

La investigación reportada en esta tesis es parte de los programas de investigación del CICESE (Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California).

La investigación fue financiada por el SECIHTI (Secretaría de Ciencia, Humanidades, Tecnología e Innovación).

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México). El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo o titular de los Derechos de Autor.

**Centro de Investigación Científica y de Educación
Superior de Ensenada, Baja California**



**Maestría en Ciencias
en Electrónica y Telecomunicaciones
con orientación en Telecomunicaciones**

**Diseño de una arquitectura cliente-servidor para la
transmisión en tiempo cuasi-real de
señales de audio cardiopulmonar comprimidas**

Tesis
para cubrir parcialmente los requisitos necesarios para obtener el grado de
Maestro en Ciencias

Presenta:

Isaac Manuel Campos Guzmán

Ensenada, Baja California, México
2025

Tesis defendida por
Isaac Manuel Campos Guzmán

y aprobada por el siguiente Comité

Dr. Salvador Villareal Reyes
Director de tesis

Dr. José Eleno Lozano Risk

Dr. Roberto Conte Galván

Dra. Rosa Reyna Mouriño Pérez

M. C. Christian Yair Soto Olivares



Dra. María del Carmen Maya Sánchez
Coordinadora del Posgrado en Electrónica y
Telecomunicaciones

Dra. Ana Denise Re Araujo
Directora de Estudios de Posgrado

Resumen de la tesis que presenta **Isaac Manuel Campos Guzmán** como requisito parcial para la obtención del grado de Maestro en Ciencias en Electrónica y Telecomunicaciones con orientación en Telecomunicaciones

Diseño de una arquitectura cliente-servidor para la transmisión en tiempo cuasi-real de señales de audio cardiopulmonar comprimidas

Resumen aprobado por:

Dr. Salvador Villareal Reyes
Director de tesis

En México, la introducción de tecnologías avanzadas en el ámbito de la atención médica de la medicina podría traducirse en un aumento significativo en el acceso a servicios de salud básicos, especialmente en regiones rurales o con escasa disponibilidad de profesionales especializados. Esto podría tener un efecto positivo en mejorar la calidad de vida de los habitantes de estas áreas, al facilitar el acceso a servicios médicos esenciales mediante dispositivos de consulta básica que permiten diagnósticos más accesibles y oportunos, los cuales incluyen el estetoscopio. Este es un dispositivo ubicuo, ya que es utilizado por todo el personal médico para llevar a cabo evaluaciones básicas (auscultaciones) del estado de la salud del paciente. Por esta razón, su inclusión en los sistemas de telemedicina es relevante para brindar evaluaciones físicas remotas. El presente trabajo de investigación tiene como objetivo el diseño de una arquitectura de comunicaciones para la transmisión de señales de audio cardiopulmonar comprimido en tiempo cuasi-real, empleando una arquitectura híbrida P2P y cliente-servidor que considere un esquema de cifrado básico para el flujo de información. La arquitectura desarrollada incluyó el diseño de un protocolo de señalización, la integración de técnicas transversales de NAT y la implementación de los programas de encardio y decardio, previamente desarrollados en un trabajo de investigación anterior. Por ello, se generó un streaming de audio cardiopulmonar comprimido de mejor ancho de banda que algunas empresas, como lo es Littmann™. Dando como resultado un consumo de datos máximo de 65.73 kbps, lo que representa 1.9 veces menor que las soluciones brindadas por Littmann. Asimismo, se evaluó la calidad del audio para determinar si es adecuado para hacer un diagnóstico médico a distancia con apoyo de profesionales de la salud. Se comparaba la tele-auscultación del sistema propuesto contra las plataformas convencionales y el usuario (especialista) contesta una encuesta basada en la prueba MUSHRA. La conclusión que brindó el personal de la salud de manera unánime fue que el sistema EAC tiene una transmisión de audio con calidad diagnóstica.

Palabras clave: tele-auscultación, audio cardiopulmonar, arquitectura P2P híbrida, códec de audio, consumo de datos, estetoscopio electrónico, e-salud.

