

## Escalabilidad de múltiples teleoperaciones hápticas bilaterales sobre red 5G privada

Hernandez-Goberti, Fernando.<sup>a,\*</sup>, Gomez-Barquero, David.<sup>a</sup>

<sup>a</sup>Instituto de Telecomunicaciones y Aplicaciones Multimedia, Universitat Politècnica de València, Edificio 8G, C/ Vera, S/N, 46022 Valencia, España.

### Resumen

Este trabajo estudia cuántas sesiones concurrentes de teleoperación bilateral pueden sostenerse sobre una red privada 5G en condiciones indoor. Se comparan dos modalidades: (i) teleoperación cinestésica con realimentación de fuerza y (ii) teleoperación vibrotáctil con reconocimiento de gestos. Para cada modalidad se evalúan dos patrones multi-sesión ( $N \rightarrow 1$  y  $1 \rightarrow N$ ) y dos políticas QoS (baseline y priorización de comando/retorno). La escalabilidad se cuantifica con latencia/jitter en percentiles altos y con el umbral  $N^*$  (máximo  $N$  antes de superar límites operacionales). En cinestésica,  $N^*$  pasa de 24/16 (baseline) a 30/22 (priorizada). En vibrotáctil,  $N^*$  pasa de 13/8 (baseline) a 16/11 (priorizada), con latencias base mayores que en cinestésica.

*Palabras clave:* Teleoperación bilateral, Interfaz háptica, Redes 5G privadas, Internet táctil, Visión por computador.

### Scalability of multiple bilateral haptic teleoperations over 5G private network

#### Abstract

This study investigates how many concurrent bilateral teleoperation sessions can be sustained over a private 5G network under indoor conditions. Two modalities are compared: (i) kinesthetic teleoperation with force feedback and (ii) vibrotactile teleoperation with gesture recognition. For each modality, two multi-session patterns ( $N \rightarrow 1$  and  $1 \rightarrow N$ ) and two QoS policies (baseline and command/return prioritization) are evaluated. Scalability is quantified using latency/jitter at high percentiles and the threshold  $N^*$  (maximum  $N$  before exceeding operational limits). In kinesthetic teleoperation,  $N^*$  increases from 24/16 (baseline) to 30/22 (prioritized). In vibrotactile teleoperation,  $N^*$  increases from 13/8 (baseline) to 16/11 (prioritized), with higher base latencies than in kinesthetic teleoperation.

*Keywords:* Bilateral teleoperation, Haptic interface, 5G private networks, Tactile internet, Computer vision.

## 1. Introducción y motivación

La teleoperación bilateral háptica se perfila como una tecnología habilitadora para tareas remotas de alta precisión, pero su despliegue real está condicionado por restricciones estrictas de latencia y jitter, especialmente en escenarios dentro del dominio ofrecido por la recientemente acuñada *Tactile Internet* (TI). A medida que se incrementa de una única sesión a configuraciones multi-sesión (colaboración y tele-supervisión), emergen cuellos de botella de red y de planificación de tráfico que pueden degradar la estabilidad, la transparencia y la seguridad percibida del lazo. En paralelo, distintas modalidades de realimentación (cinestésica frente a vibrotáctil) imponen perfiles de tráfico y tolerancias temporales diferentes, lo que obliga a evaluar la escalabilidad de forma comparativa y reproducible. En este contexto, las redes privadas 5G ofrecen un banco

de pruebas idóneo para estudiar políticas de priorización y su impacto en capacidad operativa.

El objetivo de este trabajo es cuantificar, de forma comparativa y reproducible, la capacidad operativa de teleoperación bilateral sobre una red privada 5G cuando se incrementa el número de sesiones concurrentes, y determinar el umbral  $N^*$  a partir de percentiles altos de latencia/jitter bajo dos patrones multi-sesión y dos políticas QoS. En concreto, las contribuciones son:

- (C1) Definición experimental de  $N^*$  con umbrales operacionales absolutos.
- (C2) Comparación de dos modalidades bajo el mismo marco de flujos de datos y patrones multi-sesión.
- (C3) Cuantificación del impacto de priorización de flujos sobre la progresión de latencia y desplazamiento de  $N^*$ .

## 2. Contexto y trabajo relacionado

La literatura reciente coincide en que el rendimiento de la teleoperación háptica no depende solo de la estabilidad del controlador, sino de cómo los *impairments* de red (retardo, jitter y pérdidas) se traducen en degradación perceptiva, de *Quality of Control* (QoC) y, en última instancia, de calidad de experiencia. Estudios psicofísicos y neurocognitivos han mostrado que retrasos en la realimentación de fuerza alteran la percepción de propiedades mecánicas (rigidez, masa o amortiguamiento) y afectan la integración sensoriomotora; además, existe evidencia de que los usuarios pueden discriminar y adaptar su comportamiento ante retardo frente a ruido, lo que motiva definir umbrales operacionales en percentiles altos y no solo en promedios Eden et al. (2024); Zhou et al. (2024). En conjunto, estos resultados respaldan el uso de criterios de ruptura que combinen *retardo* y *variabilidad*, alineados con la práctica de dimensionar el sistema para percentiles altos cuando hay concurrencia.

