología 4G fue denominada LTE (*Long Term Evolution*) o Evolución de Largo Plazo al ser diseñada con esta intención; pero la realidad acabó demostrando que la renovación tecnológica (con la llegada de 5G), ha sido requerida en un periodo de tiempo similar al de las generaciones anteriores.

Capítulo 2

Objetivos

La única manera de descubrir los límites de lo posible es aventurarse un poco más allá, hacia lo imposible.

The only way of discovering the limits of the possible is to venture a little way past them into the impossible.

Arthur C. Clarke [1]

El objetivo principal de este TFG es aportar una visión completa sobre la gestión de la señalización en el núcleo de redes móviles 5G Stand-Alone (5G SA).

Tras haber realizado un estudio general de la historia de las comunicaciones móviles desde los primeros pasos de la telefonía y las comunicaciones inalámbricas hasta la actualidad, comenzaremos ahora analizando las diferencias entre las redes 5G NSA (*Non Stand Alone*) y las redes 5G SA (*Stand Alone*) según han sido definidas y estandarizadas por 3GPP. Analizaremos las diferentes áreas del ecosistema 5G a nivel de dispositivos móviles, red de acceso radio y núcleo de red; para obtener una visión amplia y generalizada de la nueva tecnología 5G.

Continuaremos adentrándonos en el núcleo de red 5G SA con el análisis tanto de su arquitectura basada en servicios (*SBA: Service Based Arquitecture*), como de las diferentes funciones de red (*NF: Network Function*) que lo conforman. Analizaremos de manera específica las implicaciones de los escenarios de roaming o itinerancia en esta arquitectura. Ofreceremos una visión amplia y generalizada del núcleo de red 5G SA.

Finalmente, profundizaremos en todos los aspectos relativos a la gestión de la señalización a nivel de enrutamiento y seguridad en el núcleo de red 5G SA. Analizaremos las nuevas funcionalidades que el núcleo de red 5G SA ofrece, y el impacto específico de ellas para los escenarios de roaming. Desarrollaremos algunos de los escenarios más relevantes de la señalización extremo a extremo, y analizaremos la adopción de 5G en la actualidad por parte de los diferentes países y operadores móviles.

[1] Arthur C. Clarke (1917 - 2008), escritor y científico británico. Segunda ley sobre el avance científico en *Profiles of the Future: An Inquiry Into the Limits of the Possible (1973)* [48].

Capítulo 3

Estudio teórico

Toda idea revolucionaria parece provocar tres etapas de reacción. Pueden resumirse en las expresiones:
(1) Es completamente imposible;
(2) Es posible, pero no vale la pena hacerlo;
(3) ¡Dije que era una buena idea en todo momento!

Every revolutionary idea seems to evoke three stages of reaction. They may be summed up by the phrases:
(1) It's completely impossible;
(2) It's possible, but it's not worth doing;
(3) I said it was a good idea all along!

Arthur C. Clarke [1]

3.1 Introducción a 5G

La quinta generación de redes móviles celulares *5G* promete habilitar nuevos casos de uso disruptivos, no posibles con las generaciones anteriores. Los saltos generacionales anteriores estuvieron basados principalmente en el aumento de las velocidades de datos. Éste es también un aspecto fundamental del salto a 5G, con velocidades de datos de hasta 20 Gbps (utilizando la banda de ondas milimétricas de 26 GHz.); pero además de las mejoras y nuevas tecnologías de acceso radio (*5G New Radio*), 5G también incorpora un núcleo de red completamente nuevo (*5G Core*), que permite nuevas capacidades como baja latencia o segmentación de red (*network slicing*), necesarias para habilitar muchos de los nuevos casos de uso prometidos en 5G. 5G representa un cambio de generación más ambicioso que el introducido por 3G o 4G.

La ITU-R (*ITU Radiocommunication Sector*) definió el conjunto de requisitos *IMT-2020* en 2015 para las redes, dispositivos y servicios 5G [50]. Incluye tres áreas de aplicación principales que requieren las capacidades nuevas o mejoradas de 5G:

- ***Enhanced Mobile Broadband (eMBB)***: Hace referencia al caso de uso 5G que pone como objetivo el ofrecer picos de velocidades de descarga de 20 Gbps, considerablemente mayor que el objetivo de velocidad de descarga pico de 4G LTE Advanced Pro que es de 3 Gbps. Siendo estas

[1] Arthur C. Clarke (1917 - 2008), escritor y científico británico. *The Promise of Space (1968)* [49].

velocidades teóricas en casos ideales e irrealizables, las velocidades reales deberían estar cerca de los 10 Gbps en el caso de implementar todas las características de 5G. Es por tanto el caso de uso más obvio.

- ***Ultra Reliable Low Latency Communications (URLLC)***: Se trata del caso de uso 5G que requiere a la tecnología ofrecer una fiabilidad de conectividad del 99.99 % y tiempos de latencia extremadamente bajos (por debajo de 1 milisegundo). En estos casos, las capacidades de eMBB no son necesariamente requeridas, y el foco está en la fiabilidad y rapidez de conexión a bajas tasas de datos. Las latencias de 1 milisegundo e inferiores, pueden dar soporte a casos de uso como conducción autónoma con vehículos conectados, automatización industrial y muchos otros.
- ***Massive Machine Type Communications (mMTC)***: En este tipo de caso de uso 5G, se definen requisitos para poder conectar un millón de dispositivos por kilómetro cuadrado. Dispositivos de baja potencia, bajo coste, baja complejidad y con una vida útil de batería de hasta diez años. Esta categoría da soporte a casos de uso como automatización del hogar incluyendo sensores y actuadores, o sistemas de monitorización de máquinas. Aquí la inmediatez de la comunicación o las altas tasas de datos no son requisitos indispensables.

eMBB, URLLC y mMTC son las tres categorías fundamentales de casos de uso que las redes 5G habilitan. Representan las diferentes capacidades de comunicación requeridas para hacer realidad los casos de uso que demandan alta velocidad de datos, comunicación altamente fiable con baja latencia, y despliegues masivos de dispositivos conectados.

3GPP trabajó en los estándares y declaró la primera especificación 5G-NR completada en 2017, tras aceptar la propuesta de acelerar la implentación NSA (*non-stand-alone*) de 5G NR (*5G New Radio*), permitiendo que las primeras pruebas de concepto y lanzamientos se produjeran en 2019 en lugar de en 2020. Al igual que 4G LTE, 5G NR utiliza técnicas de modulación OFDM y varias mejoras tecnológicas ya introducidas en LTE-Advanced o LTE-Advanced Pro como *Massive MIMO*, *Carrier-Aggregation* o *Beamforming* (Conformación de haces).

5G NR está separada en dos rangos de frecuencias. El rango 1 incluye las bandas sub-6 GHz con algunas de ellas tradicionalmente usadas por las generaciones anteriores; y el rango 2 que incluye bandas desde 24.25 GHz a 71 GHz. Estas bandas de ondas milimétricas (*mmwave*) tienen menor alcance pero mayor ancho de banda disponible que las bandas del rango 1. Dentro de ambos rangos, 5GNR define múltiples bandas de frecuencia disponibles [51].

La especificación de 5G NR para 5G SA, así como la especificación relativa al núcleo de la red 5GC (*5G Core*), mantuvo su plan de disponibilidad inalterado, y fue introducido con la Release 15 de 3GPP en Junio de 2018. La segunda fase del desarrollo de 5G por parte de 3GPP se produce en la Release 16 completada en Junio de 2020, y concluye con Release 17 completada en Junio de 2022. El nuevo núcleo de red pretende ofrecer el requisito de muy baja latencia requerido en varios de los nuevos casos de uso como ciudades inteligentes, conducción autónoma, etc...

Los primeros despliegues de 5G NR para ser utilizados en 5G-NSA se produjeron en 2019, mientras que los primeros despliegues de 5G NR SA y 5G Core para una experiencia 5G SA extremo a extremo se producen en 2021 principalmente en Estados Unidos y Asia; como veremos en más detalle en la sección 6.1.

A fecha de realización de este trabajo, en Noviembre de 2022, existen múltiples despliegues de tecnología 5G en la parte radio (5G NR), que utilizan el modelo temporal híbrido 5G NSA (*Non Stand Alone*), en el que el núcleo de la red sigue siendo 4G LTE. Son pocos los países (principalmente China y Estados Unidos) que tienen despliegues 5G SA (*Stand Alone*) en producción en la actualidad, y en cualquier caso

se encuentran en fases preliminares, todavía con capacidades limitadas. Son mayoría los países en los que 5G SA todavía no ha sido desplegado comercialmente, incluyendo a España.

Será la renovación del núcleo de la red a 5G Core quien traga consigo todo el potencial real de 5G.

3.2 Especificaciones 3GPP

El desarrollo de las especificaciones 5G se produce en tres releases 3GPP: 15, 16 y 17. Como hemos comentado, la release inicial 15, comenzó detallando los requisitos de 5G NSA opción 3, y continuó con la versión inicial del núcleo de red 5GC. La release 16 añadió madurez a la definición del núcleo de red 5G, dejando la release 17 para mejoras adicionales y nuevos ámbitos de aplicación como por ejemplo 5G por satélite.

En cada release 3GPP, múltiples documentos son producidos bajo la nomenclatura de TS (*Technical Specification*) o TR (*Technical Report*). Los documentos tienen un número identificativo de cinco dígitos, en el que los dos primeros identifican una serie de documentos con temática específica. De este modo, por ejemplo, la serie 29 (29xxx) se utiliza para protocolos de señalización, la serie 33 (33xxx) se utiliza para aspectos de seguridad y la serie 23 (23xxx) trata sobre la realización técnica.

3GPP también dispone de múltiples grupos de estandarización TSG (*Technical Specification Group*), de manera que cada grupo es responsable del desarrollo de un subconjunto de especificaciones. Existen tres grupos principales:

- **TSG TAN** (*Technical Specification Group Radio Access Network*): Responsable de las especificaciones radio.
- **TSG SA** (*Technical Specification Group System Aspects*): Se encarga de los aspectos relativos a servicios y sistemas, y de la arquitectura en general de los sistemas basados en especificaciones 3GPP. Por tanto, es también responsable de la coordinación entre los distintos grupos de especificación.
- **TSG CT** (*Technical Specification Group Core Network and Terminals*): Se encarga tanto de la especificación de las capacidades e interfaces de los terminales a nivel físico y lógico; como del núcleo de red de los sistemas 3GPP.

Cada uno de estos grupos está dividido en varios grupos de trabajo WG (*Working Group*), con subtareas más específicas. De este modo, el CT-WG4 trata específicamente los protocolos del núcleo de red, o el SA-WG3 trata específicamente los temas de seguridad y privacidad.

Las especificaciones 3GPP están disponibles de manera gratuita cuatro veces durante cada año, alineadas con las reuniones trimestrales de los diferentes *Technical Specification Group (TSG)*. 3GPP utiliza un sistema de releases paralelas, de manera que los desarrolladores puedan trabajar con una versión estable para la implementación de funcionalidades, a la vez que permite añadir nuevas capacidades en releases subsiguientes. Van evolucionando incorporando peticiones de cambios CR (*Change Requests*) una vez son aprobados. Estas CRs pueden ser propuestas por cualquier miembro de la organización 3GPP.

Dentro de una release concreta, 3GPP utiliza el concepto de *stage* o etapa, método reutilizado de ITU-T:

- **Stage 1**: Hace referencia a la descripción del servicio desde el punto de vista de un usuario del servicio.

- **Stage 2**: Se corresponde con el análisis lógico, obteniendo una arquitectura abstracta de los elementos funcionales y los flujos de información entre ellos a través de los puntos de referencia entre entidades funcionales.

- **Stage 3**: Es la implementación concreta de la funcionalidad y de los protocolos que aparecen en las interfaces físicas entre elementos físicos sobre los que los elementos funcionales han sido mapeados.

En algunos casos, 3GPP también realiza estudios de viabilidad, que quedan recogidos en *Technical Reports* (TRs), y pueden ser considerados como un *Stage 0*. Además, algunas de las especificaciones de *Stage 3* requieren especificaciones posteriores para pruebas, que por tanto se definen como *Stage 4*.

3.2.1 3GPP Release 15

Se trata de la primera release que trata 5G. Su desarrollo comienza en 2017 y se completa en Junio de 2019.

Después de la entrega inicial a finales de 2017 de las nuevas especificaciones de radio NR Non-Stand-Alone (NSA) para 5G, durante 2018 3GPP trabajo en el primer conjunto completo de estándares 5G de la versión 15 de 3GPP.

Si bien las especificaciones iniciales permitieron sistemas de radio 5G NSA integrados en redes LTE de generaciones anteriores, el alcance de la versión 15 se amplió para cubrir 5G SA, con un nuevo sistema de radio complementado con una red central de próxima generación.

También incluye mejoras en LTE e, implícitamente, en EPC (*Evolved Packet Core*). Este punto de referencia crucial permite a los proveedores avanzar rápidamente con el diseño de chips y la implementación inicial de la red durante 2019.

Además, en release 15 se empezaron a realizar los primeros estudios que abarcaban temas tan diversos como el servicio de prioridad multimedia, los servicios de capa de aplicación V2X (*Vehicle-to-everything*), el acceso satelital 5G NTN (*Non-Terrestrial Network*), soporte de red de área local en 5G, convergencia inalámbrica y alámbrica para 5G, posicionamiento y ubicación de terminales, comunicaciones en dominios verticales, automatización de redes y nuevas técnicas de radio. Se iniciaron o avanzaron estudios adicionales sobre seguridad, códecs , servicios de transmisión, interfuncionamiento de LAN, división de redes e IoT.

Otras actividades se centraron en ampliar la aplicabilidad de la tecnología 3GPP a los sistemas de acceso por radio no terrestres, desde satélites y estaciones base aerotransportadas hasta aplicaciones marítimas, incluidas las comunicaciones de barco a tierra y de barco a barco. El trabajo también avanzó en la nueva funcionalidad de radio móvil profesional PMR (*Professional Mobile Radio*) para LTE, y mejoró los servicios orientados al ferrocarril desarrollados originalmente con la tecnología de radio GSM y que ahora se acercan al final de su vida útil.

Contenido de la Release 15:

- 5G NR.

- El sistema 5G - Fase 1.

- mMTC e Internet of Things (IoT).

- *Vehicle-to-everything* (V2X) Fase 2.

- Interoperabilidad de Comunicaciones MCC (*Mission Critical Communications*) con las redes legadas.
- WLAN y uso de espectro no licenciado.
- *Slicing*: Redes lógicas extremo a extremo.
- Exposición de APIs. Acceso de terceras partes a los servicios 5G.
- SBA.
- Mejoras LTE adicionales.
- FRMCS (*Future Railway Mobile Communication System*). Servicios orientados al ferrocarril.

3.2.2 3GPP Release 16

Su desarrollo comienza en Junio de 2018 y se completa en Junio de 2020

La versión 16 es una versión fundamental del proyecto 5G, puesto que completa la presentación de la propuesta por parte de 3GPP a la solución 5G requerida por IMT-2020.

En general incluye las mejoras necesarias para aportar madurez a los diferentes componentes introducidos en release 15. Son múltiples las mejoras introducidas en la parte de acceso radio, pero aquí nos centraremos en la relativas a la parte de núcleo de red, como objetivo fundamental de este trabajo. Algunas de las funcionalidades más importantes introducidas en la release 16 son:

- Mejoras del sistema 5G para nuevos servicios verticales como automatización industrial, URLLC, o V2X.
- FRMCS (*Future Railway Mobile Communication System*) Fase 2.
- Acceso Satelital en 5G.
- Acceso NR a espectro no licenciado.
- Convergencia fijo-móvil 5G *5G Wireless Wireline Convergence (5WWC).*
- Mejoras en la analítica de red.
- Mejoras en *network slicing.*
- Mejoras en *Service Based Architecture* (SBA).
- Mejoras en los servicios de localización.

Entre estos puntos, son de especial importancia para el desarrollo de este trabajo las mejoras introducidas en SBA. La release 16 introduce la entidad de red SCP y las siguientes características principales:

- Soporte de modelos de comunicación indirecta de Servicios NF/NF a través de una función intermediaria SCP (Servicio Proxy de Comunicación).
- Conjunto de servicios NF/NF que permite la agrupación de instancias NF o de instancias de servicio NF equivalentes. Los Servicios NF/NF dentro de un conjunto de Servicios NF/NF pueden compartir los mismos datos de contexto, mejorando así la resiliencia para procesar cualquier transacción.

- El mecanismo de vinculación (*binding*) mejora la flexibilidad y la eficiencia de la arquitectura basada en servicios al permitir que el productor de NF indique dinámicamente que el consumidor de NF, para un contexto particular; debe vincularse a una instancia de servicio de NF, a una instancia de NF, a un conjunto de servicios de NF o a un conjunto de NF para transacciones posteriores dependiendo de las políticas locales u otros criterios.

3.2.3 3GPP Release 17

Su desarrollo comienza en Diciembre de 2019 y se completa en Septiembre de 2022.

La tercera y última release del sistema 5G continúa añadiendo mejoras para completar la madurez del 5GS. Al igual que con las release anteriores, muchas de las mejoras están relacionadas con la parte de acceso radio NR, aunque aquí nos centraremos principalmente en las relacionadas con el núcleo de red 5GC. Algunas de estas funcionalidades son:

- Introducción del soporte de acceso para NTN (*Non-Terrestrial Networks*).
- Componentes Satelitales de la arquitectura 5G.
- Automatización de red para 5G - fase 2.
- *Edge Computing* en 5G.
- Servicios basados en proximidad en 5GS.
- *Network slicing* - fase 2.
- Mejoras en servicios V2X.
- Servicios de Localización LCS (*Location Services*) en 5GC.
- Mejoras en SBA.

En la relativo a la señalización en el 5GC, la release 17 introduce mejoras para las comunicaciones en interfaces de roaming entre diferentes operadoras PLMN. Se incluyen mejoras en las interfaces de roaming entre vNSSF y hNSSF, entre vNRF y hNRF, y entre vSEPP y hSEPP:

- El NSSF en la vPLMN puede invocar el servicio *Nnssf_NSSelection* proporcionado por el NSSF en la hPLMN.
- Se pueden implementar múltiples NRF en una vPLMN y/o en una hPLMN. Se actualiza el interfaz entre ellos para admitir la modificación o eliminación de la suscripción en hNRF si se implementan múltiples NRF en hPLMN.
- El interfaz entre vSEPP y hSEPP se actualiza en función de los requisitos de GSMA sobre las mejoras en el roaming.

3.3 5G NSA versus 5G SA

3GPP en la Release 15, la primera que describe 5G, define dos modos principales para redes 5G:

- **5G Non Standalone (NSA)**: En el que el acceso radio LTE existente, junto con el núcleo de red EPC de 4G, se utilizan como ancla para la gestión de movilidad y cobertura, añadiendo la portadora 5G. Esta opción permite a las operadoras ofrecer servicios 5G en un periodo de tiempo más corto y a menor coste. Se ofrecen diferentes opciones en las que se combinan los accesos radio de LTE y NR.
- **5G Standalone (SA)**: Incluye la nueva red radio 5G NR, y un núcleo de red completamente nuevo que incorpora de manera nativa un nuevo conjunto de capacidades como segmentación de red (*network slicing*), separación de plano de usuario y plano de control CUPS, «cloudificación», soporte multi-Gbps o ultra-baja latencia.

3GPP define diferentes opciones de despliegue de soluciones 5G dependiendo del enfoque utilizado. Éstas fueron publicadas en tres versiones consecutivas del conjunto de especificaciones de la release 15.

- **Release 15 versión temprana**: Completada en Diciembre 2017, introduce NSA utilizando conectividad dual entre LTE y 5G NR con el plano de control establecido en el «nodo máster» LTE y conectado al núcleo de red 4G EPC. Ésta es denominada como opción 3 o EN-DC por Conectividad Dual E-UTRA - NR.
- **Release 15 versión principal**: Completada en Junio-Septiembre de 2018, añadió las especificaciones de red para el núcleo 5G (5GC), diseñado para trabajar con 5G NR SA y denominada opción 2. También definió una actualización de LTE para soportar conectividad con el núcleo de red 5G (5GC), llamada opción 5 o *enhanced LTE* (eLTE).
- **Release 15 versión tardía**: Completada en Marzo-Junio 2019, añade modos NSA adicionales conectados al núcleo 5G en los que se usan tanto 5G NR como 4G LTE. La opción 7 o ngEN-DC (Conectividad Dual E-UTRA - NG de siguiente generación) utiliza LTE como sistema «máster»; y la opción 4 o NE-DC (Conectividad Dual NR - E-UTRA) utiliza NR como nodo Máster.

Estas opciones se muestran en la figura 3.1 [52].

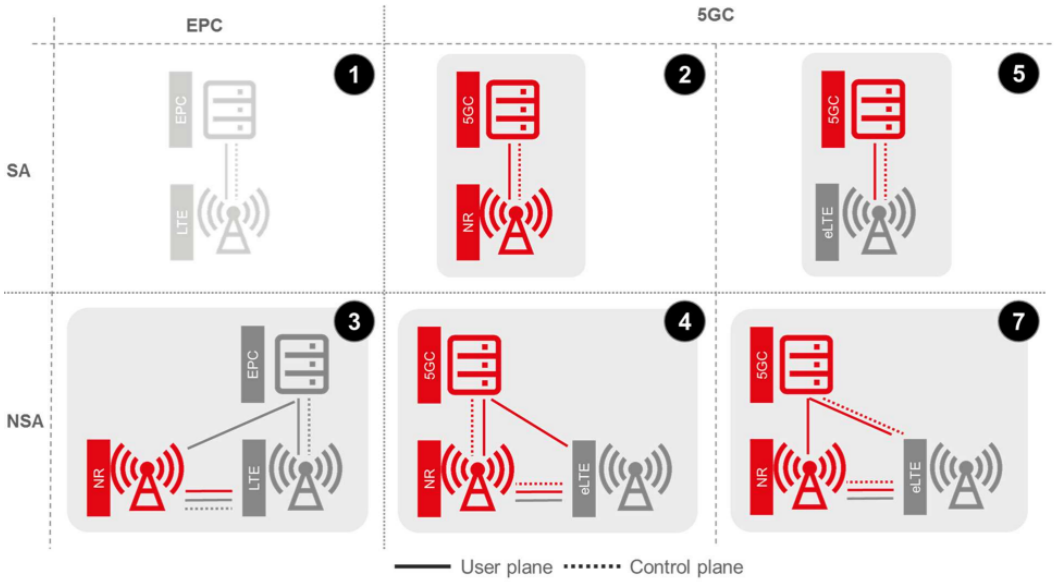

Figura 3.1: Opciones definidas por 3GPP para despliegues 5G

Las capacidades que distinguen las distintas opciones son el uso de conectividad dual, la tecnología de acceso radio que ejerce como nodo máster, y el núcleo de red utilizado.

Las opciones que utilizan conectividad dual se agrupan bajo el término NSA (*Non StandAlone*) para indicar que las tecnologías de acceso radio NR y LTE son usadas simultáneamente. Las opciones en las que se utiliza una sola tecnología de acceso se denominan SA (*Standalone*).

Desde que se desarrollaron las especificaciones de estas opciones, se preveía que las operadoras móviles comenzarían sus despliegues 5G utilizando la opción 3 del modo NSA, permitiéndoles reutilizar la funcionalidad existente en sus núcleos de red 4G LTE. Es por esto que fue la primera en ser publicada por 3GPP.

3.3.1 5G NSA

Las diferentes opciones NSA (3/3a/3x, 4/4a y 7/7a/7x) se basan en el uso de conectividad dual entre NR y LTE.

El concepto de conectividad dual se describe inicialmente en las especificaciones de LTE para DC intra-E-UTRAN en 3GPP TS 36300 [43], para casos de uso específicos de 4G. Al introducir 5G y con ello la necesidad de obtener conectividad dual entre las 2 generaciones, se describe la conectividad dual entre E-UTRAN y NR en 3GPP TS 37340 [53] aunque utilizando los conceptos fundamentales ya presentes en la especificaciones intra-LTE.

Se establece el concepto de MeNB (*Master eNB*) y SeNB (*Secondary eNB*) para los nodos involucrados en una conectividad dual. Al hablar de conectividad dual entre LTE y NR, los términos se generalizan como MN (*Master Node*) y SN (*Secondary Node*). En los escenarios de DC, el dispositivo de usuario se conecta a un MN y a un SN. El MN y el SN están conectados entre ellos por una interfaz de red, y al menos el MN está conectado al núcleo de red.

Los escenarios NSA se describen como MR-DC (*Multi-RAT Dual Connectivity*) por el hecho de que se utilizan dos tipos de acceso radio simultáneamente para proveer conectividad a un dispositivo de usuario.

Las figuras 3.2 y 3.3 muestran la conectividad para los planos de control y de usuario entre el MN y el SN para un dispositivo de usuario determinado, según 3GPP TS 37.340 [53].

En cuanto al **plano de control**, y en el caso de ofrecer *Multi-RAT Dual Connectivity* MR-DC sobre 4G EPC (EN-DC u opción 3), la entidad del núcleo de red involucrada es el MME (*Mobility Maanagement Entity*) de 4G. El interfaz S1-MME es terminado en el MN (4G eNB), mientras que el MN y el SN se conectan a través del interfaz X2-C.

En el caso de de Multi-RAT Dual Connectivity (MR-DC) sobre 5G Core (5GC) (NE-DC u opción 4, y NGEN-DC u opción 7), el elemento de núcleo de red involucrado es el AMF de 5G. El interfaz NG-C es terminado en el MN, mientras que el MN y el SN se conectan a través del interfaz Xn-C.

En cuanto al **plano de usuario**, en el caso de ofrecer MR-DC sobre 4G EPC (EN-DC u opción 3), la entidad del núcleo de red involucrada es el SGW (*Serving Gateway*) de 4G. El interfaz entre MN y SN es X2-U, mientras que el interfaz S1-U se establece entre el MN, el SN o ambos; y el SGW.

En el caso de de MR-DC sobre 5GC (NE-DC u opción 4, y NGEN-DC u opción 7), el elemento de núcleo de red involucrado para el plano de usuario es el UPF de 5G. El interfaz entre MN y SN es Xn-U, mientras que el interfaz NG-U se establece entre el MN, el SN o ambos; y el UPF.

En el plano de usuario, la conectividad utilizada depende de la portadora utilizada por el UE, ya que podría estar utilizando únicamente el MN, únicamente el SN, o ambos al mismo tiempo.

Es importante resaltar que mientras que el plano de control del nodo secundario es siempre enviado a través del nodo máster, el plano de usuario del nodo secundario puede enviarse a través del nodo máster o estar conectado al elemento del núcleo de red directamente.

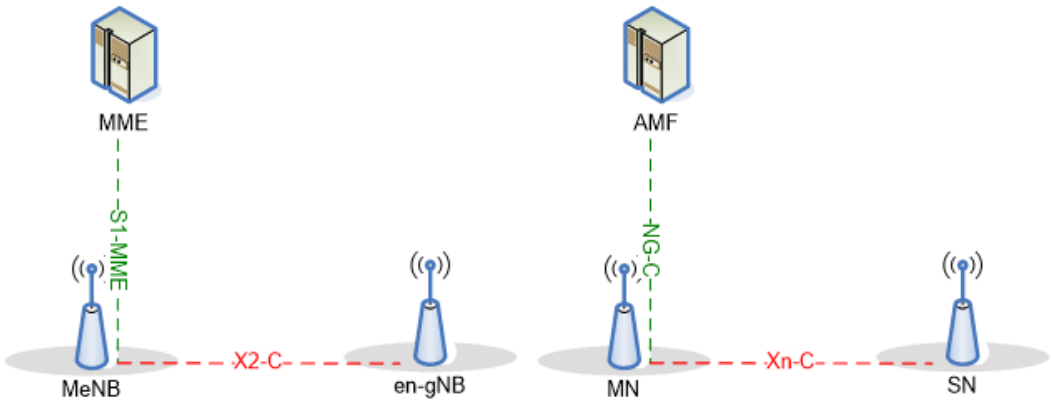

Figura 3.2: Conectividad de plano de control para EN-DC (izquierda) y MR-DC con 5GC (derecha)

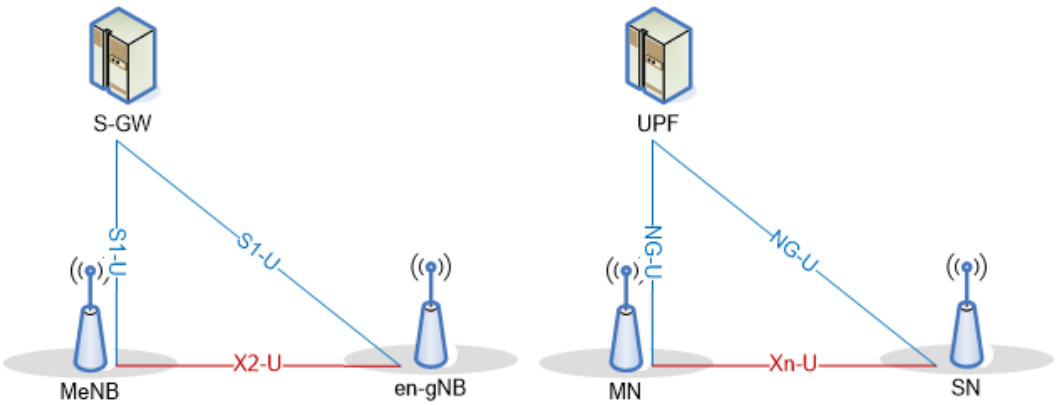

Figura 3.3: Conectividad de plano de usuario para EN-DC (izquierda) y MR-DC con 5GC (derecha)

3.3.1.1 Opción 3: EN-DC

En la opción 3, la conectividad dual (DC) es implementada sobre LTE utilizando el núcleo de red 4G EPC. Esta opción se denomina EN-DC (E-UTRAN - NR Dual Conectivity), incluyendo primero en la nomenclatura la red de acceso que actúa como ancla y MN (Master Node), que en este caso es LTE E-UTRAN.

Esta opción aprovecha los despliegues 4G existentes y es capaz de crear nuevos puntos de acceso 5G rápidamente, permitiendo la creación de nuevas aplicaciones y servicios 5G aunque de manera limitada.

Por otro lado, el modo 5G NSA opción 3 no introduce ningún cambio en la arquitectura y procedimientos de roaming existentes. Es función del operador visitado (VPLMN), el permitir que los roamers visitando su red utilicen 5G NSA Opción 3 o solo permitirles utilizar SA/LTE (Opción 1).

3.3.1.1.1 Diferencias entre las opciones 3/3A/3X En la opción 3, no existe conexión entre el 5G gNB y el 4G EPC, considerando la premisa de que el núcleo de red EPC no debe ser impactado. Es el modo de menor impacto puesto que el interfaz X2 entre LTE eNB y gNB se mantiene de la misma manera que estaba definido para la DC intra-LTE. Los cambios son únicamente necesarios en el 4G eNB.

En la opción 3a, el gNB implementa el interfaz S1-U hacia el EPC para el plano de control, pero no el interfaz X2.

La opción 3x es una combinación de la opción 3 y la opción 3a, en la que ambos interfaces (S1-U y X2) están disponibles. El impacto en la red es relativamente pequeño y se ha convertido en la opción mayoritariamente aceptada y utilizada para conectividad NSA [54].

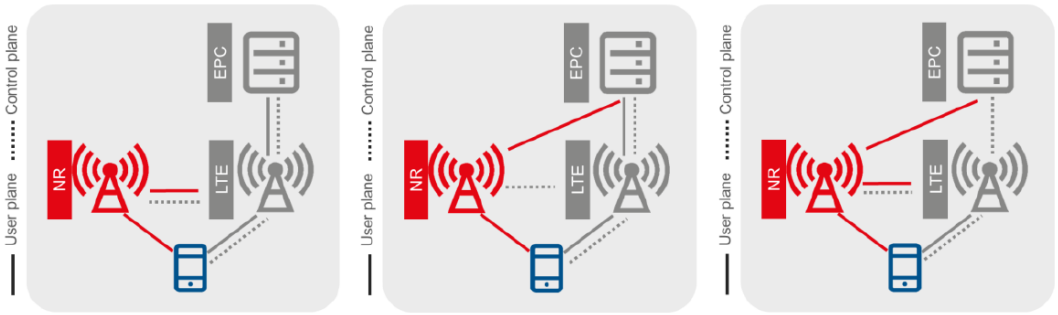

Figura 3.4: Modos de conectividad NSA de las opciones 3, 3a y 3x

3.3.1.2 Opción 4: NE-DC

En la opción 4/4a, la conectividad dual (DC) es implementada utilizando NR como ancla, con el NR gNB actuando como MN y el LTE eNB actuando como SR.

El núcleo de red es 5GC. El interfaz Xn entre gNB y LTE eNB utiliza procedimientos y protocolos diseñados específicamente para este uso. Esta opción se denomina NE-DC (NR - E-UTRAN Dual Connectivity), incluyendo primero en la nomenclatura la red de acceso que actúa como ancla y MN (Master Node), que en este caso es 5G NR.

3.3.1.2.1 Diferencias entre las opciones 4/4A En la opción 4, no existe interfaz directo entre ng-eNB y el 5GC. La información fluye a través del interfaz Xn.

En la opción 4a, no existe interfaz Xn entre el ng-eNB y el gNB. El ng-eNB está conectado al 5GC a través del interfaz NG-U.

3.3.1.3 Opción 7: NGEN-DC

En la opción 7, la conectivididad dual (DC) es implementada utilizando el LTE eNB como MN (en este caso un ng-eNB), conectado al núcleo de red 5G y utilizando el 5G gNB como SN. El eNB «evolucionado» y el gNB se conectan a través del interfaz Xn utilizando procedimientos y protocolos diseñados específicamente para este uso.

Esta opción se denomina NGEN-DC (NG-RAN-E-UTRA - NR Dual Connectivity), incluyendo primero en la nomenclatura la red de acceso que actúa como ancla y MN (Master Node), que en este caso es ng-eNB LTE. Su despliegue está más relacionado con la necesidad de aportar mayor capacidad que con ampliar cobertura. Permite al operador desplegar rápidamente servicios de amplia cobertura que requieren características del 5GC (como *Mobile Edge Computing* o *Network Slicing*), reutilizando todo la capacidad y cobertura existente de LTE además de combinarlo con las portadoras 5G NR para obtener mayores capacidades y velocidades.

3.3.1.3.1 Diferencias entre las opciones 7/7A/7X En la opción 7, no existe interfaz entre gNB y el 5GC. La información fluye a través de Xn, la conexión del 5G NR hacia el 5GC para el plano de usuario se realiza a través del LTE ng-eNB.

En la opción 7a, no existe interfaz Xn y el gNB está conectado al 5GC a través del interfaz NG-U.

La opción 7x es una combinación de las opciones 7 y 7a.

3.3.2 5G-SA

3.3.2.1 Opción 1: 4G LTE-EPC

La opción 1 se incluye en estas opciones propuestas por 3GPP a modo de mostrar por completo las diferentes opciones de integración, pero no hace referencia a 5G sino a 4G. Se trata de un 4G LTE eNB conectado al 4G EPC, es decir, la red precedente 4G.

3.3.2.2 Opción 5: 4G LTE sobre 5GC

La opción 5 representa una opción intermedia en la que se despliega el 5GC, pero se utiliza el acceso 4G LTE como único modo de acceso. Utilizando estaciones base 4G LTE actualizadas como ng-eNB, se pueden ofrecer algunos de los beneficios de 5G NR en conjunto con el núcleo de red 5GC. Su atractivo principal es la posibilidad para el operador de desplegar rápidamente servicios de amplia cobertura que requieren características del 5GC (como *Mobile Edge Computing* o *Network Slicing*), reutilizando todo la capacidad y cobertura existente de LTE. Esto podría hacer que las primeras inversiones en 5GC fueran más atractivas.

Esta enfocado a operadoras que prefieran comenzar desplegando el núcleo de red 5G antes que la parte de acceso radio NR. En la realidad, la mayoría de las operadoras han comenzado su inversión 5G al contrario, introduciendo primero la parte radio 5G NR antes que 5GC y por tanto utilizando la opción 3 NSA. El uso de *Mobile Edge Computing* o *Network Slicing* no se encuentra entre los primeros casos de uso o entre los más urgentes, y es por ello que finalmente la opción 3 NSA ha sido ampliamente adoptada como punto de entrada para despliegues 5G.

3.3.2.3 Opción 2: 5G SA

La opción 2 es la opción pura 5G SA. Es la única opción para operadoras que quieran hacer un despliegue completamente desde cero de 5G. Tiene soporte completo de todas las capacidades de 5G a nivel de eMBB, mMTC y URLLC. Necesita espectro disponible en múltiples bandas para cubrir todos estos casos de uso así como ofrecer una amplia huella de cobertura 5G. Se introduce el núcleo de red 5GC al que solo se conectan las estaciones base 5G NR.

Dentro del modo SA, también se contempla un escenario de conectividad dual entre NR y NR (NR-DC). Múltiples operadoras tienen planes de utilizar bajas frecuencias (bandas de 700/800/900 MHZ, por ejemplo) para construir su red 5G, con una banda de frecuencia alta (banda de frecuencia milimétrica *mmWave*) para suplementar la capacidad de la red. Debido a las muy diferentes características de la transmisión radio entre las bajas y las altas frecuencias, no es realista utilizar la tecnología de *Carrier Aggregation* (CA) a través de una conexión dual de baja y alta frecuencia. Por este motivo, 3GPP propone la técnica de conexión dual NR-NR.

3.3.3 Estrategias de migración entre las diferentes opciones

La GSMA recomienda diferentes estrategias de migración entre las diferentes opciones, manteniendo como objetivo final el desarrollo de la opción 2 en el que 5G SA es desplegado, y teniendo en cuenta que la opción 3x para NSA y la opción 2 para SA se han convertido en el consenso de la industria.

La mayoría de las operadoras parten de la opción 1, puesto que ya disponen de redes 4G implementadas. Desde esta situación, se puede evolucionar directamente a la opción 2 en la que se hace un despliegue desde cero en paralelo de la red 5G SA tanto a nivel radio como de núcleo de red; o se puede optar por la opción 3 NSA, reutilizando las capacidades de la red 4G existente y añadiendo poco a poco las capacidades de 5G NR. La gran mayoría de las operadoras han comenzado introduciendo la opción 3 NSA por la facilidad y rapidez de su despliegue [54].