Abstract of the thesis presented by **Isaac Manuel Campos Guzmán** as a partial requirement to obtain the Master of Science degree in Electronics and Telecommunications with orientation in Telecommunications

Design of a client-server architecture for quasi-real-time transmission of compressed cardiopulmonary audio signals

Abstract approved by:

Dr. Salvador Villarreal Reyes
Thesis Director

In Mexico, the introduction of advanced technologies in the field of medical care could lead to a significant increase in access to basic health services, particularly in rural regions or areas with limited availability of specialized professionals. This could have a positive impact on improving the quality of life for residents in these areas by facilitating access to essential medical services through basic consultation devices that enable more accessible and timely diagnoses, including the stethoscope. The stethoscope is a ubiquitous device, as it is used by all medical personnel to perform basic evaluations (auscultations) of a patient's health status. For this reason, its inclusion in telemedicine systems is relevant for providing remote physical assessments. The present research aims to design a communication architecture for the transmission of compressed cardiopulmonary audio signals in near real-time, employing a hybrid P2P and client-server architecture that incorporates a basic encryption scheme for the information flow. The developed architecture included the design of a signaling protocol, the integration of NAT traversal techniques, and the implementation of the encardio and decardio programs, which were previously developed in a prior research project. As a result, a compressed cardiopulmonary audio streaming system was generated with better bandwidth efficiency compared to some companies, such as Littmann™. The system achieved a maximum data consumption of 65.73 kbps, which is 1.9 times lower than the solutions provided by Littmann. Additionally, the audio quality was evaluated to determine whether it is suitable for remote medical diagnosis with the support of healthcare professionals. The proposed system's tele-auscultation was compared against conventional platforms, and the user (specialist) completed a survey based on the MUSHRA test. The unanimous conclusion provided by healthcare professionals was that the EAC system delivers audio transmission with diagnostic-quality sound.

Keywords: tele-auscultation, cardiopulmonar audio, hybrid P2P architecture, audio codec, data consumption, electronic stethoscope, e-health.

Dedicatoria

**Para todas aquellas personas que sostienen en
sus almas el mundo de mi existencia, que me
acompañan en cada nuevo camino y se
encuentran en las pequeñas cosas que dan vida
y sobre todo brindan ese amor...
incondicionalmente eterno a los corazones, que
a pesar de que se encuentran a la distancia, no
olvidan nunca cómo se siente estar en casa,
puesto que siempre los llevo conmigo.**

(Mis papás: Domy y Saúl)

(Mis hermanos: Omar, Jona y Mariel)

**(Mis hermanos del alma: Edder, Karla, Stephania,
Karen y Liz)**

Agradecimientos

Agradezco a la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) por brindar el apoyo económico para la realización de mis estudios de maestría.

Agradezco al Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California (CICESE) por abrirme sus puertas para realizar un futuro mejor.

Mi más sincero agradecimiento a mi asesor y director, Dr. Salvador Villarreal Reyes. Por su confianza, mucha paciencia, la oportunidad y orientación en todo este trayecto llamado posgrado y también a su fabulosa familia.

De la misma forma, agradecer a cada miembro de mi comité de tesis, el M. C. Christian Soto, Dra. Rosa Mouriño, Dr. Roberto Conte y Dr. José Lozano, por todos sus consejos que contribuyeron enormemente a la mejora de este trabajo de investigación.

Agradezco al laboratorio ARTS, que más que un laboratorio es una gran familia de la cual es un honor ser parte. Gracias a todos mis compañeros de laboratorio por sus consejos y aligerar la vida en el mar: Mari, Eduardo, Miguel, Zeydel, Neto y a las integrantes honorarias Johana y Eli.

Asimismo, a todos mis amigos que a la distancia existen y hacen del mundo un lugar mejor: Edder, Stephania, Karla, Karen, Naian, Moño, Roberto, Sofia, Caro, Alo, Isabel, Víctor, Brayans e Itzel. De igual modo, a los amigos que me brindaron las olas y calmaron las tormentas: Liz, Charlien, Erik, Andrea, Jules, Daniel e Itzel.

Estoy agradecido con todo el personal médico que brindó su valiosa ayuda y conocimientos para evaluar este proyecto de investigación. Gracias a los doctores, Omar, Jaciel, Laila, Cynthia, Carolina A., Fernanda, Elizabeth, Daira, Carolina B., Fátima y, por último, pero no menos importante, a mi hermano Jonathan.