En el contexto del *Tactile Internet*, los trabajos de estandarización y requisitos han refinado el marco de referencia para fijar objetivos temporales. El informe de IEEE 1918 sobre TI discute escenarios, arquitectura y supuestos técnicos, e incluso matiza que los rangos de latencia aceptables varían según el tipo de teleoperación y dinámica del entorno (Holland et al. (2019)). Más recientemente, ITU-T Y.3149 establece requisitos de aseguramiento QoS para TI (incluyendo objetivos muy exigentes de jitter y latencia en el caso ideal) y refuerza la necesidad de mecanismos de garantía extremo a extremo (International Telecommunication Union (2025)). A nivel de interoperabilidad de señalización, RFC 9695 define el *top-level media type haptics*, facilitando descripción y transporte de datos hápticos en ecosistemas multimedia, lo cual es relevante cuando se combinan flujos críticos y auxiliares (Muthusamy et al. (2025)).

La definición de una capacidad operativa  $N^*$  requiere conectar QoS con un criterio de calidad funcional. En esa dirección, trabajos recientes sobre *tactile cyber-physical systems* muestran una correlación directa entre QoC y las métricas de red (latencia, jitter y pérdidas), y proponen usar QoC como indicador operacional para anticipar degradación de tareas y errores de seguimiento (Polachan et al. (2022)). En conjunto, esta línea de trabajos motiva definir  $N^*$  como el mayor número de sesiones concurrentes que mantiene los percentiles altos dentro de umbrales temporales coherentes con la operabilidad del lazo, y no solo con requisitos ideales del TI.

Finalmente, la escalabilidad también está condicionada por el volumen/ritmo del tráfico háptico. Además de los codecs estandarizados, se han seguido explorando técnicas perceptuales para reducir la tasa de paquetes manteniendo calidad, incluyendo evaluaciones recientes sobre *perceptual deadband* (Hinterseer et al. (2022)). Además, revisiones recientes de protocolos y marcos de háptica en redes (incluyendo metaverso/teleoperación) resumen cómo distintos mecanismos de transporte y planificación buscan controlar latencia y jitter bajo condiciones de carga (Faisal et al. (2025)). Sin embargo, sigue habiendo una brecha en estudios experimentales compactos que comparen, con un criterio operacional explícito ( $N^*$ ), la influencia conjunta de *modalidad háptica*, *patrón multi-sesión* y *política QoS* sobre una infraestructura 5G controlada. Este trabajo se sitúa en esa intersección, aportando una comparación reproducible basada

en percentiles y umbrales que reflejan tanto requerimientos de red como restricciones perceptivas y operacionales.

## 3. Sistema y modalidades de estudio

El banco de pruebas implementa teleoperación bilateral maestro-esclavo sobre una red privada 5G (Figura 1), con separación explícita entre (i) adquisición/interpretación de intención del operador, (ii) ejecución de control y adquisición de estado en el robot remoto, y (iii) renderizado háptico local. Las mediciones se realizan bajo cobertura NR n78 (TDD) en iTEAM-UPV usando un portador de 100 MHz en el rango 3500–3600 MHz y utilizando un despliegue 5G SA alineado con 3GPP Release 17 con equipamiento radio/edge de Nokia y cores gestionados por Open5GS. Para garantizar reproducibilidad y permitir escalado multi-sesión sin necesidad de desplegar múltiples robots físicos, el sistema soporta la reproducción de trazas con marcas temporales y la virtualización de instancias (gemelos digitales) para generar carga representativa en la red.

Para implementación, se distinguen dos roles lógicos:

- **Nodo remoto de robótica (RN):** aloja el manipulador, el efector final y los procesos de control/telemetría. La adquisición de estado y la publicación de variables se realizan mediante interfaces deterministas de alta frecuencia.
- **Nodo de operador (ON):** captura la intención (movimiento, gestos) y genera los comandos. Según la modalidad, también renderiza fuerza o vibrotáctil y ejecuta módulos de percepción (detección de mano/gestos).

### 3.1. M1: teleoperación bilateral cinestésica

La modalidad cinestésica emplea un dispositivo háptico de realimentación de fuerza **Touch X** como interfaz maestro. El Touch X proporciona posicionamiento de alta resolución y capacidad de fuerza continua para interacción háptica; sus especificaciones públicas reportan resolución nominal del orden de 0.055 mm (“> 450 dpi”) y fuerza máxima de hasta 7.9 N en condiciones nominales. El esclavo remoto es un manipulador colaborativo **UR5e** con 6 GDL, alcance 850 mm y carga útil 5 kg, con repetibilidad de pose  $\pm 0.03$  mm según ISO 9283. Como efector final se utiliza la pinza paralela **OnRobot RG2**, que ofrece un recorrido máximo de 110 mm y funciones integradas de detección de agarre.

