

XVI Jornadas de Ingeniería Telemática

JITEL 2023

La Salle – Universitat Ramon Llull

Actas de las XVI Jornadas de Ingeniería Telemática (JITEL 2023), Barcelona (España), 8-10 de noviembre de 2023.

Reducción de la latencia en redes 5G-NR

Felix Delgado-Ferro, Jorge Navarro-Ortiz, Natalia Chinchilla-Romero, Juan Manuel Lopez-Soler
Department of Signal Theory, Telematics and Communications
University of Granada
Av. del Hospicio, 1, 18012 Granada.
felixdelgado@ugr.es, jorgenavarro@ugr.es, nataliachr@ugr.es, juanma@ugr.es

Este trabajo presenta un estudio que tiene como objetivo reducir la latencia en función de los parámetros más relevantes de la capa física 5G *New Radio (NR)*. En primer lugar, se explican los conceptos teóricos fundamentales, relacionados con el *subcarrier spacing (SCS)* y la estructura de las tramas. Después se describe el entorno de experimentación, compuesto por una red 5G basada en *Amarisoft* y un equipo de usuario con modem Quectel. Por último, se presentan resultados preliminares que muestran cómo afectan los parámetros seleccionados. Como conclusión principal, el *subcarrier spacing* y el uso de tramas flexibles son los parámetros fundamentales para reducir la latencia en 5G.

Palabras Clave*—5G New Radio, latencia, reducción, numerología y *subcarrier spacing

I. INTRODUCCIÓN

La ITU (*International Telecommunication Union*) definió las recomendaciones IMT-2020 (*International Mobile Telecommunications*) [1]. Considerando estas recomendaciones, el 3GPP (*3rd Generation Partnership Project*) definió tres posibles servicios que incluyen todos los casos de uso conocidos actualmente [2]. Estos servicios son: *enhanced Mobile Broadband (eMBB)*, *massive Machine Type communications (mMTC)* y *Ultra-Reliable Low Latency Communications (URLLC)*.

Entre los principales retos de las redes 5G, destacan los servicios URLLC debido a sus estrictos requisitos. Estos requisitos consisten en proporcionar una transmisión de datos con una fiabilidad del 99,99 % y una latencia por debajo del milisegundo. Existen múltiples casos de uso que deben ser soportados por este tipo de servicio [3]. Además, los avances actuales, que pretenden satisfacer estos requisitos, se han centrado en aumentar el ancho de banda -generalmente centrado en el enlace descendente-. Sin embargo, las necesidades del tráfico están cambiando. Esto implica que las nuevas aplicaciones requieren una mayor flexibilidad en la asignación de recursos debido al uso de más recursos del enlace ascendente. En consecuencia, la duplexación por división en el tiempo (TDD) es cada vez más popular y significativa, ya que permite aumentar la flexibilidad y la eficiencia espectral [4]. Por estas razones, en nuestro trabajo nos centraremos en la configuración TDD para reducir la latencia, siguiendo las especificaciones técnicas de la *Release 15* [5].

El resto del documento se organiza como sigue. La Sección II proporciona una revisión de los trabajos relacionados. A continuación, en la Sección III se presentan los fundamentos teóricos de la capa física 5G NR. La Sección IV describe el montaje y explica el diseño experimental, mientras que la Sección V muestra el análisis en términos de latencia. Finalmente, la Sección VI presenta las conclusiones de este trabajo.

II. TRABAJOS RELACIONADOS

Esta sección presenta las contribuciones relacionadas con nuestro trabajo. Principalmente, estas contribuciones se centran en el análisis del rendimiento utilizando mecanismos de optimización sobre la capa física 5G NR.