No me alcanzaría todo el agradecimiento del mundo para darle las gracias a mi familia por todo su apoyo, consejos, confianza y motivación para mejorar día a día. Gracias, mamá, papá, Jona, Mariel, Omar, Karla, Stephania, Edder, Karen y Liz por ser las personas que son y ayudarme a ser la persona que soy. Asimismo, gracias a todo el amor perruno de toyo, kenda y luna que siempre se atraviesan corriendo por la mente.

Tabla de contenido

	Página
Resumen en español.....	ii
Resumen en inglés.....	iii
Dedicatoria	iv
Agradecimientos.....	v
Lista de figuras.....	x
Lista de tablas	xii
Capítulo 1. Introducción	1
1.1 Antecedentes	4
1.1.1 El estetoscopio electrónico	4
1.1.2 El servicio de tele-auscultación en despliegues de telemedicina	5
1.1.3 Consideraciones de la arquitectura del Internet para un servicio de tele-auscultación	6
1.2 Planteamiento del problema	6
1.3 Hipótesis.....	7
1.4 Objetivos	8
1.4.1 Objetivo general.....	8
1.4.2 Objetivos específicos.....	8
1.5 Metodología	9
1.6 Organización de la tesis.....	10
Capítulo 2. El estetoscopio electrónico y la tele-auscultación.....	12
2.1 El estetoscopio electrónico y sus características	12
2.2 Rendimiento acústico en los estetoscopios	14
2.3 Ventajas del estetoscopio electrónico	15
2.4 Tele-auscultación	16
2.4.1 Integración de la tele-auscultación en plataformas de telemedicina.....	17

2.4.2	Plataformas especializadas para tele-auscultación.....	17
2.4.3	Tele-auscultación en plataformas de videoconferencia convencionales	18
2.4.4	Trabajos previos	20
2.5	Discusión del servicio de tele-auscultación.....	22
Capítulo 3.	Conectividad en la red: La importancia de la comunicación en la era digital.....	24
3.1	Arquitectura de red	24
3.1.1	Arquitectura par-a-par (P2P).....	24
3.1.2	Arquitectura cliente-servidor	25
3.1.3	Comparativa entre la arquitectura cliente-servidor y P2P.....	26
3.2	Impacto del NAT en la conectividad P2P y desafíos de la red.....	28
3.3	Métodos de conectividad para atravesar el NAT	32
3.4	Ejemplo de una arquitectura para comunicaciones híbridas – WebRTC.....	37
3.5	Componentes de seguridad en la red	41
3.6	Conclusiones del capítulo.....	43
Capítulo 4.	Arquitectura de comunicaciones propuesta	44
4.1	Módulo cliente clínica	46
4.1.1	Submódulo de captura de audio cardiopulmonar, codificación y empaquetado.....	47
4.1.1.1	Estetoscopio	47
4.1.1.2	Tarjeta de audio	49
4.1.1.3	Encardio – codificación de audio.....	50
4.1.1.4	Encardio - empaquetado de audio.....	50
4.1.1.5	Encardio - servidor UDP/IP	51
4.1.2	Submódulo de aplicación de escritorio	52
4.2	Módulo de red.....	54
4.2.1	Diseño de la arquitectura de señalización y comunicaciones.....	54
4.2.1.1	Bloque de registro	56
4.2.1.2	Bloque de inicio de llamada	57

4.2.1.3	Bloque de fin de llamada.....	59
4.2.1.4	Bloque de mantenimiento de rutas	60
4.3	Módulo de cliente especialista.....	60
4.3.1	Submódulo de aplicación de escritorio	61
4.3.2	Submódulo de desempaquetamiento, decodificación, y reproducción del audio	62
4.3.2.1	Decardio – Cliente UDP/IP.....	62
4.3.2.2	Decardio – Desempaquetamiento	63
4.3.2.3	Decardio – Decodificación y búfer circular	63
4.3.2.4	Salida de audio	63
4.4	Módulo de interfaz gráfica	64
Capítulo 5.	Implementación de la cama de pruebas y resultados experimentales	65
5.1	Descripción de la cama de pruebas.....	65
5.2	Pruebas de establecimiento de sesión y transmisión de datos	67
5.2.1	Resultados de pruebas para el escenario 1.....	68
5.2.2	Resultados de pruebas para el escenario 2.....	69
5.2.3	Resultados de pruebas para el escenario 3.....	70
5.3	Análisis del consumo de datos y la tasa de entrega de paquetes (PDR) en los escenarios propuestos.....	71
5.3.1	Consumo de datos.....	72
5.3.2	Tasa de entrega de paquetes (PDR)	73
5.4	Evaluación del sistema propuesto por especialistas de la salud.....	73
Capítulo 6.	Discusión de resultados.....	77
6.1	En relación con el establecimiento de la sesión.....	77
6.2	En relación con el consumo de datos y la tasa de entrega de paquetes	78
6.3	En relación con la evaluación realizada por el personal de la salud	78
Capítulo 7.	Conclusiones, aportaciones y trabajo futuro	80
7.1	Conclusiones.....	80