El Touch X proporciona realimentación de fuerza cinestésica de traslación a la mano del usuario cuando se detecta interacción/contacto con el objeto en el entorno remoto. En cada instante, el RN estima la fuerza de interacción  $F_{env}$  a partir del estado de la pinza y la detección de contacto/agarre por error entre posición actual y objetivo del efector, y envía esta magnitud en el flujo de retorno junto con el estado (pose/velocidad). El ON renderiza dicha fuerza sobre el Touch X aplicando un campo de fuerzas en el espacio cartesiano del dispositivo:

$$F_{cmd} = \text{sat}(G \cdot F_{env}),$$

donde  $G$  es una ganancia de escalado y  $\text{sat}(\cdot)$  limita la fuerza al rango seguro del dispositivo. Cuando no se detecta contacto, se renderiza  $F_{cmd} \approx 0$ , evitando perturbaciones innecesarias.

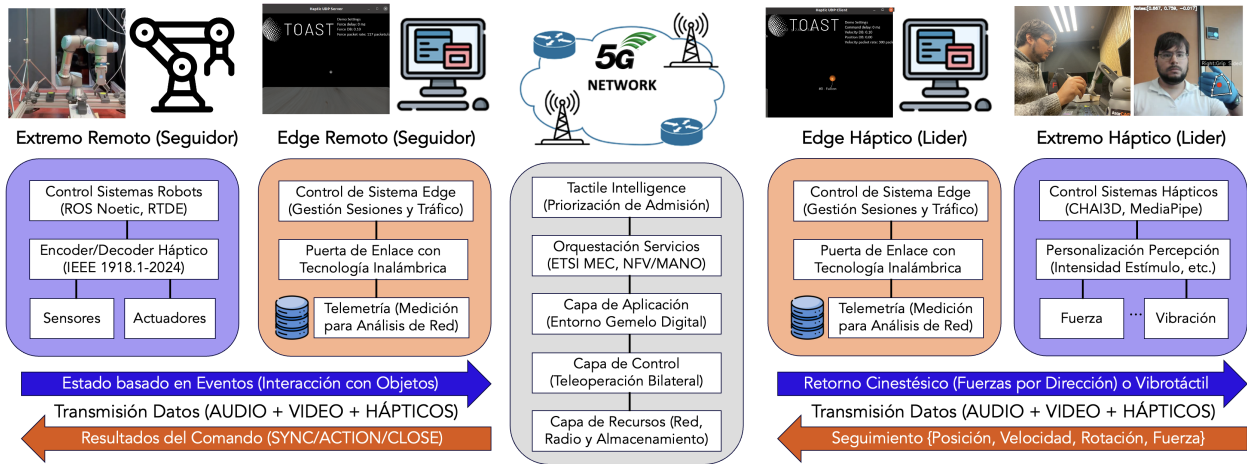


Figura 1: Arquitectura experimental UR5e+RG2 ↔ Red 5G ↔ (Touch X || RealSense+TactGloves) con flujos considerados como críticos (← comando/seguimiento y → retorno/estado) y auxiliares (telemetría/señalización)

Para renderizado háptico y virtualización se emplea la herramienta **CHAI3D**, un framework abierto para el desarrollo rápido de escenas hápticas y simulación interactiva. La interfaz con UR5e se basa en **RTDE** (Real-Time Data Exchange), diseñada para sincronizar aplicaciones externas con el controlador sin romper sus propiedades de tiempo real; puede generar mensajes de salida hasta la frecuencia del lazo de control (500 Hz en robots e-Series), omitiendo paquetes si faltan recursos y enviando el estado más reciente.

La arquitectura de control se organiza de forma que:

- El **maestro** genera consignas de posición/velocidad (y eventos de control discretos) a tasa configurada.
- El **esclavo** ejecuta seguimiento y retorna estado (pose, velocidad, estados de control y variables de renderizado).
- La **capa háptica local** (en ON) mantiene estabilidad del bucle local de fuerza para pasividad.

### 3.2. M2: teleop. vibrotáctil con reconocimiento de gestos

La modalidad vibrotáctil utiliza guantes **bHaptics TactGlove DK2** para renderizado táctil discreto (vibración) en dedos y muñeca. El fabricante describe 12 actuadores hápticos y control de *waveforms* de 11 ms mediante su SDK. Como entrada se emplea una cámara RGB (720p@30fps) para capturar la mano del operador y extraer una representación de pose/landmarks. En esta modalidad, el retorno es por patrones de vibración definidos por índices de actuadores, intensidades y duraciones de estímulos generados a partir de eventos y estados del sistema remoto (contacto/agarre y nivel de esfuerzo estimado).

La detección de mano y landmarks se implementa con **MediaPipe Hands**, que estima un esqueleto de mano (21 landmarks) a partir de una única cámara RGB. En operación, el pipeline típico se compone de:

1. **Detección de palma y encuadre** para localizar la mano.
2. **Estimación de landmarks** que produce coordenadas normalizadas y/o en espacio mundo.
3. **Clasificación de gesto/agarre**: a partir de features (relaciones geométricas entre landmarks, aperturas relativas, orientación), se clasifica la intención (p.ej., abrir/cerrar pinza, selección de fuerza/estímulo).