Ali *et al.* [6] presentan un estudio del *Physical Uplink Shared Channel (PUSCH)* teniendo en cuenta parámetros como los diferentes *Subcarrier Spacing (SCS)*, las técnicas de modulación (QPSK o QAM) y el número de antenas de la *Base Station (BS)* y del *user equipment (UE)*. Mhedhbi *et al.* [7] se centraron en aprovechar las características NR de las tramas y evaluar el impacto de las configuraciones radio y el ancho de banda reservado para el tráfico crítico en los servicios URLLC. Chinchilla-Romero *et al.* [8] evaluaron la ganancia de rendimiento al incluir 5G a un escenario con multi-conectividad y lo extrapolaron a un posible aumento en redes 5G, basándose en sus requisitos y numerología. Por último, Mohamed *et al.* [9] abordaron la gestión de los recursos radio empleando mini-slots definidos para las redes 5G NR. Estos permiten reducir la latencia de retransmisión a costa de reducir el rendimiento del enlace.

En cambio, nuestro trabajo presenta un análisis de la latencia en función del ancho de banda dedicado, el *Subcarrier Spacing (SCS)*, la estructura de las tramas y *slots*. Además, la experimentación se realiza empleando

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

una estación base real; mientras que los trabajos citados anteriormente se realizan en escenarios simulados.

III. CAPA FÍSICA DE 5G NR

La pila de protocolos de 5G NR consta de varias capas que trabajan juntas para permitir la comunicación entre dispositivos y equipos de red. Las principales capas son: *Physical (PHY), Medium Access Control (MAC), Radio Link Control (RLC), Packet Data Convergence Protocol (PDCP), Radio Resource Control (RRC)* y *Service Data Adaptation Protocol (SDAP)* [10]. Además, existen protocolos de capa superior que permiten comunicaciones extremo a extremo.

La capa física de 5G NR está diseñada a partir de la capa física de 4G LTE. Del mismo modo, ambas utilizan OFDM con prefijo cíclico (CP) para el enlace descendente (DL). En 5G NR, OFDM también puede utilizarse para el enlace ascendente (UL), aunque se establecen dos mecanismos nuevos: *Peak to Average Power Ratio (PAPR)* y OFDM con precodificación por transformada discreta de Fourier (DFT-s-OFDM) [10]. Además, NR admite dos rangos de frecuencias: FR1 (410 MHz - 7125 MHz) y FR2 (24250 MHz - 52600 MHz). Cada rango de frecuencias admite anchos de banda de canal específicos [11].

El ancho de banda (BW) de NR se compone de un conjunto de *Resource Blocks (RB)*. Un RB es un conjunto de 12 subportadoras. Una subportadora -en el dominio de la frecuencia- o un símbolo OFDM -en el dominio del tiempo- se denomina *Resource Element (RE)*.

A. Numerología

En 5G [12], la distancia entre portadoras se especifica como $\Delta f_{ref} = 15KHz$. NR define múltiples numerologías OFDM. Cada numerología está determinada por el valor μ, que está directamente relacionado con el *subcarrier spacing* ($SCS = 2^{\mu} * \Delta f_{ref}$) para un enlace ascendente o descendente y el prefijo cíclico. La tabla I muestra las múltiples numerologías disponibles para FR1.

Tabla I
NUMEROLOGÍA EN NR PARA FR1 [11][12].

μ	SCS $[KHz]$	CP	$N_{slots}^{subframe,\mu}$	N_{symb}^{slot}
0	15	Normal	1	14
1	30	Normal	2	14
2	60	Normal/Extended	4	14/12

Cabe destacar que la numerología se introduce en 5G NR, mientras que en 4G LTE el *subcarrier spacing* es siempre $SCS_{LTE} = 2^{\mu} * \Delta f_{ref} = 15KHz$ (caso $\mu = 0$).