7.2	Aportaciones	81
7.3	Trabajo futuro	82
	Literatura citada	83
	Anexos	90

Lista de figuras

Figura	Página
Figura 1: Distribución de médicos especialistas en México.	1
Figura 2: Niveles de rendimiento acústico de diferentes estetoscopios comerciales, obtenido de: (3M Littmann®, 2021).	15
Figura 3: Esquema del sistema de transmisión de audio cardiopulmonar desarrollado en (Olachea, 2019).	21
Figura 4: Funcionamiento básico del NAT.	29
Figura 5: Funcionamiento del mecanismo de traducción de puertos y direcciones de red (NAPT).	30
Figura 6: Descripción gráfica de los principales tipos de NAT: A) NAT de cono completo; B) NAT de cono restringido; C) NAT de cono con puertos restringidos; D) NAT simétrico.	32
Figura 7: Diagramas que muestran la secuencia de intercambio de paquetes utilizada en el mecanismo UDP Hole Punching.	34
Figura 8: Diagrama del mecanismo de conexión reversa.	35
Figura 9: Diagrama del mecanismo de <i>Relaying</i>	36
Figura 10: Modelo de comunicación empleando WebRTC.	38
Figura 11: Pila de protocolos utilizada en WebRTC.	40
Figura 12: Diagrama de secuencia general del protocolo DTLS version 1.3.	42
Figura 13: Diagrama que muestra el proceso de encriptación y desencriptación con AES-CBC.	43
Figura 14: Integración de la arquitectura de envío de audio cardiopulmonar propuesto como módulo de un sistema de telemedicina.	44
Figura 15: Diagrama de la arquitectura propuesta para la transmisión de audio cardiopulmonar a través de Internet.	45
Figura 16: Diagrama de bloques del submódulo de captura de audio cardiopulmonar, codificación y empaquetado.	47
Figura 17: Formato de la carga útil para la transmisión simultánea de varios gránulos de audio codificado sobre UDP/IP.	51
Figura 18: Diagrama a bloques del esquema desarrollado en (Olachea, 2019) para transmitir audio cardiopulmonar.	52
Figura 19: Diagrama a bloques del proceso de empaquetado de paquete UDP/IP por el módulo cliente clínica con destino al módulo cliente especialista.	53

Figura 20: Formato de paquete UDP/IP por el módulo cliente clínica con destino hacia el módulo cliente especialista.....	53
Figura 21: Arquitectura de comunicaciones propuesta en este trabajo de investigación.....	55
Figura 22: Diagrama de secuencia del protocolo de señalización inicial propuesto.	57
Figura 23: Diagrama de secuencia de la solicitud de llamada y establecimiento de comunicación en modo P2P.	58
Figura 24: Diagrama de secuencia para establecer la comunicación con el mecanismo de <i>Relaying</i> bajo el paradigma de comunicación cliente-servidor.	59
Figura 25: Diagrama de secuencia de la solicitud de finalización de llamada.	59
Figura 26: Diagrama de secuencia del bloque de mantenimiento de rutas, el cual emplea el mecanismo Ping-Pong.	60
Figura 27: Diagrama a bloques del proceso del submódulo del cliente especialista para su reenvío hacia la aplicación decardio.....	61
Figura 28: Diagrama de bloques del sistema decardio.	62
Figura 29: Diseño de la interfaz gráfica. A) Cliente clínica; B) Cliente especialista.....	64
Figura 30: Escenarios para evaluar el sistema propuesto. A) Escenario 1. B) Escenario 2. C) Escenario 3.	66
Figura 31: Mecanismo de ping-pong entre cliente-servidor de señalización.	67
Figura 33: Registro de mensajes del cliente clínica. (Escenario 1).....	68
Figura 34: Registro de mensajes del cliente especialista. (Escenario 1)	68
Figura 35: Registro de mensajes del cliente clínica. (Escenario 2).....	69
Figura 36: Registro de mensajes del cliente especialista. (Escenario 2)	69
Figura 37: Registro de mensajes del cliente clínica. (Escenario 3).....	70
Figura 38: Registro de mensajes del cliente especialista. (Escenario 3)	70
Figura 39: Captura de pantalla de una llamada entre los clientes.....	71
Figura 40: Herramientas utilizadas para la evaluación de los sistemas convencionales.	74
Figura 41: Características de las plataformas evaluadas.	75
Figura 42: Calidad diagnóstica por plataforma.	76
Figura 43: Preferencias de plataformas para el envío de audio cardiopulmonar.....	76