El módulo háptico vibrotáctil transforma el estado del sistema remoto (eventos de contacto, anchura de pinza, fuerza estimada o estados discretos) en estímulos táctiles. Con el SDK del fabricante es posible controlar patrones en actuadores específicos, así como su duración e intensidad temporal.

## 4. Metodología experimental

### 4.1. Modelo de comunicación, flujos y políticas QoS

En ambas modalidades se consideran tres clases de flujos (por sesión  $s$ ):

- **CMD** (*comando/seguimiento*, uplink): consignas continuas y eventos discretos hacia el RN.
- **RET** (*retorno/estado*, downlink): estado del robot y variables necesarias para feedback (fuerza o estímulos).
- **SIG** (*señalización/telemetría*, bidireccional): establecimiento de sesión, sincronización, logging y métricas.

Denotamos por  $d_{\text{CMD}}^{(s)}(k)$  y  $d_{\text{RET}}^{(s)}(k)$  la latencia extremo-extremo (E2E) para el paquete  $k$  en cada flujo, y por  $j_{\text{CMD}}^{(s)}$  y  $j_{\text{RET}}^{(s)}$  el jitter (como variación intercuantil). La medición se implementa con marcas temporales en emisor y receptor, y sincronización de relojes por NTP.

Se comparan dos políticas de priorización:

1. **Baseline**: sin priorización explícita (best effort en colas).
2. **Prioritizada**: marcado y planificación diferenciada de flujos críticos.

Para la política priorizada, los flujos CMD y RET se marcan como tráfico de baja latencia mediante **DiffServ EF**. El PHB EF está diseñado para proporcionar *baja pérdida, baja latencia y bajo jitter* asegurando que el agregado EF recibe servicio a una tasa configurada. En equipos de red, esto es implementado con una *cola de prioridad estricta* (strict priority) donde el tráfico prioritario se atiende antes que el resto mientras existe backlog. El resto de tráfico (SIG/telemetría) se mantiene como best effort y se limita mediante *shaping* para evitar ráfagas.

#### 4.2. Patrones multi-sesión y generación de carga

Para comparar escalabilidad bajo condiciones controladas se definen dos patrones para cada modalidad de teleoperación por separado (Figura 2):

- $N \rightarrow 1$  (**multi-operador / un robot**): múltiples ON (uno físico y resto duplicados) reciben RET y pueden solicitar control. La autoridad de comando se arbitra (mediante tokens), de forma que sólo un operador emite CMD efectivo en cada instante, evitando conflictos.
- $1 \rightarrow N$  (**un operador / multi-instancia**): una sesión física se complementa con  $N-1$  sesiones emuladas que reproducen trazas para generar carga representativa sin necesidad de  $N$  robots o dispositivos físicos.

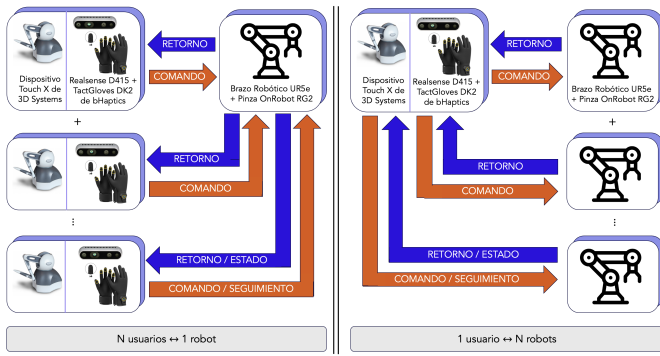


Figura 2: Patrones para teleoperación bilateral háptica y escalabilidad multi-sesión: flujos CMD/RET y patrones  $N$  usuarios-1 robot y 1 usuario- $N$  robots.

Definimos  $N^*$  como el máximo número de sesiones concurrentes para el que la calidad temporal de los flujos críticos (CMD/RET) permanece por debajo de umbrales operativos:

$$N^* = \max \{ N : Q_p(d_{\text{crit}}) \leq \tau_d \wedge Q_p(j_{\text{crit}}) \leq \tau_j \}, \quad (1)$$

donde  $Q_p(\cdot)$  es el percentil  $p95$ ,  $d_{\text{crit}}$  y  $j_{\text{crit}}$  se calculan sobre CMD y/o RET (según modalidad), y  $\{\tau_d, \tau_j\}$  son umbrales fijados por requisitos de lazo y seguridad operacional ( $\{60, 10\}$  ms para cinestésica y  $\{150, 20\}$  ms para vibrotáctil).

Para las sesiones emuladas, se adopta un enfoque *trace-driven*, con cada traza conteniendo marcas temporales de envío/recepción y tamaños de PDU por flujo, permitiendo instanciar  $N-1$  sesiones emuladas con desfases temporales controlados y tasas por sesión constantes.