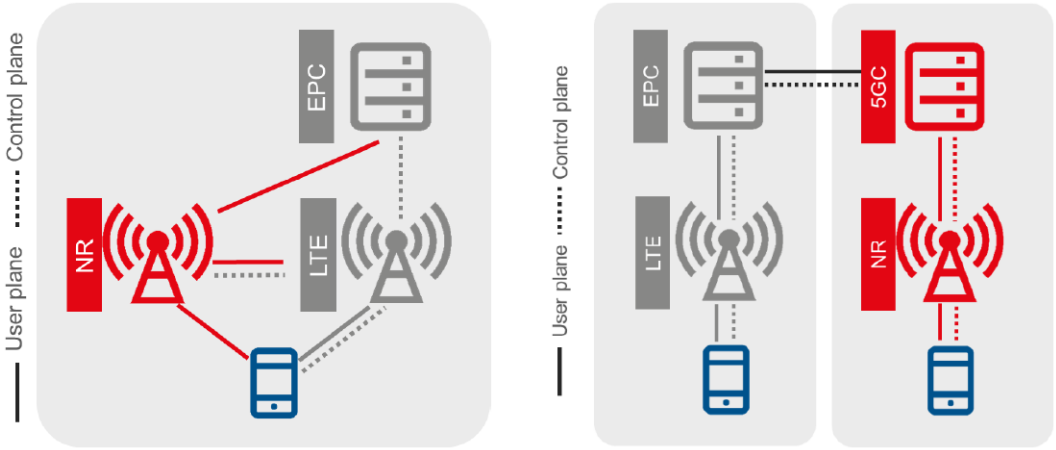

Figura 3.5: Arquitectura de alto nivel de NSA Opción 3x (izda.) y SA opción 2 (derecha)

Partiendo de la opción 3 NSA, las operadoras tienen la opción de evolucionar hacia una situación intermedia antes de adoptar completamente la opción 2 de 5G SA puro. Se trata de un escenario en el que se mezclan los modos NSA y SA, implementando interoperabilidad entre los dos núcleos de red. Hay múltiples estrategias posibles, pero GSMA recoge éstas como las más plausibles de ser implementadas:

- Una opción es pasar desde la opción 3 a la opción 7 y la opción 5, añadiendo el nuevo núcleo de red 5G y actualizando LTE a eLTE pero manteniendo el ancla o nodo máster en la red LTE. Esto permite a las operadoras continuar desplegando NR solo cuando casos de uso específicos son necesarios. Esta opción se muestra en la figura 3.6 [55].

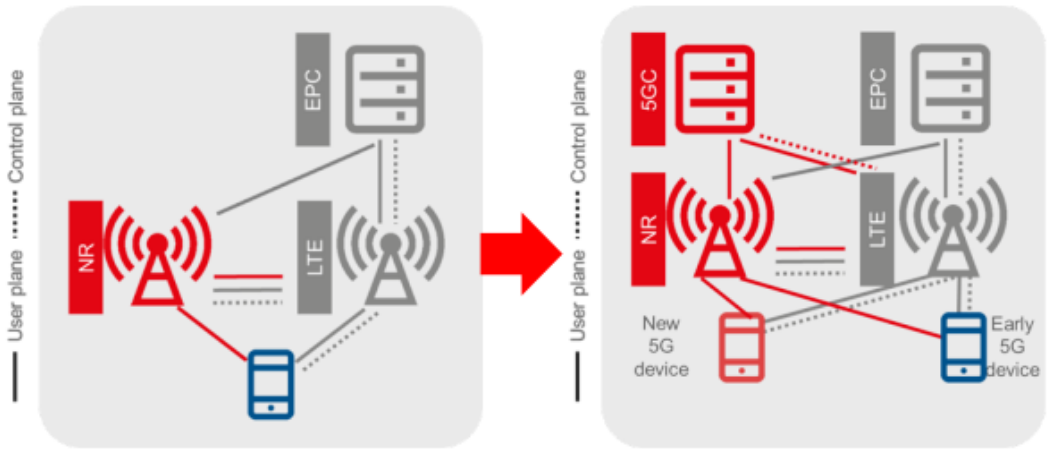

Figura 3.6: Migración desde NSA Opción 3 a NSA opción 7 y SA opción 5

- Otra opción es pasar desde la opción 3 directamente a la opción 2 SA, manteniendo ambas en paralelo. Esto implicaría que la operadora haya avanzado en su despliegue NR. Los dispositivos no

solo deben soportar 5G NSA opción 3, sino también SA. Esta opción se muestra en la figura 3.7 [55].

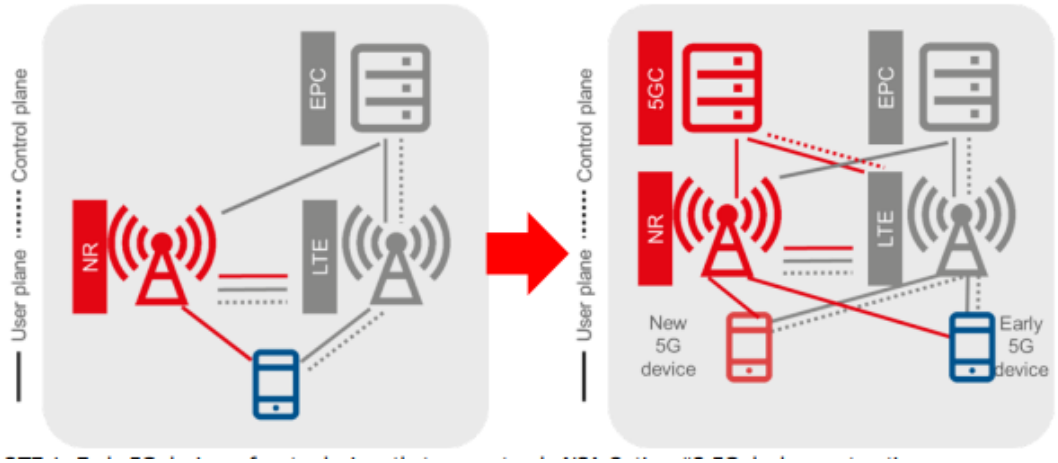

Figura 3.7: Migración desde NSA Opción 3 a NSA opción 3 y SA opción 2

- Por último, otra opción es migrar desde la opción 3 a una solución mixta de opción 4 y opción 2. Requiere dispositivos que soporten NSA 5GC (opción 4) y SA 5GC (opción 2), y se mantiene la conexión con 4G EPC para los dispositivos iniciales 5G (que solo soportan NSA opción 3). Requiere un mayor despliegue de NR puesto que NR es utilizado como nodo máster en las zonas en las que la opción 4 es utilizada. Esta opción se muestra en la figura 3.8 [55].

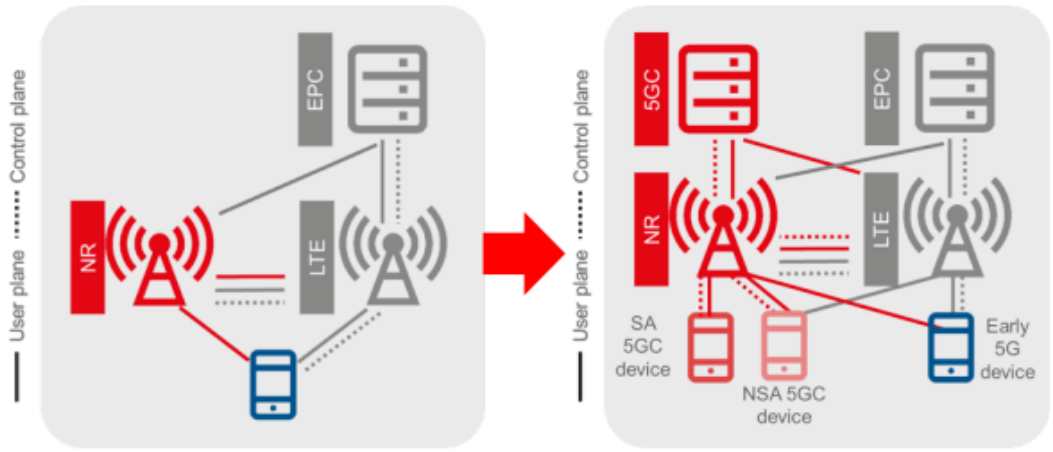

Figura 3.8: Migración desde NSA Opción 3 a NSA opción 4 y SA opción 2

3.3.4 Conclusión

El núcleo de red 5G ofrece múltiples beneficios y nuevas capacidades en comparación con el núcleo de red 4G EPC y es esencial para permitir a las operadoras evolucionar sus servicios habilitando los nuevos modelos de negocio y casos de usos avanzados que 5G puede ofrecer.

Siendo la opción 3 el modelo fundamental NSA, y la opción 2 el modelo final SA; las opciones intermedias 5, 7 y 4 son importantes en las diferentes estrategias de migración y para conectar los servicios del núcleo de red 5G con las diferentes áreas de cobertura 4G y 5G de manera transparente permitiendo continuidad del servicio.

Sin embargo, esto también implica que los diferentes proveedores de elementos de red y de dispositivos móviles, así como todo el ecosistema de la industria, deben soportar el desarrollo de estas múltiples opciones para su uso comercial.

3.4 Áreas del ecosistema 5G

La tecnología 5G se compone de 3 áreas fundamentales que veremos en esta sección. Por un lado están los dispositivos móviles de usuario, que deben ser modernizados para soportar la nueva tecnología en sus diferentes opciones. El siguiente componente es la nueva red de acceso radio llamada 5G NR que debe soportar las nuevas tecnologías de modulación radio y las múltiples bandas de frecuencias introducidas con el estándar 5G. Por último, la tecnología 5G implementa un núcleo de red completamente nuevo llamado 5GC (5G Core), con nuevas capacidades, tecnologías y protocolos.

Este trabajo se centra en el núcleo de red 5G (y en concreto en la relativo a la señalización), pero en está sección daremos una visión de alto nivel de estas tres áreas.

3.4.1 Dispositivos 5G

La modernización de los dispositivos móviles hacia 5G, debe ir relacionada con la adopción por parte de las operadoras de las diferentes opciones de despliegue de 5G propuestas por 3GPP.

Un dispositivo o UE (*User Equipment*) que requiera trabajar en una red que implementa la opción NSA, deberá soportar todos los protocolos requeridos por la implementación NSA y podría opcionalmente soportar los protocolos requeridos por SA. Del mismo modo, un dispositivo que necesita trabajar en una red que implementa la opción SA, deberá soportar todos los protocolos requeridos por la opción SA, pudiendo soportar opcionalmente los protocolos requeridos por la opción NSA.

Dado que el paso inicial de las operadoras ha estado relacionado con la adopción de NSA opción 3, en la que los eNBs de LTE actúan como nodo máster, también ha sido éste el enfoque de los desarrolladores de dispositivos móviles. La mayoría de fabricantes de dispositivos móviles comenzaron su actualización a 5G poniendo en el mercado dispositivos compatibles con NSA opción 3, pero no compatibles con la opción SA y los nuevos protocolos del nuevo 5GC. Esto también ha permitido a los fabricantes de dispositivos móviles posponer el desarrollo relacionado con SA y el nuevo núcleo de red 5GC, que requiere mayor tiempo de desarrollo y mayor coste.

De manera similar a cómo ocurrió con las generaciones anteriores 3G y 4G, con 5G también se está produciendo un despliegue progresivo de dispositivos móviles con soporte 5G. Es evidentemente un requisito indispensable que la base de clientes vaya adquiriendo este tipo de dispositivos para que las operadoras pueden ofrecer sus nuevos servicios sobre 5G.

Aunque muchos dispositivos móviles sean comercializados como «5G compatible», es importante entender el soporte exacto, pues puede ser solo compatible con NSA (la mayoría actualmente), o con SA. Por otro lado, también pueden ser compatibles con alguna o algunas de las bandas de frecuencia 5G pero no con todas. El soporte a 5G en la banda sub-6GHz es relativamente sencillo para los fabricantes (bandas similares a las tecnologías precedentes), y permite un tiempo de disponibilidad comercial menor. Por otra parte, esta banda no está optimizada para algunos de los nuevos casos de uso más allá de eMBB. Son las bandas milimétricas (*mmWave*) las que están optimizadas para el soporte de todos los servicios y casos de uso 5G, pero debido a la gran atenuación de su señal radio y a las dificultades de manejar frecuencias tan altas, su soporte puede ser complejo para los fabricantes de dispositivos; requiriendo mayor tiempo de desarrollo y mayor coste.

Por tanto, más allá de identificar dispositivos móviles compatibles o no con 5G, será necesario especificar las capacidades técnicas realmente soportadas: NSA, SA, sub6GHz, *mmWave* o VoNR (*Voice over New Radio*).

3.4.2 5G New Radio

El segundo área fundamental de la tecnología 5G es la red de acceso radio, denominada 5G NR (*New Radio*), a diferencia del LTE (*Long Term Evolution*) de 4G. 5G NR está separada en dos rangos de frecuencia. El rango 1 incluye las bandas sub-6 GHz con algunas de ellas tradicionalmente usadas por las generaciones anteriores; y el rango 2 que incluye bandas desde 24.25 GHz a 71 GHz. Estas bandas de ondas milimétricas (*mmwave*) tienen menor alcance pero mayor ancho de banda disponible que las bandas del rango 1. Dentro de ambos rangos, 5G NR define múltiples bandas de frecuencia disponibles [51].

A nivel de tecnologías de modulación, 5G NR reutiliza y expande muchas de las tecnologías ya incluidas en LTE-Advanced y LTE-Advanced Pro.

- **Carrier Aggregation (CA)** o Agregación de portadora permite combinar múltiples portadoras al mismo tiempo para obtener un mayor ancho de banda combinado.
- **Massive MIMO (mMIMO)**, o MIMO masivo, mejora las capacidades de MIMO ya utilizadas con anterioridad en comunicaciones inalámbricas. Combina más antenas para una mejor eficiencia espectral. mMIMO proporciona una gran cantidad de antenas en la estación base 5G (es decir, el gNB), lo que mejora la conectividad general y proporciona mejores velocidades. MIMO requiere algoritmos particulares que identifican dónde enfocar la energía. El termino masivo hace referencia a la cantidad de antenas, lo que ayuda a mejorar el rendimiento y la eficiencia de los datos.
- **Beamforming**, o formación de haces, consiste en la aplicación de múltiples elementos radiantes que transmiten la misma señal en una longitud de onda y fase idénticas, que se combinan para crear una sola antena con un flujo más largo y más específico. Se forma reforzando las ondas en una dirección específica. Cuantos más elementos radiantes componen la antena, más estrecho es el haz. Un elemento de la formación de haces son los lóbulos laterales. Éstos son esencialmente radiaciones no deseadas de la señal que forma el lóbulo principal en diferentes direcciones. Una ingeniería deficiente de los conjuntos de antenas daría como resultado una interferencia excesiva por parte del lóbulo lateral de una señal formada por haces. Cuantos más elementos radiantes componen la antena, más enfocado está el haz principal y más débiles son los lóbulos laterales.

Massive MIMO y Beamforming trabajan de manera coordinada y son fundamentales para la obtención de los anchos de banda ofrecidos por 5G.

3.4.2.1 gNB

En cuanto al elemento de la red 5G NR que actúa como estación base, 5G introduce la nomenclatura de gNB, o gNodeB (*next generation Node B*), como evolución del eNodeB de 4G. El gNB permite al UE o dispositivo de usuario 5G conectar con el núcleo de red 5G utilizando la interfaz aérea de 5G NR, ofreciendo la terminación de tráfico de plano de control y de plano de usuario al dispositivo de usuario.

Los gNB o estaciones base 5G utilizan la tecnología NR (*New Radio*), implementan soluciones SDR (*Software Defined Radio*) con soporte de varias opciones MIMO (*Multiple Inputs Multiple Outputs*) incluyendo 2x2, 4x4, 8x8 y opciones de *Massive MIMO* para capacidades mayores. También soportan despliegues tanto en bandas Sub-6GHz como en bandas de ondas milimétricas (*mmWave*).

Entre sus múltiples funciones, destacan estas tres tareas fundamentales:

- **MC (*Mobility Control*)**: Control y gestión de la movilidad de los usuarios a nivel de plano de control y señalización.
- **RRM (*Radio Resource Management*)**: Gestión de los Recursos Radio disponibles.
- **SM (*Session Management*)**: Control de las sesiones en plano de usuario establecidas por los usuarios conectados al gNB.

A nivel funcional y de comunicación, el gNodeB se compone principalmente de *Central Unit* (CU) y *Distributed Unit* (DU). La Unidad Central se divide entre el plano de control hacia AMF y SMF utilizando la interfaz N2; y el plano de usuario hacia UPF utilizando la interfaz N3. CU se encarga de las capas más altas de la pila de protocolos, mientras que DU se encarga de las capas más bajas, aunque existe flexibilidad a la hora de disponer las diferentes funciones del gNB en la Unidad Central o en las Unidades Distribuidas dependiendo de diferentes casos de uso. Existe un único CU por gNB, pero un CU puede controlar múltiples DUs. El interfaz entre CU y BU es denominado F1 y es un interfaz abierto especificado por 3GPP permitiendo conectar CU y DUs de diferentes proveedores [56].

Uno de los motivos principales de esta separación está relacionado con la virtualización y la voluntad de ofrecer Software Defined Radio (SDR) en el gNB. La virtualización ofrece mayor flexibilidad y reducción de costes pero, al menos por ahora, es complicado virtualizar las capas más bajas. Sin embargo, las capas más altas de la pila de protocolos pueden manejarse en un CU como solución software virtualizada.

Tanto el gNB como el UE 5G tienen su pila de protocolos estrechamente ligada a la pila de protocolos de la red radio 5G NR, con el objetivo de obtener una estrecha interoperabilidad entre 5G NR y 5GC. Cuando un UE desea conectar con una red de datos (como Internet), la red 5G establece un túnel extremo a extremo entre el UE y el UPF (*User Plane Function*), que representa una sesión PDU (*Protocol Data Unit*).

Del mismo modo, se establece un enlace único de señalización entre el UE y la red 5G para intercambiar mensajes de control. En la interfaz aérea, el túnel de datos extremo a extremo se denomina *Data Radio Bearer* (DRB), y el enlace de señalización se denomina *Signaling Radio Bearer* (SRB). Estas portadoras radio son esencialmente túneles de capa 2.

Los mensajes de señalización con el protocolo NAS (*Non-Access Stratum*) entre UE y el núcleo 5G (entre UE y AMF/SMF) se intercambian a través de una asociación única por UE sobre la interfaz N2; mientras que la sesión de datos PDU toma la forma de un túnel GTP (*GPRS Tunneling Protocol*) entre gNB y el UPF.

El protocolo de capa 1 hace referencia a la capa física *5G NR Physical Layer (PHI)*. La capa 2 incluye *Media Access Control (MAC)*, *Radio Link Control (RLC)* y *Packet Data Convergence Protocol (PDCP)*, mientras que la capa 3 hace referencia a la capa *Radio Resource Control (RRC)*. La figura 3.9 resume la pila de protocolos de la red radio 5G.

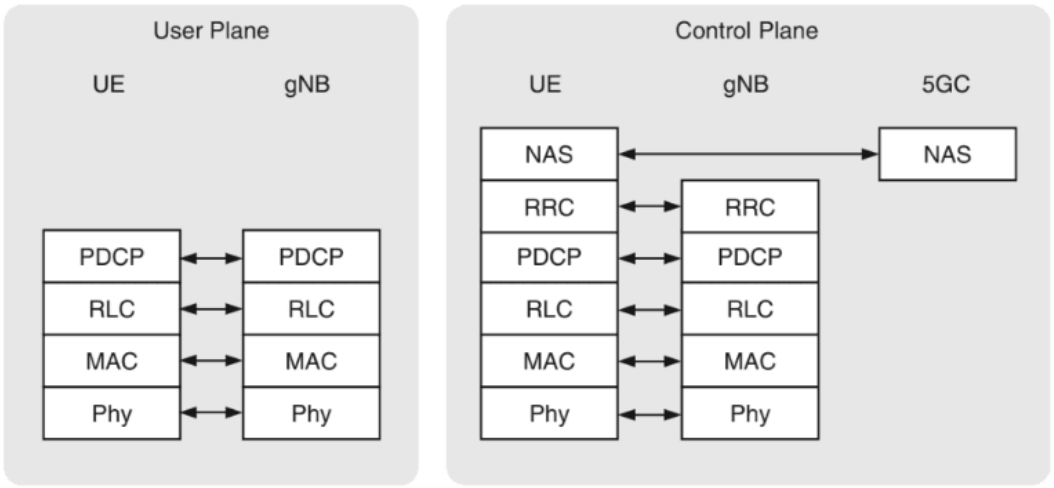

Figura 3.9: Protocolos de la red de acceso radio 5G NR

Por último, aún siendo un elemento 4G, es importante mencionar el elemento radio ng-eNB (*Next generation eNodeB*), que consiste en una versión mejorada de un eNodeB 4G en la que se soporta la conexión de un dispositivo de usuario 5G con el núcleo de red 5G utilizando la interfaz aérea 4G LTE. Es especialmente requerido en la transición y en el despliegue inicial de 5G NR en los modos 5G NSA. En muchas áreas geográficas en las que no exista cobertura 5G pero sí cobertura 4G, el ng-eNB permite al abonado 5G tener acceso a los servicios de la red 5G conectándole al núcleo 5G utilizando la interfaz aérea 4G LTE.

3.4.3 5G Core

Por último, el tercer área fundamental de la tecnología 5G es el nuevo núcleo de red 5G llamado 5GC (*5G Core*). Introduce un nuevo modelo de arquitectura basado en servicios (SBA: *Service Based Architecture*), con un conjunto de funciones y entidades de red completamente nuevo. Mientras que muchas de ellas tienen un rol parecido al de elementos similares de tecnologías precedentes, otras implementan conceptos completamente nuevos como por ejemplo las relacionadas con el descubrimiento automatizado del resto de funciones de red y sus servicios ofrecidos (NRF: *Network Repository Function*), o las relacionadas con la segmentación lógica de la red (NSSF: *Network Slice Selection Funtion*).

En lo relacionado con el plano de control y señalización, 5GC introduce HTTP2 como protocolo proveniente del sector de las tecnologías de la información, a diferencia de los protocolos de señalización específicos del sector de telecomunicaciones utilizados en las tecnologías precedentes: SS7 en 2G y 3G, y Diameter en 4G.

Por otro lado, mientras que estos protocolos de señalización de generaciones anteriores incorporaban diferentes fallos de seguridad revelados a lo largo de los años, 5GC introduce el concepto de seguridad por diseño o *security-by-design* ofreciendo nuevos mecanismos de autenticación y autorización, así como el envío de señalización cifrada utilizando TLS mutuo entre cada par de funciones de red que establecen comunicación de señalización entre ellas.

El 5G Core es el objetivo fundamental de este trabajo, en el que nos centraremos en los siguientes capítulos.

Capítulo 4

Desarrollo

Toda tecnología lo suficientemente avanzada es indistinguible de la magia.

Any sufficiently advanced technology is indistinguishable from magic.

Arthur C. Clarke [1]

4.1 Arquitectura del sistema 5G

La release 15 de 3GPP fue la primera en introducir el núcleo de red 5G. En esta versión inicial, varios de los elementos específicos para la gestión de la señalización a nivel de enrutamiento y seguridad como SCP o SEPP no habían sido aún introducidos o suficientemente desarrollados. Es por eso que nos centraremos en la release 16, como release más madura y release principalmente considerada por las operadoras como release objetivo en sus despliegues iniciales.

La arquitectura del sistema 5G se define para admitir la conectividad de datos y servicios que permiten utilizar técnicas como, por ejemplo, virtualización de funciones de red, redes definidas por software o *network slicing*. La arquitectura del sistema 5G aprovecha las interacciones basadas en servicios entre las funciones de red del plano de control CP (*Control Plane*). Algunos principios y conceptos clave del sistema 5G son:

- Separar las funciones del plano de usuario UP (*User Plane*) de las funciones del plano de control (CP), permitiendo escalabilidad independiente, evolución e implementaciones flexibles. Esto permite, por ejemplo, escoger entre ubicación centralizada o ubicación distribuida (remota).
- Modularizar el diseño de funciones para permitir una división de red flexible y eficiente.
- Definir los procedimientos (es decir, el conjunto de interacciones entre las funciones de la red) como servicios, de modo que sea posible su reutilización.
- Permitir que cada función de red (NF) y sus servicios de función de red interactúen con otra NF y sus servicios de función de red directa o indirectamente a través de un SCP (*Service Comunication Proxy*, o Proxy de comunicación de servicio).

[1] Arthur C. Clarke (1917 - 2008), escritor y científico británico. Tercera ley sobre el avance científico en *Profiles of the Future: An Inquiry Into the Limits of the Possible (1973)* [48].

- Minimizar las dependencias entre la Red de Acceso (AN) y la Red Núcleo (CN). La arquitectura se define con una red central convergente con una interfaz AN - CN común que integra diferentes tipos de acceso, por ejemplo acceso 3GPP y acceso no 3GPP.
- Ofrecer un marco de autenticación unificado.
- Admitir NFs sin estado (*stateless*), donde el recurso de cómputo está desacoplado del recurso de almacenamiento.
- Exposición de las capacidades soportadas por cada NF.
- Permite acceso simultáneo a servicios locales y centralizados. Para admitir servicios de baja latencia y acceso a redes de datos locales, las funciones UP se pueden implementar cerca de la red de acceso.
- Soporta itinerancia (roaming) tanto con el modelo HR (*Home Routed* o tráfico enrutado a la red doméstica), como con LBO (*Local Breakout*, o tráfico de ruptura local en la PLMN visitada).

La figura 4.1 muestra la arquitectura de referencia del sistema 5G en el caso nacional (sin itinerancia o roaming) utilizando la representación basada en puntos de referencia, en la que se muestran las interacciones entre las diferentes funciones de red [57]:

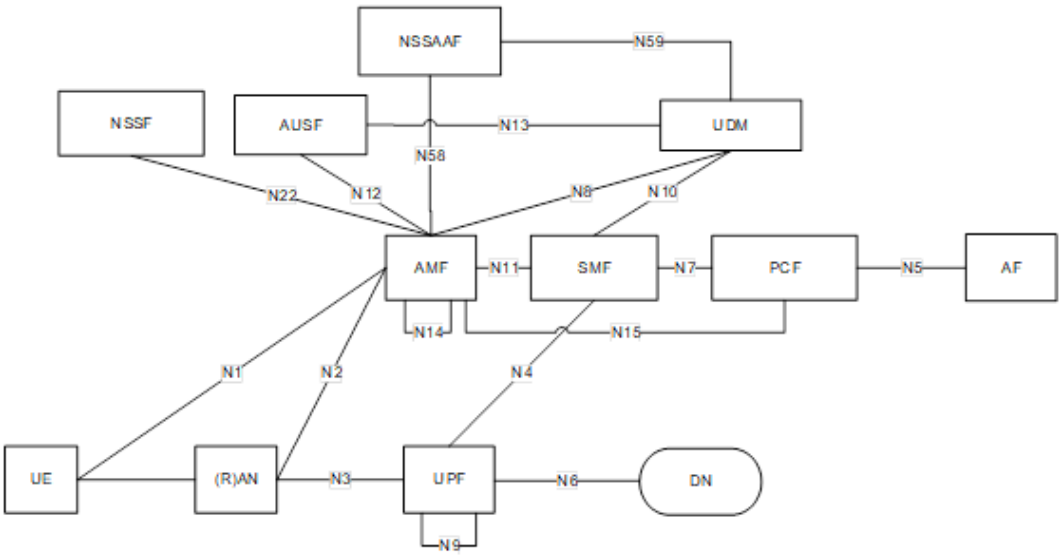

Figura 4.1: Arquitectura del sistema 5G en representación basada en puntos de referencia.

Con el objetivo de ofrecer mayor claridad, algunas interacciones no son mostradas en este diagrama. El UDSF, UDR, NEF, NRF o NWDAF no se muestran, puesto que todas las NFs mostradas pueden interactuar con ellos en caso de ser necesario.

4.1.1 SBA (Service Based Architecture)

Un cambio fundamental del sistema 5G (5GS) y el núcleo de red 5G (5GC) con respecto a cualquiera de las generaciones precedentes 2G/3G/4G, es el uso de una arquitectura basada en servicios. Cada NF es

capaz de exponer los servicios que ofrece al resto de la red, para que estos puedan ser descubiertos de manera automática por otra NF consumidora que necesite utilizarlos.

En la arquitectura del sistema 5G, las funciones de red del plano de Control (CP) deben estar basadas en SBA.

Es por esto que, además de la representación de la arquitectura basada en puntos de referencia de la sección anterior, 5G también incluye una representación de la arquitectura basada en SBA como la de la figura 4.2 [57]:

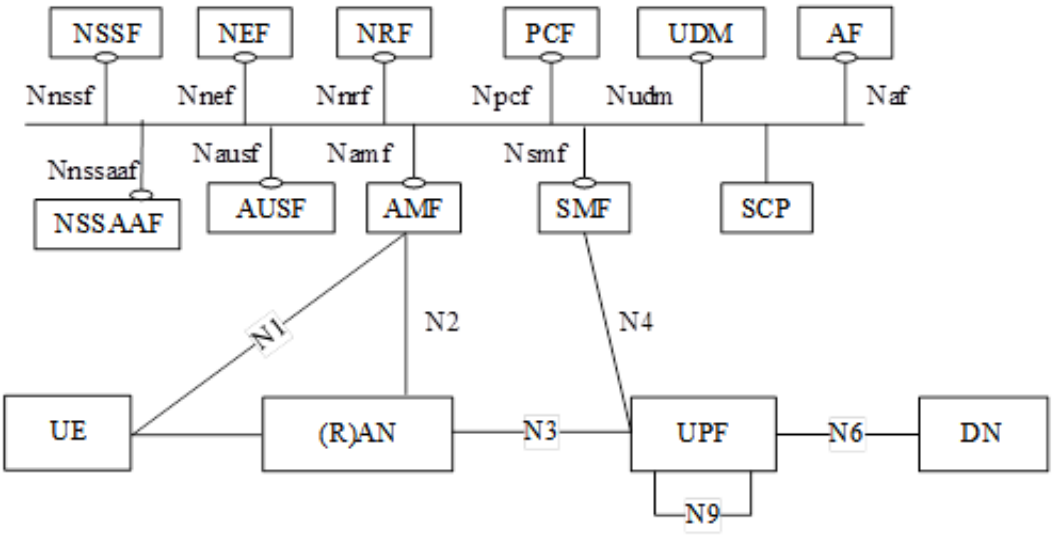

Figura 4.2: Arquitectura del sistema 5G en representación basada en SBI.

4.1.1.1 SBI (Service Based Interfaces)

Una interfaz basada en servicios representa cómo una NF dada proporciona o expone el conjunto de servicios que ofrece. Ésta es la interfaz donde se invocan las operaciones del servicio NF.

Las funciones de red dentro del plano de control 5GC solo utilizarán interfaces basadas en servicios (SBI) para sus interacciones. Las NF y los servicios de NF pueden comunicarse directamente, lo que se conoce como comunicación directa, o indirectamente a través del SCP.

La arquitectura del sistema 5G contiene las siguientes SBIs:

- **Namf**: Interfaz basada en servicios expuesto por el AMF.
- **Nsmf**: Interfaz basada en servicios expuesto por el SMF.
- **Nnef**: Interfaz basada en servicios expuesto por el NEF.
- **Npcf**: Interfaz basada en servicios expuesto por el PCF.
- **Nudm**: Interfaz basada en servicios expuesto por el UDM.
- **Naf**: Interfaz basada en servicios expuesto por el AF.
- **Nnrf**: Interfaz basada en servicios expuesto por el NRF.
- **Nnssaaf**: Interfaz basada en servicios expuesto por el NSSAAF.
- **Nnssf**: Interfaz basada en servicios expuesto por el NSSF.

- **Nausf**: Interfaz basada en servicios expuesto por el AUSF.
- **Nudr**: Interfaz basada en servicios expuesto por el UDR.
- **Nudsf**: Interfaz basada en servicios expuesto por el UDSF.
- **N5g-eir**: Interfaz basada en servicios expuesto por el 5G-EIR.
- **Nnwdaf**: Interfaz basada en servicios expuesto por el NWDAF.
- **Nchf**: Interfaz basada en servicios expuesto por el CHF.
- **Nucmf**: Interfaz basada en servicios expuesto por el UCMF.

4.1.1.2 Marcos de servicios de NF

La arquitectura basada en servicios soporta el marco de servicio de NF (*NF Service Framework*) que incluye los siguientes mecanismos:

- **Registro y cancelación del servicio NF**: para que el NRF conozca las instancias NF disponibles y los servicios admitidos.
- **Descubrimiento de servicios NF**: para permitir que un consumidor de servicios NF descubra instancias de productores de servicios NF que brindan los servicios de NF esperados.
- **Autorización de servicios de NF**: para garantizar que el Consumidor de servicios de NF (cNF) esté autorizado a acceder al servicio de NF proporcionado por el Productor de servicios de NF (pNF).
- **Comunicación entre servicios**: los consumidores de servicios de NF y los productores de servicios de NF pueden comunicarse directa o indirectamente a través de un proxy de comunicación de servicios (SCP). Que una NF use comunicación directa o comunicación indirecta a través de un SCP se basa en la configuración de la NF.

4.1.1.3 Protocolos sobre SBA

La pila de protocolos utilizada en las interfaces basadas en servicios se muestra en la figura 4.3 [58]:

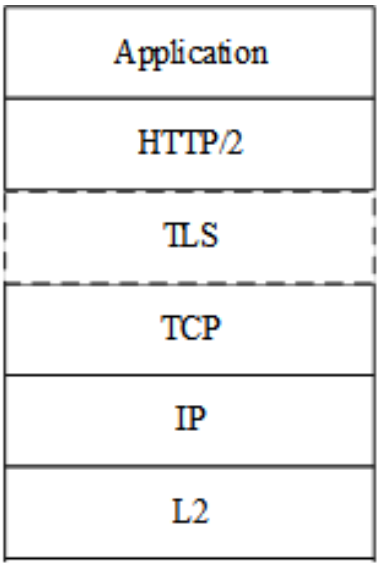

Figura 4.3: Pila de protocolos en las interfaces SBI.

Las interfaces basadas en servicios utilizan el protocolo HTTP/2 [59] con JSON [60] como protocolo de serialización de la capa de aplicación. Para la protección y seguridad en la capa de transporte, todas las NF de 3GPP deben admitir TLS y TLS se utilizará dentro de una PLMN si la seguridad de la red no se proporciona por otros medios.

Analizaremos el uso del protocolo HTTP2 en más detalle en la sección 5.2.3.

A nivel de protocolo de transporte, las interfaces SBI utilizan TCP [61] como protocolo de transporte requerido por HTTP2. Al utilizar TCP como protocolo de transporte, una conexión HTTP2 se mapea a una conexión TCP. Si una NF no registra ningún número de puerto específico en el NRF, entonces deberá estar preparada para recibir conexiones en los puertos por defecto, es decir, en el puerto 80 TCP para URIs de tipo «http» y en el puerto 443 para URIs de tipo «https».

Las *Uniform Resource identifier (URI)* son secuencias compactas de caracteres que identifican un recurso físico y abstracto. Su sintaxis y sus procesos asociados también están definidos por IETF [62].

4.1.1.4 Lenguaje de definición de Interfaces

En la definición hecha por parte de 3GPP de las distintas interfaces SBI del sistema 5G, se utiliza la especificación OpenAPI [63]. OpenAPI es una especificación para archivos de interfaz legibles por máquina para describir, producir, consumir y visualizar servicios web RESTful. Su desarrollo es supervisado por la iniciativa OpenAPI, un proyecto de colaboración de código abierto de la fundación Linux.

Por su parte, REST es una arquitectura software de interfaz que utiliza HTTP para obtener datos o indicar la ejecución de operaciones sobre datos. En la actualidad se usa en el sentido más amplio para describir cualquier interfaz entre sistemas que utilice directamente HTTP para obtener datos o indicar la ejecución de operaciones sobre los datos, en cualquier formato (XML, JSON, etc) sin las abstracciones adicionales de los protocolos basados en patrones de intercambio de mensajes.

REST utiliza estos conceptos clave o restricciones:

- Un protocolo cliente/servidor sin estado. Cada mensaje HTTP contiene toda la información necesaria para comprender la petición. Como resultado, ni el cliente ni el servidor necesitan recordar ningún estado de las comunicaciones entre mensajes.
- Un conjunto de operaciones bien definidas que se aplican a todos los recursos de información. HTTP define un conjunto de operaciones: POST, GET, PUT, DELETE, PATCH, OPTIONS.
- Una sintaxis universal para identificar los recursos (elementos de información). En un sistema REST, cada recurso es direccionable únicamente a través de su identificador global URI.
- El uso de hipermedia, tanto para la información de la aplicación como para las transiciones de estado de la aplicación. La representación de este estado en un sistema REST es típicamente HTML o XML.

La implementación de una API se denomina RESTful cuando cumple las restricciones que impone la arquitectura REST.

4.1.2 NFs (Network Functions)

La arquitectura del sistema 5G (5GS) se compone de las siguientes funciones de red (NFs):

- Authentication Server Function (AUSF).

- Access and Mobility Management Function (AMF).
- Data Network (DN), es decir, servicios del operador, acceso a Internet o servicios de terceras partes.
- Unstructured Data Storage Function (UDSF).
- Network Exposure Function (NEF).
- Network Repository Function (NRF).
- Network Slice Specific Authentication and Authorization Function (NSSAAF).
- Network Slice Selection Function (NSSF).
- Policy Control Function (PCF).
- Session Management Function (SMF).
- Unified Data Management (UDM).
- Unified Data Repository (UDR).
- User Plane Function (UPF).
- Application Function (AF).
- User Equipment (UE).
- (Radio) Access Network ((R)AN).
- 5G-Equipment Identity Register (5G-EIR).
- Network Data Analytics Function (NWDAF).
- CHarging Function (CHF).

También contiene las siguientes «entidades» de red:

- Service Communication Proxy (SCP).
- Security Edge Protection Proxy (SEPP).

Nos centraremos específicamente en NRF, SCP, SEPP y BSF como funciones y entidades de red especialmente encargados del control de la señalización HTTP2 a nivel de enrutamiento y seguridad, aunque daremos aquí un análisis de alto nivel de todas ellas.

4.1.2.1 NRF (Network Repository Function)

Mientras que muchas de las funciones de red del núcleo de red 5G son similares a funciones de red en tecnologías precedentes que tenían un conjunto de funciones similar, en el caso de NRF se trata de una función de red completamente nueva en 5G, sin equivalente en ninguna de las redes precedentes.

El NRF es un repositorio de perfiles de NFs en la red, que permite la automatización en el descubrimiento de NF productor por parte de cualquier NF consumidor. En todas las tecnologías precedentes, los elementos de red eran configurados manualmente con la red de elementos de red vecinos con los que requería interactuar. El 5GC introduce el NRF como habilitador de un nuevo enfoque de red en el que cada NF no tiene necesidad de conocer o estar configurado de antemano con el resto de NFs. NRF y la

arquitectura SBA ofrecen la capacidad de descubrir el resto de NFs, y en concreto qué NF de la red puede ofrecer los servicios necesitados por otra NF en cualquier momento, de manera dinámica y automatizada.