B. Estructura de la trama

Las transmisiones en NR se organizan en tramas flexibles. En el dominio del tiempo, la duración de la trama se define como $T_{\mathrm{f}} = (\Delta f_{\mathrm{máx}} N_{\mathrm{f}}/100) \cdot T_{\mathrm{c}} = 10$ ms [12]. Cada trama se subdivide en diez subtramas con duración $T_{\mathrm{sf}} = (\Delta f_{\mathrm{máx}} N_{\mathrm{f}}/1000) \cdot T_{\mathrm{c}} = 1$ ms

Además, cada subtrama consta de $N_{slots}^{subtrama,\mu} \in \{1, 2, 4, 8, 16\}$ *slots*, o ranuras temporales, dependiendo de la SCS o numerología seleccionada (ver Tabla I). La Fig. 1 muestra los casos de valores de SCS de 15; 30 y 60 de numerologías posibles en NR para FR1. Además, cada *slot* consta de 14 símbolos OFDM (independientemente del SCS) precedidos de un prefijo cíclico.

C. Configuración del Slot

NR admite el funcionamiento tanto FDD como TDD empleando la misma estructura de trama. En TDD, que permite la adaptación flexible del tráfico, cada símbolo OFDM de un *slot* se puede clasificar como *"downlink", üplink"* o *"flexible"*.

El gNB debe proporcionar estas configuraciones al *UE* mediante el parámetro *TDD-UL-DL-ConfigurationCommon*. A continuación, el UE establece el formato *slot-by-slot* para un número especificado de *slots*. La información proporcionada por el parámetro *TDD-UL-DL-ConfigurationCommon* es la siguiente:

- Configuración del *subcarrier spacing* de referencia μ_{ref} mediante el parámetro *referenceSubcarrierSpacing* (see Table I).
- *Pattern 1* y opcional *Pattern 2* proporcionan la periodicidad (P) del patrón DL-UL en [ms], el número de *slots* DL completos y consecutivos al principio de cada patrón DL-UL (n_{slots}^{DL}), el número de símbolos DL consecutivos al principio del *slot* que sigue al último *slot* DL (n_{symb}^{DL}), el número de *slot* UL completas y consecutivas al final de cada patrón DL-UL (n_{slots}^{UL}); y número de símbolos UL consecutivos al final del *slot* que precede al primer *slot* UL (n_{symb}^{UL}).

En [5] se presentan los formatos de *slot* para prefijo cíclico normal que se establecen mediante *Pattern 1*. Si la configuración sólo utiliza este patrón, entonces la periodicidad (P) incluye $n_{slots} = P{*}2^{\mu_{ref}}$ *slots*, donde μ_{ref} es la configuración del *subcarrier spacing*. La periodicidad (P) del patrón DL-UL cambia en función de la numerología de referencia (μ_{ref}). El cálculo P/20 identifica la posición del primer símbolo de una trama par que debe coincidir con el primer símbolo de cada periodo para que estén sincronizados. Esto se resume en una tabla II.

El *Pattern 2* puede utilizarse opcionalmente. En este caso, la periodicidad es ($P' = P + P_2$). El número de *slots* para el *Pattern 2* seguirá la misma lógica que el *Pattern 1*. El uso de dos patrones permite transmitir diferentes cantidades de *slots* de enlace descendente y ascendente en

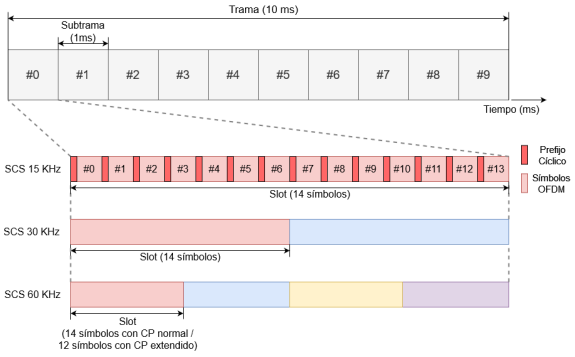


Figura 1. Estructura de la trama en NR para FR1

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

tramas consecutivas -sin necesidad de modificar el patrón del UE-. El usuario espera que esto se transmita en dos tramas consecutivas ($P' = P + P_2 = 20\text{ms}$) y que el ajuste del *subcarrier spacing* μ sea mayor o igual que el valor de referencia μ_{ref} para ambos patrones. La Fig. 2 muestra un ejemplo de la configuración cuando se utilizan ambos patrones.