Lista de tablas

Tabla	Página
Tabla 1: Plataformas de telesalud existentes en el mercado y sus características.....	17
Tabla 2: Características de aplicaciones 3M™ Littmann® y EKO.	18
Tabla 3: Características de plataformas de videoconferencia.	19
Tabla 4: Configuraciones del programa con los códecs de audio.	22
Tabla 5: Comparación entre las diferentes arquitecturas de red.	27
Tabla 6: Rango de direccionamientos establecidos por la IANA.....	28
Tabla 7: Funcionamiento de las técnicas transversales del NAT en relación con el tipo de NAT presente en cada extremo.....	37
Tabla 8: Características de las tarjetas de audio “Sound Blaster Play! 3” y “Sound BlasterX G1”	49
Tabla 9: Consumo de datos obtenido por cada escenario propuesto.	72
Tabla 10: Relación de entrega de paquetes.....	73

Capítulo 1. Introducción

Los sistemas de telesalud o telemedicina surgen debido a la necesidad de médicos especialistas que existe a nivel mundial, un factor que afecta particularmente a México. En el 2023, la Organización para la Cooperación y el Desarrollo Económicos (OECD, por sus siglas en inglés) menciona en sus datos que México cuenta con 2.5 médicos por cada mil habitantes. Esto representa un reto significativo para la prestación adecuada de servicios de salud a toda la población, debido a que la OECD tiene su estándar de contar con un promedio de 3.7 médicos por cada mil habitantes. Adicionalmente, en México no solo la cantidad de médicos especialistas representa un problema relevante, ya que existe una distribución geográfica desigual en el país donde hay menos profesionales en áreas rurales y remotas (González et al., 2015; Heinze-Martin et al., 2018).

Lo expuesto anteriormente se ilustra en la Figura 1, donde se observa que la mayor concentración de médicos especialistas se encuentra en estados como Ciudad de México, Nuevo León, Jalisco y Baja California, que cuentan con más de 100 médicos especialistas por cada 100,000 habitantes. En contraste, estados como Chiapas, Guerrero, Oaxaca y Veracruz tienen menos de 80 especialistas por cada 100,000 habitantes.