El NRF, o función de repositorio de red, ofrece la siguiente funcionalidad:

- Mantiene el perfil de NF de las instancias NF disponibles, así como sus servicios soportados.
- Mantiene el perfil de SCP de las instancias SCP disponibles.
- Función de descubrimiento de servicios. Recibe la solicitud de descubrimiento de NF por parte de una instancia de NF o SCP; y proporciona la información de las instancias de NF descubiertas a la instancia de NF o SCP.
- Soporta el descubrimiento de SCP por parte de otras instancias de SCP.
- Notifica sobre instancias NF y SCP recién registradas, actualizadas, o dadas de baja (junto con sus posibles servicios NF) al consumidor del servicio NF o SCP que se hubiera suscrito.
- Mantiene el estado de salud de NFs y SCP.

La figura 4.4 muestra la arquitectura de referencia 5G, con foco en el NRF y la interacción entre vNRF y hNRF en escenarios de roaming [64]:

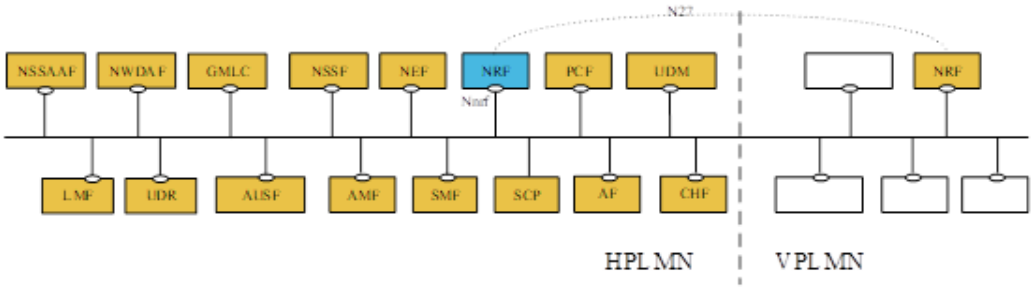

Figura 4.4: Arquitectura del sistema 5G con foco en NRF.

El NRF no es generalmente mostrado en las representaciones por punto de referencia del sistema 5G debido a que el NRF interactúa con todas y cada una de las NFs del 5GC. Particularizando en el escenario de roaming, se identifica el punto de referencia N27 entre vNRF y hNRF.

En el contexto de ofrecer soporte a *Network Slicing*, en función de la implementación de la red, se pueden implementar múltiples NRF en diferentes niveles:

- **Nivel PLMN**: El NRF se configura con información para toda la PLMN.
- **Nivel de segmento compartido**: El NRF se configura con información perteneciente a un conjunto de segmentos (*slices*) de red .
- **Nivel específico de segmento**: El NRF se configura con información perteneciente a un S-NSSAI concreto.

En el contexto de roaming, se pueden implementar múltiples NRF en las diferentes redes:

- **NRFs en la PLMN visitada**: Conocidos como vNRF y configurados con información para la PLMN visitada.

- **NRFs en la PLMN doméstica**: Conocidos como hNRF y configurados con información para la PLMN doméstica, a los que hace referencia y consultará el vNRF a través del interfaz N27.

El NRF mantiene registrados los perfiles de las diferentes NFs que componen un 5GC específico. El perfil de NF de la instancia NF mantenida en un NRF incluye múltiples parámetros, entre los que cabe destacar la siguiente información:

- Identificador de instancia de NF (*nfInstanceId*).
- Tipo de NF.
- Identificador de la PLMN.
- Identificador(es) de segmento (*slice*) de red S-NSSAI, e identificación de *Network Slice Instance (NSI)*.
- FQDN o dirección IP del NF.
- Información de capacidad de la NF.
- Información de prioridad del NF.
- Identificador del conjunto de NF.
- Identificador de conjunto de servicios de NF de la instancia de servicio de NF.
- Información de autorización sobre servicio específico del NF.
- Si corresponde, nombres de los servicios soportados.
- Dirección(es) de extremo (*endpoints*) de cada servicio compatible.
- Identificación de los datos/información almacenados (para un perfil UDR).
- *Data Network Name (DNN)* o lista e DNNs.
- Punto final de notificación para cada tipo de notificación que el servicio NF está interesado en recibir.
- Información de ubicación para la instancia de NF, como información específica identificada por el operador. Ejemplos de dicha información pueden ser la ubicación geográfica, o el centro de datos.
- Información de nivel de carga del NF.
- Indicador de enrutamiento, para indicar el UDM y AUSF a utilizar.
- Uno o más *Global Unique AMF ID (GUAMI)*, en el caso de AMF.
- Identidad(es) del área de SMF en el caso de UPF.
- Identificador de grupo de UDM, rango(s) de SUPI, rango(s) de Generic Public Subscription Identifier (GPSI), rango(s) de identificadores de grupo interno o rango(s) de identificadores de grupo externo para UDM.
- Identificador de grupo de UDR, rango(s) de SUPI, rango(s) de GPSI o rango(s) de identificadores de grupo externo para UDR.
- Identificador de grupo de AUSF, intervalo(s) de SUPI para AUSF.

- Identificador de grupo de PCF, rango(s) de SUPI para PCF.
- Lista de dominios IP, rango(s) de direcciones IPv4 o de prefijos IPv6 de UEs, en el caso de BSF.
- Dominio SCP al que pertenece el NF.

Ademas de las funciones de red, el NRF tambien almacena el perfil de la entidad de red SCP. El perfil SCP mantenido en un NRF incluye la siguiente información:

- Identificador de SCP.
- FQDN o dirección IP de SCP.
- Indicación de que el perfil es de un SCP (por ejemplo, parámetro de tipo NF establecido en tipo SCP).
- Información de capacidad de SCP.
- Información de carga de SCP.
- Prioridad de SCP.
- Información de ubicación del SCP (localidad).
- Ubicación(es) servida(s).
- Identificador(es) relacionado(s) de segmento (*slice*) de red, como S-NSSAI e identificación de NSI.
- PLMN remotas accesibles a través de SCP.
- Direcciones de puntos finales accesibles a través del SCP.
- NFs y conjuntos de NF atendidos por el SCP.
- Dominio SCP al que pertenece el SCP. Si un SCP pertenece a más de un dominio SCP, el SCP podrá conectar estos dominios, es decir, enviar mensajes entre ellos.

El NRF ofrece los siguientes servicios disponibles para cualquier otra NF de la red 5G:

- **Nnrf_NFManagement**: Brinda soporte para el servicio de registro, desregistro y actualización a NFs, a servicios de NFs y a SCPs. Proporciona a los consumidores de servicios de NF y al SCP notificaciones sobre eventos de NF que han sido registrados/actualizados/cancelados. También proporciona al SCP notificaciones de SCP recién registrado/actualizado/desregistrado. En lo relativo al almacenamiento de perfiles, soporta las operaciones de **NFRegister** (registro de nueva NF), **NFUpdate** (actualización del perfil previamente registrado) y **NFDeregister** (desregistro de un perfil previamente registrado). En cuanto a gestión de notificaciones, soporta las operaciones de **NFStatusSubscribe** (para que una NF o SCP puedan suscribirse a notificaciones sobre cambios en determinadas NFs registradas en NRF), **NFStatusNotify** (la notificación enviada por NRF informando de estos cambios a NFs suscritas previamente) y **NFStatusUnsubscribe** (para terminar la suscripción a cambios en determinados perfiles de NF). Por último, el NRF también soporta la operación **NFListRetrieval** (para obtener una lista de las NFs y SCPs registrados en el NRF) y **NFProfileRetrieval** (para obtener el perfil concreto de una NF o SCP).

- **Nnrf_NFDiscovery**: Permite que un consumidor de servicios de NF o SCP descubra un conjunto de instancias de NF en base a un servicio de NF específico o un tipo de NF de destino. También permite que un consumidor de servicio NF o SCP descubra un servicio NF específico. También permite que un SCP descubra un SCP de siguiente salto. El consumidor utilizará un conjunto de parámetros de descubrimiento (*discovery parameters*) para permitir al NRF filtrar entre los diferentes perfiles almacenados, e identificar los que el consumidor realmente necesita. La operación incluida en este servicio es **NFDiscover**, invocada por un NF consumidor o SCP, solicitando descubrir instancias de NF (NFs objetivo o *Target NFs*).

- **Nnrf_AccessToken**: Proporciona tokens de acceso OAuth2 2.0 para la autorización de NF a NF. Expone un token *endpoint* donde los consumidores de servicios de NF pueden enviar una petición de servicio para solicitar un token de acceso. La operación de servicio es **Nnrf_AccessToken_Get**.

- **Nnrf_Bootstrapping**: Permite que los consumidores de servicios NF de la NRF conozcan los puntos finales (*endpoints*) de los servicios que soporta el NRF, mediante el uso de un *endpoint* URI independiente de la versión de API soportada por el NRF, y que no es necesario descubrir mediante el uso de un servicio *Discovery*. Este servicio se utilizará en escenarios inter-PLMN donde el NRF en una PLMN-A necesita invocar servicios de un NRF en PLMN-B, cuando no hay información preconfigurada que indique la versión de los servicios desplegados en PLMN-B. Este servicio también puede ser utilizado en escenarios intra-PLMN, para evitar configurar de forma estática en los diferentes NF información sobre las versiones del servicio desplegadas en el NRF. La operación de servicio es **Nnrf_Bootstrapping_Get**.

4.1.2.2 SCP (Service Communication Proxy)

El SCP es el elemento central de enrutamiento de señalización HTTP2 en el núcleo de red 5G y en la arquitectura SBA. No fue incluido en la release 15, pero fue añadido en la release 16, revisando los modos de comunicación entre NFs, utilizando los servicios ofrecidos por NRF y las capacidades de enrutamiento centralizado del SCP.

Las mejoras en arquitectura del sistema 5G introducidas en la release 16 tienen como resultado la definición de diferentes modelos de comunicación que las NF y los servicios de NF pueden utilizar para interactuar entre ellos. Las distintas opciones están relacionadas con la decisión de utilizar o no la nueva entidad de red SCP, como se muestra en la figura 4.5 [65].

- **Modelo A: Comunicación directa sin interacción con NRF**: No se utilizan NRF ni SCP. Los consumidores se configuran estáticamente con los perfiles NF de los productores que necesitará contactar, y se comunican directamente con un productor de su elección.

- **Modelo B: Comunicación directa con interacción con NRF**: Los consumidores descubren a los productores consultando el NRF. Con base al resultado del descubrimiento, el consumidor hace la selección. El consumidor envía la solicitud al productor seleccionado.

- **Modelo C: Comunicación indirecta sin descubrimiento delegado**: Los consumidores descubren a los productores consultando al NRF. Según el resultado del descubrimiento, el consumidor selecciona un conjunto de NF o una instancia de NF específica del conjunto de instancias de NF. El consumidor envía la solicitud al SCP e introduce en la petición la dirección del productor del servicio seleccionado. El SCP enruta la solicitud a la instancia del productor de servicios NF seleccionada.

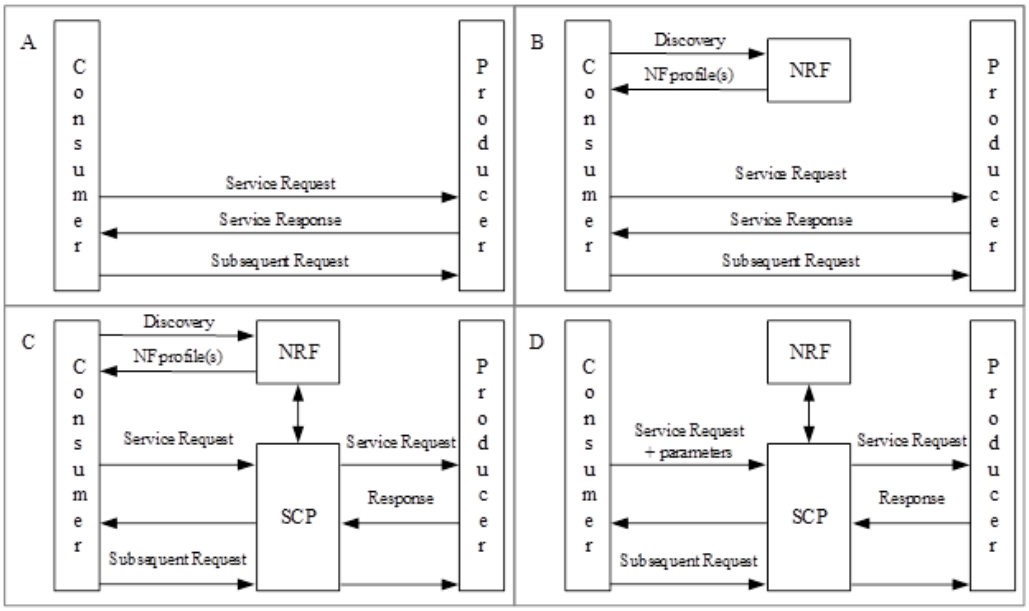

Figura 4.5: Modelos de comunicación para comunicaciones entre NFs o servicios de NFs.

- **Modelo D: Comunicación indirecta con descubrimiento delegado**: Los consumidores no hacen ningún descubrimiento o selección. El consumidor agrega los diferentes parámetros de descubrimiento y selección que sean necesarios para encontrar un productor adecuado a la solicitud de servicio, como cabeceras adicionales del mensaje HTTP2. El SCP utiliza la dirección de solicitud y los parámetros de descubrimiento y selección del mensaje de solicitud para enrutar la solicitud a una instancia de productor adecuada. El SCP puede realizar un descubrimiento con un NRF y obtener un resultado de descubrimiento.

El SCP, o proxy de comunicación de servicio, incluye las siguientes funcionalidades:

- Soporte a la comunicación indirecta entre NFs a través de SCP.
- Descubrimiento Delegado. Las NF envían sus peticiones de servicio directamente al SCP. El SCP se encargará de descubrir el NF productor correspondiente consultando al NRF, y de seleccionar el pNF objetivo al que enviar la petición de servicio recibida del consumidor.
- Servicio de reenvío y enrutamiento de mensajes a destino NF/NF.
- Reenvío y enrutamiento de mensajes a un SCP de siguiente salto.
- Seguridad de la comunicación (por ejemplo, autorización del consumidor de servicios de NF para acceder a la API del productor de servicios de NF), balanceo de carga, monitoreo, control de sobrecarga, etc.
- Interactuar opcionalmente con UDR, para resolver el identificador de grupo UDM, identificador de grupo UDR, identificador de grupo AUSF, identificador de grupo PCF, identificador de grupo CHF o identificador de grupo HSS en función de la identidad del UE, por ejemplo SUPI o IMPI/IMPU. El concepto de grupo de NF y grupo de servicios de NF agrupa los NF del plano de control equivalentes en un conjunto de NF, o agrupa varias instancias de servicios de NF en un conjunto de servicios de NF. Los servicios de NF o los NFs dentro de un conjunto de servicios de NF o de NFs; pueden compartir los mismos datos de contexto. Cuando la instancia de productor de NF no está disponible,

se selecciona otra instancia de productor de NF dentro del mismo conjunto de NF. Cuando múltiples instancias de servicio NF dentro de un conjunto de servicios NF están expuestas al consumidor de servicio NF o SCP y la indisponibilidad de la instancia de servicio NF es detectada o notificada por el NRF, el consumidor de servicio NF o el SCP selecciona otra instancia de servicio NF del mismo grupo de servicio de NF si está disponible. De lo contrario, el consumidor del servicio de NF o el SCP selecciona una instancia NF diferente dentro del mismo conjunto de NFs.

El SCP puede implementarse de manera distribuida. Más de un SCP puede estar presente en la ruta de comunicación entre los servicios NF. Para permitir que los SCP enruten mensajes a través de varios SCP (es decir, descubrimiento del SCP de siguiente salto), un SCP puede registrar su perfil en el NRF. Alternativamente, se puede utilizar configuración local estática del siguiente SCP en la ruta del mensaje.

En caso de incorporar el descubrimiento dinámico del SCP de siguiente salto, un mensaje puede atravesar varias instancias de SCP hasta llegar a su destino final. Cada NF está configurada con su(s) SCP(s) de servicio, pero un SCP puede descubrir y seleccionar un SCP de siguiente salto consultando el servicio Nnrf_NFDiscovery de NRF. Un SCP puede usar los parámetros de perfil de SCP descritos en la sección 4.1.2.1 como parámetros de descubrimiento en su mensaje Nnrf_NFDiscovery. Los parámetros que se utilizarán dependen de la implementación de la red. El NRF devuelve una lista de perfiles SCP según los parámetros de descubrimiento proporcionados.

La manera en la que el SCP utiliza la información recuperada del NRF para resolver la ruta óptima hacia un productor depende de la implementación del SCP, la configuración específica de la implementación y las políticas del operador al respecto.

Analizaremos en más detalle los flujos de señalización y capacidades del SCP en la sección 5.3.

4.1.2.3 BSF (Binding Support Function)

El *Binding Support Function (BSF)*, o función de soporte a la vinculación, tiene dos funciones principales:

- Almacena la información de vinculación (*binding info*) para una determinada sesión de PDU registrada, actualizada o eliminada por parte de NFs consumidoras como el PCF.
- Ofrece el descubrimiento de información de vinculación (por ejemplo, la información de la dirección del PCF seleccionado).

El BSF ofrece un único servicio denominado **Nbsf_Management** que se utiliza para proporcionar una funcionalidad de enlace de sesión de PDU, lo que garantiza que una solicitud de AF para una determinada sesión de PDU llegue al PCF relevante que contiene esa información de sesión de PDU, o garantiza que se seleccione la misma PCF para varias sesiones de PDU.

Por tanto, este servicio:

- Permite a los consumidores de servicios NF registrarse, actualizar y eliminar información de vinculación.
- Permite a los consumidores de servicios NF recuperar información vinculante.

Los consumidores conocidos del servicio Nbsf_Management son:

- Policy Control Function (PCF)

- Network Exposure Function (NEF)
- Application Function (AF)
- Network Data Analytics Function (NWDAF)

El PCF mantendrá actualizada la información de vinculación (*binding*) de sesiones PDU en el BSF; mientras que NEF, AF o NWDAF consumirán la información de vinculación de la sesión para poder acceder correctamente a la sesión enrutando su petición al PCF correcto [66].

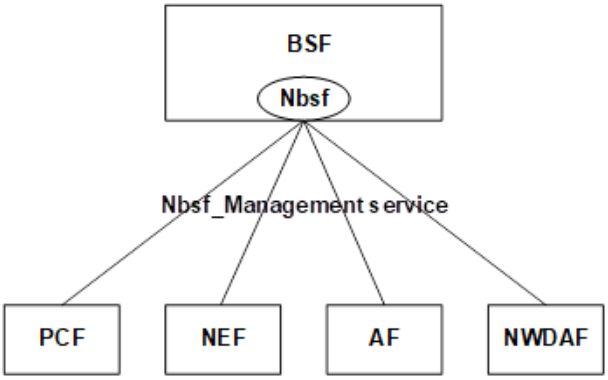

Figura 4.6: Arquitectura de referencia del servicio Nbsf_Management.

Dentro del servicio Nbsf_Management, BSF ofrece cuatro operaciones de servicio:

- **Nbsf_Management_Register**: Utilizada por PCF para registrar la información de vinculación para un UE.
- **Nbsf_Management_Deregister**: Utilizada por PCF para desregistrar la información de vinculación para un UE.
- **Nbsf_Management_Update**: Utilizada por PCF para actualizar la información de vinculación ya existente para un UE.
- **Nbsf_Management_Discovery**: Utilizada por un NEF, AF o NWDAF para descubrir la información de vinculación de una sesión concreta.

El BSF se puede implementar de forma independiente o junto con otras funciones de red, como PCF, UDR, NRF y SMF. Su uso está estrechamente relacionado con la decisión de enrutamiento a tomar, por lo que también puede implementarse junto con el SCP.

En las redes 4G la información de vinculación o *binding* se gestionaba en el *Diameter Routing Agent (DRA)* como elemento centralizado de enrutamiento de la señalización de red, utilizando el protocolo Diameter. En 5G se ha definido una función de red específicamente para esta función (BSF), pero es por esto que puede colocalizarse con el SCP, al ser éste el elemento centralizado de enrutamiento de la señalización de red en 5G utilizando HTTP2. Esta localización conjunta permite replicar una arquitectura similar a la utilizada en 4G.

Por otro lado, puesto que las redes 4G y 5G deben convivir, es necesario que el BSF implemente capacidad de acceso desde la red 4G Diameter para los casos en los que desde la red 4G se quiera acceder a una información de vinculación de sesión que no está presente en el DRA. De este modo, el BSF es el único elemento del núcleo de red 5G que además de implementar HTTP2, debe soportar el interfaz Diameter contra la red 4G para escenarios de interoperabilidad entre las dos redes.

4.1.2.4 SEPP (Security Edge Protection Proxy)

El SEPP es el elemento principal para el roaming tanto a nivel de enrutamiento de interconexión como de seguridad en el límite de la red 5G.

El SEPP, o proxy de protección perimetral de seguridad, es una entidad de red que soporta la siguiente funcionalidad:

- Actúa como un nodo proxy no transparente.
- Protege los mensajes del plano de control de la capa de aplicación entre dos NF pertenecientes a diferentes PLMN que utilizan el interfaz N32 para comunicarse entre sí.
- Realiza la autenticación mutua y la negociación de conjuntos de cifrado con el SEPP en la red remota.
- Se encarga de los aspectos de gestión de claves que impliquen la configuración de las claves criptográficas necesarias para proteger los mensajes en la interfaz N32 entre dos SEPP.
- Realiza el filtrado de mensajes y vigilancia en estas interfaces. El SEPP protege la conexión entre los consumidores de servicios y los productores de servicios desde una perspectiva de seguridad, es decir, el SEPP no duplica la autorización de servicio aplicada por los productores de servicios. Sin embargo, actuará como *firewall* de señalización para bloquear mensajes no permitidos en interfaces de roaming.
- Oculta la topología al limitar la información de topología interna visible para las partes externas (*topology hiding*).
- Como proxy inverso, el SEPP proporciona un único punto de acceso y control a las NF internas.
- El SEPP receptor puede verificar si el SEPP emisor está autorizado a utilizar el identificador de PLMN incluido en el mensaje N32 recibido.
- El SEPP debe poder diferenciar claramente entre los certificados utilizados para la autenticación de SEPP en comunicación directa y los certificados utilizados para la autenticación de intermediarios que realizan modificaciones de mensajes (interconexiones a través de un IPX); por ejemplo mediante la implementación de almacenamientos de certificados separados.
- Descarta los mensajes de señalización N32 con formato incorrecto.
- Implementa funcionalidades de limitación de velocidad (*rate limiting*) para defenderse a sí mismo y a las NF posteriores contra la señalización excesiva de plano de control. Esto incluye mensajes de señalización de SEPP a SEPP.
- Implementa mecanismos contra la suplantación de identidad (*anti-spoofing*) que permiten la validación entre capas de las direcciones e identificadores de origen y destino como por ejemplo FQDN o identificador de PLMN. Un ejemplo de un mecanismo anti-spoofing de este tipo es el siguiente: si hay una discrepancia entre las diferentes capas del mensaje o la dirección de destino no pertenece a la propia PLMN del SEPP, el mensaje se descarta.
- El SEPP puede utilizar uno o más identificadores de PLMN. En el caso de que una PLMN utilice más de un identificador de PLMN, el SEPP de esta PLMN puede utilizar la misma conexión N32 para todos los identificadores de PLMN, con cada uno de los socios remotos (*roaming partners*) de la PLMN. Si diferentes PLMN están representadas por los identificadores de PLMN admitidos

por un SEPP, el SEPP utilizará conexiones N32 separadas para cada pareja de PLMN doméstica y visitada.

El SEPP aplica la funcionalidad anterior a cada mensaje del plano de control en la señalización entre PLMNs, actuando como un servicio de retransmisión entre el productor del servicio real y el consumidor del servicio real. Tanto para el productor como para el consumidor del servicio, el resultado de la retransmisión del servicio es equivalente a una interacción directa del servicio.

Cada mensaje del plano de control en la señalización entre PLMNs enviado desde el vSEPP hasta el hSEPP, puede pasar a través de entidades intermedias *IP Exchange Service (IPX)*, como se muestra en la figura 4.7 [67].

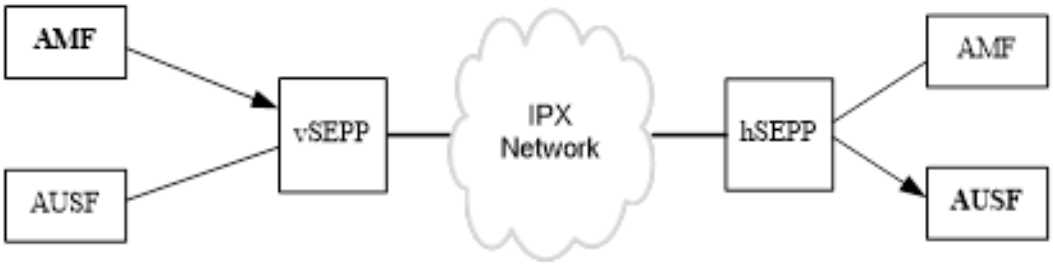

Figura 4.7: Mensaje de señalización desde AMF (vPLMN) hacia AUSF (hPLMN) atravesando los respectivos SEPPs.

El interfaz N32 se utiliza entre los SEPPs de una VPLMN y una HPLMN en escenarios de roaming. El SEPP que está en el lado del consumidor del servicio NF se denomina cSEPP y el SEPP que está en el productor del servicio NF se denomina pSEPP. El interfaz N32 se puede considerar a nivel lógico como 2 interfaces separadas tal cual se indica a continuación.

4.1.2.4.1 Interfaz N32-c N32-c es el interfaz de plano de control entre los SEPPs, que implementa el procedimiento de reconocimiento inicial (*handshake*) entre el vSEPP y el hSEPP, donde uno actúa como SEPP de inicio y el otro como SEPP de respuesta. En este procedimiento, se negocian los parámetros que se aplicarán para el reenvío real de las peticiones de servicio N32 tanto a nivel de enrutamiento como de seguridad [68].

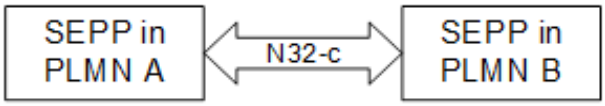

Figura 4.8: Interfaz N32-c.

4.1.2.4.2 Interfaz N32-f N32-f es el interfaz de reenvío de peticiones de servicio entre los SEPPs, relativas a la comunicación entre el consumidor del servicio NF y el productor del servicio NF entre diferentes PLMNs. Se establece después de aplicar la protección de seguridad acordada en el procedimiento de reconocimiento inicial efectuado sobre el interfaz N32-c[68].

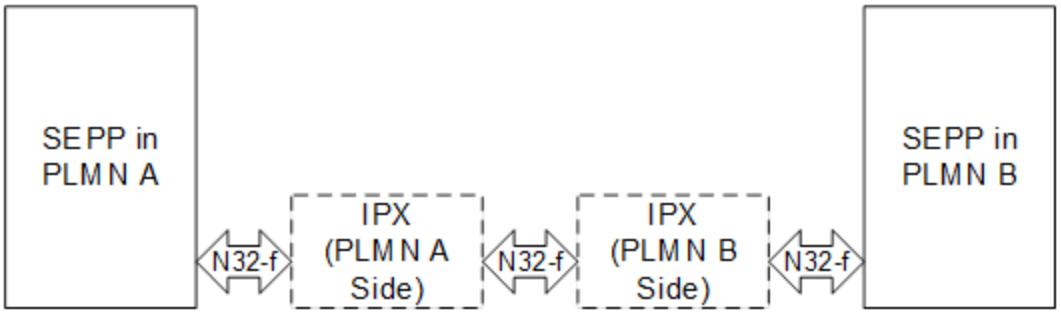

Figura 4.9: Interfaz N32-f.

Los SEPPs negocian como parte del reconocimiento inicial (*handshake*), cuál es el protocolo de seguridad a utilizar. SEPP soporta dos opciones: TLS y PRINS. La funcionalidad de protección de seguridad de la capa de aplicación *Application Layer Security (ALS)* del N32-f se utiliza solo si el protocolo para la seguridad de interconexión N32 (PRINS) se negocia entre los SEPPs.

El interfaz N32-f proporciona las siguientes funcionalidades de protección de seguridad de la capa de aplicación:

- Protección de mensajes sobre la información intercambiada entre el consumidor del servicio NF y el productor del servicio NF entre diferentes PLMNs mediante la aplicación de mecanismos de seguridad de la capa de aplicación.
- Reenvío del mensaje protegido de la capa de aplicación desde un SEPP en una PLMN a un SEPP en otra PLMN. Dicho reenvío puede involucrar proveedores IPX en la ruta.
- Si los proveedores IPX están en la ruta de SEPP en PLMN A a SEPP en PLMN B, el reenvío en la interfaz N32-f puede implicar la inserción de instrucciones de modificación de contenido que el SEPP receptor aplica después de verificar la integridad de dichas instrucciones de modificación.

Si TLS es la política de seguridad negociada entre los SEPPs, entonces el interfaz N32-f implicará solo el reenvío de los mensajes HTTP/2 de los productores de servicios de NF y los consumidores de servicios de NF sin reformatear al transitar por los diferentes SEPPs de PLMN y opcionalmente de los IPX.

4.1.2.5 Otras NFs

4.1.2.5.1 AMF La función de Access and Mobility Management Function (AMF), o Gestión de Acceso y Movilidad, incluye la siguiente funcionalidad:

- Terminación de interfaz de acceso *Radio Access Network (RAN)* en el plano de control *Control Plane (CP)* del interfaz N2.
- Terminación de *Non Access Stratum (NAS)* (interfaz N1), cifrado de NAS y protección de integridad.
- Gestión de registros.
- Gestión de conexiones.
- Gestión de la accesibilidad.
- Gestión de la movilidad.

- Interceptación legal (para eventos de AMF y para implementar el interfaz con el sistema de *Lawful Interception (LI)*.
- Proporcionar transporte para mensajes cortos SMS entre UE y SMF.
- Proxy transparente para enrutar mensajes SMS.
- Autenticación de acceso.
- Autorización de acceso.
- Proporcionar transporte para mensajes SMS entre UE y SMSF.
- Funcionalidad de anclaje de seguridad *Security Anchor Function (SEAF)*.
- Gestión de servicios de localización *LoCation Services (LCS)* para servicios regulatorios.
- Proporcionar transporte para mensajes de servicios de ubicación entre UE y *Location Management Function (LMF)*, así como entre RAN y LMF.
- Asignación de identificador de portador de *Evolved Packet System (EPS)* para interfuncionamiento con el EPS de 4G.
- Notificación de eventos de movilidad UE.
- Aprovisionamiento de parámetros externos (parámetros de comportamiento esperado de UE o parámetros de configuración de red).
- Compatibilidad con autenticación y autorización específicas de segmento (*slice*) de red.
- Recopilación de datos de carga y soporte de la interfaz de carga.
- Opcionalmente, funcionalidad adicional para admitir redes de acceso que no sean 3GPP.

AMF también puede incluir la siguiente funcionalidad específica para escenarios de roaming:

- Normalización de informes de acuerdo con los acuerdos de roaming entre VPLMN y HPLMN (por ejemplo, cambiar la granularidad de ubicación en un informe desde el nivel de celda a un nivel que sea apropiado para la HPLMN).
- Generación de información de cobro/contabilidad para reportes que son enviados a la HPLMN.

4.1.2.5.2 SMF El *Session Management Function (SMF)*, o función de gestión de sesión, incluye la siguiente funcionalidad:

- Gestión de sesiones: Establecimiento, modificación y liberación de sesión, incluido el mantenimiento del túnel entre el UPF y el nodo de acceso AN.
- Asignación y gestión de direcciones IP de UE (incluida la autorización opcional). La dirección IP del UE puede recibirse desde un UPF o desde una red de datos externa.
- Funciones *Dynamic Host Configuration Protocol (DHCP)* tanto DHCPv4 como DHCPv6 para servidor y cliente.

- Funcionalidad para responder a solicitudes de *Address Resolution Protocol (ARP)* y/o solicitudes de vecino IPv6 basadas en información de caché local para las PDU de Ethernet. El SMF responde al ARP y/o a la solicitud de vecino IPv6 proporcionando la dirección *MAC* correspondiente a la dirección IP enviada en la solicitud.
- Selección y control de la función UPF para plano de usuario.
- Configuración de la dirección del tráfico en el UPF para enrutar el tráfico al destino adecuado.
- Terminación de interfaces hacia funciones de control de políticas.
- Interceptación legal (para eventos de SMF y para implementar el interfaz con el Sistema LI).
- Recopilación de datos de carga y soporte de interfaces de carga.
- Control y coordinación de la recogida de datos de tarificación en el UPF.
- Notificación de datos de enlace descendente.
- Determinar el modo *Session and Service Continuity (SSC)* de una sesión.
- Soporte de compresión de cabecera.
- Aprovisionamiento de parámetros externos (parámetros de comportamiento esperado de UE o parámetros de configuración de red).
- Actuar como V-SMF con las siguientes funcionalidades de roaming:
  - Manejar la aplicación local para aplicar *Quality of Service (QoS)* y *Service Level Agreement (SLA)* (SMF en la VPLMN).
  - Interfaz de carga y recopilación de datos de carga (SMF en la VPLMN).
  - Interceptación legal (SMF en la VPLMN para eventos SMF e implementación del interfaz al Sistema LI).
- Soporte para interacción con redes de datos DN *Data Networks* externas para transporte de señalización para autenticación/autorización de sesión de PDU por DN externo.
- Instruye a UPF y NG-RAN para realizar transmisión redundante en interfaces N3/N9.

Además de las funcionalidades del SMF descritas anteriormente, el SMF también puede incluir la siguiente funcionalidad específica para escenarios de roaming:

- Normalización de reportes según acuerdos de roaming entre VPLMN y HPLMN.
- Generación de información de cobro/contabilidad para reportes que son enviados a la HPLMN.

4.1.2.5.3 UPF El User Plane Function (UPF), o función de plano de usuario, incluye la siguiente funcionalidad:

- Punto de anclaje para movilidad Intra-RAT e Inter-RAT (cuando corresponda).
- Asignación de dirección o prefijo IP de UE en respuesta a la solicitud de SMF.
- Punto de sesión de PDU externo de interconexión a la red de datos DN.

- Enrutamiento y reenvío de paquetes (por ejemplo, compatibilidad con el clasificador de enlace ascendente para enrutar los flujos de tráfico a una instancia de una red de datos, o compatibilidad con el punto de ramificación para admitir sesiones de PDU de alojamiento múltiple).
- Inspección de paquetes. Por ejemplo, detección de aplicaciones basada en la plantilla de flujo de datos de servicio y los *Packet Flow Detection (PFD)* opcionales recibidos del SMF.
- Parte de plano de usuario de la aplicación de reglas de políticas, por ejemplo, redirección de tráfico.
- Interceptación legal.
- Informes de uso de tráfico.
- Manejo de QoS para el plano de usuario, por ejemplo cumplimiento de la tasa de UL/DL, o marcado de QoS en DL.
- Verificación de tráfico de enlace ascendente.
- Marcado de paquetes a nivel de transporte en los enlaces ascendente y descendente.
- Almacenamiento en *buffer* de paquetes de enlace descendente y activación de notificación de datos de enlace descendente.
- Funcionalidad para responder a solicitudes de ARP y/o solicitudes de vecino IPv6 basadas en información de caché local para las PDU de Ethernet. La UPF responde al ARP y/o a la solicitud de vecino IPv6 proporcionando la dirección MAC correspondiente a la dirección IP enviada en la solicitud.
- Duplicación de paquetes en sentido descendente y eliminación en sentido ascendente en la capa GTP-U.
- Funcionalidad *Inter PLMN UP Security (IPUPS)* para aportar seguridad en el plano de usuario.

4.1.2.5.4 PCF El PCF, o función de control de políticas (PCF) incluye la siguiente funcionalidad:

- Soporta un marco de políticas unificado para gobernar el comportamiento de la red.
- Proporciona reglas de política a la(s) función(es) del plano de control para hacerlas cumplir.
- Accede a la información de suscripción relevante para las decisiones de política en un repositorio de datos unificado (UDR).

4.1.2.5.5 NEF El NEF, o función de exposición de red, soporta la siguiente funcionalidad independiente:

- Exposición de capacidades y eventos. Las capacidades y los eventos de NF pueden ser expuestos de forma segura por NEF, por ejemplo a funciones de aplicación de terceros.
- Provisión segura de información desde una aplicación externa a la red 3GPP. Proporciona un medio para que las funciones de la aplicación proporcionen información de forma segura a la red 3GPP, por ejemplo comportamiento esperado de UE, o información específica del servicio. En ese caso, el NEF puede autenticar, autorizar y ayudar a limitar las funciones de la aplicación.

- Traducción de información interna-externa. Se traduce entre la información intercambiada con el AF y la información intercambiada con la función de red interna. En particular, NEF maneja el enmascaramiento de la información confidencial de la red y del usuario para los AF externos de acuerdo con la política de la red.
- NEF recibe información de otras funciones de la red (en función de las capacidades expuestas por otras funciones de la red). NEF almacena la información recibida como datos estructurados utilizando una interfaz estandarizada para un repositorio de datos unificado (UDR).
- Exposición de analíticas. Los análisis NWDAF pueden ser expuestos de forma segura por NEF para terceros.
- Recuperación de datos de una parte externa por NWDAF. Los datos proporcionados por la parte externa pueden ser recopilados por NWDAF a través de NEF para fines de generación de análisis. NEF maneja y reenvía solicitudes y notificaciones entre NWDAF y AF.
- Recopilación de datos de carga y soporte de interfaces de carga.

Para exposición externa de servicios relacionados con UE(s) específicos, el NEF reside en la HPLMN. Dependiendo de los acuerdos del operador, el NEF en la HPLMN puede tener interfaz(es) con NF(s) en la VPLMN.