Tabla II
PERIODICIDAD DE TRANSMISIÓN APLICABLE (BASADA EN [5])

$P[ms]$	μ	SCS $[KHz]$	$P/20$
0,625	3	120	32
1,25	2/3	60/120	16
2,5	1/2/3	30/60/120	8
5	-	Not described	4
10	-	Not described	2

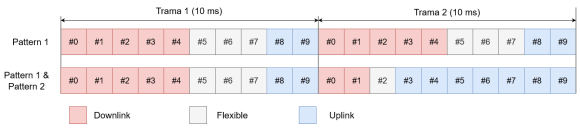


Figura 2. Ejemplo de configuración del *Pattern 1* y *Pattern 2*

Además, mencionar que a partir de la *Release 16* se ha definido un nuevo parámetro *tdd-UL-DL-ConfigDedicated* que mejorará las prestaciones en servicios URLLC [5].

IV. EXPERIMENTOS

La experimentación para este trabajo se ha realizado en el *testbed* que se muestra en la Fig. 3. Concretamente, se han realizado los experimentos en la conexión inalámbrica 5G NR. Esta conexión, que aparece destacada en la Fig. 3, se establece empleando una gNB, desplegada con el software de *Amarisoft* con version 2023-03-17, y un módem Quectel RM500Q-GL (v1.6) que permite la conexión mediante 5G, conectado al Intel NUC que actúa como *Customer Provided Equipment (CPE)*.



Figura 3. Testbed para conectividad de baja latencia

Cabe mencionar que la red 5G NR está deplegada en modo *standalone (SA)*. Por otro lado, debido a tener que transmitir utilizando bandas licenciadas, este radioenlace estará contenido dentro de una jaula de Faraday para no contaminar el espectro radioeléctrico.

Previo al proceso de experimentación, se decidieron los parámetros de configuración de la capa física 5G NR que se modificarían para comprobar su impacto y poder caracterizar la conexión en términos de latencia. Estas configuraciones aparecen en la Tabla III.

Para los valores correspondientes a la estructura de la trama se ha optado por emplear configuraciones definidas en [5]. Estos se describen en la Tabla IV.

Tabla III
CONFIGURACIONES DE LOS EXPERIMENTOS

Parámetro	Acrónimo	Valores
Ancho de banda	BW	20/40 MHz
Subcarrier Spacing	SCS	15/30 KHz
Estructura de la trama	FS	1/2/3/4

Tabla IV
CONFIGURACIONES DE LAS TRAMAS

FS	Patterns	n^{DL}_{slots}	n^{DL}_{symb}	n^{UL}_{symb}	n^{UL}_{slots}
1	*Pattern 1*	7	2	2	2
2	*Pattern 1*	7	6	4	2
3	*Pattern 1*	6	2	2	3
4	*Pattern 1*	3	6	2	4
	Pattern 2	4	0	0	0

Este trabajo se ha centrado en el estudio de la latencia en función de cada uno de los parámetros citados previamente. Debido a las múltiples configuraciones posibles se ha establecido una configuración por defecto para poder comparar de forma coherente. Esta configuración es: $BW = 20$ MHz; $SCS = 30$ KHz y $FS = 2$. A parte de estas configuraciones, se han dejado estáticas el número de antenas (MIMO 2x1) y la banda de frecuencia empleada para TDD (n41).

V. RESULTADOS PRELIMINARES

Como se ha comentado, los resultados se centran principalmente en el análisis de la latencia, dado que es unos de los indicadores críticos para el despliegue en el futuro de servicios eMBB y URLLC.