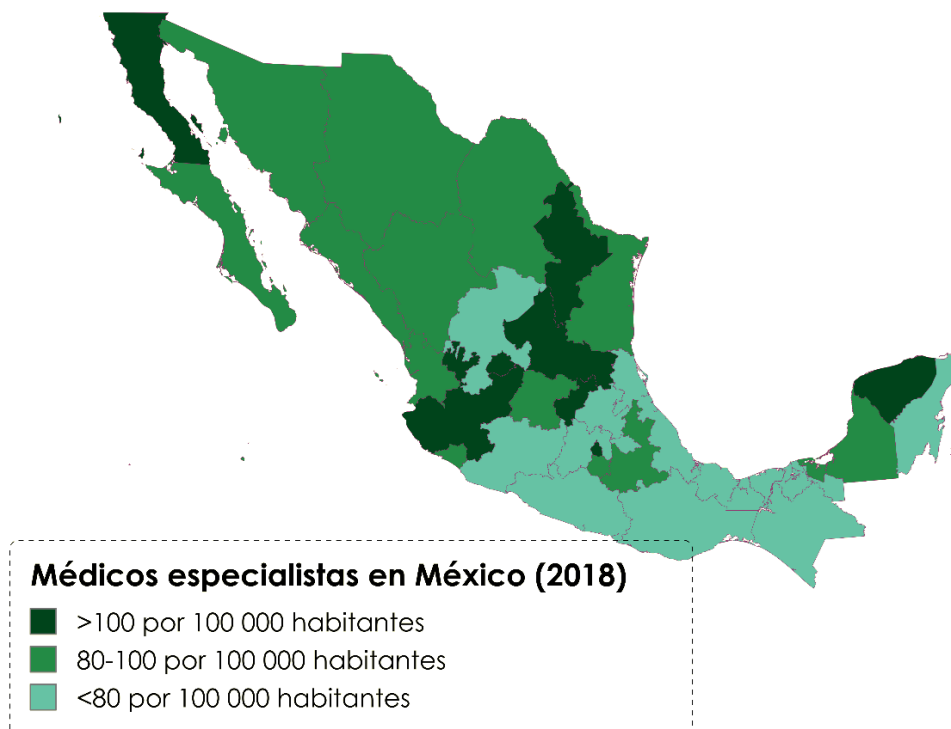


Figura 1: Distribución de médicos especialistas en México.

Una solución a esta limitación es el uso de plataformas de telesalud, las cuales permiten a los pacientes acceder a servicios médicos profesionales a distancia, independientemente de su ubicación. Tales plataformas ofrecen servicios de consulta virtual, monitoreo remoto, y un mejor seguimiento de pacientes crónicos o de poblaciones vulnerables (Agnisarman et al., 2017; Casillas et al., 2015; Khemapech et al., 2019; Magana-Rodriguez et al., 2018; Ward et al., 2015).

Se puede definir a la telesalud como el uso de tecnologías de información y comunicación (TICs) para proporcionar servicios de salud a distancia (Salud, 2024; *Telehealth*, 2020). La telesalud puede incluir servicios como consultas médicas virtuales, monitoreo remoto de pacientes, y la educación a distancia de profesionales de la salud. Asimismo, también apoyan a los familiares o cuidadores de personas que se encuentran en estado de vulnerabilidad y tienen dificultades para acceder a servicios de salud de forma presencial.

Los servicios de salud electrónicos (e-salud) tienen el potencial de mejorar significativamente el diagnóstico de los pacientes y la accesibilidad y calidad del tratamiento (Miah et al., 2017). Estos sistemas digitales permiten también la prestación de una gama más amplia de servicios de salud para mejorar el bienestar y la calidad de vida de las poblaciones desfavorecidas o desatendidas.

Existen diversos artículos que exponen el empleo de plataformas de telesalud o telemedicina (Uscher-Pines y Mehrotra, 2014) que señalan que las plataformas de telesalud contribuyen a disminuir el uso de equipos de protección personal y mejorar el acceso a la atención médica. Asimismo, Pettingil y Tewes (2020) destacan que la telesalud permite reducir costos, mejorar la eficiencia y aumentar la comodidad para los pacientes al evitar traslados y tiempos de espera. Debido a las mejoras potenciales en el área de la salud, las empresas privadas han tenido interés en este ámbito en los últimos años.

Google y Amazon han identificado en la telemedicina un sector emergente con alto potencial para transformar las industrias del futuro (Pettingil y Tewes, 2020). Empresas como Amwell están mejorando la experiencia del paciente al reducir los tiempos de espera de 121 minutos en modelos tradicionales a solo 2-3 minutos. Paralelamente, más de 275 empresas en Estados Unidos han invertido mucho en este campo para mejorar el sistema de atención médica y mejorar su accesibilidad y eficiencia (Amwell, 2023).

En el proceso de atención médica, el estetoscopio es considerado una herramienta esencial para los profesionales de la salud y se utiliza en consultorios médicos y hospitales de todo el mundo (Bedingham et al., 2016; Swarup y Makaryus, 2018). Se trata de un instrumento médico que se utiliza para escuchar

los sonidos del corazón, los pulmones y otros órganos del cuerpo. En los años recientes, como resultado de la amplia utilización de las TIC's dentro del ámbito médico, se ha observado un notable aumento en la relevancia de la transmisión en tiempo real de las señales de audio cardiopulmonar capturadas por estetoscopios digitales. Esto se debe a que la auscultación digital se muestra como una herramienta prometedora para elevar la calidad de la atención médica en áreas rurales (Pereira et al., 2016). Por esta razón, uno de los servicios de telesalud que es esencial incluir en un despliegue de teleconsulta es la auscultación remota, conocida como tele-auscultación.