4.1.2.5.6 UDM El Unified Data Management (UDM), o gestión unificada de datos, incluye soporte para la siguiente funcionalidad:

- Generación de credenciales de autenticación 3GPP *Authentication and Key Agreement (AKA).*
- Manejo de identificación de usuario. Por ejemplo, almacenamiento y gestión de *Subscription Permanent Identifier (SUPI)* para cada suscriptor en el sistema 5G.
- Soporte de desocultación del identificador de suscripción protegido por privacidad *Subscription Concealed Identifier (SUCI).*
- Autorización de acceso basada en datos de suscripción. Por ejemplo, restricciones de roaming.
- Gestión de registro de NF de servicio del UE. Por ejemplo, almacenamiento del AMF de servicio para el UE y almacenamiento del SMF de servicio para la sesión de PDU del UE.
- Soporte para la continuidad del servicio/sesión. Por ejemplo manteniendo la asignación SMF/DNN (*Data Network Name*) de las sesiones en curso.
- Soporte de entrega MT-SMS.
- Funcionalidad de interceptación legal. Especialmente en caso de roaming saliente (de mis abonados en otras PLMNs), donde UDM es el único punto de contacto para LI.
- Gestión de suscripciones.
- Gestión de SMS.
- Compatibilidad con el aprovisionamiento de parámetros externos (parámetros de comportamiento de UE esperado o parámetros de configuración de red). Para proporcionar esta funcionalidad, el UDM utiliza datos de suscripción (incluidos los datos de autenticación) que pueden almacenarse en UDR, en cuyo caso un UDM implementa la lógica de la aplicación y no requiere un almacenamiento

de datos de usuario interno. En este caso, varios UDM diferentes pueden servir al mismo usuario en diferentes transacciones.

El UDM se ubica en la HPLMN de los abonados que atiende, y accede a la información de la UDR ubicada en la misma PLMN.

4.1.2.5.7 AUSF El *Authentication Server Function (AUSF)*, o función del servidor de autenticación, soporta la siguiente funcionalidad:

- Autenticación para acceso 3GPP y para acceso no confiable que no es 3GPP.

4.1.2.5.8 N3IWF El *Non-3GPP InterWorking Function (N3IWF)* utilizado para el acceso no confiable que no es 3GPP incluye la siguiente funcionalidad:

- Compatibilidad con el establecimiento de túneles IPsec con el UE. N3IWF finaliza los protocolos IKEv2/IPsec con el UE a través del interfaz NWu y retransmite a través del interfaz N2 la información necesaria para autenticar al UE y autorizar su acceso a la red central 5G.
- Terminación de las interfaces N2 y N3 hacia 5G Core Network para el plano de control y el plano de usuario, respectivamente.
- Retransmisión de señalización de NAS (N1) de plano de control de enlace ascendente y enlace descendente entre el UE y AMF.
- Manejo de señalización N2 desde SMF (retransmitida por AMF) relacionada con sesiones de PDU y QoS.
- Establecimiento de *IPsec Security Association (IPsec SA)* para admitir el tráfico de sesión de PDU.
- Retransmisión de paquetes de plano de usuario de enlace ascendente y enlace descendente entre el UE y el UPF.
- Hacer cumplir la QoS correspondiente al marcado de paquetes N3, teniendo en cuenta los requisitos de QoS asociados a dicho marcado recibido a través de N2.
- Marcado de paquetes en el plano de usuario N3 en el enlace ascendente.
- Anclaje de movilidad local dentro de redes de acceso no 3GPP que no son de confianza.
- Apoyo a la selección de AMF.

4.1.2.5.9 AF El *AF*, o función de aplicación, interactúa con la red central 3GPP para proporcionar estos servicios:

- Influencia de la aplicación en el encaminamiento del tráfico.
- Acceder a la función de exposición de la red (NEF).
- Interactuar con el marco de políticas para el control de políticas.
- Interacciones desde la red IMS hacia el 5GC.

Según la implementación del operador, las funciones de aplicación que el operador considera confiables pueden interactuar directamente con las funciones de red relevantes. Las funciones de aplicación a las que el operador no permita un acceso directo a las funciones de red, deberán utilizar el marco de exposición externo a través del NEF para interactuar con las funciones de red pertinentes.

4.1.2.5.10 UDR El *Unified Data Repository (UDR)*, o repositorio de datos unificado, soporta la siguiente funcionalidad:

- Almacenamiento y recuperación de datos de suscripción por parte del UDM.
- Almacenamiento y recuperación de datos de pólizas por parte del PCF.
- Almacenamiento y recuperación de datos estructurados para su exposición.
- Datos de la aplicación, incluidas las descripciones de flujo de paquetes PFD *Packet Flow Detection* para la detección de aplicaciones, o información de solicitud de AF para varios UE.
- Almacenamiento y recuperación de identificadores de grupo de NF correspondientes al identificador del suscriptor. Ppor ejemplo, *IMS Private ID (IMPI)*, *IMS Public User Identity (IMPU)*, *SUPI*.

El repositorio de datos unificado se encuentra en la misma PLMN que los consumidores de servicios de NF que almacenan y recuperan datos de él mediante el interfaz SBI Nudr. Nudr es una interfaz intra-PLMN. Las implementaciones pueden optar por desplegar UDR y UDSF conjuntamente.

4.1.2.5.11 UDSF El *Unstructured Data Storage Function (UDSF)* es una función opcional que soporta la siguiente funcionalidad:

- Almacenamiento y recuperación de información como datos no estructurados por cualquier NF.

Los datos estructurados en esta especificación se refieren a datos para los cuales la estructura se define en las especificaciones 3GPP. Los datos no estructurados se refieren a datos para los cuales la estructura no está definida en las especificaciones 3GPP y están abiertas a implementaciones privadas diferentes por parte de suministradores de equipamiento de red y/o operadoras. Las implementaciones pueden optar por desplegar UDSF y UDR conjuntamente.

4.1.2.5.12 SMSF El *Short Message Service Function (SMSF)* soporta la siguiente funcionalidad para admitir SMS a través de NAS:

- Comprobación de datos de suscripción de gestión de SMS y realización de la entrega de SMS en consecuencia.
- Gestión de *Short Message - Relay Protocol (SM-RP)* y *Short Message - Control Protocol (SM-CP)* con el UE.
- Retransmitir el SM desde UE hacia el elemento del núcleo de red correspondiente: *Short Message Service - Gateway Mobile Switching Centre (SMS-GMSC)*, *Inter Working Mobile Switching Centre (IWGMSC)* o *SMS-Router*.
- Retransmitir el SM desde el elemento del núcleo de red correspondiente (*SMS-GMSC/IWMSC/SMS-Router*) hacia el UE.
- Generación del *Call Detailed Record (CDR)* relacionado con el SMS.
- Interceptación legal.
- Interacción con AMF y SMS-GMSC para el procedimiento de notificación en el caso de que el UE no esté disponible para la transferencia de SMS. SMSF notifica a SMS-GMSC para informar a UDM cuando el UE no está disponible para el servicio SMS.

4.1.2.5.13 NSSF El *Network Slice Selection Function (NSSF)*, o función de selección de segmento (*slice*) de red soporta la siguiente funcionalidad:

- Seleccionar el conjunto de instancias de segmentos de red (*Network Slices*) que sirven al UE.
- Determinar el *Network Slice Selection Assistance Information (NSSAI)* permitido y, si es necesario, el mapeo a los *Single - Network Slice Selection Assistance Information (S-NSSAI)* suscritos.
- Determinar el NSSAI configurado y, si es necesario, el mapeo a los S-NSSAI suscritos.
- Determinar el conjunto de AMF que se usará para servir al UE o, en base a la configuración, una lista de AMF(s) candidatos, posiblemente consultando al NRF.

4.1.2.5.14 5G-EIR El *5G-Equipment Identity Register (5G-EIR)* es una función de red opcional que soporta la siguiente funcionalidad:

- Comprobar el estado de *Permanent Equipment Identifier (PEI)* para comprobar que no ha sido prohibido.

4.1.2.5.15 NWDAF El *NWDAF* representa la función lógica de análisis de red gestionada por el operador. El NWDAF incluye la siguiente funcionalidad:

- Recopilación de datos de NF y AF.
- Recopilación de datos de *Operations, Administration and Maintenance (OAM).*
- Registro de servicios NWDAF y exposición de metadatos a NFs/AFs.
- Aprovisionamiento de información analítica para NFs y AF.

3GPP define el soporte de NWDAF y el interfaz SBI Nnwdaf. La funcionalidad específica que ofrezca el NWDAF es decisión de cada operador o proveedor, no está definida por los estándares 3GPP.

4.1.2.5.16 NSSAAF El *Network Slice Specific Authentication and Authorization Function (NSSAAF)*, o función de autenticación y autorización específica de segmento de red, soporta la siguiente funcionalidad:

- Compatibilidad con un servidor *Authentication, Authorization and Accounting (AAA)*, AAA-S. Si el AAA-S pertenece a un tercero, el NSSAAF puede comunicarse con el AAA-S a través de un proxy AAA (AAA-P).

4.2 Itinerancia o Roaming en 5G

La arquitectura del sistema 5G se adapta a las particularidades de los escenarios de roaming, en los que 2 redes 5G diferentes deben interactuar. Existen dos modelos en función de las tareas que la red doméstica o red *Home* (HPLMN) permita que sean controladas por la red visitada (VPLMN):

- **HR (*Home Routed*)**: La red doméstica HPLMN mantiene el control de las sesiones de usuario en la red visitada. Este es el modelo que es utilizado en la generaciones anteriores 2G/3G/4G por la necesidad de los operadores de tener control total sobre sus abonados. Como parte negativa, este tráfico debe transitar entre ambas redes, aumentando latencias en las comunicaciones. Esto puede ser un problema para múltiples casos de uso de la tecnología 5G donde la ultra-baja latencia es requisito indispensable.

- **LBO (*Local Breakout*)**: En este escenario se da mayor control a la red visitada VPLMN para gestionar al *roamer* recibido. La red doméstica tiene menor control del usuario y el tráfico se gestiona localmente en la red visitada sin necesidad de transitar internacionalmente hasta la red *home*. Puede ser imprescindible para casos de uso donde la ultra-baja latencia sea requisito obligatorio. Sin embargo, este modelo también fue propuesto en generaciones anteriores como 4G, pero nunca fue llevado a la práctica.

La figura 4.10 muestra la arquitectura de referencia del sistema 5G en el caso de roaming o itinerancia en modo HR (*Home Routed*) utilizando la representación basada en SBI (Interfaces basadas en Servicios), en la que se muestran los servicios ofrecidos por cada NF [57].

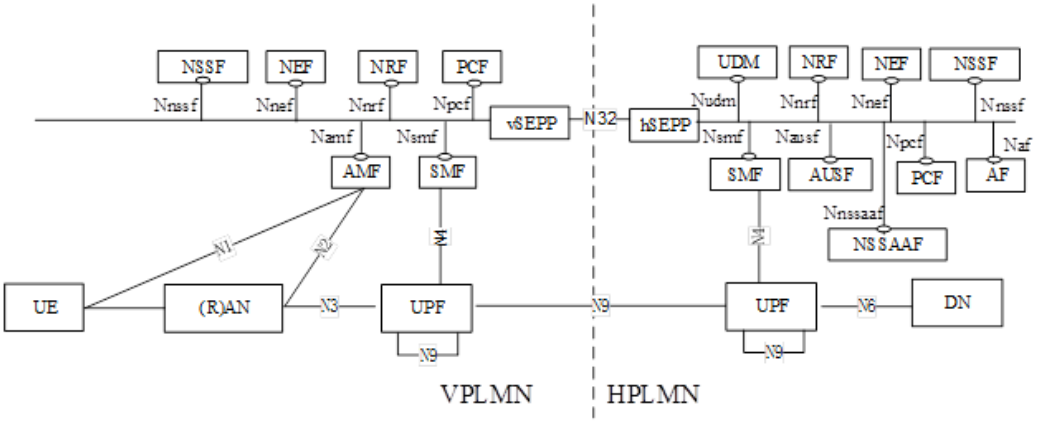

Figura 4.10: Arquitectura del sistema 5G Roaming HR en representación de SBI.

Se observa que en el escenario *Home Routed* el tráfico de datos de usuario transita entre red VPLMN y red HPLMN por el interfaz N9 entre vUPF y hUPF.

La figura 4.11 muestra la arquitectura de referencia del sistema 5G en el caso de roaming o itinerancia en modo HR (*Home Routed*) utilizando la representación basada en puntos de referencia, en la que se muestran las interacciones entre las diferentes funciones de red [57].

La figura 4.12 muestra la arquitectura de referencia del sistema 5G en el caso de roaming o itinerancia en modo LBO (*Local Breakout*) utilizando la representación basada en SBI (Interfaces basadas en Servicios), en la que se muestran los servicios ofrecidos por cada NF [57].

En la arquitectura LBO, el PCF de la red VPLMN puede interactuar con el AF para generar reglas PCC (*Policy and Charging Control*), para los servicios que sean ofrecidos desde la red VPLMN. El PCF en la red visitada utiliza políticas configuradas localmente en función del acuerdo de roaming establecido con el operador HPLMN como base para la generación de reglas PCC. El PCF en la red VPLMN no tiene acceso a la información de políticas de abonado definidas en la HPLMN.

El SCP puede ser utilizado para la comunicación indirecta entre NFs, aunque no se muestra en la arquitectura de roaming por simplicidad.

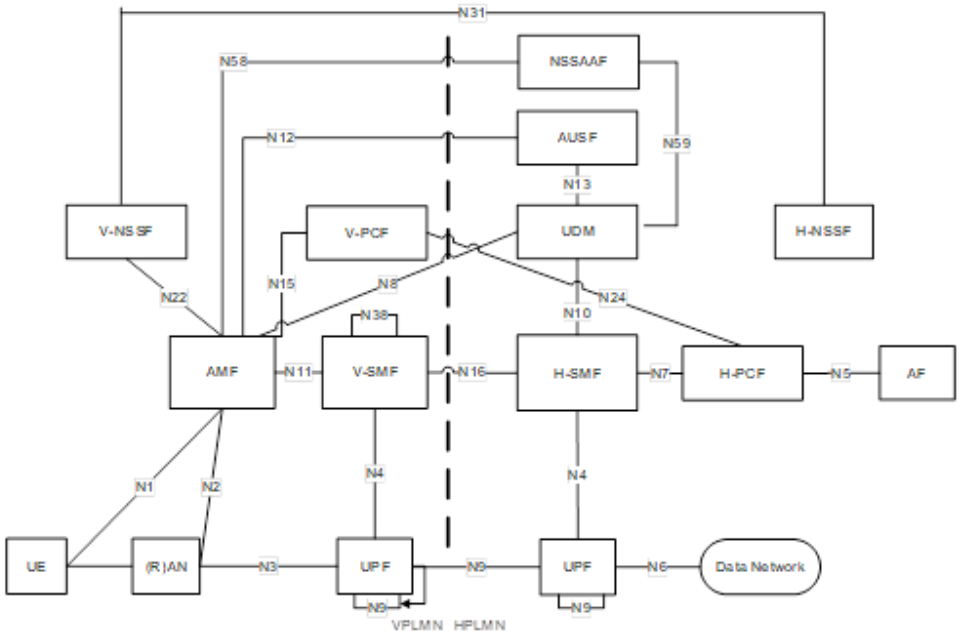

Figura 4.11: Arquitectura del sistema 5G Roaming HR en representación basada en puntos de referencia.

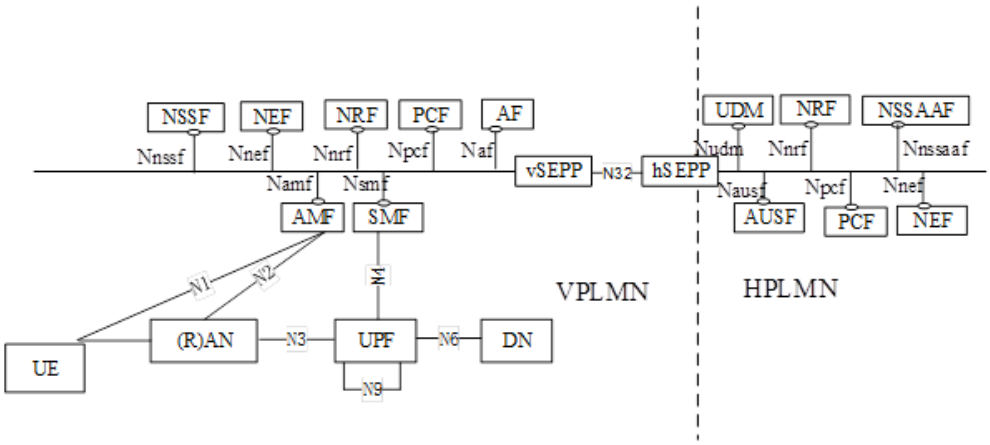

Figura 4.12: Arquitectura del sistema 5G Roaming LBO en representación de SBI.

La figura 4.13 muestra la arquitectura de referencia del sistema 5G en el caso de roaming o itinerancia en modo LBO (*Local Breakout*) utilizando la representación basada en puntos de referencia, en la que se muestran las interacciones entre las diferentes funciones de red [57].

El NRF y el SEPP no se muestran en la arquitectura basada en puntos de referencia por simplicidad.

Tanto en el escenario de LBO como en HR, existe un interfaz entre vNRF y hNRF con el objetivo de securizar las conexiones y ocultar la topología interna de la red en las interfaces inter-PLMN. Esto se muestra en la figura 4.14 [57].

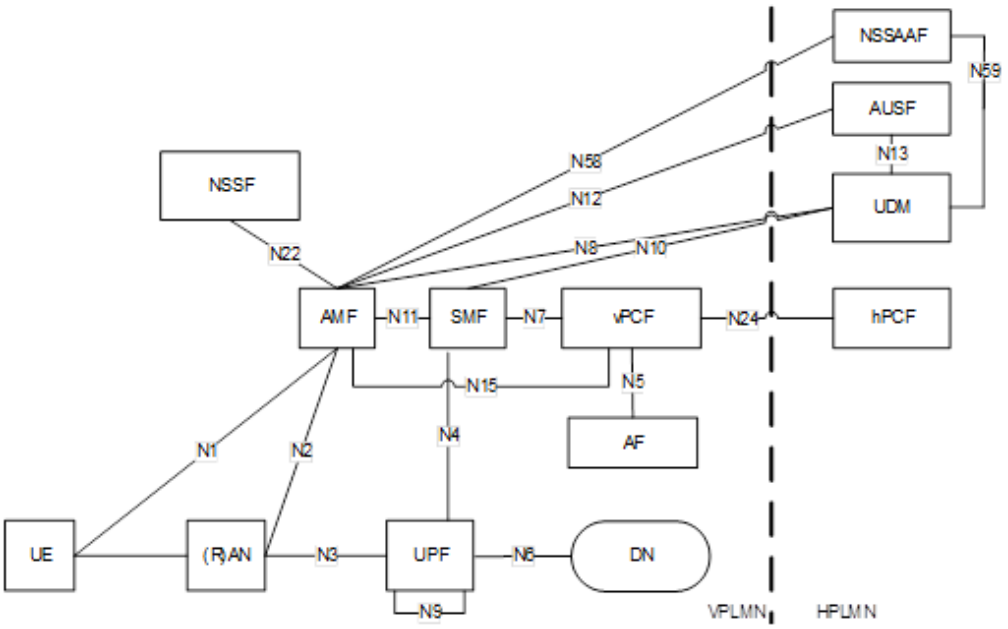

Figura 4.13: Arquitectura del sistema 5G Roaming LBO en representación basada en puntos de referencia.

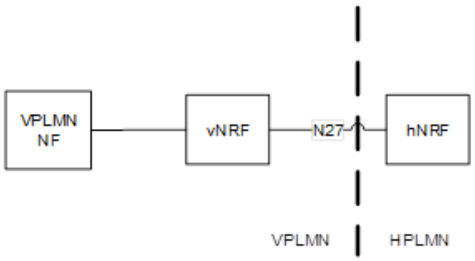

Figura 4.14: Arquitectura del sistema 5G Roaming para NRF en representación basada en puntos de referencia.

Como en los casos anteriores, los SEPPs a ambos lados de la frontera entre PLMNs no se muestran en la figura para mayor claridad.

4.3 Conclusiones

Este capítulo aporta los conceptos fundamentales necesarios para entender el núcleo de red del sistema 5G y específicamente el funcionamiento de su plano de control. Nos hemos centrado en las funciones que gestionan específicamente el enrutamiento y la seguridad de la señalización HTTP2; y profundizaremos en su funcionamiento y implementación en el siguiente capítulo.

Capítulo 5

Resultados

Averiguo lo que el mundo necesita. Luego, voy y lo invento.

I find out what the world needs. Then, I go ahead and invent it.

Thomas Edison [1]

5.1 Introducción

Una vez analizada la arquitectura del núcleo de red 5G SA y sus diferentes funciones y entidades de red, especialmente las relativas al plano de control de señalización; en este capítulo ofreceremos los resultados de este estudio.

Analizaremos el cambio que supone la utilización del protocolo HTTP2 con respecto a los protocolos de señalización de generaciones anteriores (SS7 y Diameter), analizando cómo es la implementación real del enrutamiento y seguridad de la señalización en el núcleo de red 5G.

Exploraremos las nuevas capacidades del 5GC en cuento a seguridad por diseño (*security by design*) y segmentación de red (*network slicing*). Describiremos el salto tecnológico relacionado con la implementación de funciones basadas en microservicios e infraestructura en la nube, como paso adicional a la virtualización de funciones de red que ya fue introducida con la tecnología 4G.

Expondremos en detalle la casuística especial de los escenarios de roaming en 5G SA, así como los temas pendientes de estandarización y consenso entre la industria del sector; y estudiaremos los casos de uso principales en cuanto a señalización extremo a extremo en la red 5G.

5.2 Comparativa de la señalización 5G con generaciones anteriores

El núcleo de la red 5G incorpora el uso del protocolo de señalización HTTP2, a diferencia del protocolo Diameter utilizado en 4G EPC, y del protocolo SS7 utilizado en las tecnologías 2G y 3G.

[1] Thomas Edison (1847 - 1931), inventor, científico y empresario norteamericano. *American Greats (2000). Robert A. Wilson y Stanley Marcus, p. 70* [69].

En telecomunicaciones, la señalización hace referencia al uso de señales para controlar las comunicaciones. Hace referencia al establecimiento, control y gestión de circuitos de telecomunicaciones y de la red. Los sistemas de señalización se han categorizado tradicionalmente en base a diferentes características.

Señalización en banda y señalización fuera de banda:

- **Sistemas de señalización en banda**: En los que la información de control de la llamada se transmite por el mismo canal físico, o en la misma banda de frecuencia que la llamada telefónica. Un ejemplo fuera del sector de la telefonía sería el protocolo *Simple Mail Transfer Protocol (SMTP)* de correo en el que los mensajes de control son enviados en el mismo flujo que el contenido del mensaje. En telefonía, la señalización en banda fue utilizada por el sistema de señalización Multi-Frecuencia (MF) introducido por Bell en Estados Unidos después de la Segunda Guerra Mundial. Éste fue posteriormente estandarizado por *International Telegraph and Telephone Consultative Committee (CCITT)* (ahora ITU-T) como R1 (*Regional System No. 1*) y su versión europea como R2 (*Regional System No. 2*) durante los años 1960. La señalización en banda también fue utilizada por el sistema SS5 (*Signaling System No. 5*) en los años 1970 para llamadas internacionales de larga distancia IDDD (*International Direct Distance Dialing*). Estos sistemas son precursores tecnológicos de los tonos *Dual-Tone Multi-Frequency Signaling (DTMF)* introducidos en 1963, que utilizan el mismo principio fundamental pero se usan principalmente para señalizar información de direccionamiento y de control desde el dispositivo de usuario hacia la red utilizando un total de ocho frecuencias.
- **Sistemas de señalización fuera de banda**: En los que existe un canal separado dedicado a señalización, diferente del utilizado para la llamada telefónica en sí. En telefonía, la señalización fuera de banda ha sido utilizada desde la introducción de SS6 (*Signaling System No. 6*) en los años 1970, como uno de los primeros métodos de señalización fuera de banda (y por canal común) para troncales de comunicaciones entre centros de conmutación internacionales. Posteriormente en 1980 con SS7 (*Signaling System No. 7*) se introdujo el estándar de señalización más ampliamente utilizado en las redes telefónicas mundiales, reemplazando a los sistemas de señalización en banda predecesores.

Además, en telecomunicaciones también se distingue entre señalización por canal asociado y señalización por canal común.

- **Señalización por canal asociado**: Utiliza un canal de señalización dedicado a cada canal de portadora.
- **Señalización por canal común**: Utiliza un canal de señalización que aglutina la información de señalización relativa a múltiples canales de portadora que, por tanto, tienen un canal de señalización común.

En el núcleo de las redes móviles, tanto en 5G como en generaciones anteriores, se utiliza señalización fuera de banda por canal común.

Por último, también se puede distinguir entre señalización de línea y señalización de registro:

- **Señalización de línea**: Hace referencia a la información relacionada con el estado de la línea o el canal; como supervisión, tono de línea o tono de espera.
- **Señalización de registro**: La relacionada con información de direccionamiento; como el número llamado y llamante, o el registro de un suscriptor en la red. En los comienzos de la telefonía, esto era llevado a cabo por personas.

En este trabajo, nos centraremos fundamentalmente en señalización de registro y de control del núcleo de la red 5G.

5.2.1 2G, 3G y el protocolo SS7

El comienzo de las redes de telefonía móvil digital o 2G alrededor de 1990 con GSM, llegó con SS7 como protocolo de señalización en el núcleo de red. Este protocolo llevaba siendo utilizado casi una década en sistemas de telefonía fija, desde su estandarización por parte de ITU-T en 1981 en la serie de recomendaciones Q.7XX.

El estándar SS7 define el protocolo y los procedimientos mediante los cuales los elementos de la red de telefonía conmutada pública (la PSTN) intercambian información sobre una red digital para efectuar el enrutamiento, establecimiento y control de llamadas. La definición de ITU para SS7 permite variantes nacionales tales como *American National Standards Institute (ANSI)* usado en Norteamérica y *European Telecommunications Standards Institute (ETSI)* ITU usado en Europa [70].

SS7 provee una estructura universal para señalización de redes de telefonía, mensajería, interconexión y mantenimiento de redes. Se ocupa del establecimiento de una llamada, intercambio de información de usuario, enrutamiento de llamada, estructuras de abonado diferentes y soporta servicios de Redes Inteligentes o IN (*Intelligent Networks*).

El modelo de capas de SS7 responde al modelo de capas propuesto por OSI de ISO, en el cual, en sus niveles superiores se puede distinguir claramente el empleo de protocolos de nivel de aplicación específicos de la red móvil en sus vertientes GSM, GRPS y UMTS (MAP, CAP, INAP) y para la parte de control de llamadas de telefonía digital conmutada (ISUP); como se muestra en la figura 5.1.

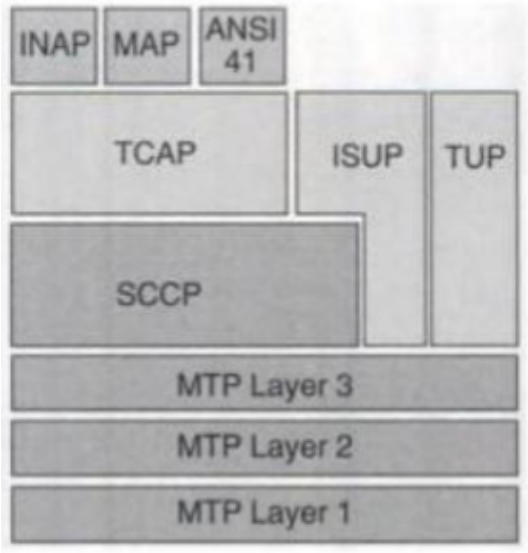

Figura 5.1: Pila de protocolos SS7.

SS7 utiliza el concepto de SP *Signaling Point* (Punto de señalización), que constituye un nodo de red con capacidad de gestión de mensajes de control, y que son identificados por un PC *Point Code* (Código de Punto).

SS7 utiliza diversos protocolos. En concreto, para redes de telefonía móvil GSM, GPRS y UMTS se usan:

- **MTP1 (*Message Transfer Part Level 1*)**: El nivel físico de la comunicación. Las interfaces físicas definidas incluyen E-1 en Europa (2048 Mbps, con 32 canales o *timeslots* de 64 Kbps) o T-1 en América (1544 Mbps, con 24 canales de 64 Kbps), como sistemas de portadora desarrollados para la transmisión digital de múltiples llamadas telefónicas simultáneas con multiplexación por división en el tiempo.

- **MTP2 (*Message Transfer Part Level 2*)**: Cumple las tareas del nivel de enlace, asegurando la transmisión extremo a extremo del mensaje y ofreciendo control de flujo, detección de errores y chequeo de secuencia. También retransmite mensajes no confirmados.

- **MTP3 (*Message Transfer Part Level 3*)**: Posee una dirección de punto de acceso (*Point Code*) que permite el enrutamiento de mensajes de señalización hasta un *endpoint* o punto final. La unidad de mensaje MTP transmitida es denominada MSU (*Message Signal Unit*).

- **SCCP (*Signaling Connection Control Part*)**: es un protocolo de capa de red que ofrece capacidades extendidas de enrutamiento y control de flujo. Aunque MTP ofrece capacidades de enrutamiento basado en *Point Code*, SCCP permite enrutamiento basado en *Point Code* y Número de Subsistema (*Subsystem Number*); o *Global Title* como dirección que identifica un destino concreto y tiene estructura similar a un número de teléfono. La unidad de mensaje transmitida es denominada UDT (*Unit Data*).

- **TCAP (*Transaction Capabilities Application Part*)**: Facilita el uso de múltiples diálogos concurrentes entre los mismos subsistemas de los mismo elementos de red, utilizando diferentes identificadores de transacción (*Transaction IDs*) para identificarlos; de un modo similar a como el protocolo TCP utiliza puertos para multiplexar conexiones entre las mismas direcciones IP.

- **MAP (*Mobile Application Part*)**: Aporta la capa de aplicación en la redes GSM, GPRS y UMTS para ofrecer servicios a los usuarios. Esto incluye servicios de movilidad (gestión de la localización o autenticación), servicios de manejo de llamadas (enrutamiento, chequeo de disponibilidad del llamado o gestión de llamadas en roaming), servicios suplementarios, SMS, etc.

- **ISUP (*ISDN -Integrated Services Digital Network- User Part*)**: Se encarga de los mensajes de señalización de llamadas en las redes públicas de telefonía conmutada (PSTN), compatible con ISDN o RDSI.

- **INAP (*Intelligent Network Application Part*)**: Protocolo de señalización de aplicación utilizado en la arquitectura de red inteligente IN (*Intelligent Networks*) para servicios de valor añadido como prepago, mensajería unificada o control de fraude; pensados inicialmente para su uso en telefonía fija.

- **CAP (*CAMEL Application Part*)**: Protocolo de señalización de aplicación utilizado en la arquitectura de red inteligente IN (*Intelligent Networks*) para servicios de valor añadido como prepago, mensajería unificada o control de fraude. Como evolución de INAP, está pensado especialmente para telefonía móvil GSM y para escenarios de roaming, facilitando la interoperabilidad entre las diferentes redes a nivel mundial.

SIGTRAN (*SIGnaling TRANsport*) es el nombre del grupo de trabajo de la IETF que desarrolló una serie de protocolos que permiten transportar señalización de control de telefonía pública SS7 por redes IP (RFC 2719 presentada en 1999). SIGTRAN es la evolución de SS7, que define los adaptadores y una capacidad de transporte básico donde se mezclan protocolos SS7 y de paquetes para ofrecer a los usuarios lo mejor de ambas tecnologías. El objetivo principal fue brindar una arquitectura de comunicación de

mensajes de señalización sobre el protocolo IP, brindando una interconexión entre redes SS7 y redes IP, sin necesidad de desarmar las infraestructuras y arquitecturas existentes.

La familia de protocolos SIGTRAN incluye:

- **SCTP (*Stream Control Transmission Protocol*)**: El protocolo más significativo de SIGTRAN. Es el protocolo de nivel de transporte, alternativa a TCP y UDP. Es el protocolo básico en la pila SIGTRAN que proporciona servicio de transporte entre capas en IP (RFC 2960, actualizado en RFC 3309). SCTP es utilizado por uno de los siguientes protocolos de usuario de adaptación de capa.
- **M3UA (*MTP3 User Adaptation Layer*)**: Adaptación de la capa MTP3 de SS7 (RFC 4666), por lo que puede ser utilizada por TCAP/SCCP como capa inferior.
- **M2PA (*MTP2 Peer-to-Peer User Adaptation Layer*)**: Adaptación de la capa MTP2 de SS7 entre *peers* (RFC 4165). Cada *endpoint* M2PA representa un nodo SS7 con su propio *Point Code*. Utiliza los procedimientos de gestion del nivel MTP3 de SS7.
- **M2UA (*MTP2 User Adaptation Layer*)**: Adaptación de la capa MTP2 de SS7 (RFC 3331). En M2UA solo el cliente representa a un nodo SS7 con un *Point Code*. Utiliza procedimientos de gestión propios. En la práctica, la mayoría de las operadoras móviles implementaron M2PA en lugar de M2UA por la mayor compatibilidad con los implementaciones SS7 previas.
- **SUA (*SCCP User Adaptation Layer*)**: Adaptación de la capa SCCP, no utilizada en la práctica por ninguna red de telefonía móvil.

Para cada uno de ellos, la pila SS7 se sustituye en una de sus capas bien definidas con un reemplazo de transporte de paquetes. Al movernos hasta las capas superiores de la pila, más conceptos del legado SS7 pueden ser eliminados y reemplazados con paquetes flexibles y las capacidades de enrutamiento IP. Como SIGTRAN es un estándar de la industria, permite a los clientes interactuar en un entorno multi-fabricante.

En realidad Sigtran es una solución de compromiso para poder señalizar redes inteligentes y movilidad empleando la pila TCP/IP, por lo que se añaden los cuatro niveles inferiores de este modelo de capas por debajo del modelo de capas puro de SS7.

La arquitectura de SS7 consta de tres componentes esenciales:

- **SSP (*Service Switching Point*)** o Punto de Conmutación de Servicios: Constituyen el origen de los requerimientos de servicios y envían mensajes a la red de señalización para establecer las llamadas o características de acceso de servicio requeridas por un abonado.
- **STP (*Signaling Transfer Point*)** o Punto de Transferencia de Señales: Realizan la función de enrutamiento de mensajes dentro de la Red SS7.
- **SCP (*Service Control Point*)** o Punto de Control de Servicio: Gestiona la red, la base de datos de las operaciones y los servicios suplementarios.

5.2.1.1 Enrutamiento SS7

Comparado con sus predecesores de señalización en banda como SS5; SS7 mejoraba la eficiencia de red, puesto que en la señalización en banda el canal de voz era utilizado durante el establecimiento de llamada lo que lo hacía indisponible para tráfico de voz real. En llamadas de larga distancia, el camino de la llamada

podía atravesar múltiples nodos, lo que reducía significativamente su capacidad utilizable. Con SS7 la conexión no es establecida entre los puntos finales hasta que todos los nodos del camino han confirmado su disponibilidad. Si el extremo remoto está ocupado, el llamante obtiene una señal de «ocupado» sin consumir un canal de voz.

Como comentábamos anteriormente, el elemento fundamental encargado del enrutamiento de mensajes dentro de la red SS7 es el STP. El STP no fue inicialmente un elemento estandarizado por GSM o 3GPP, pero fue finalmente introducido en todas las redes SS7 para gestionar de una manera unificada y centralizada las decisiones de enrutamiento de la red SS7.

El STP ofrece enrutamiento a nivel MTP3 (basado en *Point Code*), por el que identifica en enrutamiento de salida para un código de punto de destino (*Destination Point Code*) presente en la MSU recibida. También ofrece enrutamiento a nivel SCCP (basado en *Global Title*), para el que realiza traducción de Global Title GTT (*Global Title Translation*), que sería el equivalente al enrutamiento IP. Se examina la dirección de destino CGT (*Called Global Title*) y se decide cómo identificarlo en la red telefónica. Diferentes parámetros del mensaje pueden ser analizados además del CGT para tomar esta decisión como el plan de numeración NPI (*Numbering Plan Indicator*), el tipo de traducción TT (*Translation Type*) o la naturaleza de la dirección NAI (*Nature of Address Indicator*). El STP implementa tablas de enrutamiento con las diferentes combinaciones de estos parámetros, obteniendo una decisión final que puede identificar uno o varios *Point Codes* MTP3, o las direcciones IP de destino en caso de utilizar Sigtran.

El STP también cubre funciones de enrutamiento complementarias como la conversión entre variantes ITU y ANSI, reparto de carga, enrutamiento basado en origen, detección de bucles o transformación del mensaje antes de ser enviado.

5.2.1.2 Seguridad SS7

El protocolo SS7 fue diseñado inicialmente en 1975 por AT&T. Por aquel entonces los requisitos de seguridad eran muy diferentes de los actuales. El hecho de introducir señalización fuera de banda por canal común en un sistema digital, ofrecía mejoras de seguridad significativas con respecto a los sistemas de señalización precedentes que utilizaban señalización en banda. En aquellos, la señalización era enviada por medio de tonos especiales por la línea telefónica ocasionando graves problemas de seguridad cuando los usuarios descubrían que ellos podían simular esos tonos en sus propios terminales y controlar la red. Se utilizaban pequeñas cajas con equipamiento electrónico llamadas «*blueboxes*». SS7 solucionaba esta problemática de seguridad al utilizar un canal de señalización separado de los canales de datos de usuario.

La introducción de SCTP con SIGTRAN en el año 2000 aporta una mejora en la seguridad al proveer confiabilidad, control de flujo y secuenciación. Permite *multihoming*, en el que uno o los dos extremos de la comunicación pueden tener más de una dirección IP. Esto permite reaccionar en forma transparente a fallos en la red. Aporta transmisión confiable tanto para flujos de datos ordenados como desordenados. Permite selección de camino (*path*) y monitorización para seleccionar el camino primario de transmisión de datos y comprobar su conectividad. Aporta mecanismos de validación y reconocimiento para proteger contra ataques por inundación. También mejora los mecanismos de detección de errores.

Sin embargo, era cuestión de tiempo que el antiguo protocolo SS7 quedara desfasado a nivel de seguridad. En 2008, múltiples vulnerabilidades de SS7 fueron publicadas en las que se muestra como se podía rastrear a usuarios móviles. Se utilizaban *IMSI-catchers* como dispositivos que interceptaban la señalización SS7 para obtener información del número de IMSI del abonado y rastrear la información de su localización. La especificación GSM requiere que el dispositivo se autentique ante la red, pero no requiere que la red se autentique ante el dispositivo. Este agujero de seguridad es bien conocido y explotado por

el *IMSI-catcher*, que actúa como una estación base falsa e identifica todos los números de IMSI de todas las estaciones móviles en un determinado área geográfico. Fuerza al dispositivo a no utilizar encriptación de llamada o a utilizar encriptación muy sencilla fácilmente vulnerable; haciendo que la llamada sea fácil de interceptar y convertir a audio. Aunque los estándares 3G y 4G incorporan autenticación mutua tanto en el dispositivo como en la red y eliminan este ataque, algunos ataques sofisticados permiten forzar al dispositivo a «bajar» a 2G, donde la autenticación mutua no es requerida.