## 5. Resultados

En teleoperación **cinestésica** (Tablas 1-2 y Figura 3), se observa que la priorización de CMD/RET (marcado EF y cola de prioridad estricta) incrementa la capacidad en  $\approx 25\%$  para  $N \rightarrow 1$  y  $\approx 37.5\%$  para  $1 \rightarrow N$  respecto al baseline, manteniendo la latencia  $p95$  del lazo crítico en el rango  $\approx 40-60$  ms en el entorno de  $N^*$ . En particular, para  $N \rightarrow 1$  y  $N=24$  (cercano al baseline), la priorización reduce la latencia  $p95$  de CMD de 55 a 46 ms y de RET de 58 a 50 ms, junto con una caída del

jitter  $p95$  de  $\approx 10$  a  $\approx 7$  ms en ambos flujos, retrasando la violación del umbral operacional y permitiendo alcanzar  $N^* \approx 30$ . Análogamente, en  $1 \rightarrow N$  para  $N=16$  (punto de ruptura en baseline), la priorización desplaza CMD de 55 a 43 ms y RET de 60 a 48 ms, con jitter  $p95$  pasando de  $\approx 10$  a  $\approx 6$  ms, y mantiene el lazo operable hasta  $N=22$ .

Tabla 1: Resultados en formato “Latencia (Jitter)” de patrón  $N \rightarrow 1$  en el caso actual de estudio de teleoperación bilateral cinestésica.

$N$	Baseline (BE) [ms]		QoS priorizada (EF) [ms]	
	CMD p95	RET p95	CMD p95	RET p95
1	24.3 (3.0)	26.2 (3.0)	24.4 (3.0)	26.3 (3.0)
5	28.2 (4.0)	30.7 (4.0)	27.6 (3.5)	29.5 (3.5)
10	33.5 (5.0)	35.3 (5.0)	31.9 (4.0)	33.7 (4.0)
15	38.4 (6.0)	41.1 (6.0)	35.3 (5.0)	38.4 (5.0)
20	46.7 (8.0)	49.8 (8.0)	41.5 (6.0)	45.2 (6.0)
24	55.8 (10.0)	58.2 (10.0)	46.7 (7.0)	50.6 (7.0)
30	—	—	54.4 (9.0)	58.7 (9.0)

Tabla 2: Resultados en formato “Latencia (Jitter)” de patrón  $1 \rightarrow N$  en el caso actual de estudio de teleoperación bilateral cinestésica.

$N$	Baseline (BE) [ms]		QoS priorizada (EF) [ms]	
	CMD	RET	CMD	RET
1	24.3 (3.0)	26.2 (3.0)	24.4 (3.0)	26.3 (3.0)
5	32.7 (5.0)	35.8 (5.0)	30.1 (4.0)	33.5 (4.0)
10	42.6 (7.0)	45.3 (7.0)	36.4 (5.0)	40.9 (5.0)
16	55.2 (10.0)	60.5 (10.0)	43.1 (6.0)	48.2 (6.0)
20	—	—	50.7 (7.0)	56.9 (7.0)
22	—	—	55.4 (8.0)	60.4 (8.0)

En teleoperación **vibrotáctil** con reconocimiento de gestos (Tablas 3-4 y Figura 4), el comportamiento temporal viene determinado por la cadena percepción→CMD→RET, donde la etapa de visión y estimación de mano introduce una latencia base adicional sobre el transporte. En las mediciones se registraron latencias medias E2E del orden de 95 ms con TCP y 80 ms con UDP, observándose en este último caso una mayor variabilidad temporal asociada a la naturaleza no orientada a conexión. La campaña multi-sesión se ejecutó íntegramente en interior bajo condiciones homogéneas, obteniéndose directamente la capacidad operativa  $N^*$  para ambos patrones: en *baseline* se alcanzaron  $N^* \approx 13$  y  $N^* \approx 8$  para  $N \rightarrow 1$  y  $1 \rightarrow N$ , respectivamente, mientras que la QoS priorizada de los flujos críticos desplazó el punto de ruptura a  $N^* \approx 16$  y  $N^* \approx 11$ , confirmando que proteger CMD/RET como agregado de baja latencia/jitter incrementa la concurrencia sostenible, especialmente en  $1 \rightarrow N$  donde domina la contención entre sesiones.

### 5.1. Escalabilidad multi-sesión de modalidad cinestésica

En el patrón  $N \rightarrow 1$ , el tráfico agregado crece de forma moderada al existir un único esclavo, pero aumentan la variabilidad temporal y el ensanchamiento de colas en instantes de alta actividad (cambios rápidos de consigna y transiciones). En baseline, a partir de  $N^* \approx 24$  los percentiles altos de latencia y jitter superan el umbral operacional; el punto de ruptura se identifica por: (a) incremento no lineal del jitter, (b) aparición de variabilidad/colas y pérdidas intermitentes (especialmente cuando parte

del tráfico se transporta por UDP), y (c) degradación apreciable en estabilidad/seguimiento (sobreoscilaciones y necesidad de correcciones por el operador). La QoS priorizada desplaza el punto de ruptura a  $N^* \approx 30$ , reduciendo el crecimiento de jitter en alta carga y manteniendo la latencia p95 del lazo crítico en el rango  $\approx 40\text{--}60$  ms en el entorno de  $N^*$ .