A. *Ancho de banda*

La comparativa de la latencia al emplear las configuraciones de ancho de banda establecidas en la Tabla III se muestran en la Fig. 4. Observando esta *Función de Distribución Acumulada (CDF)*, podemos comprobar como el aumento del BW mejora la latencia de la comunicación en media, pero el rango de latencias obtenidas son similares [6-7.5 ms].

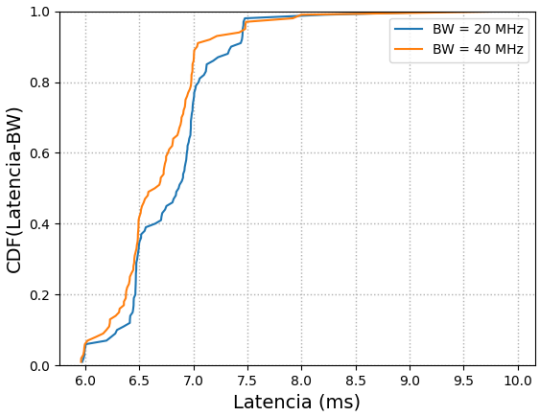


Figura 4. CDF de la latencia en función del BW

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)

B. Subcarrier Spacing (SCS)

El *Subcarrier Spacing (SCS)* es un parámetro muy significativo a la hora de mejorar la latencia, como se aprecia en la Fig. 5. Además, la numerología y el uso de tramas flexibles en 5G, empleando diferentes periodicidades y *subcarrier spacing*, conlleva una mejora significativa de la latencia frente a la configuración estándar de 4G o 5G *Non Standalone (NSA)*. Esta mejora implica una reducción de la latencia de aproximadamente el 50 % y existen SCS superiores que deben tener un comportamiento aún mejor.

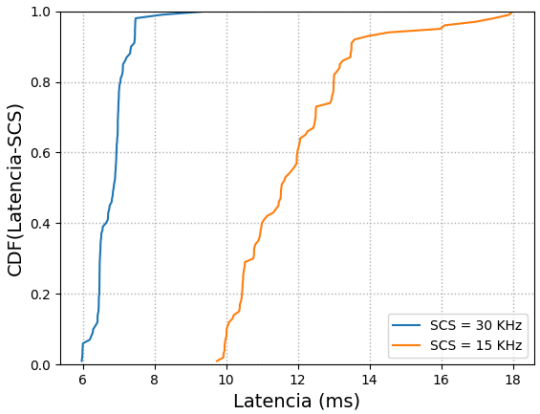


Figura 5. CDF de la latencia en función del *Subcarrier Spacing*

C. Estructura de Trama

Las estructuras de las tramas presentan un rango de latencias comprendido entre 6 y 7.5 ms. Cabe destacar que el caso con $FS = 3$ reduce significativamente la latencia frente al resto y su rango se encuentra entre 6 y 7 ms, como se puede comprobar sobre la Fig. 6. Esto implica que la latencia en este caso es más constante y menor debido a la asignación de más recursos al UL y menos al DL.

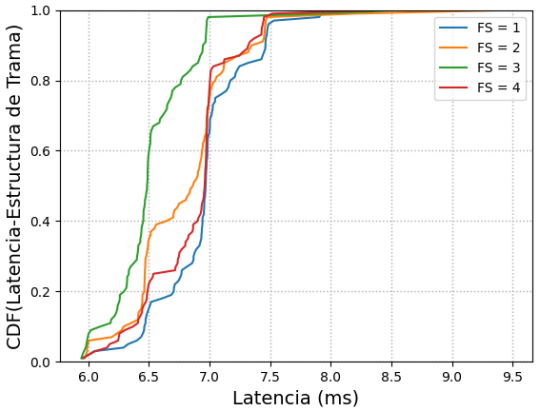


Figura 6. CDF de la latencia en función de la estructura de trama

VI. CONCLUSIONES Y TRABAJOS FUTUROS

En este trabajo se presenta un estudio de la latencia en la capa física 5G NR. De acuerdo a los resultados, se concluye que el *subcarrier spacing* y la estructura de trama afectan significativamente a la latencia; también concluimos que al emplear tramas flexibles, usando mayor numerología, la latencia puede reducirse abruptamente (≈50 %). Por tanto, se considera que, aunque actualmente en este trabajo no se presentan latencias inferiores al milisegundo, estas latencias serán posibles con las futuras mejoras introducidas en Release 16.