Otro ejemplo de vulnerabilidad de SS7 es el relacionado con la autenticación de dos factores utilizando mensajes SMS, y por tanto el acceso no autorizado a, por ejemplo, cuentas bancarias. Los perpetradores instalan *malware* en los ordenadores o dispositivos comprometidos, y configuran reenvíos para los números de teléfono de la víctima a líneas telefónicas controladas por ellos, donde reciben las llamadas o los SMS de confirmación.

Múltiples vulnerabilidades fueron siendo reportadas, muchas de ellas específicas de escenarios de roaming. Las operadoras fueron implementando diferentes soluciones para bloquear estos ataques. Finalmente, en 2014, la GSMA introdujo el documento FS.11, en el que se detallan las diferentes vulnerabilidades, se exponen los riesgos potenciales y se ofrecen un conjunto de recomendaciones de seguridad para que las operadoras desplieguen seguridad adicional en sus redes SS7, con el fin de mitigarlas. La GSMA define tres categorías de amenazas:

- **Categoría 1**: Mensajes prohibidos. Mensajes que en condiciones normales solo deberían ser recibidos dentro de una misma red, o entre redes con acuerdos bilaterales para dicho intercambio.
- **Categoría 2**: Mensajes no autorizados. Mensajes sobre un usuario, que solo deberían ser enviados por parte de su red local.
- **Categoría 3**: Mensajes de localización sospechosos. Mensajes sobre un usuario, que solo deberían ser enviados por la red visitada en la que se encuentre.

Bajo estas recomendaciones, diferentes proveedores de red empezaron a ofrecer soluciones de *firewall* de señalización SS7 (SS7-FW), que las operadoras empezaron a desplegar junto a sus STPs para ofrecer una solución combinada de enrutamiento y seguridad para la señalización SS7.

5.2.2 4G y el protocolo Diameter

Aunque SS7 se mantuvo tanto para la generación 2G como para 3G, el núcleo de red 4G EPC incorpora un cambio a nivel de protocolo de señalización con la aparición de Diameter.

Diameter es un protocolo de red, diseñado para ofrecer un marco de trabajo que ofrezca servicios AAA para aplicaciones que involucran acceso a redes o aplicaciones móviles IP. Su desarrollo inicial en 2003 estuvo basado en el protocolo precedente *Remote Authentication Dial-In User Service (RADIUS)* introducido en 1997. El origen del nombre Diameter no está relacionado con un acrónimo, es simplemente un juego de palabras representando al diámetro como el doble del radio, haciendo énfasis en el hecho de evolucionar y mejorar el protocolo anterior. Mientras que RADIUS utiliza típicamente transporte UDP, Diameter utiliza protocolos de transporte fiable como SCTP o TCP. A diferencia del modelo cliente-servidor implementado por RADIUS, Diameter es *peer-to-peer*, permite negociación de capacidades, es más fácil de extender a nuevos comandos y atributos y tiene un espacio de direcciones mayor para AVPs (*Attribute Value Pairs*, o parejas atributo-valor), entre otras mejoras.

Los servicios AAA soportados por Diameter son:

- **Autenticación**: La autenticación es el proceso con el que se verifica la identidad de quien envía información y también de quien la recibe. Para lograr la autenticación y averiguar que alguien es quien dice ser, se utilizan las identidades que los usuarios presentan a la red a través de credenciales. Existen autenticación de usuario, autenticación de equipo, autenticación de mensaje, autenticación unilateral, autenticación mutua, autenticación de servidor, etc.

- **Autorización**: La autorización es el proceso posterior a la autenticación, en el que se le asigna ciertos privilegios al poseedor de un credencial particular. Los privilegios se asocian al perfil del usuario del terminal. Los privilegios pueden ser permitir el acceso a una serie de recursos como bases de datos, archivos, tiempo de uso de un procesador, ejecución de ciertas instrucciones, etc.

- **Contabilidad**: Es el proceso de recolección de información sobre el uso de recursos con el fin de realizar otras funciones como planificación de capacidad, auditorías, facturación y asignación de costes. La auditoría consiste en el chequeo periódico para determinar la firmeza de la información y de las políticas de gestión, en especial sobre la seguridad.

A nivel de pila de protocolos, Diameter utiliza SCTP o TCP como protocolo de transporte sobre el puerto 3868, e IP como protocolo de red. La implementación base de Diameter aporta capacidades básicas como establecimiento, liberación y monitorización de la conexión Diameter, y permite el uso de aplicaciones específicas DIAMETER sobre ella. Esta pila de protocolos se muestra en la figura 5.2.

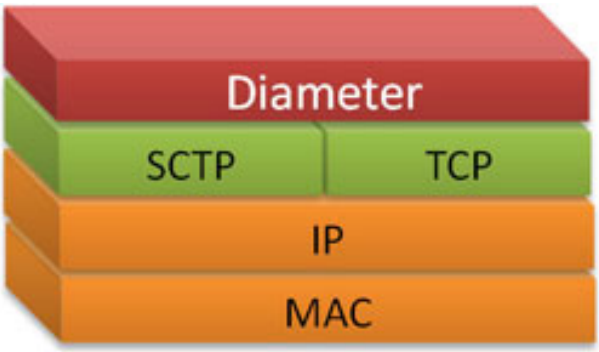

Figura 5.2: Pila de protocolos Diameter.

El flujo de mensajes en una conexión Diameter empieza con el establecimiento de la conexión. El agente Diameter que inicia la conexión envía un mensaje de Solicitud de Intercambio de Capacidades (CER), y el receptor responde con un mensaje de respuesta de intercambio de capacidades (CEA). En este punto la conexión ya está establecida y lista para el intercambio de mensajes de aplicación. Cualquiera de los dos extremos puede enviar solicitudes de dispositivo «perro guardián» (DWR) y el otro responderá con una respuesta (DWA). Cuando uno de los dos extremos desee finalizar la conexión, enviará un mensaje de solicitud de desconexión (DPR), que será respondido con un mensaje (DPA) para dar por finalizada la conexión.

Cada aplicación Diameter está definida por un identificador de la aplicación y puede añadir nuevos códigos de comando y/o nuevos AVPs. Estas aplicaciones son desarrolladas como extensiones sobre el protocolo base Diameter y por tanto se van diseñado a medida que se necesitan. 3GPP define múltiples aplicaciones Diameter para las diferentes interfaces entre funciones de red del núcleo de red 4G EPC y de la red IMS. Por ejemplo la aplicación Diameter s6a entre MME y HSS para el intercambio de datos de suscripción y autentificación, o la aplicación Diameter Gx entre el PCEF y el PCRF para calidad de servicio y políticas.

5.2.2.1 Enrutamiento Diameter

Cada mensaje Diameter contiene una cabecera en la que, entre otra información, encontramos el código de comando y una cantidad variable de AVPs. Los AVPs hacen referencia a una representación de datos llamados atributos, usada en esquemas en los que se requiere que las estructuras de datos permitan extensiones flexibles sin modificar el código. Cada AVP contiene un código de AVP que lo identifica de manera única, un campo de longitud de AVP identificando su longitud total en octetos, tres flags que identifican la presencia o no del identificador de proveedor (flag V), si el AVP es obligatorio (flag M), o la necesidad de cifrado extremo a extremo (flag P); y por último los datos del AVP. Las cabeceras y AVPs más relevantes a nivel de enrutamiento son:

- **Application-ID**: Identificador de cada aplicación Diameter definido por *Internet Assigned Numbers Authority (IANA)*. El protocolo base no necesita identificador pues debe ser soportado por todas las aplicaciones. Por ejemplo, la aplicación s6a para la señalización 4G LTE entre MME y HSS relacionado con el registro de usuarios tiene el application-ID «16777251».
- **Hop-by-hop ID**: Utilizado para la correlación entre solicitudes y respuestas. A medida que el mensaje pasa de un salto a otro, se cambia este identificador; pero en cada respuesta se envía el mismo número que se encontró en ese campo en la solicitud que generó dicha respuesta.
- **End-to-End ID**: Identifica los extremos de la comunicación Diameter y se mantiene constante durante los diferentes saltos del mensaje en su transito desde la parte originante hasta el receptor final.
- **Destination Host AVP**: Identificador del nodo específico de destino, que puede estar presente en las peticiones Diameter (si el *host* destino concreto es conocido), y que no debe estar presente en las respuestas Diameter.
- **Destination Realm AVP**: Identificador del dominio administrativo del destino. Se utiliza en los nodos de enrutamiento como rutas. No debe estar presente en las respuestas Diameter.
- **Origin Host AVP**: Identificador del nodo específico de origen, que debe estar presente en todos los mensajes.
- **Origin realm AVP**: Identificador del dominio administrativo del origen, que debe estar presente en todos los mensajes.
- **User-Name**: En 4G LTE identificará el IMSI del abonado relacionado con el mensaje Diameter.
- **Session-ID**: Todos los paquetes Diameter con el mismo identificador de sesión o *Session-Id* son considerados parte de la misma sesión.

El enrutamiento Diameter es llevado a cabo por agentes Diameter, que poseen las tablas de enrutamiento. Existen cuatro tipos de agentes Diameter:

- **Relay Agent**: Son los agentes Diameter que aceptan y enrutan los mensajes de otros nodos hacia su destino, en función de la información que contiene el mensaje y las tablas de enrutamiento. No necesitan analizar el contenido del mensaje, solo analizan los campos necesarios para el enrutamiento. Modifican la parte del mensaje relativa al enrutamiento para el correcto enrutamiento extremo a extremo, pero no modifican el contenido del mensaje.

  - **Proxy Agent**: Es similar a un agente Relay con toma de decisiones sobre la base de ciertas políticas de acceso. Necesita analizar los mensajes que transcurren por él y puede generar mensajes de rechazo en caso de violación de las políticas.

- **ReDirect Agent**: Actúa como individuo intermediario para la transformación de $dominios administrativos) a direcciones de servidores con las tablas de enrutamiento de un grupo determinado. Un Diameter R$ $Es el encargado de la compatibilidad, realiza la traducción entre Diameter y otros protocolos AAA, como por ejemplo RADIUS$

En la red de señalización Diameter de 4G EPC, el elemento de red encargado del enrutamiento es el DRA (*Diameter Routing Agent*). En la mayoría de los casos, las capacidades de *Redirect* y *Translation* no son requeridas, por lo que suelen actuar como un agent *Proxy*. En caso de enrutamientos de escenarios sencillos pueden actuar como un agente *Relay*.

El elemento de red 4G EPC encargado del enrutamiento Diameter en el interfaz de interconexión y roaming es el DEA (*Diameter Edge Agent*) por nomenclatura establecida por GSMA. En muchas ocasiones se implementa conjuntamente con el DRA e implementa las mismas capacidades de enrutamiento pero con requisitos mayores a nivel de seguridad.

La información fundamental utilizada por un DRA/DEA para sus decisiones de enrutamiento es la obtenida de los AVPs *Destination-Realm* y *Destination-Host*; así como *Origin-Realm* y *Origin-Host* en caso de querer implementar enrutamientos diferentes a un mismo destino en base al origen (*OBR: Origin-Based Routing*). Estos AVPs tienen la siguiente estructura [71]:

- **realm**: epc.mnc<MNC>.mcc<MCC>.3gppnetwork.org
- **host**: <host>.epc.mnc<MNC>.mcc<MCC>.3gppnetwork.org

En ellos, la combinación de *Mobile Network Code (MNC)* y *Mobile Contry Code (MCC)* identifica unívocamente a una operadora a nivel mundial. En los casos en los que el *destination-host* está presente, el DRA/DEA enrutará el mensaje directamente a ese *host* concreto. En los casos en los que no se incluya *destination-host*, el DRA identificará el destino en base al análisis de diferentes parámetros como *destination-realm*, *application-id*, código de comando o IMSI (*User-name* AVP).

Es importante resaltar que Diameter es un protocolo de señalización transaccional, por lo que una vez definido el enrutamiento para el mensaje de solicitud, el mensaje de respuesta será enrutado automáticamente, deshaciendo el camino recorrido por la solicitud, mediante el uso del las cabeceras Diameter de *end-to-end-id* y especialmente *hop-by-hop-id*. El enrutamiento de la solicitud y la respuesta siempre es simétrico. Ésto representa un cambio importante con respecto a SS7 donde cada mensaje es enrutado de manera independiente (ya sea solicitud o respuesta), e incluso son posibles escenarios asimétricos donde la solicitud enviada del origen al destino sigue un camino distinto del de la respuesta enviada desde el destino al origen.

5.2.2.2 Seguridad Diameter

Aunque el estándar Diameter soporta el uso de IPSec o TLS para ofrecer seguridad, éstos no han sido realmente utilizados en la práctica en el núcleo de red EPC 4G de las operadoras; ni en escenarios nacionales ni en escenarios de roaming internacionales.

La implementación más común de DIAMETER en EPC utiliza SCTP en lugar de TCP aportando mayor confiabilidad, control de flujo y secuenciación, y el uso de *multihoming*, del mismo modo que es utilizada en SIGTRAN SS7.

La aplicación Diameter s6a relacionada con el registro de usuarios en el interfaz entre Mobility Management Entity (MME) y Home Subscriber Server (HSS) está definida en la especificación 3GPP 29272 [72]. En ella, se realizó una adaptación de los mensajes utilizados en las redes SS7 con el protocolo MAP para la comunicación entre *Mobile Switching Center (MSC)* y *Home Location Register (HLR)*. Los parámetros necesarios fueron adaptados a sus correspondientes AVPs de Diameter y se mantuvo una estructura de códigos de operación similar para solicitudes de autenticación de usuario, registro y localización de usuario, actualización, cancelación de registro, etc. Por un lado, esto facilita la interoperabilidad entre las redes 2G/3G y 4G para escenarios como *Circuit-Switched Fallback (CSFB)* en los que un abonado cambia con frecuencia de registro entre las diferentes redes, e incluso mantiene registros combinados simultáneos en ambas redes (*Combined Attach*). Como parte negativa, una gran mayoría de las vulnerabilidades a nivel de señalización encontradas a lo largo de los años en el protocolo SS7, aplican también a la señalización Diameter de EPC; y particularmente a sus escenarios de roaming. Incluso, la flexibilidad introducida por Diameter como estructura abierta a nivel de AVPs, en comparación con la rigidez en cuanto a parámetros y estructura de mensaje de SS7; la hace más propicia a ataques en determinados escenarios.

Por estos motivos, de un modo similar al ocurrido con SS7, en 2017 la GSMA introdujo el documento FS.19, en el que se detallan las diferentes vulnerabilidades, se exponen los riesgos potenciales y se ofrecen un conjunto de recomendaciones de seguridad para que las operadoras desplieguen seguridad adicional en sus redes Diameter, con el fin de mitigarlas. La GSMA mantiene las tres categorías de amenazas identificadas para SS7, las expande para Diameter, e incluso las agrega para contemplar escenarios de ataque combinados entre ambas tecnologías.

Bajo estas recomendaciones, diferentes proveedores de red empezaron a ofrecer soluciones de *firewall* de señalización Diameter (Diameter-FW), que las operadoras empezaron a desplegar junto a sus DRAs para ofrecer una solución combinada de enrutamiento y seguridad para la señalización Diameter. Para contemplar los escenarios de seguridad combinados entre ambas tecnologías, se introducen como solución unificada de *firewall* de señalización (Sig-FW) soportando tanto SS7 como Diameter.

5.2.3 HTTP2 desde el sector TI hacia el sector Telco

Con el salto tecnológico hacia 5G, se decidió dejar atrás los protocolos de señalización SS7 o Diameter que las tecnologías precedentes utilizaban en el plano de control del núcleo de red, y apostar por HTTP2. Ya no se utilizan protocolos específicos del mundo de las telecomunicaciones y de la telefonía, y se escoge un protocolo del mundo de las tecnologías de la información (TI) y de Internet desarrollado por *Internet Engineering Task Force (IETF)* [59].

HTTP2 ofrece una expresión optimizada de la semántica del protocolo *HyperText Transfer Protocol (HTTP)*, que permite un uso más eficiente de los recursos de la red y reducción en latencia mediante la introducción de encabezado de compresión. También permite múltiples intercambios simultáneos en la misma conexión y un procesado más eficiente de mensajes al utilizar codificación binaria. Es por estos motivos que las especificaciones de 5G seleccionan HTTP2 como protocolo para la señalización del núcleo de red en SBI.

HTTP2 utiliza los mismos esquemas de URI «http» y «https» que utiliza HTTP/1.1. Comparte también los mismos números de puerto predeterminados: 80 para URI de tipo «http» y 443 para URI de tipo «https». Como resultado, las implementaciones que procesan solicitudes de URI de recursos de destino necesitan descubrir primero si el servidor ascendente (el par inmediato al que el cliente desea establecer una conexión) admite HTTP2. El identificador «h2» identifica el protocolo en el que HTTP2 utiliza TLS, y se utiliza en la negociación previa de conexión TLS al identificar qué protocolo de nivel de aplicación hará uso de la conexión TLS.

En el caso de utilizar HTTP2 sobre TCP en texto claro, el identificador es «h2c» y se utiliza en la negociación de actualización (*upgrade*) desde HTTP/1.1 hacia HTTP2. Aún así, en el caso de las funciones de red 5G, clientes y servidores ya saben de antemano que el otro extremo soporta HTTP2, con lo que los clientes pueden enviar un prefacio de conexión que indica el conocimiento previo de que el servidor soporta HTTP2 para, directamente, enviar tramas HTTP2 al servidor. Esto es aplicable al envío de señalización HTTP2 sobre TCP, puesto que en el caso de utilizar TLS, esta negociación se realiza durante la fase previa de negociación TLS (*TLS handshake*).

Como veremos más adelante, 5G introduce el concepto de seguridad por diseño, por lo que las diferentes funciones de red se comunican entre ellas utilizando HTTP2 sobre TLS.

En HTTP2, un flujo (*stream*) es una secuencia de tramas bidireccional, intercambiadas entre el cliente y el servidor dentro de una conexión HTTP2. Una conexión HTTP2 puede tener múltiples flujos concurrentes abiertos, con cada uno de los dos extremos intercambiando tramas de múltiples flujos. Estos flujos pueden ser terminados por cualquiera de los dos extremos y se identifican con un número entero creciente asignado por el extremo que inicia el flujo. Los iniciados por el cliente contienen un número impar, mientras que los iniciados por el servidor tendrán un número par. El extremo que recibe un identificador de flujo no esperado debe responder con un error de conexión.

HTTP2 soporta diferentes tipos de tramas:

- **Trama de cabeceras (HEADERS)**: Utilizada para abrir un nuevo flujo.
- **Trama de datos (DATA)**: Transporta los datos a intercambiar e identifica a un flujo concreto.
- **Trama de prioridad (PRIORITY)**: Especifica la prioridad del flujo, anunciada por el extremo que envía.
- **Trama de reinicio (RST_STREAM)**: Permite la terminación inmediata de un flujo, identificando códigos de error.
- **Trama de parámetros (SETTINGS)**: Contiene parámetros de configuración que afectan a cómo se comunican los extremos.
- **Trama de promesa de envío (PUSH_PROMISE)**: Para notificar al otro extremo con antelación de flujos que pretende iniciar.
- **Trama de ping (PING)**: Mecanismo para medir el tiempo de mínimo de trayecto desde un extremo.
- **Trama de terminación (GOAWAY)**: Utilizada para iniciar la terminación de una conexión por parte de un extremo, dejando de aceptar nuevos flujos, aunque terminando de procesar los flujos previamente establecidos. Identifica el código de error por el que esta terminación se produce.
- **Trama de actualización de ventana (WINDOW_UPDATE)**: Utilizada para implementar control de flujo.
- **Trama de continuación (CONTINUATION)**: Utilizada para continuar un secuencia de fragmentos de bloque de cabeceras.

Para el envío de peticiones y respuestas en HTTP2 la gran mayoría de funcionalidades, capacidades y especificaciones de HTTP/1.1 se mantienen. De manera específica para HTTP2, el cliente que envía una petición HTTP, lo hace en un nuevo flujo, utilizando un identificador de flujo no utilizado previamente. El servidor envía la respuesta HTTP2 utilizando el mismo flujo que la petición.

Mientras que HTTP/1.x utiliza un mensaje de comienzo de línea para indicar el URI objetivo, el método de la petición y el código de estado de la respuesta; HTTP2 utiliza campos de pseudo-cabecera que comienzan con el carácter «:» para este propósito.

Las pseudo-cabeceras de HTTP2 en una petición son:

- **:method** incluye el método de la petición HTTP (GET, PUT, POST, DELETE, PATCH, ...).
- **:scheme** define la parte de esquema del URI objetivo (HTTP, HTTPS u otros).
- **:authority** contiene la parte de autoridad del URI objetivo, que reemplaza al campo «*host*» de HTTP/1.x.
- **:path** incluye la parte de ruta o *path* para las peticiones hacia el URI objetivo.

La pseudo-cabecera de HTTP2 en una respuesta es:

- **:status** contiene el código de estado HTTP que debe estar presente en todas las respuestas.

Las conexiones HTTP2 son persistentes. Los clientes no cerrarán las conexiones HTTP2 hasta que no tengan más comunicaciones que enviar al servidor. Tampoco abrirán múltiples conexiones HTTP2 hacia un determinado *host* y puerto; aunque las renovarán ante determinadas circunstancias como renovación del material de claves de la conexión TLS, reemplazar conexiones con errores o reemplazar conexiones que agotan los recursos del espacio de identificadores de flujo. Un cliente también puede mantener múltiples conexiones con una misma IP y puerto en TLS utilizando deferentes indicadores de nombre de servidor SNI (*Server Name Indicatio*n), o para ofrecer diferentes certificados TLS de lado cliente.

El intercambio de una petición y respuesta HTTP2 consume completamente un identificador específico de trama. Cuando los identificadores de trama HTTP2 de una determinada conexión HTTP2 se agotan, uno de los extremos HTTP2 debe establecer una nueva conexión HTTP2.

En cuanto al uso de TLS, las implementaciones de HTTP2 deben utilizar TLS versión 1.2 o superiores para ofrecer HTTP2 sobre TLS, bloqueando versiones TLS anteriores. Se debe soportar la extensión de TLS denominada *Server Name Indication (SNI)*, para que los clientes HTTP2 identifiquen el nombre dominio del servidor objetivo en la negociación TLS.

Además de HTTP2, en 5G se requiere la utilización de JSON [60] como protocolo de serialización para transporte de datos; y de la especificación OpenAPI [63] como lenguaje de definición de interfaces SBI.

La estructura típica del mensaje HTTP se muestra en la figura 5.3 [67].

5.3 Enrutamiento y Seguridad de Señalización en 5G Core

Con la adopción del protocolo HTTP2 por parte de 3GPP para el desarrollo de SBI en el núcleo de red 5G, además de reutilizar múltiples cabeceras HTTP básicas, se incorporan unas cabeceras customizadas específicas de 3GPP que toda función de red debe soportar.

Las cabeceras estándar reutilizadas de HTTP2 especialmente importantes para el enrutamiento en 5GC son:

- **Accept**: Utilizada para indicar los tipos de contenido que son aceptables en la respuesta.

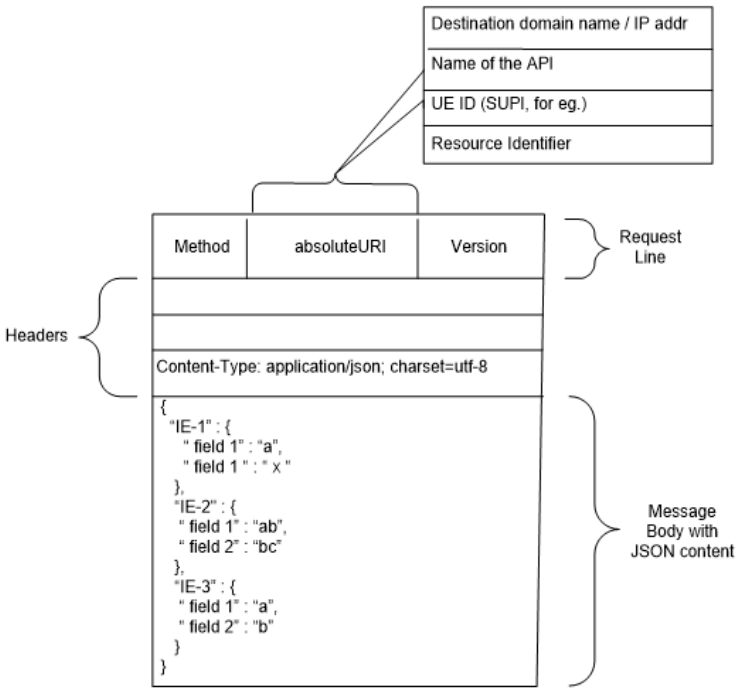

Figura 5.3: Estructura típica del mensaje HTTP.

- **Content-type**: Utilizada para indicar el tipo de contenido que contiene el mensaje de petición.
- **User-Agent**: Utilizada para identificar el tipo de NF del cliente HTTP2.
- **Via**: Debe ser introducida por un HTTP proxy como SCP o SEPP cuando transmiten una petición HTTP.
- **Authorization**: Debe ser incluida si el acceso basado en autorización OAuth 2.0 es utilizado.
- **Location**: Utilizada en algunas respuestas para referir a un recurso específico.

Las cabeceras customizadas especialmente importantes para el enrutamiento de mensajes en el núcleo de red 5G son:

- **3gpp-Sbi-Message-Priority**: Se utiliza para especificar la prioridad del mensaje HTTP2 en SBI.
- **3gpp-Sbi-Callback**: Se usa para indicar si un mensaje HTTP/2 es una devolución de llamada (*callback*), como por ejemplo, una notificación. Se incluirá en los mensajes HTTP POST para los *callback* hacia los consumidores de servicios NF en otra PLMN a través de SEPP; y en los mensajes HTTP POST para *callbacks* en comunicación indirecta a través de SCP.
- **3gpp-Sbi-Target-apiRoot**: Utilizada por un cliente HTTP para indicar la *apiRoot* del URI de destino cuando se comunica indirectamente con el servidor HTTP a través de un SCP. SCP también utiliza este encabezado para indicar la *apiRoot* del URI de destino, si se selecciona o se vuelve a seleccionar un nuevo servidor HTTP y no se incluye una cabecera de ubicación en la respuesta. También puede ser utilizado por un cliente HTTP hacia su SEPP local para indicar la *apiRoot* del URI de destino hacia el servidor HTTP en otra PLMN. También se puede usar entre los SEPPs para indicar la *apiRoot* del URI de destino hacia el servidor HTTP en otra PLMN, cuando se usa la seguridad TLS con el encabezado *3gpp-Sbi-Target-apiRoot* entre los SEPPs.

- **3gpp-Sbi-Routing-Binding**: Se utiliza en una solicitud de servicio para señalar información vinculante para dirigir la solicitud de servicio a un servidor HTTP que tiene el contexto de recurso de servicio NF objetivo.

- **3gpp-Sbi-Binding**: Se utiliza para señalizar información vinculante relacionada con un recurso de servicio NF a un consumidor futuro (cliente HTTP) de ese recurso.

- **3gpp-Sbi-Discovery-***: Las cabeceras que comienzan con el prefijo *3gpp-Sbi-Discovery-* se utilizan en el modo de comunicación indirecta para el descubrimiento y la selección de un productor adecuado por parte del SCP. Dichos encabezados pueden incluirse en cualquier mensaje SBI e incluir información que permita a un SCP encontrar un productor adecuado según los parámetros de detección delegados incluidos del consumidor.

- **3gpp-Sbi-Producer-Id**: Se usa en una respuesta del SCP al consumidor de servicios de NF, cuando se usa comunicación indirecta, para identificar al productor de servicios de NF.

- **3gpp-sbi-originating-network-id**: Introducida en la release 17, se añadió para poder identificar el PLMN-ID de origen de la NF que envía el mensaje (cNF o pNF), especialmente en escenarios de roaming.

- **3gpp-Sbi-Access-Token**: Utilizada en una respuesta enviada por el SCP a una NF consumidora para proporcionar un token de acceso que poder utilizar en las subsiguientes peticiones.

Las diferentes decisiones de enrutamiento dentro del plano de control del núcleo de red 5G se realizan en base a los valores de las pseudo-cabeceras, cabeceras estándar de HTTP2, y cabeceras HTTP2 customizadas por 3GPP.

5.3.1 Registro y descubrimiento de servicios de las funciones de red

La comunicación de señalización entre las distintas NFs de la red 5G se basa en el descubrimiento automático del resto de NFs de la red y de los diferentes servicios que cada una ofrece. En la tecnologías precedentes las diferentes funciones eran configuradas de manera manual con conexiones estáticas a los diferentes elementos de red vecinos con los que tenían que interactuar. Esto queda automatizado en 5G debido a los mecanismo de registro y descubrimiento de servicios de las funciones de red.

Aunque los procesos descritos en esta sección muestran la comunicación directa entre una NF y el NRF, al introducir el SCP en la red, estos mensajes fluyen entre NF y NRF a través del SCP, con ciertas diferencias según estemos en modelo C o D tal cual veíamos en la sección 4.1.2.2.

5.3.1.1 Registro de servicios de NF

Es iniciado por una función de red que registrará su perfil y el conjunto de servicios ofrecidos en el NRF, con el objetivo de que puedan ser descubiertos por otras funciones de red consumidoras que puedan requerir dichos servicios.

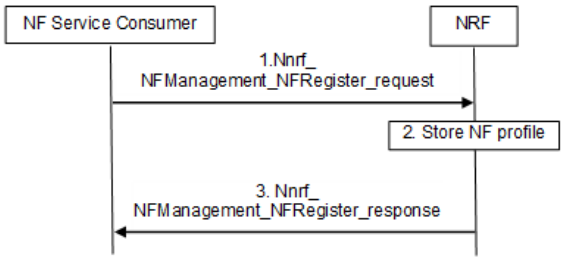

Figura 5.4: Procedimiento de registro de servicios de NF.

En este procedimiento la función de red que desea registrar su perfil y servicios en el NRF, envía un mensaje de solicitud Nnrf_NFManagement_NFRegister a NRF para informarle de su perfil de NF. La NF iniciará este procedimiento cuando se vuelva operativa por primera vez. Entre los parámetros más relevantes del perfil se encuentran la dirección de direccionamiento (como su FQDN o dirección IP), así como los detalles de los diferentes servicios ofrecidos (como veíamos en la sección 4.1.2.1).

El NRF almacenará el perfil NF recibido y lo marcará como disponible, para posteriormente enviar la respuesta de reconocimiento Nnrf_NFManagement_NFRegister.

5.3.1.2 Actualización de servicios de NF

Es iniciado por una función de red que previamente ha registrado su perfil y el conjunto de servicios ofrecidos en el NRF, y quiere actualizar algunos de los parámetros de ese perfil.

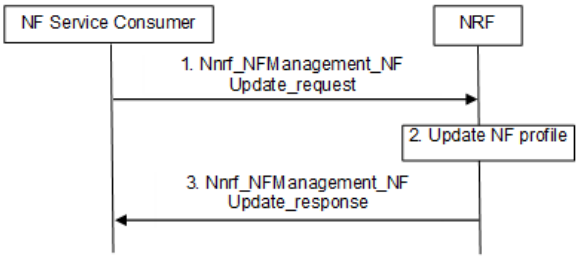

Figura 5.5: Procedimiento de actualización de servicios de NF.

En este procedimiento la función de red que desea actualizar su perfil y servicios en el NRF, envía un mensaje de solicitud Nnrf_NFManagement_NFUpdate a NRF para informarle de los cambios en su perfil de NF. El NRF actualizará el perfil NF recibido, para posteriormente enviar la respuesta de reconocimiento Nnrf_NFManagement_NFUpdate.

5.3.1.3 Desregistro de servicios de NF

Es iniciado por una función de red que previamente ha registrado su perfil y el conjunto de servicios ofrecidos en el NRF, y quiere indicar su indisponibilidad, por ejemplo debido a un apagado o a una desconexión de la red.

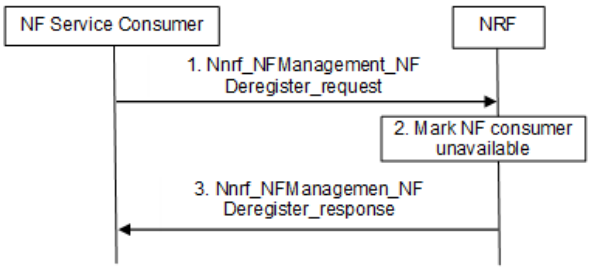

Figura 5.6: Procedimiento de desregistro de servicios de NF.

En este procedimiento la función de red que desea desregistrar su perfil y servicios del NRF, envía un mensaje de solicitud Nnrf_NFManagement_NFDeregister a NRF para informarle de su indisponibilidad. El NRF actualizará el perfil NF correspondiente marcándolo como «indisponible», para posteriormente enviar la respuesta de reconocimiento Nnrf_NFManagement_NFDeregister. El NRF también podría eliminar completamente el perfil.

5.3.1.4 Descubrimiento de servicios de NF

Una vez las diferentes NFs de la red se han ido registrando en el NRF, publicando sus servicios ofrecidos y su información de disponibilidad; el resto de funciones de la red puede descubrir dinámicamente qué NFs están disponibles y qué servicios ofrecen consultando al NRF. Las NFs podrán descubrir estos servicios especificando el tipo de NF objetivo y el nombre de servicio requerido. Éstos serán los parámetros de búsqueda principales, pero la NF consumidora podrá incluir muchos otros parámetros de petición (query-parameters) para ser más preciso en la obtención de una NF productora apropiada. Por ejemplo, puede indicar la localización preferida, o un SUPI-IMSI específico.

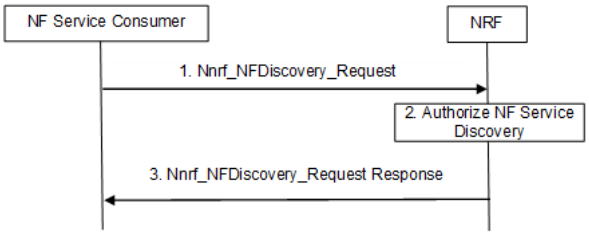

Figura 5.7: Procedimiento de descubrimiento de servicios de NF.

En este procedimiento la función de red consumidora que desea descubrir a una NF productora para un servicio concreto, envía un mensaje de solicitud Nnrf_NFDiscovery_Request a NRF para realizar esta petición. El NRF autoriza la petición e identifica si la NF consumidora que envía esta petición está autorizada para descubrir el NF objetivo. En caso afirmativo, el NRF identifica una instancia o conjunto de instancias de NF que cumplen los parámetros de descubrimiento recibidos para posteriormente enviar la respuesta de reconocimiento Nnrf_NFDiscovery_Request conteniendo la información de estos perfiles.

5.3.1.5 Selección de NF productora

Una vez la NF consumidora ha recibido la respuesta del NRF conteniendo un listado de perfiles de NF compatibles con los requisitos del servicio a utilizar, la NF consumidora seleccionará entre los perfiles recibidos, el de la NF productora a utilizar. En el caso de utilizar la arquitectura del modelo D, con descubrimiento delegado, será el SCP el que realice esta selección en lugar de la NF consumidora.

Al recibir esta lista de perfiles, estos pueden ser almacenados en memoria caché y podrán ser utilizados mientras estén dentro del tiempo de validez reportado en la respuesta del NRF. La NF consumidora o el SCP podrán utilizar diferentes criterios para escoger entre los perfiles NF disponibles en base a la información de estos perfiles como por ejemplo prioridad y capacidad (criterio por defecto que siempre debe ser tenido en cuenta), o parámetros adicionales como por ejemplo nivel actual de carga de la pNF.

5.3.2 Mecanismos de enrutamiento en el núcleo de red 5G

Los mecanismos de enrutamiento en la arquitectura basada en servicios toman como referencia los procedimientos descritos en la RFC 7230 [73] para enrutamiento de mensajes HTTP1.1, pero con adaptaciones relacionadas con la utilización de HTTP/2 y particularidades específicas del sistema 5G.

Para identificar el destino de un mensaje se utiliza el URI objetivo (*target URI*) del mensaje. Al enviar un mensaje de petición de servicio, éste comienza con una trama «HEADERS» que contiene los campos de pseudo-cabeceras que identifican el destino objetivo de la petición. La pseudo-cabecera «:method» siempre está presente, y salvo en los mensajes de tipo «CONNECT» o «OPTIONS», también lo están las pseudo-cabeceras «:scheme», «:authority» y «:path». Esta última incluye la información de ruta (*path*) y también los componentes de parámettros de petición (*query parameters*) de la URI objetivo.

En el enrutamiento de mensajes HTTP/2 dentro de una PLMN, el componente «authority» del URI objetivo debe contener el «uri-host» que podrá ser la FQDN del servicio de NF objetivo, o la dirección IP del servicio NF objetivo. Esta FQDN no necesita contener el identificador de PLMN. Opcionalmente también puede contener un puerto, especialmente necesario en caso de utilizar puertos no estándar para http (80) o https (443).

Para el enrutamiento de mensajes entre diferentes PLMNs, donde es necesario acceder al servicio de NF correcto en la PLMN correcta, la pseudo cabecera «:authority» del mensaje HTTP/2 debe contener un «uri-host» formado por la FQDN del servicio de NF objetivo o la parte FQDN de una URI en el caso de ser una petición de tipo *callback*. En ambos casos, la FQDN debe contener el identificador de PLMN. Este identificador contiene el dominio de red doméstica (*Home Network Domain*) que tiene la siguiente estructura:

5gc.mnc<MNC>.mcc<MCC>.3gppnetwork.org

Los campos *<MNC>* y *<MCC>* identifican inequívocamente a la PLMN. El MCC (*Mobile Country Code*) identifica a un país y siempre tiene 3 dígitos, mientras que el MNC (*Mobile Network Code*) identifica a cada operadora dentro de un país y puede tener 2 o 3 dígitos. La combinación de ambos se denomina HNI (*Home Hetwork Identifier*). Para la formación de esta FQDN, ambos campos tendrán siempre 3 dígitos. En el caso de que el MNC solo contenga dos dígitos significativos, se inserta un «0» a la izquierda del MNC para cumplir con la codificación en tres dígitos requerida en el enrutamiento entre PLMNs para identificar puntos finales (*endpoints*) de servicios de NF.