En el patrón  $1 \rightarrow N$ , la tasa agregada crece aproximadamente linealmente con  $N$  y el umbral de servicio queda condicionado por contención entre sesiones (competencia por ancho de banda/colas) más que por la dinámica del robot. En baseline, para  $N > 16$  se observa incremento simultáneo de jitter, pérdidas intermitentes y variabilidad en la latencia de retorno, manifestándose en episodios de degradación del seguimiento y una ejecución menos continua del control (micro-paradas o correcciones tardías). La QoS priorizada desplaza el punto de ruptura a  $N^* \approx 22$ , con una mejora relativa mayor que en  $N \rightarrow 1$ , al proteger explícitamente los flujos críticos frente a la congestión inducida por sesiones paralelas.

## 5.2. Escalabilidad de modalidad vibrotáctil con gestos

En indoor, la modalidad vibrotáctil muestra una latencia E2E mayor que la cinestésica debido al coste de percepción y reconocimiento de gestos (detección de mano/landmarks, clasificación y empaquetado de comandos), además del transporte y el retorno de estado/háptica. En las mediciones indoor reportadas para el sistema vibrotáctil, la latencia media E2E con TCP se sitúa en torno a 95 ms, mientras que UDP reduce la media a  $\approx 80$  ms a costa de un jitter más alto (arribos menos predecibles). Este resultado motiva el uso de combinaciones mixtas de protocolo (TCP para control crítico y UDP para retorno táctil cuando se prioriza latencia), o bien el uso de UDP con mecanismos de protección a nivel de aplicación si se busca minimizar la cola.

Tabla 3: Resultados en formato “Latencia (Jitter)” de patrón  $N \rightarrow 1$  en el caso actual de estudio de teleoperación bilateral vibrotáctil.

$N$	Baseline (BE) [ms]		QoS priorizada (EF) [ms]	
	CMD p95	RET p95	CMD p95	RET p95
1	85.4 (5.0)	92.7 (6.0)	85.9 (4.0)	92.1 (5.0)
3	92.3 (6.0)	98.1 (7.0)	90.8 (4.8)	96.9 (5.8)
5	99.7 (7.0)	106 (8.0)	95.5 (5.5)	102 (6.5)
8	110 (8.0)	118 (9.0)	102 (6.5)	110 (7.5)
10	118 (9.0)	126 (10.0)	108 (7.0)	116 (8.0)
13	128 (10.0)	138 (12.0)	116 (8.0)	124 (9.0)
16	—	—	126 (9.5)	134 (10.5)

Tabla 4: Resultados en formato “Latencia (Jitter)” de patrón  $1 \rightarrow N$  en el caso actual de estudio de teleoperación bilateral vibrotáctil.

$N$	Baseline (BE) [ms]		QoS priorizada (EF) [ms]	
	CMD p95	RET p95	CMD p95	RET p95
1	85.4 (6.0)	92.7 (7.0)	85.9 (5.0)	92.1 (6.0)
3	100 (8.0)	110 (9.0)	96.2 (6.5)	106 (7.5)
5	112 (10.0)	124 (11.0)	104 (7.5)	116 (8.5)
8	130 (13.0)	146 (14.0)	116 (9.0)	130 (10.0)
11	—	—	128 (10.5)	142 (11.5)

Bajo baseline, el patrón  $N \rightarrow 1$  alcanza  $N^* \approx 13$ , mientras que el patrón  $1 \rightarrow N$  alcanza  $N^* \approx 8$ , reflejando que la

contención agregada penaliza antes cuando un único operador gestiona múltiples instancias. Con QoS priorizada (marcado EF para los flujos equivalentes a CMD/RET, y limitación de telemetría), el punto de ruptura se desplaza a  $N^* \approx 16$  y  $N^* \approx 11$  respectivamente. La ganancia relativa (23–38 %) es comparable a la observada en la modalidad cinestésica, sugiriendo que la priorización es especialmente efectiva en regímenes de alta concurrencia donde el rendimiento viene dominado por colas.

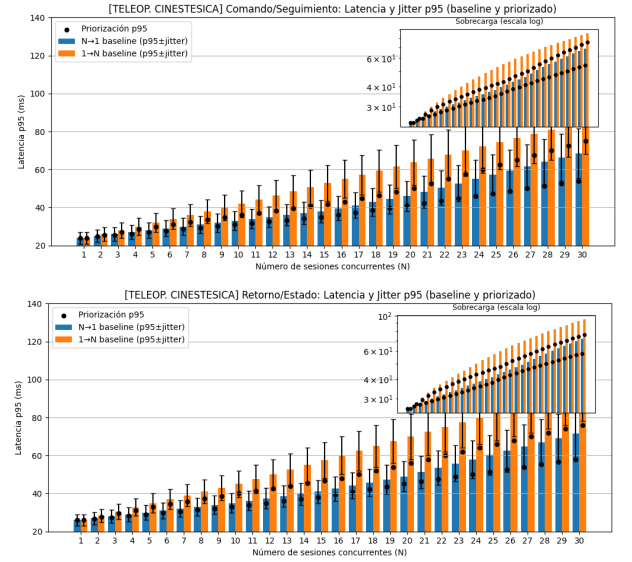


Figura 3: Latencia y Jitter en CMD/RET vs  $N$  para patrones  $N \rightarrow 1$  y  $1 \rightarrow N$  en modalidad cinestésica (M1) con políticas baseline y priorizadas.