Respecto a los trabajos futuros, se están considerando emplear dispositivos que nos permitan aumentar el *subcarrier spacing*, emplear FR2 para aumentar el ancho de banda y reducir la periodicidad de las tramas. Y, por último, realizar un estudio de cómo estos factores afectan también al *throughput* en DL y UL.

AGRADECIMIENTOS

Este trabajo está financiado por el Ministerio español de Economía y Transformación Digital (TSI-063000-2021-28); el Ministerio español de Ciencia e Innovación (PID2019-108713RB-C53 y PID2022-137329OB-C43); y el Ministerio de Universidades español (Beca FPU: 20/02621)

REFERENCIAS

[1] Recommendation ITU-R M.2083: IMT Vision - "Framework and overall objectives of the future development of IMT for 2020 and beyond", Sept. 2015.

[2] Technical Report TR38.912 V15.0.0, "5G; Study on scenarios and requirements for next generation access technologies,"3GPP, Sept. 2019.

[3] J. Navarro-Ortiz, P. Romero-Diaz, S. Sendra, P. Ameigeiras, J. J. Ramos-Munoz and J. M. Lopez-Soler, .A Survey on 5G Usage Scenarios and Traffic Models,ïn *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 905-929, Secondquarter 2020, doi: 10.1109/COMST.2020.2971781

[4] GSMA, "5G TDD Synchronisation Guidelines and Recommendations for the Coexistence of TDD Networks in the 3.5 GHz Range,"Technical document, 2020.

[5] Technical Specification TS38.213, V15.3.0, "5G; NR; Physical layer procedures for control"3GPP, Oct. 2018.

[6] M. Ali and Y. Kabalci, "Throughput Analysis over 5G NR Physical Uplink Shared Channels,"*2020 2nd Global Power, Energy and Communication Conference (IEEE-GPECOM-2020)*, Nov. 2020, doi: 10.1109/GPECOM49333.2020.9247906

[7] M. Mhedhbi, M. Morcos, A. Galindo-Serrano and S. Elayoubi, "Performance Evaluation of 5G Radio Configurations for Industry 4.0,"*15th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob 2019)*, Oct 2019, Barcelone, Spain. HAL id: hal-03687385

[8] L. Chinchilla-Romero, J. Prados-Garzon, P. Muñoz, P. Ameigeiras and J. M. Lopez-Soler, ÜRLLC Achieved Data Rate through Exploiting Multi-Connectivity in Industrial Private 5G Networks with Multi-WAT RANs,"2023 IEEE Wireless Communications and Networking Conference (WCNC 2023), Glasgow, United Kingdom, 2023, pp. 1-6, doi: 10.1109/WCNC55385.2023.10119085.

[9] A. Mohamed, A. Quddus, P. Xiao, B. Hunt and R. Tafazolli, "5G and LTE-TDD Synchronized Coexistence with Blind Retransmission and Mini-Slot Uplink,"*2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, Antwerp, Belgium, 2020, pp. 1-5, doi: 10.1109/VTC2020-Spring48590.2020.9128796.

[10] Technical Specification TS38.300, V16.4.0, "5G; NR and NG-RAN Overall description; Stage-2"3GPP, Jan. 2021.

[11] Technical Report TR38.912 V15.0.0, "5G; Study on scenarios and requirements for next generation access technologies,"3GPP, Sept. 2019.

[12] Technical Specification TS38.211, V16.2.0, "5G; NR; Physical channels and modulation"3GPP, Jul. 2020.

This work is licensed under a Creative Commons 4.0 International License (CC BY-NC-ND 4.0)