Como ejemplo para una PLMN con MCC 345 y MNC 12 se identifica el siguiente dominio de red doméstica:

5gc.mnc012.mcc345.3gppnetwork.org

A la hora de identificar NFs específicas, y poniendo como ejemplo el NRF para el caso en el que una NF no haya sido pre-configurada con la FQDN del NRF de su red doméstica, un NF puede construir el FQDN del NRF de su PLMN utilizando la etiqueta «nrf» y siguiendo este formato:

nrf.5gc.mnc<MNC>.mcc<MCC>.3gppnetwork.org

En escenarios entre PLMNs, el vNRF podrá comunicarse con el hNRF construyendo una URI objetivo del siguiente modo:

- La pseudo-cabecera :authority debe contener la FQDN del NRF tal cual acabamos de describir.
- La pseudo-cabecera :scheme debe ser «https».
- El puerto debe ser el puerto por defecto para el esquema https, es decir 443.
- No se debe usar ningún otro prefijo opcional.

Como ejemplo, el API raíz (*APIroot*) de los servicios de NRF para una PLMN con MCC 345 y MNC 12 debe ser:

https://nrf.5gc.mnc012.mcc345.3gppnetwork.org

Siguiendo este modelo, las diferentes instancias de funciones de red quedarán identificadas añadiendo una parte *host* específica, como por ejemplo:

https://scp12.5gc.mnc456.mcc789.3gppnetwork.org

https://sepp34.5gc.mnc456.mcc789.3gppnetwork.org

Para el enrutamiento entre PLMNs, el SEPP es el encargado final del enrutamiento comunicándose con el SEPP de la red remota. Para permitir protección TLS entre el SEPP y las funciones de red dentro de una PLMN, 3GPP define dos opciones, el uso de FQDNs telescópicas (introducido en release 15), o el uso de una cabecera específica denominada «3ggp-sbi-target-api-root», introducida en release 16 y ampliamente identificada como solución de futuro a implementar. La cabecera «3ggp-sbi-target-api-root» se utiliza para identificar el destino real del mensaje, en los escenarios en los que el mensaje se envía primeramente a una función intermedia que actúa como proxy (como puede ser el caso de SCP o SEPP). De este modo, el URI objetivo identificado por la pseudo-cabecera «:authority» identifica el destino inmediato del mensaje (el SCP o el SEPP por ejemplo), mientras que la cabecera «3ggp-sbi-target-api-root» identifica el servicio de NF objetivo final del mensaje.

5.3.2.1 Comunicación directa vs. comunicación indirecta

Las comunicaciones entre NFs dentro del núcleo de red 5G se pueden realizar según cuatro modelos, como veíamos en la sección 4.1.2.2. Estos cuatro modelos pueden coexistir, y a grandes rasgos podemos distinguir entre comunicación directa (donde NF consumidora y NF productora se interconectan directamente entre ellas), y comunicación indirecta donde el SCP centralizará estas comunicaciones. Estos dos enfoques se identifican en la figura 5.8 [57].

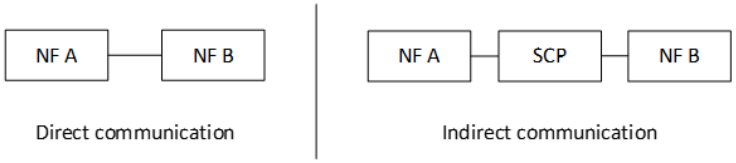

Figura 5.8: Comunicaciones entre servicios NF a NF.

5.3.2.2 Enrutamiento basado en el SCP

En esta sección nos centraremos de manera específica en las capacidades de enrutamiento ofrecidas por el SCP como elemento central de routing en el núcleo de red 5G.

Un SCP puede ser conocido por el NF (pre-configurado) o no. Si la NF conoce el SCP, la NF se configurará con un esquema, autoridad y, opcionalmente, un prefijo específico de implementación del SCP. El esquema puede ser «http» o «https». Si el esquema es «https», la autoridad deberá contener un FQDN y no una dirección IP literal. Si el esquema es «http», la autoridad podrá contener un FQDN o una dirección IP literal. En cualquier caso, la autoridad puede contener opcionalmente un número de puerto.

Los modelos de comunicación indirecta deben soportar el mismo nivel de seguridad que los de comunicación directa. Se utilizará TLS entre el SCP y las NF, si la seguridad de la red no se proporciona por otros medios. En la ruta de comunicación entre un consumidor de servicios NF (cNF) y un productor de servicios NF (pNF) pueden estar presentes 1 o más SCPs.

Se deben establecer conexiones HTTP(S) separadas entre el cliente HTTP (cNF) y el SCP, entre los diferentes SCPs (si hay más de un SCP en la ruta de comunicación entre cNF y pNF), y entre el SCP y el servidor HTTP (pNF).

La NF que actúa como un cliente HTTP/2, se conectará estableciendo (o reutilizando) una conexión a un SCP disponible para enviar una solicitud HTTP/2. Al reenviar una solicitud a otro SCP, el SCP se conectará estableciendo (o reutilizando) una conexión con el SCP del próximo salto. Cuando el SCP reenvía la solicitud al servidor HTTP, el SCP (que actúa como un cliente HTTP/2) se conectará a un servidor de autoridad para el recurso de destino. Para conectarse a una autoridad que no esté en la misma PLMN, el SCP se conectará al SEPP.

Para comunicaciones indirectas, con o sin descubrimiento delegado, al enviar una solicitud al SCP, el cliente HTTP establecerá las pseudo-cabeceras de la siguiente manera:

- **:scheme** establecido en «http» o «https».
- **:authority** conteniendo el FQDN o dirección IP del SCP (si el esquema es «http»), o el FQDN del SCP (si el esquema es «https»).
- **:path** incluyendo los componentes de ruta y parámetros de consulta (*query parameters*) del URI de destino.

El cliente HTTP incluirá la *apiRoot* de un servidor de autoridad para el recurso de destino, si está disponible, en la cabecera «3gpp-Sbi-Target-apiRoot».

Al reenviar una solicitud a otro SCP, el SCP reemplazará el *apiRoot* del SCP recibido en el URI de solicitud del mensaje entrante, por el *apiRoot* del SCP del siguiente salto. El SCP incluirá una cabecera «3gpp-Sbi-Target-apiRoot» conteniendo el *apiRoot* de un servidor de autoridad para el recurso de destino, si se recibió la cabecera «3gpp-Sbi-Target-apiRoot» en la solicitud original. El SCP fijará las pseudo-cabeceras, modificando el campo «:authority» al FQDN o dirección IP del SCP de siguiente salto (si el esquema es «http»), o al FQDN del SCP (si el esquema es «https»).

Al enviar una solicitud al servidor HTTP, el SCP deberá reemplazar el *apiRoot* del SCP recibido en el URI de la petición del mensaje entrante, por el *apiRoot* de la instancia de NF objetivo. Si la cabecera «3gpp-Sbi-Target-apiRoot» se recibió en la solicitud original, el SCP deberá utilizarla como el *apiRoot* de la instancia de NF objetivo en el caso de que el SCP no re-seleccione un servidor HTTP diferente.

En cualquier caso, el SCP deberá eliminar esta cabecera del mensaje reenviado y finalmente entregado al servidor HTTP. El SCP fijará las pseudo-cabeceras, modificando el campo «:authority» al FQDN o dirección IP de la instancia de NF objetivo (si el esquema es «http»), o al FQDN de la instancia de NF objetivo (si el esquema es «https»).

Para comunicaciones indirectas con o sin descubrimiento delegado, el cliente HTTP debe incluir la cabecera «3gpp-Sbi-Target-apiRoot» establecida en el apiRoot de un servidor de autoridad para el recurso de destino, si está disponible, en las solicitudes que envía al SCP.

5.3.2.3 Comunicación indirecta sin descubrimiento delegado

La comunicación entre NF consumidora y NF productora en el escenario en que se utiliza un SCP para ofrecer comunicación indirecta, pero no para el descubrimiento de NF productora ; se produce según el diagrama de la figura 5.9 [74].

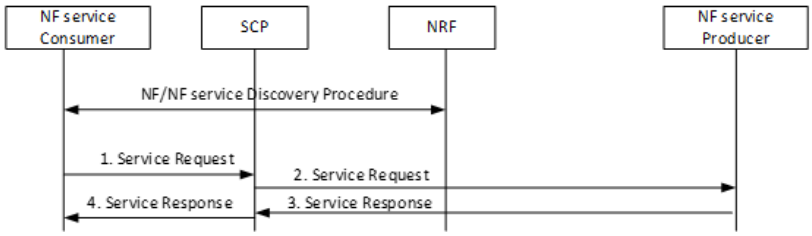

Figura 5.9: Procedimiento de comunicación indirecta sin uso de descubrimiento delegado.

- Para la comunicación indirecta sin descubrimiento delegado, una solicitud de servicio enviada al SCP para crear un recurso debe incluir un encabezado «3gpp-Sbi-Target-apiRoot» establecido en la *apiRoot* de la instancia de servicio NF seleccionada del productor de servicio NF, cuando el consumidor de servicio NF ha seleccionado una instancia de servicio NF específica.

- Después de que se haya creado un recurso, las solicitudes de servicio subsiguientes enviadas al SCP y dirigidas al recurso, incluirán un encabezado «3gpp-Sbi-Target-apiRoot» establecido en el *apiRoot* recibido anteriormente en la cabecera de ubicación (*Location*) de las respuestas de servicio del productor de servicios de NF.

- Las notificaciones o devoluciones de llamada (*callbacks*) enviadas a través del SCP incluirán un encabezado «3gpp-Sbi-Target-apiRoot» establecido en la *apiRoot* de la URI de notificación o devolución de llamada (es decir, esquema «http» o «https», la cadena fija «://» y autoridad (*host* y puerto opcional).

En este escenario, el SCP no interviene en el proceso de descubrimiento de pNF, para el que la cNF se comunica directamente con el NRF. Sin embargo, una vez el cNF ya ha descubierto los perfiles de pNF ofrecidos por el NRF, y ha efectuado la selección del pNF objetivo al que enviar sus peticiones de servicio; enviará estas peticiones al SCP, que se encargará del enrutamiento del mensaje hacia el pNF. En esta petición enviada desde la cNF al SCP, el cNF incluirá la dirección del pNF (FQDN o IP) en la cabecera «3gpp-sbi-target-apiroot», mientras que la pseudo-cabecera «:authority» contendrá la dirección (FQDN o IP) del SCP.

5.3.2.4 Comunicación indirecta con descubrimiento delegado

La comunicación entre NF consumidora y NF productora en el escenario en el que se utiliza un SCP para ofrecer comunicación indirecta, pero también para ofrecer descubrimiento de NF productora delegado, se produce según el diagrama de la figura 5.10 [74].

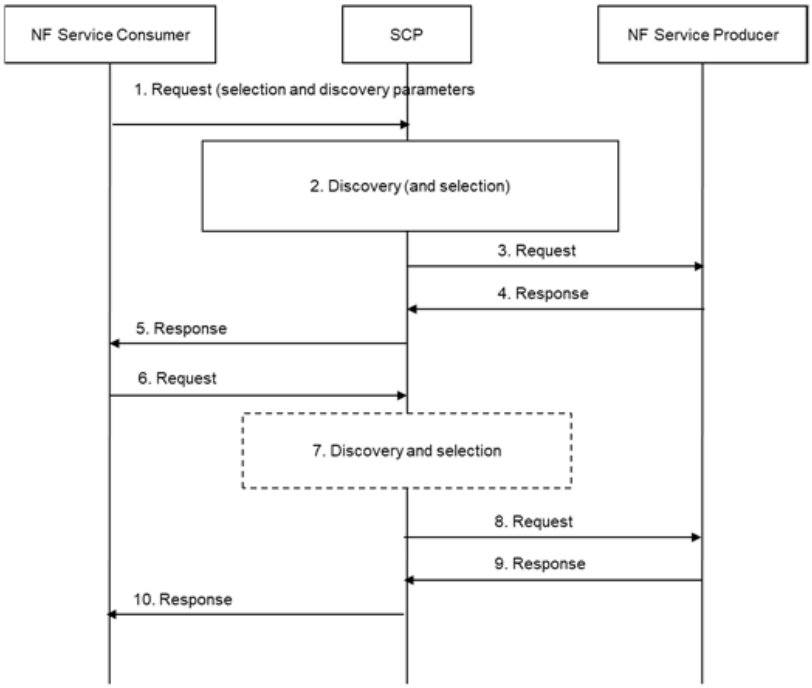

Figura 5.10: Descubrimiento delegado dentro de la misma PLMN.

1. El consumidor del servicio NF tiene la intención de comunicarse con un productor de servicio NF, y envía la solicitud de servicio a un SCP. La solicitud puede incluir parámetros de descubrimiento y selección necesarios para descubrir y seleccionar una instancia de productor de servicios NF. Los parámetros de descubrimiento y selección se incluyen en la solicitud del consumidor del servicio NF mediante cabeceras, de manera que el SCP no necesita analizar el cuerpo de la solicitud.

2. El SCP puede realizar el descubrimiento delegado ya sea interactuando con un NRF (utilizando el servicio Nnrf_NFDiscovery) o puede usar la información recopilada durante las interacciones anteriores con un NRF (mediante la operación del servicio Nnrf_NFDiscovery o Nnrf_NFManagement_NFStatusNotify). El SCP junto con el NRF autoriza la solicitud. El SCP selecciona el productor de servicios NF de destino. Si los parámetros de descubrimiento y selección en la solicitud incluyen una identidad de UE específica, por ejemplo SUPI o IMPI/IMPU, el SCP puede resolver el identificador de grupo de NF correspondiente a la identidad del UE y luego invocar el servicio Nnrf_NFDiscovery.

3. Si el consumidor de servicios de NF está autorizado a comunicarse con el productor de servicios de NF, el SCP reenvía la solicitud al productor de servicios de NF seleccionado de acuerdo con la configuración de *Network Slice*. Por ejemplo, controlará que las instancias de NF solo son accesibles por NFs en el mismo segmento de red.

4. El productor de servicios NF envía una respuesta al SCP. Si la solicitud crea un recurso en el productor de servicios NF, el productor de servicios NF responde con información de recursos que identifica el recurso creado.

5. El SCP enruta la respuesta al consumidor del servicio NF. Si el consumidor del servicio NF recibe una dirección de recurso, la utiliza para solicitudes posteriores relacionadas con el recurso en cuestión. De lo contrario, el procedimiento termina aquí.

6. En una operación posterior para el recurso creado, el consumidor del servicio NF se dirige al recurso a través de la dirección del recurso devuelta por el productor del servicio NF en el paso 4.

7. El SCP resuelve la dirección del productor de servicios NF y selecciona una instancia de productor de servicios NF de destino. Luego, el SCP enruta la solicitud a la instancia del productor de servicios NF seleccionada.

8. El SCP entrega la solicitud al productor de servicios NF.

9. El productor de servicios NF envía una respuesta al SCP. El productor de servicios de NF puede responder con una información de recursos actualizada diferente a la utilizada en la respuesta anterior.

10. El SCP envía una respuesta al consumidor del servicio NF. Si se actualizó la información del recurso, el consumidor del servicio NF utiliza la información del recurso recibida para solicitudes posteriores en el mismo recurso. De manera similar al manejo de la información de recursos, el productor del servicio NF también puede proporcionar una indicación vinculante y utilizarla para las solicitudes posteriores del consumidor del servicio NF.

En función de la configuración del SCP, un SCP que decida dirigirse a un SCP de siguiente salto para una solicitud de servicio puede delegar la selección de instancia de NF y/o instancia de servicio a SCP posteriores y proporcionar los parámetros de descubrimiento y selección al SCP de siguiente salto.

5.3.2.5 Vinculación consumidor-productor entre NFs

Junto con la introducción del SCP en la release 16, se introdujo también un mecanismo de vinculación entre NF consumidora y NF productora para mejorar la eficiencia de la arquitectura basada en servicios.

Cuando el consumidor del servicio NF se comunica con el productor del servicio NF, el productor puede devolver una indicación vinculante al consumidor. El consumidor almacena la indicación vinculante recibida y la utiliza para las solicitudes posteriores relacionadas con el contexto de los datos, como se muestra en la figura 5.11.

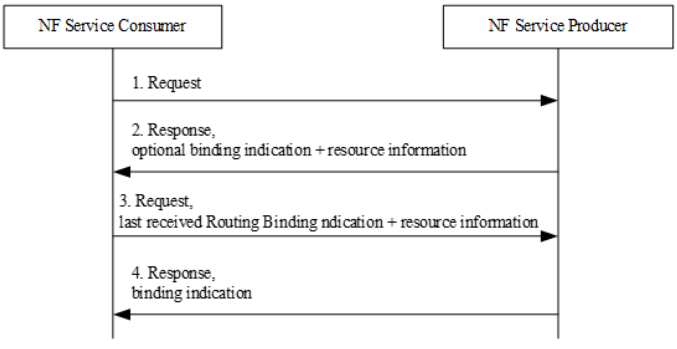

Figura 5.11: Vinculación consumidor-productor entre NFs.

1. La solicitud inicial entre cNF y pNF puede producirse de diferentes maneras:

- **Si se utiliza la comunicación directa**, el consumidor del servicio NF selecciona el productor del servicio NF y envía la solicitud al productor del servicio NF seleccionado.
- **Si se utiliza la comunicación indirecta sin descubrimiento delegado**, el consumidor del servicio NF selecciona el conjunto o la instancia del productor del servicio NF y envía la solicitud al productor del servicio NF seleccionado a través del SCP. Si el consumidor del servidor NF solo selecciona el conjunto de productores de servicios NF, entonces proporciona los parámetros de selección necesarios y el SCP selecciona la instancia del productor de servicios NF específica.
- **Si se utiliza la comunicación indirecta con descubrimiento delegado**, el consumidor del servicio de NF envía la solicitud al SCP y proporciona dentro de la solicitud de servicio al SCP los parámetros de descubrimiento y selección necesarios para descubrir y seleccionar un productor de servicios de NF.

2. El productor del servicio NF envía una respuesta al consumidor del servicio NF. En la respuesta, el productor del servicio NF puede incluir una indicación vinculante. Si el consumidor del servicio NF recibe una información de recursos y una indicación de vinculación, las utiliza para solicitudes posteriores relacionadas con el recurso en cuestión. De lo contrario, el procedimiento termina aquí.

3. El consumidor del servicio NF utiliza la indicación de vinculación y la información de recursos recibida en el paso anterior para solicitudes posteriores relacionadas con el recurso en cuestión. Si se utiliza comunicación indirecta con descubrimiento delegado, el consumidor del servicio NF incluye una indicación de vinculación de enrutamiento con el mismo contenido que la indicación de vinculación recibida. Si se utiliza la comunicación indirecta sin descubrimiento delegado, el consumidor del servicio NF también incluye la indicación de vinculación de enrutamiento con el mismo contenido que la indicación de enlace recibida, a menos que el consumidor del servicio NF realice una nueva selección. El SCP encaminará la solicitud de servicio utilizando la indicación de vinculación de enrutamiento y la información de recursos enviada desde el consumidor del servicio NF.

4. El productor de servicios NF envía una respuesta al consumidor. El productor del servicio NF puede responder con una indicación vinculante actualizada, diferente a la utilizada en la respuesta anterior.

Si un SCP recibe una indicación de vinculación de enrutamiento dentro de una solicitud de servicio o notificación y decide reenviar esa solicitud a un SCP de siguiente salto, deberá también incluir la indicación de vinculación de enrutamiento en la solicitud reenviada.

5.4 Nuevas capacidades del 5G Core

5.4.1 Seguridad por Diseño

Las redes precedentes 2G/3G/4G presentaron diferentes problemas de seguridad, que han querido ser solventados en 5G con la introducción del concepto de seguridad por diseño (*Security by Design*). Se produce un cambio de paradigma en 5G. Mientras que generaciones anteriores distinguían entre entornos «confiables» y «no confiables», en 5G se considera que toda comunicación es potencialmente no confiable, incluso comunicaciones internas. Es por tanto que se requiere el concepto de seguridad por diseño en la que toda comunicación está securizada por defecto.

Las capacidades de seguridad del sistema 5G incluyen:

- Autenticación del usuario (UE) por parte de la red y viceversa (autenticación mutua entre red y UE).
- Generación y distribución de contexto de seguridad, de manera que cada UE mantiene un contexto específico con la red de acceso radio.
- Confidencialidad y protección de la integridad en los datos del plano de usuario, basada en políticas de seguridad específicas para las sesiones *Protocol Data Unit (PDU)*.
- Confidencialidad y protección de la integridad en la señalización del plano de control, con la introducción de TLS con autenticación mutua para las comunicaciones de señalización de control entre NFs.
- Confidencialidad de la identidad de usuario, con la introducción de un nuevo identificador de usuario cifrado SUCI.
- Soporte a los requisitos de *LI* sujetos a requisitos de regulaciones regionales o nacionales, incluyendo protección de los datos LI almacenados o transferidos por una NF.

Dado el foco de este trabajo, nos centraremos en los aspectos específicos a la señalización en el plano de control del núcleo de red 5G.

5.4.1.1 TLS mutuo en el plano de control

La señalización en el núcleo de redes 5G introduce una novedad significativa en comparación con las redes anteriores 2G/3G/4G en la relativo a la seguridad de las comunicaciones. 3GPP indica que todas las funciones de red deben soportar TLS con autenticación mutua y HTTPS para comunicarse a través de *SBI*. Si bien en estas redes anteriores la señalización entre elementos de red se realiza sin cifrar, 5G incorpora por primera vez el uso de TLS. Las redes 4G Diameter también describían la utilización de TLS como una opción soportada, pero no se utilizó en ninguna implementación real, y tampoco en ningún escenario de interconexión de roaming entre operadoras. Introducir TLS supone un cambio significativo en la operativa de estas redes. La monitorización de las redes por parte de los equipos de operaciones, estaba basada en el acceso a una copia de esta señalización recogida mediante sondas y/o sistemas de monitorización basados en tecnología de «*port-mirroring*» desde los enrutadores IP de la red. Éste es uno de los motivos fundamentales por los que TLS no fue utilizado en 4G. Suponía también cambiar por completo el diseño de los sistemas de monitorización existentes, mientras que podían utilizar la alternativa TCP/SCTP que también era soportada por el estándar 4G.

Sin embargo, en 5G el uso de TLS es obligatorio en interfaces de interconexión entre operadores (entre SEPPS), y también lo es en el núcleo local de red de una PLMN en caso de que este cifrado no sea ofrecido por otros medios. Esto también implica que los sistemas de monitorización de redes precedentes basados en sondas y «*port-mirroring*» a nivel de *router* IP ya no son válidos y se necesita que la propia función de red que termina la conexión TLS deba ser la encargada de facilitar la información de la señalización transitada para ser monitorizada, o de enviarla a sistemas de monitorización externos. Es por tanto que las funciones de red como el SCP (que centraliza todo el tráfico de señalización dentro de la red nacional), o el SEPP (que centraliza todo el tráfico de interconexión y roaming), son un punto idóneo para realizar esta función y facilitar la monitorización y operativa de la red.

El uso de TLS en 5G para la señalización HTTP/2, requiere la utilización de TLS mutuo en el que no solo es necesario autenticar al servidor (como ocurre generalmente con HTTP e Internet), sino que esta autenticación debe ser mutua y bidireccional, con el servidor autenticando también al cliente. Para

el uso de HTTP2 sobre TLS en redes 5G, es necesario también la utilización del parámetro TLS *Server Name Indication (SNI)* como extensión TLS, que se enviará como parte del mensaje TLS *Client Hello* y debe incluir la FQDN del servidor, es decir la FQDN de la NF objetivo. Del mismo modo, solo son soportadas las versiones TLS v1.2 y v1.3, no estando permitidas versiones anteriores del protocolo TLS. El flujo de mensajes en la negociación TLS (*handshake*) [75] se muestra en la figura 5.12, donde ambas partes enviaran sus certificados para obtener un escenario de autenticación mutua.

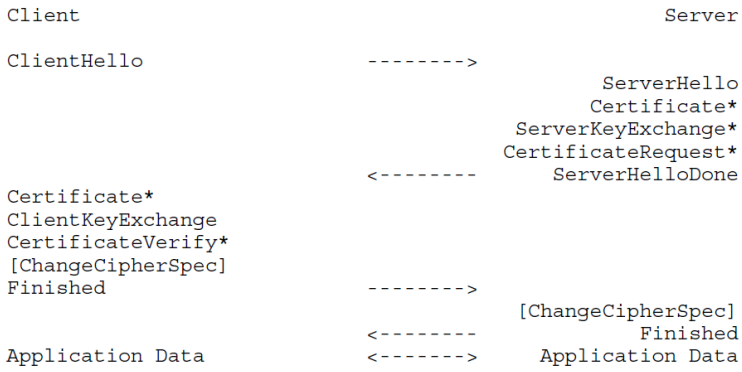

Figura 5.12: Flujo de mensajes de la negociación TLS (handshake).

Esto implica que las funciones de red 5G deben manejar el almacenamiento y gestión de certificados, claves privadas y claves públicas. Necesitan también la integración con una Autoridad de Certificación (*Certification Authority*) que se encargue de firmar sus certificados TLS tanto para el despliegue inicial como para la renovación de estos certificados cuando se aproxima su fecha de expiración. Es necesario gestionar el conjunto de autoridades de certificación confiables, a la hora de interaccionar con funciones de red de otras PLMNs (en escenarios de roaming o interconexión entre vSEPP y hSEPP), en los que cada PLMN utilizará Autoridades de Certificación diferentes.

5.4.1.2 SUPI y SUCI

El principal identificador de usuario ampliamente utilizado en generaciones anteriores es el *International Mobile Subscriber Identity (IMSI)*, que contiene 15 o 16 dígitos decimales donde los 5 o 6 primeros dígitos identifican el *MCC* y el *MNC* de la operadora PLMN a la que pertenece el usuario. En 5G se introduce el concepto de identificador *SUPI*, que también es asignado a cada subscriptor en el sistema 5G y está provisionado en el UDM/UDR. Se utiliza dentro del sistema 5G, aplicándose seguridad y privacidad específica sobre él. El SUPI puede contener diferentes tipos de parámetros aunque principalmente contendrá un IMSI (en escenarios de redes móviles 3GPP). En escenarios de redes privadas podrá contener un identificador específico de red *Network Access Identifier (NAI)*, en escenarios de redes fijas residenciales de cable un *Global Cable Identifier (GCI)*, y en escenarios de de redes fijas residenciales de banda ancha contendrá un *Global Line Identifier (GLI)*.

En cualquier caso, cuando un usuario (UE) necesita indicar su SUPI a la red (por ejemplo como parte del procedimiento de registro), el UE proporciona este SUPI en forma oculta. Para permitir los escenarios de roaming, la parte relacionada con el identificador de red doméstica HNI (el MCC y el MNC en los casos de SUPI basado en IMSI) debe ser visible. Sin embargo, los demás dígitos del SUPI irán cifrados en modo oculto. Esta versión oculta y cifrada del SUPI se denomina *SUCI*, y es un identificador con mantenimiento de privacidad, usada para transmitir la versión oculta del SUPI.

Esto previene el uso de *IMSI-catchers* en interfaces radio que afectan a redes 2G/3G/4G en los que se intenta obtener identificadores de usuario IMSI no cifrados, mediante los que se podía identificar, localizar

y monitorizar a usuarios. En estas redes se intentaba minimizar este impacto utilizando identificadores temporales (en lugar del IMSI como identificador permanente), pero en determinados escenarios esto no era posible (como en el registro inicial de usuario donde éste aun no tiene identificador temporal asignado). Estas situaciones son evitadas en 5G con el uso de la pareja SUPI-SUCI.

Un SUPI conteniendo un IMSI tendrá el formato de la figura 5.13.

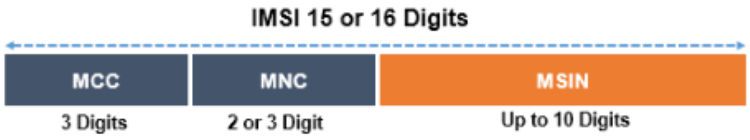

Figura 5.13: Formato de SUPI conteniendo un IMSI.

Un SUCI contiene el formato de la figura 5.14.

Figura 5.14: Formato de SUCI.

5.4.1.3 Autorización en la comunicación NF-NF

El descubrimiento y el registro de servicios de NF deben ofrecer confidencialidad, integridad y protección de reproducción. El NRF debe garantizar que las solicitudes de descubrimiento y registro de NF estén autorizadas. El descubrimiento y el registro de servicios de NF deben ocultar la topología de las NF disponibles y soportadas en un dominio administrativo (o dominio de confianza) a las entidades en diferentes dominios de confianza (por ejemplo, entre las NF de las redes visitadas y domésticas).

El procedimiento de solicitud y respuesta de servicios de NF debe soportar la autenticación mutua entre el consumidor de servicios de NF y el productor de servicios de NF. Cada NF validará todos los mensajes entrantes. Los mensajes que no sean válidos de acuerdo con la especificación del protocolo y el estado de la red serán rechazados o descartados por la NF.

En el NRF, se mantienen los perfiles de NF de la red y se reciben solicitudes de descubrimiento de NF por parte de instancias de NF, proporcionando la información de las instancias de NF detectadas. El NRF debe aplicar requisitos de seguridad cuando se recibe la solicitud de descubrimiento de NF de una instancia de NF, antes de proporcionar la información de las instancias de NF detectadas a la instancia de NF solicitante. Por un lado el NRF y las NF que solicitan el servicio se autenticarán mutuamente, y por otro el NRF podrá proporcionar autenticación y autorización a las NF para establecer una comunicación segura entre ellas.

El SCP tiene interfaces con funciones de red (NF) y SCP pares dentro de la PLMN. La interfaz entre el SCP y las NF y entre los dos SCP deberá ofrecer autenticación mutua, y todas las comunicaciones entre el SCP y las NF y entre los SCP estarán protegidas por confidencialidad, integridad y protección de reproducción.

La autorización del servicio de NF garantiza que el consumidor del servicio de NF esté autorizado a acceder al servicio de NF proporcionado por el proveedor de servicios de NF, de acuerdo, por ejemplo, con la política de NF, la política del operador de servicio o el acuerdo entre operadores en escenarios de roaming.

La información de autorización del servicio se configura como uno de los componentes en el perfil NF del productor del servicio NF. Incluirá los tipos de NF y los dominios u orígenes de NF autorizados para consumir los servicios de NF del Productor de Servicios de NF.

Debido a los acuerdos de roaming y las políticas del operador, un consumidor del servicio NF debe ser autorizado en función de la información de UE/suscriptor/roaming y el tipo de NF. La autorización del servicio consistirá en:

- Comprobar si el consumidor de servicios de NF tiene permiso para descubrir la instancia de productor de servicios de NF solicitada. Esto ocurrirá durante el procedimiento de descubrimiento de servicios de NF recibido por parte del NRF y analizado para cada NF.
- Comprobar que el consumidor de servicios de NF tiene permiso para acceder al productor de servicios de NF solicitado para consumir el servicio de NF, para cada solicitud. Esta autorización se integra dentro de la lógica del servicio específico en función de UE, suscripción o acuerdos de roaming.

El flujo de comunicación relacionado con los aspectos de autorización [67] para los modelos de comunicación indirecta se ilustra en la figura 5.15.

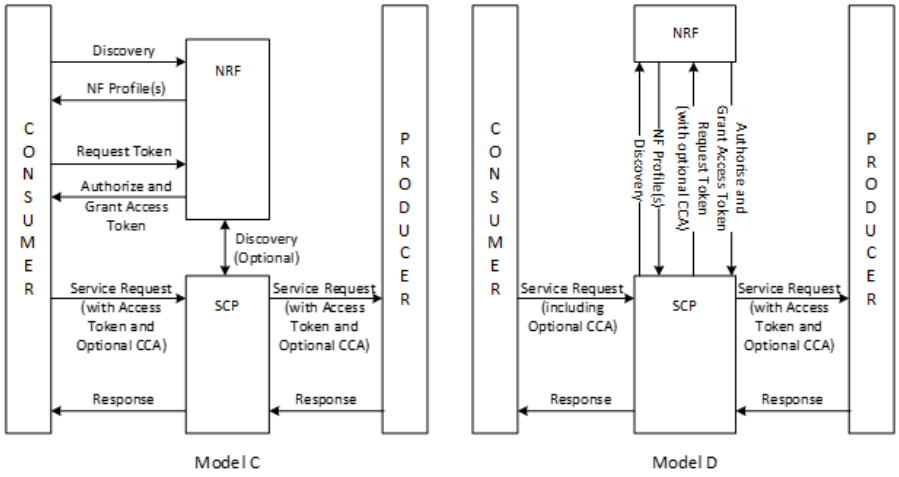

Figura 5.15: Aspectos relacionados con la autorización en los modelos de comunicación indirecta.

En la petición de descubrimiento de pNF inicial enviada por el cNF (o por el SCP en el caso de descubrimiento delegado del modelo D), el NRF solo devolverá el listado de pNF que cumplen el criterio de selección establecido en la petición original, si estos perfiles están autorizados para ser descubiertos por la cNF solicitante, en base a diferentes parámetros como identificador de PLMN, tipo de NF, «*slice*» de red, etc.

Por otro lado, después de haber descubierto el perfil de la pNF objetivo, la función cNF o el SCP (en el caso del modelo D), solicitarán al NRF un token de acceso de tipo OAuth2 2.0 para utilizar en la comunicación con dicho pNF. El NRF autorizará dicha comunicación y ofrecerá el token que será utilizado por el cNF (o por el SCP) en las subsiguientes comunicaciones y peticiones de servicio enviadas al pNF. Dichas peticiones de servicio no serán atendidas por el pNF si no incorporan un token de acceso correcto.

5.4.2 *Slicing* de red

Una novedad fundamental en el núcleo de red 5G es el soporte a *slicing* de red. Esto permite que mediante el uso de *slices* lógicas, la red puede ser dividida en subredes independientes con diferentes capacidades,

recursos, requisitos y casos de uso soportados. Esto puede permitir a las operadoras mantener una única red 5G real, pero dividirla en múltiples redes lógicas de manera que pueda ofrecer *slices* específicos dedicados a empresas, separar la red móvil de abonados de la red de comunicaciones IoT o M2M, etc.

La selección del conjunto de instancias de *Network Slice* para un UE es llevada a cabo por el primer AMF contactado durante el procedimiento de registro, interactuando con el NSSF. Esto puede conducir a un cambio de AMF si el utilizado inicialmente no es compatible con el *Network Slice* a utilizar. Una *Network Slice* concreta se identifica por el parámetro S-NSSAI.

La configuración de *Network Slice* para un UE contiene uno o más *NSSAI* configurados. La operadora doméstica puede configurar un NSSAI para algunas vPLMN específicas, o puede definir un NSSAI configurado por defecto que se aplicará en cualquier vPLMN para la cual no se haya proporcionado ninguna NSSAI específica en el UE.

Con la introducción de *Network Slicing* en 5G, se introduce una nueva función de red sin equivalentes en generaciones anteriores que es el *NSSF*. El NSSF interactúa con las diferentes funciones de red para seleccionar e identificar el *Network Slice* a utilizar para cada usuario en cada momento. Para un UE concreto, el NSSF selecciona el conjunto de instancias de segmentos de red (*Network Slices*) que pueden ser utilizadas para al UE. Manejará la información para determinar el *NSSAI* permitido o configurado, así como el mapeo a los *S-NSSAI* suscritos. El NSSF también determina el conjunto de AMFs que se usará para servir al UE o, en base a la configuración, una lista de AMF(s) candidatos, posiblemente consultando al NRF.

Pese a ser una de las áreas que mayores posibilidades ofrece, es cierto que las operadoras que comienzan a desplegar sus redes 5G-SA aun mantienen los aspectos relacionados con *Network Slicing* en fases de experimentación, con el objetivo de identificar los mejores casos de uso y modelos de negocio detrás de ellos. Parece que pueden estar principalmente enfocados a empresas y organizaciones que tengan requisitos específicos de red como valores mínimos de capacidad o tiempo de conexión garantizado, incluso en escenarios de congestión.

A fecha de realización de este trabajo, en Noviembre de 2022, los despliegues de *Network Slicing* están enfocados a pruebas experimentales, tratando de identificar nuevas fuentes de ingresos para las operadoras, con algunos ejemplos en operadoras de Holanda [76], Finlandia [77] y Kenya [78]; o entre suministradores de red como Nokia y Ericsson con Google como suministrador de dispositivos Android [79].

5.4.3 Infraestructura Cloud y aplicaciones basadas en microservicios

Si bien esta sección no trata específicamente de las funciones de red 5G, sí tiene una influencia fundamental en ella al tratar sobre la infraestructura en la que estas redes son desplegadas. En las redes 2G/3G se utilizaban principalmente despliegues con Hardware dedicado en cada una de las funciones de red. Los proveedores de equipamiento de red no solo producían el Software sino también un Hardware específico con las características y recursos específicos que cada elemento de red requería.

Parcialmente con 3G y sobretodo con 4G, se empezaron a utilizar cada vez más servidores hardware de propósito general o *Commercial Off-The-Shelf (COTS)*, y se popularizaron los despliegues virtualizados. Éstos ya eran comunes en el sector de las tecnologías de la información (IT), pero aún no habían sido utilizados en el sector de las telecomunicaciones, donde se denominan *Network Function Virtualization (NFV)*. Las operadoras comenzaron a desplegar sus infraestructuras virtualizadas privadas *Network Function Virtualization Infrastructure (NFVI)* principalmente basadas en dos entornos: *OpenStack* (como

opción de software libre y código abierto bajo términos de licencia Apache) y *VMWare ESXi* (como alternativa privada). Gran parte de las funciones del núcleo de red 4G fueron desplegadas por las operadoras como *Virtual Network Function (VNF)*.

Al evolucionar desde 4G hacia 5G se introduce un cambio abrupto en este área al requerir despliegues de tipo «*Cloud Native*» basados en microservicios, para las funciones de red 5G, es decir, despliegues *Containerized Network Function (CNF)*. Al igual que ocurrió con la virtualización en 4G, esta tecnología está ampliamente extendida en el sector IT, pero aún no ha sido utilizado en el sector de las telecomunicaciones para funciones del núcleo de red. El software de estas aplicaciones de red se despliega en contenedores, sobre un entorno de orquestación *Cloud* como *Kubernetes*.

Kubernetes, comúnmente conocido como k8s, es una plataforma de sistema distribuido de código libre para la automatización del despliegue, escalado y manejo de aplicaciones en contenedores que fue originalmente diseñado por Google y donado a la *Cloud Native Computing Foundation (CNCF)* (parte de la *Linux Foundation*). Soporta diferentes entornos para la ejecución de contenedores, y ha sido ampliamente identificado como el estándar de facto para el desarrollo de funciones de red en entorno *Cloud* para el núcleo de red 5G. Del mismo modo, es también el modelo escogido en los casos en los que las operadoras desean actualizar o renovar elementos de red de generaciones precedentes para los que también quieren migrara a un despliegue CNF.