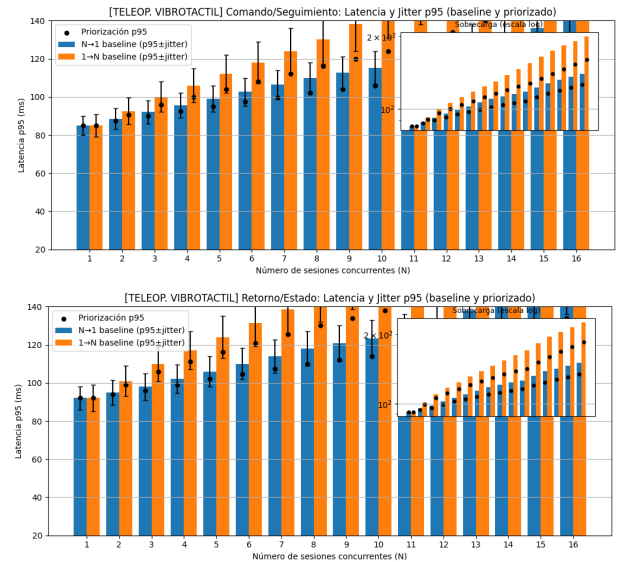


Figura 4: Latencia y Jitter en CMD/RET vs  $N$  para patrones  $N \rightarrow 1$  y  $1 \rightarrow N$  en modalidad bilateral vibrotáctil (M2) con políticas baseline y priorizadas.

## 6. Discusión

**Modalidad.** En los experimentos cinestésicos, la evolución de los percentiles altos con  $N$  muestra un incremento moderado hasta el entorno de ruptura, y la priorización desplaza de forma clara ese punto manteniendo valores operables: por ejemplo, en  $N \rightarrow 1$  a  $N=24$  la latencia p95 se reduce de 55.8→46.7 ms

(CMD) y de 58.2→50.6 ms (RET), lo que deja margen suficiente para sostener  $N=30$  con 54.4/58.7 ms. En vibrotáctil, el presupuesto temporal parte más alto incluso con pocas sesiones (p.ej.,  $N=1$ : 85.4/92.7 ms en CMD/RET), y el crecimiento con  $N$  es más acusado: en  $N \rightarrow 1$  el RET p95 aumenta de 92.7 ms a 138 ms en  $N=13$ , mientras que con priorización se mantiene en 124 ms y permite llegar a  $N=16$  (134 ms). Esta diferencia entre modalidades se refleja también en la separación CMD–RET: en vibrotáctil el RET tiende a situarse sistemáticamente por encima del CMD (p.ej.,  $N=10$  baseline: 118 vs 126 ms), lo que indica que el retorno/estado es el primer flujo en tensionarse cuando aumenta la concurrencia.

**Patrón multi-sesión.** En  $1 \rightarrow N$  se observa una aceleración de la degradación al crecer  $N$  que es visible en los saltos de p95 entre niveles: en cinestésica baseline, el RET pasa de 45.3 ms ( $N=10$ ) a 60.5 ms ( $N=16$ ), mientras que con priorización el mismo tramo queda contenido en 40.9→48.2 ms y se mantiene operable hasta  $N=22$  (60.4 ms). En vibrotáctil, el mismo patrón se acentúa: baseline incrementa RET de 124 ms ( $N=5$ ) a 146 ms ( $N=8$ ), mientras que con priorización el valor en  $N=8$  se reduce a 130 ms y se prolonga el funcionamiento hasta  $N=11$  (142 ms). Un aspecto relevante observado en ambas modalidades es que, cuando el patrón es  $1 \rightarrow N$ , la separación entre baseline y priorizada crece con  $N$  (p.ej., en cinestésica RET a  $N=16$ : 60.5 vs 48.2 ms; en vibrotáctil RET a  $N=8$ : 146 vs 130 ms), lo que sugiere que la ganancia está ligada a la contención entre sesiones y no a un desplazamiento constante.