Un sistema de tele-auscultación digital debe permitir a los profesionales de la salud escuchar los sonidos cardíacos y pulmonares de los pacientes a distancia con audio de calidad diagnóstica. Si se llega a implementar un servicio de tele-auscultación facilitará a los médicos la capacidad de tomar decisiones más ágiles y precisas al momento de efectuar diagnósticos. Además de poder realizar un tratamiento oportuno de patologías. Si bien puede resultar tentador conectar un estetoscopio electrónico (como el estetoscopio digital ONE de la empresa de Thinklabs, (Thinklabs Medical LLC, 2014)) a un sistema de teleconferencia para habilitar un servicio de tele-auscultación, este tipo de plataformas utiliza codificadores-decodificadores (códecs) que no fueron diseñados para la transmisión de audio cardiopulmonar. En particular, la gran mayoría de las plataformas de teleconferencia convencional utilizan códecs específicos para voz. Sin embargo, una señal de audio-cardiopulmonar es muy diferente a una señal de voz. Esto crea la necesidad de implementar códigos específicos para este tipo de señales, ya que las señales de audio cardiopulmonar pueden llegar a distorsionarse en las plataformas de teleconferencia convencionales y de esta forma puede dar pie a un diagnóstico erróneo por parte del especialista a distancia (Olachea, 2019; Ramírez, 2017).

Además de los códecs, es necesario tener en cuenta el ancho de banda que se requiere para la tele-auscultación. Soluciones comerciales como la empresa 3M™ Littmann® que requieren de anchos de banda grandes para transmitir audio cardiopulmonar debido a que no usan esquemas de compresión. Esto implica que para una teleconsulta con servicio de tele-auscultación se requerirá un ancho de banda significativo, el cual es importante reducir especialmente en despliegues en zonas rurales con conectividad limitada de datos.

Por lo tanto, existe la necesidad de desarrollar esquemas que permitan la transmisión de señales de audio cardiopulmonar con calidad diagnóstica, utilizando un menor ancho de banda para un despliegue en zonas rurales o zonas urbano-marginadas.

1.1 Antecedentes

1.1.1 El estetoscopio electrónico

El estetoscopio es un instrumento médico que permite auscultar los sonidos internos del cuerpo humano, especialmente los del corazón y los pulmones. Su origen se remonta al siglo XIX, cuando el médico francés René Laënnec ideó una forma de escuchar mejor los latidos de sus pacientes sin tener que apoyar directamente su oído en el pecho. Laënnec se inspiró en un juego de niños que consistía en transmitir sonidos a través de una tabla y creó el primer estetoscopio con un tubo de papel enrollado. Este invento revolucionó el diagnóstico médico y abrió las puertas a nuevos avances en la fisiología y la cardiología (Wilson, 1928).

Desde entonces, el estetoscopio ha evolucionado en diferentes tipos y modelos, adaptados a las necesidades y preferencias de los profesionales de la salud. No obstante, el estetoscopio tradicional presenta algunas limitaciones, como la habilidad y la experiencia que posea el personal médico, la interferencia del ruido ambiental y la dificultad para almacenar y compartir los sonidos auscultados.

Por ello, en los últimos años se han desarrollado estetoscopios digitales que incorporan tecnologías de procesamiento de señales, transmisión inalámbrica y análisis inteligente. Estos dispositivos permiten mejorar la calidad y la precisión de la auscultación, así como facilitar su integración con los sistemas de información clínica. Por ejemplo, de acuerdo con los datos de prueba recopilados por Cruz et al. (2019), se observó que el estetoscopio electrónico de cancelación de ruido logró mejorar la relación señal-ruido (SNR) de los sonidos del corazón y los pulmones en diversos entornos.

Algunos ejemplos de estetoscopios digitales son el Littmann 3200, el Eko Core, el Thinklabs One y el eKuore Pro (Seah et al., 2023). Estos estetoscopios mejoran la experiencia de auscultación presencial. Sin embargo, al integrarlos en un sistema de teleconsulta requieren un ancho de banda superior a 128 kbps (p. ej. el sistema ofrecido de 3M™ Littmann®) o no ofrecen un servicio de transmisión audio cardiopulmonar en tiempo real (p. ej. la aplicación ofrecida por Thinklabs, “Wave”)., sino que ofrecen un servicio en el cual los datos obtenidos se almacenan en la nube. No obstante, el servicio de teleconsulta en tiempo real es necesario, ya que de esta manera posibilita la detección temprana de enfermedades o el cuidado del paciente con una consulta de primer nivel acompañada con un especialista de la salud.