Ante esta situación, muchas operadoras están desplegando sus propias infraestructuras *Cloud* privadas, ofreciendo a los diferentes proveedores de funciones de red un modelo de despliegue *Containers as a Service (CaaS)*, en el que el proveedor de funciones de red ofrecerá sus contenedores software; para ser organizados, ejecutados, escalados y gestionados por la virtualización basada en contenedores ofrecida por la infraestructura *Cloud*. En estos escenarios los nodos *host* sobre los que se realiza el despliegue del *cluster Kubernetes* pueden ser máquinas virtuales (contenedores sobre VMs), o servidores hardware físicos COTS (contenedores sobre «*bare metal*») en función del modelo preferido por la operadora como proveedora de la infraestructura *Cloud*.

Por otro lado, hay operadoras que realizaron una fuerte inversión en su infraestructura virtualizada NFVI, y prefieren continuar amortizándola sin desplegar infraestructura *Cloud* propia; pero aun así dar el salto requerido al despliegue de funciones de red 5G de tipo CNF. En estos casos, las operadoras piden a los proveedores de funciones de red que además de la función de red basada en contenedores, aporten también la capa Kubernetes y los componentes Cloud necesarios, para ser desplegados sobre maquinas virtuales *Virtual Machine (VM)* ofrecidas por la operadora desde su infraestructura NFVI. Se trata de un despliegue de contenedores sobre VMs, con Kubernetes proporcionado por el proveedor de la NF, y desplegando un *cluster Kubernetes* específico y dedicado para dicha NF.

También hay operadoras que no poseen o no desean utilizar su infraestructura (virtualizada o *cloud*) y requieren al proveedor un despliegue de contenedores sobre servidores físicos COTS (contenedores sobre «*bare metal*»), en el que el proveedor de la NF también proporciona la capa *Kubernetes* a desplegar sobre un conjunto de servidores COTS dedicado a dicho *cluster Kubernetes* utilizado específicamente para esta NF.

Por último, las operadoras también están explorando el despliegue de funciones de red 5G basadas en contenedores sobre infraestructura *cloud* pública, donde en lugar de desplegar sus propias infraestructuras privadas, utilizan los servicios de infraestructura ofrecidos por empresas como *Amazon Web Services (AWS)*, *Microsoft Azure* o *Google Cloud Platform (GCP)*. En algunos casos se analiza el modelo *Cloud* híbrido, donde algunas NFs (o algunas instancias de una NF concreta) están desplegadas en infraestructura *Cloud* pública, mientras que otras lo están en infraestructura *Cloud* privada. En cualquier caso, estos potenciales despliegues son analizados cuidadosamente, puesto que los proveedores de *Cloud*

pública deben entender y soportar la criticidad de los servicios de telecomunicaciones y los potenciales impactos causados por degradaciones serias del servicio o por cortes del servicio. Es fundamental identificar y acordar la propiedad de la responsabilidad sobre parámetros como tiempo de disponibilidad, resiliencia o seguridad de la infraestructura subyacente a nivel de hardware, sistema operativo, motor de contenedores o servicio de orquestación de contenedores (Kubernetes); incluyendo servicios de plataforma críticos como descubrimiento de servicios *Kubernetes*, conectividad de red o almacenamiento. Lo mismo aplica en aspectos de seguridad y protección de datos sensibles en concordancia con las regulaciones legales existentes.

El despliegue real de 5G SA a nivel mundial lleva cierto retraso con respecto a sus planes iniciales. Son múltiples los motivos incluyendo evidentemente todo lo relacionado con el periodo de pandemia global de coronavirus de los últimos años. Aún así, uno de los motivos fundamentales es también la disponibilidad de infraestructuras *Cloud* por parte de las operadoras, que supone un paso fundamental previo al despliegue CNF de las diferentes funciones de red del 5GC, y que también es una tecnología completamente nueva para las operadoras.

5.4.3.1 Automatización de red

Este salto al mundo *Cloud*, supone por tanto un cambio radical en cuanto al desarrollo y despliegue de funciones de red en el sector de las telecomunicaciones. Introduce además nuevos conceptos de automatización como integración continua *Continuous Integration (CI)*, entrega continua *Continuous Delivery (CD)*, despliegue continuo *Continuous Deployment (CD)* y *testing* continuo *Continuous Testing (CT)*.

Los conceptos CI/CD están más enfocados a la parte del proveedor de la función de red, y a la manera en que los nuevos desarrollos software están disponibles para ser utilizados por la operadora. CI facilita a los desarrolladores de código el acceso a un repositorio de código centralizado y confiable, permitiendo identificar y solucionar rápidamente posibles conflictos entre diferentes cambios de código. CD es la implementación de automatización del proceso completo de liberación de software. Esto incluye la manera en que este software publicado por el desarrollador se hace llegar al operador. Se utilizan repositorios de artefactos software para que las diferentes versiones de software pueden estar disponibles automáticamente para ser desplegadas.

Los conceptos CD/CT se centran en la parte de la operadora PLMN y en la manera en que las nuevas versiones software son desplegadas y probadas en entornos reales ya sean de pruebas o de producción. CD hace referencia al despliegue automatizado de software, y CT se centra en el testing automatizado del nuevo software.

Al igual que con las otras tecnologías mencionadas en esta sección, la adopción de la metodología CI/CD y CD/CT de la cultura *DevOps* por parte del sector de telecomunicaciones implica un cambio de paradigma significativo. La adopción de estos procesos se lleva a cabo de manera gradual, prefiriendo mantener control manual de ciertas acciones durante las primeras etapas de esta adopción.

En cuanto al despliegue de las funciones de red 5G, la tecnología *Cloud* permite la introducción de capacidades de escalado automatizado (*auto-scaling*), o de resolución de incidencias automatizadas (*auto-healing*) gracias a las capacidades de orquestación de contenedores ofrecida por *Kubernetes*. En base a los valores ofrecidos por ciertas métricas como por ejemplo uso de cpu, uso de memoria o transacciones por segundo procesadas; se pueden tomar decisiones de escalado automatizado para, por ejemplo, desplegar un nuevo contenedor de la función de red. De esto modo el dimensionado de la red es flexible, estando cubiertos ante escenarios de picos de tráfico puntuales o estacionales. De un modo similar, ante la detección de contenedores en mal estado (según sus indicadores de «salud»), se puede automatizar la decisión de re-desplegar el contenedor.

5.4.3.2 Software de código abierto. Plataforma como servicio

En general, el salto al mundo *Cloud* de la tecnología 5G implica la adopción de múltiples herramientas de código abierto y proyectos de la fundación *CNCF* para el desarrollo de funciones de red 5G, en concreto para sus requisitos relacionados con la plataforma software subyacente. Mientras que en tecnologías anteriores los proveedores de funciones de red solían desarrollar de modo privado la solución extremo a extremo, es común en las funciones de red CNF como las de 5G que el proveedor de NF se centre en la funcionalidad específica de la NF, apoyándose en muchas de estas herramientas de código abierto para diferentes aspectos generales. Estas herramientas se engloban dentro de diferentes categorías como recolección de métricas, logs y eventos (Prometheus, fluentd), visualización de métricas y eventos (Grafana, Kibana, Opensearch Dashboards), bases de datos (MariaDB, MongoDB, Elastic), replicación de datos (Kafka), gestión de certificados (cert-manager), etc.

En algunos casos, este enfoque de plataforma software común sobre la que se desplieguen las NFs específicas se ofrece como un servicio, dentro del modelo *Platform as a Service (PaaS)*, como plataforma común agnóstica con respecto a distintos proveedores, basada en especificaciones y estándares comunes y siendo más eficiente a nivel de costes.

5.5 5G SA Roaming

Si bien la arquitectura y procedimientos del núcleo de red 5G en lo relativo a señalización del plan de control son comunes para escenarios nacionales (intra-PLMN) o de itinerancia/roaming (inter-PLMN); en estos últimos es necesario tener en cuanta algunos aspectos adicionales tanto a nivel de enrutamiento como de seguridad.

Cuando una cNF localizado en la red visitada (vPLMN) pretende descubrir a una NF o a un servicio de NF en la red doméstica (hPLMN), el NRF de la red visitada enviará una petición «NF Discovery» al NRF de la hPLMN, en la que deberán especificarse tanto el identificador de vPLMN como el identificador de hPLMN [74].

En el caso ser el propio cNF quien realice el descubrimiento (en caso de no utilizar descubrimiento delegado), la comunicación sigue el flujo de la figura 5.16.

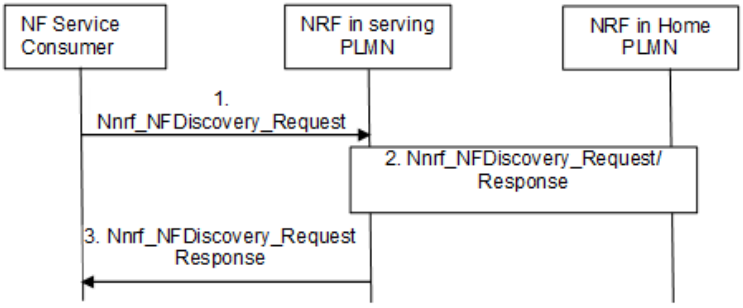

Figura 5.16: Descubrimiento de NF en escenario de roaming inter-PLMN.

Po otro lado, en el caso ser el SCP de la red visitada el que realiza el descubrimiento (en caso de utilizar descubrimiento delegado), la comunicación sigue el flujo de la figura 5.17. Es importante resaltar que el descubrimiento delegado siempre es realizado por el SCP de la red visitada, no por el SCP de la red doméstica.

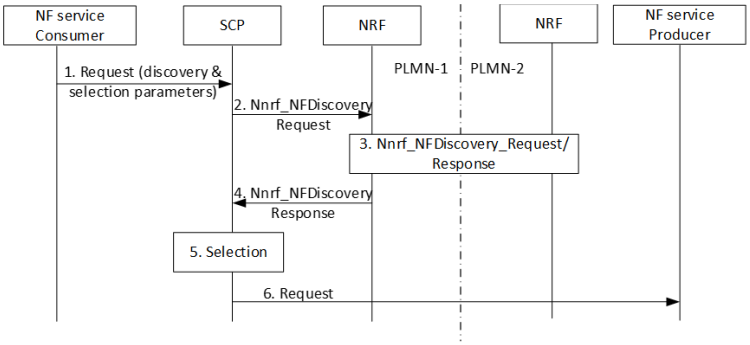

Figura 5.17: Descubrimiento delegado de NF en escenario de roaming inter-PLMN.

En este caso, la petición enviada al SCP debe contener los identificadores de PLMN originante y PLMN objetivo como parte de los parámetros de descubrimiento, necesarios para que el SCP pueda descubrir y seleccionar una instancia de pNF apropiada.

Por simplificación, estos diagramas no muestran al cSEPP ni al pSEPP, que están presentes en la frontera de ambas PLMNs. Toda comunicación entre ambas PLMNs se producirá siempre a través de los SEPPs y de la conexión N32 entre ellos descrita en la sección 4.1.2.4.

La comunicación entre PLMNs para escenarios de suscripciones o notificaciones a cambios en los perfiles de NFs, también presenta particularidades para adaptarse a la especial casuística de roaming.

Para alcanzar el servicio NF de destino correcto en la PLMN correcta, los mensajes de solicitud HTTP2 donde el componente de autoridad URI de destino (pseudo-cabecera *:authority*) designa un servidor que no está en la misma PLMN que el cliente; ésta deberá contener el FQDN (no una IP), incluyendo el ID de la PLMN objetivo, con este formato:

:authority = uri-host [: puerto], excluyendo la información [*userinfo "@"*].

El *uri-host* contendrá el FQDN del servicio NF de destino o la parte FQDN (autoridad) de un URI de devolución de llamada (*callback URI*), que en cualquiera de los dos casos deberán contener el identificador PLMN.

Para permitir la protección TLS entre el SEPP y las funciones de red dentro de una PLMN, el SEPP implementará una de estas dos opciones, basado en la política escogida por cada PLMN, e independientemente del método utilizado por la PLMN remota o en la interconexión entre ambas PLMNs:

- Certificado TLS de tipo comodín (*wildcard certificate*) para su nombre de dominio y generación de FQDN telescópico,.
- Reenvío de solicitudes HTTP2 originadas por NFs dentro de la PLMN del SEPP hacia la PLMN remota utilizando la cabecera «3gpp-Sbi-Target-apiRoot».

Ambas soluciones para la protección TLS entre el SEPP y las funciones de red dentro de una PLMN se pueden usar simultáneamente en una PLMN, por ejemplo en la fase transitoria en la que no todas las NFs de la PLMN se han actualizado para admitir la cabecera «3gpp-Sbi-Target-apiRoot» pero cuando el operador de la PLMN desea utilizar la solución basada en la cabecera «3gpp-Sbi-Target-apiRoot» con las NFs actualizadas. En este caso, el SEPP debería omitir la conversión de URI a FQDN telescópicos (y usar la solución basada en el encabezado «3gpp-Sbi-Target-apiRoot») en estos escenarios:

- Respuestas HTTP2 recibidas de la PLMN remota (por ejemplo incluyendo el FQDN del servicio NF objetivo) cuando la solicitud HTTP2 correspondiente contiene una cabecera «3gpp-Sbi-Target-apiRoot».

- Solicitudes HTTP2 recibidas de la PLMN remota (por ejemplo incluyendo los URI de devolución de llamada o *callback*) mediante políticas de SEPP basadas en el URI de destino (es decir, el FQDN de destino).

El uso de FQDNs telescópicas fue introducido en la release 15 de 3GPP, mientras que el uso de la cabecera «3gpp-Sbi-Target-apiRoot» fue introducido en la release 16. La opción basada en cabecera se considera más sencilla y solución a futuro. Es la opción por defecto que está siendo considerada por las operadoras para establecer sus primeros acuerdos de roaming 5G SA.

Un SCP que recibe una solicitud HTTP2 dirigida a un URI con autoridad de una PLMN remota deberá enrutar la solicitud HTTP2 hacia el SEPP del mismo modo que se enruta en escenarios intra-PLMN para comunicaciones indirectas, es decir, el SCP deberá reenviar la cabecera «3gpp-Sbi-Target-apiRoot» que contiene la *apiRoot* del URI objetivo en la PLMN remota, en la solicitud HTTP2 reenviada hacia el SEPP, y establecerá la *apiRoot URI* (pseudo-cabecera *:authority*) de la solicitud, al valor de *apiRoot* del SEPP.

La comunicación entre la NF y el SEPP dento de una PLMN puede ser directa o pasar por un SCP.

En cuanto a la comunicación entre los SEPPs de la red visitada y la red doméstica, ésta puede estar basada en dos mecanismos de seguridad diferentes: TLS o PRINS, como veíamos en la sección 4.1.2.4.

En caso de utilizar PRINS (tras haber sido el mecanismo de seguridad acordado entre ambos SEPPs durante el procedimiento de negociación), la cabecera «3gpp-Sbi-Target-apiRoot» no debe ser utilizada, y el *apiroot URI* contendrá el *apiroot* del SEPP remoto.

En caso de haber negociado TLS como mecanismo de seguridad, la negociación también se utilizará para indicar si el uso de la cabecera «3gpp-Sbi-Target-apiRoot» es soportado por cada uno de los SEPPs. En caso de que al menos uno de los dos SEPPs no la soporte, ésta no será utilizada en la comunicación entre SEPPs. En estos casos, si esta cabecera es usada en la comunicación entre NF y SEPP, el SEPP la eliminará antes de reenviar el mensaje al SEPP remoto por la interfaz N32-f, pero introduciendo su valor en el campo *apiroot URI (:authority)*. Si entre NF y SEPP se utiliza FQDN telescópica (o no se utiliza TLS), el SEPP deberá modificar el *apiroot URI (:authority)* al valor de la NF objetivo derivada de la FQDN telescópica, o del URI de la petición respectivamente.

5.5.1 TLS vs. PRINS

Cuando la interconexión entre dos PLMNs es directa, ésta siempre se realizará utilizando TLS como mecanismo de seguridad para la intercambio de mensajes entre ellas a través de la interfaz N32-f.

Sin embargo, en los escenarios de roaming es habitual el uso de proveedores IPX, como un modelo centralizado de interconexión en el ámbito de las telecomunicaciones para el intercambio de tráfico basado en IP entre diferentes PLMNs. Los proveedores IPX, particularizados al ámbito de la interconexión de señalización de plano de control para roaming de redes 2G/3G/4G/5G, también son denominados Hubs de Roaming (*Roaming-HUBs*) y ofrecen un marco comercial a las PLMNs por el que proveen conectividad, interoperabilidad y gestión de acuerdos de roaming en base a acuerdos de nivel de servicio, desempeño garantizado, calidad y seguridad.

La gran mayoría de las PLMNs utiliza estos servicios IPX en las redes existentes tanto 2G/3G (interconexión SS7) como 4G (interconexión Diameter), siguiendo las recomendaciones y requisitos definidos por *GSMA* al respecto en los documentos IR.34 [80] e IR.77 [81].

Sin embargo, la utilización de IPX/Roaming-HUB para la interconexión de roaming 5G SA no es tan sencilla debido a la nueva necesidad de cifrado extremo a extremo y a las soluciones definidas por 3GPP

para llevarlas a cabo. Para la interconexión de roaming 5G SA, en el caso de utilizar proveedores IPX, 3GPP define un nuevo protocolo de seguridad de nivel de aplicación denominado *PRINS*. El objetivo es mantener la seguridad extremo a extremo (entre ambas PLMNs), pero permitiendo cierto acceso al mensaje por parte del proveedor IPX de modo que pueda seguir ofreciendo sus servicios de un modo similar al de generaciones anteriores. Con PRINS, los mensajes HTTP2 se envían protegidos mediante *Javascript Object Signing and Encryption (JOSE)* y utilizando cifrado *JSON Web Encryption (JWE)* y *JSON Web Signature (JWS)*. Cada mensaje tendrá una política de protección específica, y el SEPP reformateará el mensaje HTTP2 encapsulando el mensaje completo en el cuerpo de un nuevo mensaje HTTP POST. El cuerpo del mensaje de solicitud/respuesta HTTP POST contendrá el mensaje de solicitud/respuesta HTTP2 original reformateado. El cuerpo de solicitud/respuesta HTTP POST se codificará como los cuerpos JSON «N32fReformattedReqMsg» y «N32fReformattedRspMsg» respectivamente [68]. Con este cifrado a nivel de capa de aplicación, se podrá seleccionar que elementos de información específicos del mensaje son cifrados extremo a extremo, y cuáles podrán ser accesibles por parte de intermediarios como los IPX.

El envío de mensajes entre SEPPs utilizando PRINS en el interfaz N32-f se muestra en la figura 5.18.

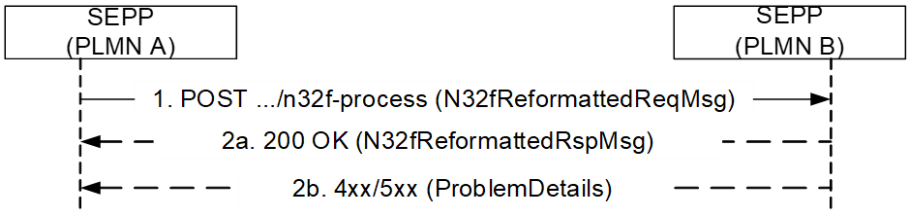

Figura 5.18: Envío de mensajes entre SEPPs en el interfaz N32-f cuando se utiliza PRINS.

5.5.1.1 Estandarización en GSMA

Pese a que a nivel teórico y desde la perspectiva de seguridad, el enfoque de PRINS tiene sentido (aportando seguridad extremo a extremo), en la práctica presenta muchos problemas para ser llevado a cabo. A nivel operativo presenta una importante complejidad, al mantener diferentes políticas de seguridad por cada tipo de mensaje o por cada PLMN remota; con diferentes conjuntos de elementos de información a cifrar en cada caso. Este control debe ser llevado a cabo en las PLMNs, mientras que los IPX únicamente podrán tener acceso al subconjunto de elementos de información que las PLMNs consideren. Uno de los motivos principales que llevaron a la utilización de IPX en escenarios 2G/3G/4G es la simplicidad. Los IPX permiten a las PLMNs simplificar enormemente su gestión de acuerdos de roaming e interconexión, al evitar gestionar cientos de conexiones directas y enrutamientos con múltiples PLMNs remotas; pasando esa tarea a los IPX. Sin embargo, el enfoque de PRINS va en la dirección contraria en este aspecto, pues introduce una gran complejidad en las PLMNs al tener que implementar PRINS, que por otro lado podrían optar por conexiones directas con las PLMNs remotas basadas únicamente en TLS.

La GSMA, en sus diferentes grupos de trabajo, busca poner de acuerdo a los diferentes operadores, IPX y proveedores de red a nivel mundial en los aspectos relativos al roaming e interconexión. Existen áreas de discrepancia que están siendo ampliamente debatidas con el objetivo de desarrollar unas pautas o recomendaciones que permitan a fabricantes y operadores comenzar con sus despliegues y acuerdos de roaming, de manera coordinada y acordada [82].

La complejidad de PRINS ha llevado a muchas operadoras a proponer alternativas para los escenarios que involucran a IPX o HUBs de roaming dentro del foro de GSMA. Con el objetivo de mantener la simplicidad de TLS pero introduciendo al IPX y respetando las especificaciones de 3GPP, GSMA

definió el concepto de SEPP «subcontratado» o «hosteado» (*Hosted SEPP*), de manera que el SEPP es desplegado en el IPX, pero sigue considerándose un escenario bilateral directo inter-PLMN. Esto se muestra en la figura 5.19.

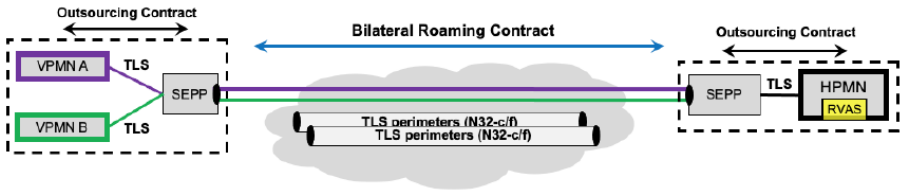

Figura 5.19: Escenario bilateral entre VPLMN y HPLMN con SEPPs «hosteados».

Existe un grupo de trabajo en GSMA denominado *5G Mobile Roaming Revisited (5GMRR)*, en el que también se está explorando la posibilidad de ofrecer TLS salto a salto (*hop-by-hop TLS*), para los escenarios de roaming que involucran a IPX. Este tipo de solución no está recogida en las especificaciones 3GPP por lo que, en caso de fructificar dentro del foro de GSMA, GSMA podría realizar una petición de cambio (*Change Request*) a 3GPP para modificar la especificación, o incluso mantenerse como una solución de ámbito privado sin estandarizar entre IPXs y sus PLMN clientes.

5.5.2 Seguridad en SEPP y Firewall de señalización

En generaciones anteriores las tareas de routing y de seguridad en el área de roaming eran realizadas por elementos de red diferentes. Los STPs internacionales (SS7) y los DEAs (Diameter) se encargaban únicamente del enrutamiento, mientras que se añadieron *firewalls* de señalización externos para cubrir las carencias en materia de seguridad que los protocolos SS7 y Diameter padecen. En 5G, ambas tareas quedan englobadas en la entidad de red SEPP.

3GPP especifica estas capacidades a nivel de seguridad que deben ser ofrecidas por el SEPP:

- Actuar como un nodo proxy no transparente (puede modificar parte del contenido de los mensajes).
- Proteger los mensajes del plano de control de la capa de aplicación entre dos NF que pertenecen a diferentes PLMN y utilizan la interfaz N32 para comunicarse entre sí.
- Realizar la autenticación mutua y la negociación de conjuntos de cifrado con el SEPP en la red de itinerancia.
- Encargarse de los aspectos de gestión de claves que impliquen la configuración de las claves criptográficas necesarias para proteger los mensajes en la interfaz N32 entre dos SEPP.
- Ocultar la topología al limitar la información de topología interna visible para las partes externas (*topology hiding*).
- Como proxy inverso, el SEPP debe proporcionar un único punto de acceso y control a las NF internas.
- El SEPP receptor debe verificar si el SEPP remitente está autorizado a utilizar el ID de PLMN en el mensaje N32 recibido.
- Diferenciar claramente entre los certificados utilizados para la autenticación de SEPP pares (bilaterales) y los certificados utilizados para la autenticación de intermediarios que realizan modificaciones de mensajes (por ejemplo mediante la implementación de almacenamientos de certificados separados).

- Descartar los mensajes de señalización N32 con formato incorrecto.
- Implementar funcionalidades de limitación de tasa de datos (*rate-limiting*) para defenderse a sí mismo y a las NF posteriores contra la señalización excesiva de plano de control. Esto incluye mensajes de señalización de SEPP a SEPP.
- Implementar mecanismos contra la suplantación de identidad (*anti-spoofing*) que permitan la validación entre diferentes capas del mensaje en cuanto a parámetros como direcciones e identificadores de origen y destino (por ejemplo FQDN o identificador de PLMN).
- Utilizar conexiones N32 separadas para cada pareja de PLMN doméstica y visitada.
- En caso de utilizar TLS entre SEPPs, el SEPP debe correlar la conexión TLS del interfaz N32-f con la conexión N32-c mediante la comparación de los identificadores de PLMN contenidos en los certificados TLS usados para establecer las conexiones N32-c y N32-f.

Además de estas capacidades definidas por 3GPP, la GSMA también elabora un conjunto de recomendaciones de seguridad para el escenario de roaming 5G SA para el plano de control y la señalización HTTP2 entre SEPPs en su documento FS.36 [83]. Este documento es del mismo perfil que los elaborados anteriormente para la interconexión SS7 y Diameter, comentados en las secciones 5.2.1.2 y 5.2.2.2 respectivamente. Está enfocado en detallar las recomendaciones de seguridad para prevenir y bloquear posibles ataques presentes en las interfaces de interconexión, relacionados con escenarios de roaming 5G SA y la señalización HTTP2.

5.6 Flujo de señalización para Registro de Usuario 5G

El procedimiento de registro brinda la funcionalidad requerida para registrar un UE/usuario con el 5GS para un acceso 3GPP. Un UE debe registrarse en la red para obtener la autorización para recibir servicios, habilitar el seguimiento de la movilidad y habilitar la accesibilidad. El UE inicia el procedimiento de Registro utilizando uno de los siguientes tipos de registro:

- Registro inicial al 5GS.
- Actualización de registro de movilidad al cambiar a una nueva área de seguimiento *Tracking Area (TA)* fuera del área de registro del UE, o cuando el UE necesita actualizar sus capacidades o parámetros de protocolo que se negocian en el procedimiento de registro.
- Actualización de registro periódico (debido a un período de tiempo predefinido de inactividad).
- Registro de emergencia.

El flujo de llamada de registro general de la figura 5.20 se aplica a todos estos procedimientos de registro, pero no es necesario que el registro periódico incluya todos los parámetros que se utilizan en el resto de escenarios de registro.

- Los pasos 1 a 7 tienen lugar en el interfaz radio, con el UE iniciando la petición de registro, la red de acceso seleccionando un AMF y el nuevo AMF solicitando la transferencia del contexto al AMF antiguo (en caso de que el UE estuviera previamente registrado en otro AMF).

- El paso 8 no está desglosado pero identifica el procedimiento de descubrimiento y selección llevado a cabo por el AMF consultando al NRF para obtener perfiles validos de AUSF en base al SUPI o SUCI. Una vez seleccionado un AUSF objetivo, el AMF efectuará la autenticación del UE correspondiente al paso 9.
- Una vez realizada la autenticación del usuario, el AMF también inicia la validación del terminal, solicitando la identidad al UE y chequeando con el 5G-EIR el PEI del dispositivo; en los pasos 11 y 12.
- Con el paso 13, el AMF realiza el descubrimiento y selección de UDM en base al SUPI, consultando al NRF. Al seleccionar el UDM objetivo, el AMF iniciará el procedimiento de registro (paso 14a), obteniendo la información del perfil del abonado aportada por el UDM incluyendo datos de suscripción de acceso y de movilidad, de selección de SMF, de contexto en SMF o información de localización (paso 14b). Además, el AMF se suscribirá a recibir notificaciones por parte del UDM ante posibles cambios de la información solicitada (paso 14c).
- Cuando el UDM almacena la información del nuevo AMF junto con el tipo de acceso (3GPP), el UDM inicia en el paso 14d, el procedimiento de des-registro del UE hacia el AMF antiguo correspondiente al mismo tipo de acceso (en caso de que el UE estuviera previamente registrado en otro AMF), quien eliminará su suscripción del UDM en el paso 14e.
- El AMF también iniciará el procedimiento de descubrimiento y selección de PCF consultado al NRF (paso 15), para posteriormente poder consultar al PCF las políticas de gestión de acceso (paso 16).
- El AMF contactará con el SMF para actualizar el contexto de sesión PDU del usuario (paso 17), para activar las conexiones de plano de usuario de dichas sesiones PDU.
- Los pasos 18 y 19 aplicarían en el caso de registrarse también para un tipo de acceso «no-3GPP».
- El AMF informará al UE de que el registro es aceptado (paso 21), notificándole múltiples parámetros como su *5G Global Unique Temporary Identifier (5G-GUTI)*, su área de registro, el estado de su sesión PDU, NSSAI permitido y muchos otros.
- El AMF inicia el establecimiento de asociación de políticas de UE contra el PCF (paso 21b).
- El UE envía un mensaje de registro completado al AMF (paso 22) una vez procesada y actualizada toda la información recibida en el paso 21.
- En el caso de que la información de suscripción de acceso y movilidad recibida en el paso 14b incluya parámetros de redirección de roaming (*Steering of Roaming*) con indicación de que el UDM sea notificado, el AMF notificará al UDM (paso 23).
- En el paso 24, el AMF también notificará al UDM sobre el soporte de servicio de voz sobre la red IMS, en caso de haber podido evaluar este soporte.
- Por último, en caso de que el UE indique el soporte del procedimiento de autenticación y autorización específico de «*slice*» de red durante su petición de registro inicial; este procedimiento se ejecuta en el paso 25.

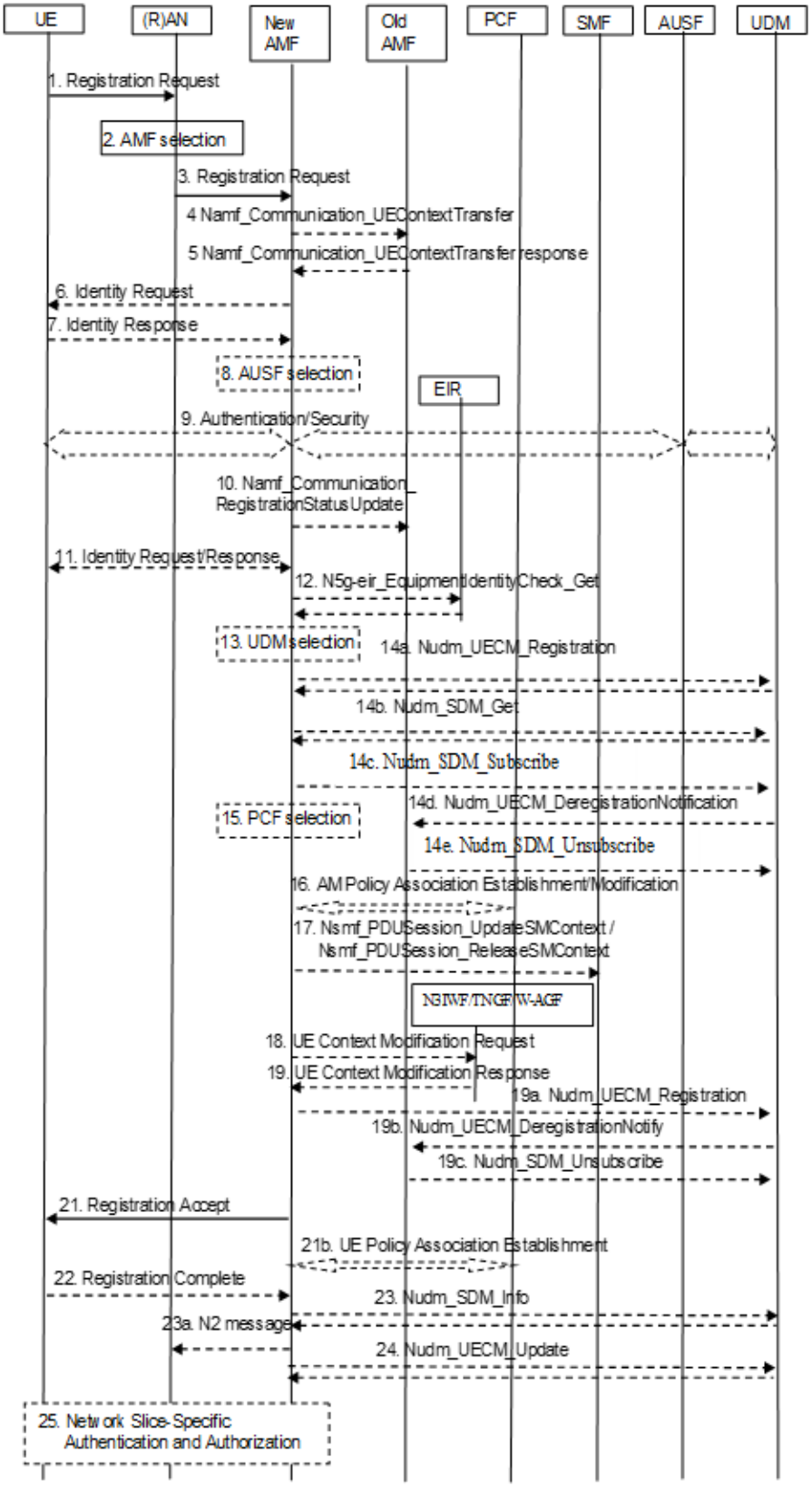

Figura 5.20: Procedimiento de registro de un UE en el sistema 5G.

Capítulo 6

Conclusiones y líneas futuras

Solo podemos visualizar un poco del futuro, pero
podemos ver que allí hay mucho por hacer.

We can only see a short distance ahead, but we can see
plenty there that needs to be done.

Alan Turing [1]

En este capítulo comenzaremos realizando un análisis de la adopción real de la tecnología 5G SA en la actualidad por parte de los diferentes países y operadores móviles a nivel mundial. Resumiremos las conclusiones obtenidas tras la realización de este trabajo, y propondremos futuras líneas de investigación derivadas de este estudio.

6.1 5G SA en la actualidad

A fecha de realización de este trabajo, en Noviembre de 2022, son muchos los países adentrados en los despliegues de tecnología 5G. Los fabricantes de dispositivos móviles llevan años ofreciendo teléfonos móviles compatibles con 5G, y las diferentes operadoras llevan años invirtiendo en esta tecnología.

Las operadoras comenzaron a anunciar los lanzamientos comerciales de sus redes 5G en 2019. Sin embargo, estos son despliegues 5G-NSA, utilizando el núcleo de red de la red 4G-LTE y usando espectro disponible en las bandas de 700 MHz. o 3,7 GHz. Aunque incorpora mejoras a nivel de velocidad y latencia, se puede considerar un paso previo hacia el 5G real. Muchos de los dispositivos anunciados como 5G también hacen referencia únicamente a 5G-NSA, en los que las mejoras con respecto a 4G son limitadas.

En muchos casos, las operadoras han lanzado 5G NSA sin cambiar los equipos de radio, pero usando tecnología *Dynamic Spectrum Sharing (DSS)* para lo que solo ha sido necesario una actualización Software. De este modo reutilizan el espectro de 3G y 4G para radio también 5G, usando el espectro sobrante en las bandas de 1,8 y 2,1 GHz. Se trata de un concepto similar al «*refarming*» usado desde hace años con redes 2G/3G/4G (redestinando espectro hacía tecnologías más modernas), pero optimizado puesto que ahora se pueden utilizar de manera simultánea. Este enfoque es especialmente eficiente al comienzo de 5G puesto que no hay muchos usuarios y de este modo se puede dar servicio 5G sin afectar a 4G.

[1] Alan Turing (1912-1954), lógico y matemático británico. *Computing Machinery and Intelligence (1950)* [84].

Por ahora, las inversiones 5G de las operadoras han estado centradas principalmente en la parte de acceso radio, donde actualizan las capacidades de sus estaciones base para ser compatibles con la nueva tecnología 5G NR (*5G New Radio*). Además, las conversaciones con los reguladores de los diferentes países sobre la disponibilidad de espectro radioeléctrico para 5G están con carácter general, en estado avanzado. En los casos en los que estas nuevas bandas de frecuencia no están disponibles, las operadoras reutilizan frecuencias de 4G, con las que limitan los beneficios de la nueva tecnología pero les permite progresar en su implementación mientras las nuevas bandas quedan disponibles. Estos despliegues, conectan la parte de acceso radio 5G con el núcleo de red pre-existente de la tecnología 4G y por tanto son sistemas 5G NSA (*5G Non Stand Alone*), que también limitan las capacidades reales de las redes 5G, pero permiten a las operadoras progresar en su implantación.

Los desarrollos 5G-SA (*Stand Alone*), incluyendo también el nuevo núcleo de red 5G Core, están mucho menos avanzados. A grandes rasgos (con algunas excepciones principalmente en China y Estados Unidos), los principales operadores a nivel mundial han estado realizando pruebas de concepto sobre esta tecnología y abriendo los procesos de adjudicación con los diferentes proveedores tecnológicos de funciones de red principalmente entre los años 2020 y 2022. Las primeras adjudicaciones se realizaron entre 2021 y 2022, con despliegues planeados mayoritariamente entre 2022 y 2024. Los países que ya tienen despliegues 5G SA se encuentran por norma general en fases preliminares, todavía con capacidades limitadas.

Los primeros despliegues 5G SA a nivel mundial se produjeron principalmente en América del Norte y la región de Asia-Pacífico. Éstos son algunos ejemplos relevantes:

- En Mayo de 2020 Telstra en Australia anuncia ser la primera operadora en introducir un 5G Core y en actualizar la red de acceso 5G para ser capaz de manejar tráfico 5G SA en Sidney. Aún así, los dispositivos compatibles con 5G SA llegarían a finales de 2020 [85].
- En Agosto de 2020, T-Mobile en Estados Unidos anuncia ser la primer red 5G SA lanzada a nivel nacional [86].
- En Enero 2021, Rogers en Canadá es la primera operadora del país en lanzar el núcleo de 5G SA Lo hace de la mano de Ericsson [87].
- En Abril de 2021 Vodafone en Alemania lanza 5G SA Core junto con Ericsson y Nokia en la banda de 3,5 GHz en varias ciudades [88].
- En Mayo de 2021, Singtel en Singapur lanza 5G SA Core con Ericsson [89].
- En Julio de 2021 TPG Telecom (Vodafone Australia) lanza 5G SA en la banda de 700 MHz con Nokia [90].
- En Julio de 2021 Singapore M1 y Samsung realizan la primera llamada *Voice over New Radio (VoNR)* sobre una red 5G SA [91].
- En Agosto de 2022 China anuncia estar cerca de tener 1 billón de usuarios 5G (englobando tanto 5G NSA como 5G SA) [92].