**Política de priorización.** El efecto experimental de la priorización se aprecia como una reducción simultánea de CMD y RET, pero con mayor impacto en RET en el entorno de ruptura, que es precisamente donde la operabilidad se compromete primero: en cinestésica  $N \rightarrow 1$  a  $N=24$  el ahorro es de 9.1 ms (CMD) y 7.6 ms (RET), mientras que en  $1 \rightarrow N$  a  $N=16$  el ahorro asciende a 12.1 ms (CMD) y 12.3 ms (RET). En vibrotáctil, aunque las reducciones absolutas en milisegundos son algo menores en  $N \rightarrow 1$  (p.ej., a  $N=13$ : RET 138→124 ms), resultan decisivas para desplazar el punto de ruptura (de  $N^* \approx 13$  a  $N^* \approx 16$ ) y evitar que los percentiles altos superen rápidamente los umbrales operacionales. En conjunto, los datos muestran que la priorización no solo “baja” las curvas, sino que suaviza la pendiente efectiva cerca de  $N^*$ , haciendo que el incremento de p95 por sesión adicional sea menor en la región crítica y ampliando la ventana de concurrencia en la que el lazo permanece utilizable.

## 7. Conclusiones

En este trabajo se ha caracterizado, bajo condiciones *indoor* comparables, la escalabilidad de dos modalidades de teleoperación bilateral sobre una red privada 5G: (i) cinestésica y (ii) vibrotáctil con reconocimiento de gestos. Los resultados confirman que la capacidad operativa  $N^*$  está fuertemente condicionada por la política QoS y por el patrón multi-sesión. En la modalidad cinestésica,  $N^*$  aumenta de  $N^* \approx 24$  a  $N^* \approx 30$

en  $N \rightarrow 1$  y de  $N^* \approx 16$  a  $N^* \approx 22$  en  $1 \rightarrow N$  al priorizar CMD/RET, manteniendo la latencia p95 del lazo crítico en el rango  $\approx 40$ – $60$  ms en el entorno de ruptura. En la modalidad vibrotáctil, aun partiendo de una latencia base superior (p95 indoor  $\approx 80$ – $95$  ms), la priorización desplaza  $N^*$  de  $N^* \approx 13$  a  $N^* \approx 16$  ( $N \rightarrow 1$ ) y de  $N^* \approx 8$  a  $N^* \approx 11$  ( $1 \rightarrow N$ ), evidenciando que proteger los flujos críticos es especialmente efectivo cuando el rendimiento está dominado por contención y colas. En conjunto, los hallazgos refuerzan la necesidad de mecanismos de aseguramiento de QoS para aplicaciones hápticas y muestran que una asignación diferenciada permite incrementar de forma tangible la capacidad multi-sesión sin comprometer el comportamiento temporal de los flujos críticos. Como trabajo futuro, ampliaremos el estudio incorporando (i) *haptic codecs* y compresión perceptual para reducir carga bidireccional sin degradar la interacción, y (ii) asignación/adaptación dinámica de *QoS Flows* (p.ej., mapeos 5QI y priorización extremo a extremo) para maximizar  $N^*$  manteniendo percentiles altos de latencia dentro de los requisitos de *Tactile Internet*.

## Agradecimientos

Este trabajo fue apoyado por el programa de investigación e innovación Horizonte Europa de la Unión Europea (HORIZON-MSCA-2022-DN-01) a través del proyecto TOAST bajo el acuerdo de subvención Marie Skłodowska-Curie n.º 101073465.

## Referencias

- Eden, J., et al., 2024. During haptic communication, the central nervous system distinguishes delay and noise and uses compensation mechanisms. *PLOS Computational Biology*.  
URL: <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1012037>
- Faisal, M., et al., 2025. Haptic network protocols: A comprehensive review and directions for next-gen metaverse applications. *ACM Computing Surveys*.  
URL: <https://dl.acm.org/doi/full/10.1145/3759459>
- Hinterseer, P., et al., 2022. Perceptual deadband for haptic data compression. *RO-MAN*.  
URL: <https://dl.acm.org/doi/10.1109/RO-MAN53752.2022.9900602>
- Holland, O., et al., 2019. The IEEE 1918.1 “tactile internet” standards working group and its standards. *Proceedings of the IEEE*.  
URL: <https://www.st.ewi.tudelft.nl/vijay/pdfs/holland2019ieee1918.pdf>
- International Telecommunication Union, 2025. Recommendation ITU-T Y.3149 (08/2025): Quality of service assurance requirements for the tactile internet. <https://www.itu.int/rec/T-REC-Y.3149/en>, accessed: 2026-02-02.
- Muthusamy, Y. K., et al., 2025. RFC 9695: The haptics top-level media type. RFC Editor.  
URL: <https://www.rfc-editor.org/rfc/rfc9695.pdf>
- Polachan, K., et al., 2022. Assessing quality of control in tactile cyber-physical systems. *EEE Transactions on Network and Service Management*.  
URL: [https://cnl.iisc.ac.in/assets/pdf/papers/Assessing\\_Quality\\_of\\_Control\\_in\\_Tactile\\_CyberPhysical\\_Systems.pdf](https://cnl.iisc.ac.in/assets/pdf/papers/Assessing_Quality_of_Control_in_Tactile_CyberPhysical_Systems.pdf)
- Zhou, T., et al., 2024. Neural dynamics of delayed feedback in robot teleoperation. *Frontiers in Neurobotics*.  
URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC11215083/>