Europa está avanzando en 5G SA con algo de retraso con respecto a Norteamérica y Asia, aunque son múltiples las adjudicaciones otorgadas durante 2021 y sobre todo 2022 por parte de las principales operadoras europeas a diferentes proveedores de equipamiento de núcleo de red 5G. Algunos ejemplos son:

- La operadora A1 en Austria selecciona Nokia en Febrero de 2021 [93].

- Bouygues Telecom en Francia selecciona Ericsson en Junio de 2022 [94].
- WindTre en Italia selecciona Ericsson en Julio de 2022 [95].
- El regulador noruego pretende disponer frecuencias en la banda de 26 MHz en 2023 [96].
- BT en Reino Unido pretende lanzar 5G SA a comienzos de 2023, y prueba por primera vez en Europa Agregación de Portadora (CA) con 4 canales de espectro en una red 5G SA en colaboración con Nokia [97].

La *Global Mobile Suppliers Association (GSA)* realiza informes mensuales con análisis sobre la adopción de la tecnología 5G a nivel mundial. Sus datos más recientes a fecha de realización de este trabajo son de Septiembre de 2022 [98]. Su informe indica que a finales de Agosto de 2022 existen 35 operadoras a nivel mundial que han desplegado y lanzado servicio 5G SA en redes públicas. La cifra sube a 111 al identificar las operadoras que están actualmente invirtiendo en 5G SA (incluyendo pruebas piloto, proyectos en fase de planificación o proyectos con despliegue limitado). Al compararlo con los despliegues 5G NSA, vemos que el número de operadores que han lanzado servicios móviles 5G (incluyendo 5G NSA) es de 222 en 89 países.

Es por esto que, pese a que la gran mayoría de operadores anunciaron el lanzamiento de su servicio 5G desde hace tiempo, éste se trata de 5G NSA, y en muchos casos reutilizando frecuencias de la tecnología 4G. Las grandes promesas tecnológicas anunciadas con la tecnología 5G solo serán cubiertas con el despliegue generalizado de 5G SA tanto a nivel radio como de núcleo de red. 5G SA traerá consigo todo el potencial real de 5G.

6.2 Conclusiones

A la conclusión de este trabajo, podemos confirmar que los objetivos establecidos han sido cubiertos.

Tras realizar un estudio general de la historia de las comunicaciones móviles desde los primeros pasos de la telefonía y las comunicaciones inalámbricas hasta la actualidad, hemos comenzado aportando una visión completa sobre la tecnología 5G, analizado las diferencias entre las redes 5G NSA (*Non Stand Alone*) y las redes 5G SA (*Stand Alone*); así como el ecosistema 5G completo basado en las especificaciones de 3GPP.

Posteriormente nos hemos adentrado de manera específica en el núcleo de red 5G SA con el análisis tanto de su arquitectura basada en servicios SBA, como de las diferentes funciones de red NF que lo conforman. Hemos analizado las implicaciones de los escenarios de roaming en esta arquitectura; y aportado una visión amplia y generalizada del núcleo de red 5G SA.

Finalmente, profundizamos en todos los aspectos relativos a la gestión de la señalización a nivel de enrutamiento y seguridad en el núcleo de red 5G SA, comparándolo con generaciones precedentes. Analizamos las nuevas funcionalidades que el núcleo de red 5G SA ofrece, y el impacto específico de ellas en los escenarios de roaming. Por último, desarrollamos algunos de los escenarios más relevantes de la señalización extremo a extremo.

Podemos concluir que el núcleo de red 5G presenta un cambio fundamental con respecto a las generaciones precedentes en lo relativo a la señalización de plano de control en el núcleo de la red. En primer lugar, se debe a la utilización del protocolo HTTP2 (en lugar de los protocolos anteriores SS7 y Diameter), así como a sus diferentes y mejoradas características de enrutamiento y seguridad.

Las especificaciones 5G también definen un nuevo modo automatizado de comunicación entre diferentes funciones de red que es posible gracias a los servicios del NRF. Las funciones de red registran su perfil y sus servicios en el NRF mediante los procedimientos de registro, actualización y desregistro de servicios. Posteriormente, cualquier función de red consumidora que necesite de los servicios de una función de red productora, podrá enviar una solicitud de descubrimiento de función de red al NRF para obtener una lista de perfiles de NF válidos (que cumplan los requisitos de descubrimiento utilizados) y realizar una selección de la NF productora concreta a utilizar. A partir de ese momento, la función de red consumidora podrá enviar sus peticiones de servicio a la función de red productora descubierta y seleccionada.

A nivel de enrutamiento dentro de la red 5G, se presentan dos modelos fundamentales como son la comunicación directa y la comunicación indirecta en función de la utilización de la entidad de red SCP. El SCP es introducido en la release 16 de 3GPP, permitiendo esta comunicación indirecta y centralizada, además de aportar funcionalidad de descubrimiento delegado para simplificar las tareas de enrutamiento en las diferentes NFs consumidoras.

Para evitar realizar múltiples descubrimientos de NF productora adicionales, el núcleo de red 5G también incorpora el mecanismo de vinculación entre consumidor y productor (*consumer-producer binding*), mediante el uso de cabeceras HTTP2 específicas entre cNF y pNF dedicadas a este fin.

A nivel de seguridad, 5G incorpora el concepto de seguridad por diseño. Esto se refleja en la necesidad de utilizar comunicaciones HTTP2 cifradas sobre TLS y con autenticación mutua entre cliente y servidor. También se definen nuevos identificadores de la identidad de los usuarios (SUCI) que permiten cifrar el contenido del identificador para que no pueda ser interceptado en las comunicaciones de acceso radio.

El «*slicing*» de red presenta una de las capacidades más prometedoras de 5G, aunque de momento con un nivel de implementación real limitada a la espera de que la industria defina casos de uso y modelos de negocio más específicos para ella.

La utilización de la tecnología *Cloud* es también nueva para el mundo de las telecomunicaciones con respecto a las generaciones anteriores, que tan solo habían adoptado tecnología virtualizada (VNF), pero no basada en contenedores (CNF).

En los escenarios de roaming entre PLMNs, 3GPP define TLS y PRINS como dos alternativas de seguridad en el interfaz de interconexión, siendo PRINS el modelo propuesto para los casos en los que un proveedor IPX intermedio está desplegado entre 2 PLMNs. Sin embargo, la complejidad técnica de PRINS, ha hecho que el grupo de trabajo 5GMRR de la GSMA estudie la posibilidad de presentar alternativas a PRINS, como puede ser el uso de TLS salto a salto (*hop-by-hop TLS*) o el ofrecimiento de servicios de SEPP «*hosteado*».

6.3 Líneas futuras

Esta sección recoge varias posibles líneas futuras de investigación, que podrían utilizar este trabajo como base.

6.3.1 Casos de uso reales de 5G SA

La adopción de 5G es un proceso a largo plazo, pero se observa que los casos de uso comienzan a ganar terreno a medida que los operadores buscan monetizar sus inversiones. El impulso fundamental necesario para lanzar su adopción y utilización, se encuentra en la disponibilidad de espectro en las diferentes

bandas y en la inversión por parte de las operadoras tanto en espectro como en infraestructuras de acceso y de núcleo de red 5G SA.

Hasta entonces, no podremos adentrarnos en algunos de los potenciales casos de uso de 5G más disruptivos, como pueden ser los despliegues de IoT masivos (para diferentes escenarios como transporte, ciudades inteligentes, etc.) o la realidad aumentada (en sectores como el educativo, industrial o de ocio).

En las fases preliminares actuales, es necesario encontrar casos comerciales con periodo de recuperación rápido, que permitan a las operadoras crecer y monetizar 5G a corto plazo.

Una posible línea de investigación futura es la relacionada con el análisis de casos de uso específicos 5G, una vez éstos vayan siendo implementados y ampliamente extendidos.

6.3.2 Apagado de redes 2G/3G

Otra posible temática para análisis futuro es el apagado de las redes 2G y 3G por parte de cada vez más operadoras a nivel mundial. El interés por parte de las operadoras en dicho proceso (por motivos evidentes de reducción de costes y optimización de red); se ve acelerado por los incipientes despliegues de 5G, la disponibilidad de esta nueva tecnología, así como por la posibilidad de reutilizar frecuencias reservadas para 3G en las tecnologías actuales 4G y 5G. Esto permite, con la misma cantidad de espectro, ofrecer mucho más caudal y hacerlo a costa de un menor consumo energético.

Los enfoques para llevar este proceso a cabo son diferentes en las distintas regiones del planeta. Si bien en África todavía existen muchos dispositivos 2G y en la mayoría de países este proceso no está cercano; en regiones como Asia, Oceanía o Norteamérica el apagado de redes 2G ya ha comenzado en múltiples países y operadoras. En Centroamérica y Sudamérica, diversos reguladores nacionales están definiendo una hoja de ruta para el apagado de 2G. En el caso de Europa, todo apunta a que las operadoras comenzarán planificando el apagado de su red 3G.

Algunos de los factores que influyen en la decisión sobre qué red apagar antes (2G o 3G) son: la huella de cobertura existente en cada tecnología, el parque de dispositivos y sus tecnologías soportadas, el nivel de adopción de la tecnología *Voice over Long Term Evolution (VoLTE)* que permite no necesitar de las redes 2G/3G para ofrecer servicios de voz, posibles imperativos regulatorios de cada país y el nivel de competencia entre operadoras dentro de cada país. Por otro lado, los dispositivos *Machine to Machine Communications (M2M)* se apoyan mayoritariamente en redes 2G y son los más difíciles de migrar a una nueva tecnología, por lo que los operadores que ofrecen estos servicios a un importante parque de clientes están optando, de momento, por mantener las redes 2G.

En lo relacionado con el servicio de voz, muchas operadoras van avanzando en la adopción de VoLTE (para la que son necesarias las redes 4G e IMS, pero que no requiere de las redes 2G/3G), pero en muchas otras aun existe un uso muy amplio de comunicaciones de voz basadas en conmutación circuitos ofrecida por las redes 2G/3G.

Bajo este escenario, el apagado de la red 3G planteado en múltiples operadoras europeas permite utilizar 4G como red principal (mientras se va evolucionando hacia la adopción de 5G), y seguir utilizando la red 2G en caso de necesitar utilizar servicios de voz sobre circuitos o con dispositivos solo compatibles con 2G (como en algunos servicios M2M). Por ejemplo, el operador Vodafone está llevando este proceso a cabo en los diferentes países en los que opera, comenzando con Holanda en 2020, Italia y República Checa en 2021, y Reino Unido y España en 2022 ([99]).

Extendiéndonos a nivel mundial, en la mayoría de los casos las operadoras mantienen este objetivo con un plazo de realización de varios años, que permita la amplia adopción de la tecnología VoLTE y la completa modernización del parque de dispositivos a las tecnologías más recientes. Sin embargo, muchas

operadoras se han adentrado ya en este proceso como por ejemplo la operadora Claro en Colombia ([100]) o la operadora Personal en Paraguay ([101]). Un paso más allá van las diferentes operadoras de Sudáfrica, que a propuesta del regulador nacional, pretenden apagar completamente ambas tecnologías (e incluso prohibir las licencias de dispositivos 2G), para el año 2025 ([102]).

Estos procesos (llevados a cabo a diferentes velocidades a nivel mundial), así como sus diferentes implicaciones (por ejemplo en lo referente a las comunicaciones de roaming entre redes con distintos niveles de adopción de estos apagados), pueden ser objeto de análisis futuro.

6.3.3 5G satélite NTN

La release 17 de 3GPP (finalizada en Junio de 2022) incluye por primera vez especificaciones para la implementación de 5G vía satélite, o redes 5G no terrestres NTN (*Non-Terrestrial Networks*). Una de sus aplicaciones es la cobertura en zonas rurales y con alta dispersión geográfica con el objetivo de abaratar costes donde no serían rentables despliegues ordinarios. Otro uso es la posibilidad de ofrecer una red de apoyo o seguridad con la que nunca se pierda cobertura. Múltiples operadoras a nivel mundial han comenzado pruebas sobre esta tecnología y avanzado en proyectos de implementación en el corto periodo de tiempo desde la finalización de las especificaciones (Junio de 2022) y la fecha de realización de este TFG (Noviembre 2022).

La empresa Starlink anunció en Agosto de 2022 planes conjuntos con la operadora T-Mobile en Estados Unidos para ofrecer una red de retorno (*backhaul*) 5G a través de sus satélites, así como ayudar en la cobertura de zonas rurales [103].

La empresa estadounidense Omnispace colabora con la operadora filipina Smart Communications para explorar la posibilidad de desplegar conectividad por satélite para 5G NTN según especificaciones 3GPP [104].

La compañía estadounidense AST SpaceMobile tiene planes de desplegar una constelación de unos 100 satélites *Low Earth Orbit (LEO)* equipados con la tecnología necesaria para proporcionar cobertura 5G, 4G y 2G en terminales normales sin requerir en ellos ninguna modificación. Su primer satélite prototipo fue lanzado el 10 de Septiembre de 2022. Entre las operadoras que participarán en las pruebas se encuentran Vodafone y Orange [105].

En España, la compañía Hispasat se adhirió al 3GPP para impulsar el desarrollo de estas especificaciones y está desarrollando servicios 5G interconectados vía satélite [106] junto con operadoras como Masmovil. Del mismo modo, la start-up española Sateliot está enfocada en servicios de conectividad dual para 5G *Narrowband Internet of Things (NB-IoT)*, integrando la red satélite con las redes terrestres existentes para conectividad IoT. Sus primeros pilotos pre-comerciales de la mano de la operadora Telefónica, están planeados para finales del año 2022 [107].

La empresa OneWeb también tiene planes con la Agencia Espacial Europea y con la operadora AT&T [108], y la start-up Lynk Global también trabaja con operadoras en África Central, Mongolia y diversas naciones del Pacífico y del Caribe.

Google confirmó en Septiembre de 2022 que la versión de su sistema operativo Android 14 será compatible con comunicaciones satelitales [109], y Apple anunció en Septiembre de 2022 que el iPhone 14 incluirá un servicio de mensajería de texto de emergencias propio que permitirá la comunicación en zonas remotas sin cobertura de red móvil celular. Lo hace de la mano de la empresa estadounidense Globalstar y sus satélites LEO [110].

Éste es, sin duda, un tema interesante a explorar como futura línea de investigación sobre 5G.

6.3.4 5G Advanced

Si bien la release 17 de 3GPP incluía por primera vez las especificaciones para las redes 5G no terrestres, la release 18 (con planes de finalización para finales de 2023 o principios de 2024), pretende incluir las especificaciones de 5G Advanced.

5G Advanced promete ser el próximo hito en la era 5G y una evolución de la última generación de tecnología móvil. Tiene como objetivo habilitar una gama de servicios que incluyen *eXtended Reality (XR)* y el metaverso, junto con beneficios para IoT. Deberá estar diseñada para hacer frente a las demandas que XR impondrá en cuanto a velocidades de datos y latencia, lo que podría requerir recursos de radio masivos. Otro tema a tratar es el de *Integrated Sensing And Communications (ISAC)*, relacionado con la conducción autónoma de vehículos. También se esperan contribuciones relacionadas con la optimización del rendimiento energético de la *RAN*, y aumentar la eficiencia mediante automatización [111].

La operadora Verizon ha indicado que invertirá activamente en 5G Advanced, y que lo considera un motor fundamental para la adopción real de *network slicing*. También lo considera fundamental para el soporte adicional requerido en dispositivos IoT y soluciones de *Artificial Intelligence (AI)* [112].

La tecnología 5G Advanced será una imporante línea futura de investigación, una vez que el conjunto de sus especificaciones haya finalizado y las operadoras a nivel mundial hayan comenzado a implementarlo en sus redes.

6.3.5 6G

Por último, si bien aun nos encontramos en los primeros pasos de la tecnología 5G, es interesante identificar qué requisitos podrá tener la siguiente generación de redes móviles o 6G. En estos momentos, se trata de identificar tendencias tecnológicas de la actualidad, y la dirección que puedan ir tomando en los próximos años.

La Comisión Europea ha seleccionado a Nokia para liderar la segunda fase de un proyecto de 6G ideado para crear una plataforma pre-estandarizada y una vista general del sistema. Éste comenzará en Enero de 2023 y tendrá una duración de dos años y medio. La segunda fase de este proyecto de investigación denominado *Hexa-X-II* y financiado por la CE representa «la cadena de valor completa para futuras soluciones de conectividad». Nokia liderará un consorcio de 44 empresas y organizaciones, incluidos proveedores de equipos de red, operadores, proveedores verticales y de tecnología, que trabajarán junto con destacados institutos europeos de investigación de comunicaciones. Ericsson fue nombrado gerente técnico del proyecto y diferentes entidades serán responsables de diferentes paquetes de trabajo que incluyen evolución e innovación de radio, dispositivos futuros e infraestructura flexible, gestión y valores de redes inteligentes, o requisitos y ecosistema.

Nokia también dirigió la primera fase del programa, que se presentó en 2020 enfocada en los casos de uso, el desarrollo de tecnologías subyacentes y la definición de la arquitectura potencial para la tecnología 6G [113].

Uno de los aspectos tecnológicos más mencionados en 6G es el de la integración completa de los mundos digital, físico y humano. Las expectativas a nivel de desempeño podrían implicar cientos de gigabits por segundo y latencias por debajo del milisegundo. También habría expectativas renovadas a nivel de confiabilidad, seguridad o sostenibilidad.

La empresa Samsung elaboró un informe en 2020 con su visión sobre 6G, donde vaticinaba que la comercialización de 6G dará comienzo en 2028, con comercialización masiva a partir de 2030. En sus casos de uso identificados, habla de *XR* realmente inmersiva, visores holográficos móviles de alta

resolución, 8K, realidades mixtas o réplicas digitales (*digital replicas* o *digital twins*). A nivel de requisitos tecnológicos estimaban velocidades de hasta mil gigabits por segundo de pico y latencias de hasta tan solo 100 microsegundos [114].

Las operadoras de telecomunicaciones, agrupadas en la asociación *Next Generation Mobile Networks Alliance (NGMN)*, pretenden asegurar que la infraestructura, servicios y dispositivos de las redes de nueva generación cumplan con los requisitos de las operadoras y por ende, satisfagan la demanda y expectativas de los usuarios. En Febrero de 2022 [115], publicaron un documento con casos de uso de 6G en los que recogen 4 categorías:

- **Comunicación humana mejorada**: Como experiencia inmersiva XR con telepresencia holográfica e interacción multimodal: audio, video, gusto, tacto, olfato, etc.
- **Comunicación de máquina mejorada**: Como comunicación e interacción robótica.
- **Servicios de habilitación**: Como posicionamiento 3D de alta precisión, mapeo interactivo con réplicas digitales y mundos virtuales, protección automática, salud e industria inteligente.
- **Evolución de la red**: Como la inteligencia artificial nativa (IA) expuesta como un servicio *Artificial Intelligence as a Service (AIaaA)*, la eficiencia energética y la expansión de cobertura.

En cualquier caso, será ITU-T y ETSI quienes vayan definiendo los requisitos y capacidades de la próxima generación de redes móviles 6G. Del mismo modo que hicieron con IMT-2020 (comúnmente conocido como 5G), ya han comenzado a trabajar en IMT-2030 (*International Mobile Telecommunication-2030*). En principio se prevé que no reemplazará a la infraestructura existente, y por tanto servirá como una actualización, más que como una generación móvil completamente nueva.

El análisis de la tecnología 6G es, por tanto, la línea futura de investigación fundamental derivada de este trabajo.

Bibliografía

[1] M. Cheney and R. Uth, *Tesla: Master of Lightning*. Metrobooks, 2001. [Online]. Available: https://books.google.es/books?id=PoAIAAAACAAJ

[2] A. E. Evenson, *The Telephone Patent Conspiracy of 1876: The Elisha Gray - Alexander Bell Controversy*. McFarland Publishing, 2000.

[3] U. Congress, Ed., *Expressing the sense of the House of Representatives to honor the life and achievements of 19th Century Italian-American inventor Antonio Meucci, and his work in the invention of the telephone*, no. H.Res. 269, 2002. [Online]. Available: https://www.congress.gov/bill/107th-congress/house-resolution/269

[4] ABC. (2002) El congreso de ee.uu. reconoce que antonio meucci inventó el teléfono. [Online]. Available: https://www.abc.es/sociedad/abci-congreso-reconoce-antonio-meucci-invento-telefono-200206180300-107433_noticia.html

[5] A. Labs. At&t labs. our history. [Online]. Available: https://www.research.att.com/sites/labs/our-legacy

[6] P. Sheldon Hochheiser. (2012) Your engineering heritage: The foundations of mobile and cellular telephony. [Online]. Available: https://insight.ieeeusa.org/articles/your-engineering-heritage-the-foundations-of-mobile-and-cellular-telephony/

[7] M. Shi, *Technology Base of Mobile Cellular Operators in Germany and China*, U. de TU Berlín, Ed.

[8] G. T. Yamamoto, *Mobilized Marketing and the Consumer*, B. . E. Business Science Reference, Ed.

[9] N. Telemuseum. (2005) La historia de los teléfonos móviles en noruega. [Online]. Available: https://web.archive.org/web/20070728183835/http://telemuseum.no/mambo/content/view/29/1/

[10] H. Gruber, *The Economics of Mobile Telecommunications*, C. U. Press, Ed.

[11] R. o. K. DongBack Seo, Hansung University, *Evolution and Standardization of Mobile Communications Technology*, I. G. Information Science Reference, Ed.

[12] A. P. Yuste, "Proceso de implantación de la telefonía móvil," 2002. [Online]. Available: https://www.etsist.upm.es/estaticos/catedra-coitt/web_socioeconomica/articulos.htm

[13] C. Radio-television and T. Commission. (1990) Telecom decision crtc 90-26. [Online]. Available: https://crtc.gc.ca/eng/archive/1990/dt90-26.htm

[14] T. Mobility. (2006) Autotel service across british columbia. [Online]. Available: https://archive.ph/20060313232315/http://www.telusmobility.com/clientcare/pcs_west/autotel/coverage.shtml

[15] D. H. Ring, "Mobile telephony - wide area coverage," 1947. [Online]. Available: https://www.theatlantic.com/technology/archive/2011/09/the-1947-paper-that-first-described-a-cell-phone-network/245222/

[16] T. Sato, "El camino de los teléfonos móviles 1g/2g a 3g," pp. 15–16. [Online]. Available: https://www.rf-world.jp/bn/RFW02/samples/p015-016.pdf

[17] S. K. . A. Lugn, "The launch of nmt," 2022. [Online]. Available: https://www.ericsson.com/en/about-us/history/changing-the-world/the-nordics-take-charge/the-launch-of-nmt

[18] C. Blog, "Cuando el móvil era el teléfono automático de vehículos (parte ii)," 2008. [Online]. Available: https://blog.cnmc.es/2008/12/03/cuando-el-movil-era-el-telefono-automatico-de-vehiculos-parte-ii/

[19] V. K. Garg, *Wireless Communications and Networking - A volume in The Morgan Kaufmann Series in Networking*, B. . E. Business Science Reference, Ed.

[20] C. S. Daniel Garcia-Swartz, "Cra insights: The economics of 5g," 2021. [Online]. Available: https://media.crai.com/wp-content/uploads/2021/05/05145230/Insights-The-Economics-of-5G-article-6-Open-or-Closed-System-May2021.pdf

[21] A. Russell, *The Story of Vodafone*, Vodafone, Ed., 2011.

[22] J. Howell-Jones, "Vodafone closes analog network," 2001. [Online]. Available: https://www.computing.co.uk/news/1830513/vodafone-closes-analogue-network

[23] E. Pais. Cierra moviline, el primer servicio popular de telefonía móvil. [Online]. Available: https://elpais.com/diario/2004/01/22/ciberpais/1074741868_850215.html

[24] Telefónica. Hechos significativos del año 1993. [Online]. Available: https://www.telefonica.com/es/wp-content/uploads/sites/4/2021/07/1993_hechos_significativos.pdf

[25] D. E. Borth. Mobile telephone. [Online]. Available: https://www.britannica.com/technology/mobile-telephone

[26] Y.-B. Lin, "Cellular digital packet data," 1997. [Online]. Available: https://ieeexplore.ieee.org/document/609885

[27] ETSI, "Gsm 03.02 v3.1.4 phase 1," 1992.

[28] ——, "Etsi ts 101 625 v7.0.0," 1999.

[29] ——, "Gsm 03.60 v7.0.0 release 1998 - network architecture," 1999.

[30] E. 3GPP, "3gpp tr 50.049 v4.0.1 release 4," 2001.

[31] G. Hill, *The Cable and Telecommunications Professionals' Reference*, T. Focal Press and F. Group, Eds., 2007.

[32] CDG. cdmaone. [Online]. Available: http://www.cdg.org/technology/cdmaone.asp

[33] Motorola. Motorola milestones. [Online]. Available: https://www.motorola.com/us/about/motorola-history-milestones

[34] Arib, *PERSONAL DIGITAL CELLULAR TELECOMMUNICATION SYSTEM ARIB STANDARD*, 1991.

[35] I. M. . L. E. D. Timothy K. Forde, "Exclusive sharing & virtualization of the cellular network," 2022.

[36] E. 3GPP, "3gpp ts 23.110 v3.4.0 release 99," 2000.

[37] ——, "3gpp ts 125.306 v11.8.0 release 11," 2014.

[38] R. Staff, "Qualcomm halts umb project, sees no major job cuts," 2008. [Online]. Available: https://www.reuters.com/article/marketsNews/idUSN1335969420081113?rpc=401&

[39] S. Acharya, "El seminario mundial de radiocomunicaciones de la uit se centra en las tecnologías de comunicaciones del futuro," 2010. [Online]. Available: https://www.itu.int/net/pressoffice/press_releases/2010/48.aspx

[40] E. 3GPP, "3gpp ts 29.513 requirements for evolved utra (e-utra) and evolved utran (e-utran)," 2020.

[41] ——, "3gpp ts 23.402 architecture enhancements for non-3gpp accesses," 2020.

[42] ——, "3gpp ts 23.401 v16.12.0 release 16," 2021.

[43] ——, "3gpp ts 36.300 eutra and eutran overall description, stage 2," 2020.

[44] ——, "3gpp ts 36.912 feasibility study for further advancements for e-utra (lte-advanced) release 9," 2010.

[45] ——, "Lte-advanced pro ready to go," 2015. [Online]. Available: https://www.3gpp.org/news-events/3gpp-news/1745-lte-advanced_pro

[46] V. M. Cenre. Vodafone launches 4.9g in sydney?s west. [Online]. Available: https://www.vodafone.com.au/media/vodafone-launches-4-9g-in-sydneys-west

[47] W. Forum. Technical specification library. [Online]. Available: https://wimaxforum.org/TechSpec

[48] A. C. Clarke, *Profiles of the Future: An Inquiry Into the Limits of the Possible. Revised edition.*, R. Holt and Wilson, Eds.

[49] ——, *The promise of space*, Harper and Row, Eds.

[50] S. Acharya, "La uit define la perspectiva y la hoja de ruta para el desarrollo de la tecnología móvil 5g," 2015. [Online]. Available: https://www.itu.int/net/pressoffice/press_releases/2015/27.aspx

[51] E. 3GPP, "3gpp ts 38.101 release 17," 2022.

[52] GSMA, "Operator requirements for 5g core connectivity options," 2019.

[53] E. 3GPP, "3gpp ts 37.340 release 15, sección 4.3," 2022.

[54] GSMA, "5g implementation guidelines: Nsa option 3," 2020.

[55] ——, "Road to 5g: Introduction and migration," 2018.

[56] E. 3GPP, "3gpp ts 38.470 release 16," 2022.

[57] ——, "3gpp ts 23.501 release 16 system architecture for the 5g system," 2022.

[58] ——, "3gpp ts 29.500 release 16 technical realization of the sba," 2021.

[59] IETF, "Rfc 7540 http2," 2015.

[60] ——, "Rfc 8259 json," 2017.

[61] ——, "Rfc 793 tcp," 1981.

[62] ——, "Rfc 3986 uri," 2005.

[63] O. Initiative. Openapi 3.0.0 specification. OpenAPI Initiative. [Online]. Available: https://www.openapis.org/

[64] E. 3GPP, "3gpp ts 29.510 release 16 network function repository services," 2021.

[65] ——, "3gpp ts 21.916 release 16 summary of rel-16 work items," 2022.

[66] ——, "3gpp ts 29.521 release 16 binding support management service," 2021.

[67] ——, "3gpp ts 33.501 release 16 security architecture and procedures," 2021.

[68] ——, "3gpp ts 29.573 release 16 plmn interconnection," 2021.

[69] R. Wilson and S. Marcus, *American Greats*. PublicAffairs, 2000. [Online]. Available: https://books.google.es/books?id=VngseK0IoEoC

[70] L. Dryburgh and J. Hewett, *Signaling System No. 7 (SS7/C7): Protocol, Architecture, and Services*, ser. Cisco Press networking technology series. Cisco, 2005. [Online]. Available: https://books.google.es/books?id=lKO1PNwI9tkC

[71] E. 3GPP, "3gpp ts 23.003 numbering, addressing and identification," 2021.

[72] ——, "3gpp ts 29.272 epc mme interfaces based on diameter," 2014.

[73] IETF, "Rfc 7230 http/1.1: Message syntax and routing," 2014.

[74] E. 3GPP, "3gpp ts 23.502 procedures 5gs," 2021.

[75] IETF, "Rfc 5246 tls v1.2," 2008.

[76] K. Majithia. (2022) Vodafone slices network in the netherlands. [Online]. Available: https://www.mobileworldlive.com/featured-content/top-three/vodafone-slices-network-in-the-netherlands/

[77] ——. (2022) Nokia, telia finland bring sa 5g to fwa. [Online]. Available: https://www.mobileworldlive.com/featured-content/top-three/nokia-telia-finland-bring-5g-sa-to-fwa/

[78] C. Donkin. (2022) Safaricom eyes enterprise boost after slicing trial. [Online]. Available: https://www.mobileworldlive.com/featured-content/top-three/safaricom-eyes-enterprise-boost-after-slicing-trial/

[79] A. Morris. (2022) Ericsson, nokia serve up 5g network slices on android. [Online]. Available: https://www.mobileworldlive.com/featured-content/top-three/ericsson-nokia-serve-up-5g-network-slices-on-android/

[80] GSMA, "Ir.34 guidelines for ipx provider networks v17.0," 2021.

[81] ——, "Ir.77 inter-operator ip backbone security req. for service and inter-operator ip backbone providers v5.0," 2019.

[82] ——, "Ng.113 5gs roaming guidelines v5.0," 2021.

[83] ——, "Fs.36 5g interconnect security v2.2," 2022.

[84] A. Turing, "Computing machinery and intelligence," p. 460, 1950.

[85] Ericsson. (2020) Telstra network now 5g standalone capable end-to-end with ericsson technology. [Online]. Available: https://www.ericsson.com/en/news/2020/5/telstra-5g-standalone-sa-ready-with-ericsson

[86] C. Update. (2020) T-mobile claims ?world first? standalone 5g nationwide launch. [Online]. Available: https://www.commsupdate.com/articles/2020/08/05/t-mobile-claims-world-first-standalone-5g-nationwide-launch/

[87] Ericsson. (2021) The deployment of canada?s first national 5g standalone network. [Online]. Available: https://www.ericsson.com/en/cases/2022/rogers-and-ericsson

[88] C. Update. (2021) Vodafone launches 5g sa in 170 german cities and municipalities. [Online]. Available: https://www.commsupdate.com/articles/2021/04/13/vodafone-launches-5g-sa-in-170-german-cities-and-municipalities/

[89] Ericsson. (2021) Singtel - the first and most powerful 5g standalone network in singapore. [Online]. Available: https://www.ericsson.com/en/cases/2022/singtel-and-ericsson

[90] Nokia. (2021) Nokia and tpg telecom launch world?s first live 5g standalone 700mhz service in australia. [Online]. Available: https://www.nokia.com/about-us/news/releases/2021/07/05/nokia-and-tpg-telecom-launch-worlds-first-live-5g-standalone-700mhz-service-in-australia/

[91] C. Update. (2021) M1 partners samsung to support voice over 5g new radio (vonr) on 5g sa network. [Online]. Available: https://www.commsupdate.com/articles/2021/07/26/m1-partners-samsung-to-support-voice-over-5g-new-radio-vonr-on-5g-sa-network/

[92] J. Waring. (2022) China 5g user base set to top 1b. [Online]. Available: https://www.mobileworldlive.com/asia/asia-analysis/china-5g-user-base-set-to-top-1b/

[93] Nokia. (2021) Nokia wins multi-year 5g radio and core contract with a1 austria. [Online]. Available: https://www.nokia.com/about-us/news/releases/2021/02/18/nokia-wins-multi-year-5g-radio-and-core-contract-with-a1-austria/

[94] Ericsson. (2022) End-to-end ericsson 5g to power bouygues telecom?s digitalization drive in france. [Online]. Available: https://www.ericsson.com/en/news/2022/6/end-to-end-ericsson-sa-5g-for-bouygues-telecom

[95] ——. (2022) Windtre selects ericsson 5g core to power standalone network. [Online]. Available: https://www.ericsson.com/en/press-releases/3/2022/windtre-selects-ericsson-5g-core-to-power-standalone-network

[96] A. Morris. (2022) Norway eyes 26ghz, 1500mhz for 5g. [Online]. Available: https://www.mobileworldlive.com/featured-content/top-three/norway-eyes-26ghz-1500mhz-for-5g/

[97] ——. (2022) Bt, nokia claim 4 ca first in europe. [Online]. Available: https://www.mobileworldlive.com/featured-content/top-three/bt-nokia-claim-4-ca-first-in-europe/

[98] GSA. (2022) 5g-market snapshot september 2022. [Online]. Available: https://gsacom.com/paper/5g-market-snapshot-september-2022/

[99] A. Cuesta. (2022) Vodafone se dispone a apagar su 3g en el reino unido. [Online]. Available: https://www.mobileworldlive.com/spanish/vodafone-se-dispone-a-apagar-su-3g-en-el-reino-unido/

[100] C. Colombia. (2022) Claro invita a sus clientes a cambiar sus teléfonos 2g. [Online]. Available: https://www.claro.com.co/institucional/cambiar-telefonos-2g/

[101] P. Paraguay. (2022) Apagado de las redes 2g/3g. [Online]. Available: https://www.personal.com.py/roaming/volte.html

[102] K. Majithia. (2022) South africa plots 2g, 3g shutdown by 2025. [Online]. Available: https://www.mobileworldlive.com/featured-content/top-three/south-africa-plots-2g-3g-shutdown-by-2025/

[103] M. DeGrasse. (2022) Musk, t-mobile to announce connectivity tie-up. [Online]. Available: https://www.mobileworldlive.com/featured-content/home-banner/musk-t-mobile-to-announce-connectivity-tie-up/

[104] M. Carroll. (2022) Smart launches second satellite collaboration. [Online]. Available: https://www.mobileworldlive.com/featured-content/asia-home-banner/smart-launches-second-satellite-collaboration/

[105] A. S. Mobile. (2022) Building the first and only space-based cellular broadband network. [Online]. Available: https://ast-science.com/spacemobile-network/

[106] Hispasat. (2022) Satélite, edge computing y 5g. [Online]. Available: https://blog.hispasat.com/es/articulo/114/satelite-edge-computing-y-5g-innovacion-para-una-conectividad-mas-eficiente

[107] Sateliot. (2022) Telefónica to connect iot devices via satellite with 5g technology. [Online]. Available: https://sateliot.space/en/news-sateliot-space/telefonica-to-connect-iot-devices-via-satellite-with-5g-technology/

[108] OneWeb. (2022) 5g to be showcased over oneweb satellites in new global connectivity pilot. [Online]. Available: https://oneweb.net/resources/5g-network-networks-be-showcased-over-oneweb-satellites-new-global-connectivity-pilot

[109] A. Morris. (2022) Google signals satellite compatibility in android 14. [Online]. Available: https://www.mobileworldlive.com/featured-content/home-banner/google-signals-satellite-compatibility-in-android-14/

[110] Reuters. (2022) Apple picks globalstar for emergency satellite service on iphone 14. [Online]. Available: https://www.reuters.com/technology/apple-picks-globalstar-satellite-service-iphone-14-series-2022-09-07/

[111] 3GPP. (2021) 3gpp release 18. [Online]. Available: https://www.3gpp.org/specifications-technologies/releases/release-18

[112] K. Majithia. (2022) Verizon strategy boss turns attention to 5g-advanced. [Online]. Available: https://www.mobileworldlive.com/featured-content/top-three/verizon-strategy-boss-turns-attention-to-5g-advanced/

[113] M. Robuck. (2022) Nokia to lead second phase of eu 6g project. [Online]. Available: https://www.mobileworldlive.com/featured-content/top-three/nokia-to-lead-second-phase-of-eu-6g-project/

[114] ——. (2022) 6g the next hyper connected experience for all. [Online]. Available: https://cdn.codeground.org/nsr/downloads/researchareas/20201201_6G_Vision_web.pdf

[115] ——. (2022) 6g use cases and analysis. [Online]. Available: https://www.ngmn.org/wp-content/uploads/220222-NGMN-6G-Use-Cases-and-Analysis-1.pdf

[116] L. Lamport, *LaTeX: A Document Preparation System, 2nd edition.* Addison Wesley Professional, 1994.

[117] Ericsson, "5g sa deployment: Moving beyond embb," 2022. [Online]. Available: https://www.ericsson.com/49f649/assets/local/reports-papers/mobility-report/documents/2022/063022-emr-june-2022-5g-sa-deployment-article-web.pdf

[118] J. T. J. Penttinen, *5G Second Phase Explained: The 3GPP Release 16 Enhancements*, W. Publishers, Ed.

[119] IETF, "Rfc 6733 diameter base protocol," 2012.

[120] ——, "Rfc 3589 diameter command codes for 3gpp release 5," 2003.

[121] GSMA, "Ng.132 report 5g mobile roaming revisited (5gmrr) phase 1 v1.0," 2022.

Índice alfabético

Apéndice A

Herramientas y recursos

Las herramientas utilizadas para la elaboración de este trabajo han sido:

- PC compatible
- Sistema operativo Windows
- Procesador de textos LaTeX[116]
- Recursos de organizaciones estandarizadoras como 3GPP, GSMA, ETSI-ITU, ANSI, TIA o IEEE
- Recursos bibliográficos sobre la materia
- Experiencia profesional en el sector de las telecomunicaciones y redes móviles.

Universidad de Alcalá
Escuela Politécnica Superior

ESCUELA POLITECNICA
SUPERIOR

Universidad
de Alcalá