1.1.2 El servicio de tele-auscultación en despliegues de telemedicina

Plataformas como Teladoc Health, Amwell, Doxy, VSee, Zoom for healthcare, Amd Global Telemedicine, entre otras, han adquirido gran relevancia e impulso en los últimos años, mostrando ser una alternativa viable y efectiva para la atención médica a distancia, sobre todo en países que cuentan con una infraestructura con el suficiente ancho de banda (Khemapech et al., 2019). Estas plataformas presentan diversas funcionalidades como la programación de citas, herramientas de videoconferencia seguras, formularios médicos digitales, generación de recetas, entre otros.

En lo que respecta al uso de estetoscopios digitales en estos sistemas, solamente VSee y Amd Global Telemedicine soportan su uso. Sin embargo, estos sistemas transmiten la señal de audio cardiopulmonar sin comprimir, lo que incrementa el ancho de banda requerido para establecer una teleconsulta. Esto podría aumentar el costo requerido para desplegar el servicio en zonas rurales y urbano-marginales, e incluso podría hacer que su despliegue no sea viable.

Como se mencionó anteriormente, aunque puede ser tentador habilitar el servicio de tele-auscultación conectando un estetoscopio electrónico a la sesión de videoconferencia de las otras plataformas de telemedicina, esto no se recomienda debido a que los códecs utilizados pueden deformar la señal y dar pie a diagnósticos erróneos (Castorena, 2012; Olachea, 2019; Ramírez, 2017). El principal problema radica en que los códecs están diseñados para la compresión del audio generado por el tracto vocal humano, pero no para la transmisión de sonidos fisiológicos como los ruidos cardíacos y pulmonares, que requieren una mayor fidelidad de audio. Por otro lado, transmitir el audio cardiopulmonar sin comprimir requeriría elevar el ancho de banda necesario para establecer una teleconsulta al menos a 128 kbps (3M™ Littmann®, 2010; Olachea, 2019), lo cual puede ser no viable en entornos con conexiones de Internet de baja calidad o inestables, como es el caso de zonas rurales o remotas.

Por lo tanto, es necesario desarrollar una plataforma que considere todas estas características para el correcto despliegue de un servicio de tele-auscultación, tal que ancho de banda requerido para habilitar el servicio sea menor a 128 kbps y se cumpla con los criterios de seguridad de la norma de Privacidad de la Ley de Responsabilidad y Portabilidad del Seguro (HIPAA, por sus siglas en inglés).

1.1.3 Consideraciones de la arquitectura del Internet para un servicio de tele- auscultación

Actualmente, debido al agotamiento de las direcciones de red del protocolo de Internet versión 4.0, se hace uso de la técnica de “traducción de direcciones de red” (NAT, por sus siglas del inglés) para el aprovisionamiento de conectividad. Esto implica que para establecer un flujo de datos entre una fuente (p. ej. un estetoscopio electrónico) y un resumidero de información (p. ej. un consultorio remoto) de manera directa, es necesario implementar técnicas que permitan “traspasar” el NAT (Ford et al., 2008; Poulin et al., 2012; Wagner, 2015). La implementación de estas técnicas es necesaria, debido a que las computadoras que se encuentran detrás de un NAT no pueden establecer una comunicación directa entre sí.

Las técnicas del traspaso de NAT son utilizadas en arquitecturas de Par-a-Par (P2P, de sus siglas en inglés), como es el caso del estándar de videoconferencia WebRTC (Loreto y Romano, 2014; Uberti, 2012). Sin embargo, con el mejor de nuestro conocimiento, no existe un estándar equivalente a WebRTC para la transmisión de audio cardiopulmonar. Por ello, el desarrollo de un estándar equivalente a WebRTC requiere desde la selección de códecs, hasta el diseño de arquitectura que permitan atravesar el NAT, esto sin dejar de lado los esquemas de protección de información requeridos por la HIPAA.

1.2 Planteamiento del problema

Dentro del Grupo de Investigación ARTS se han desarrollado temas de investigación, teniendo como punto central la e-Salud (p. ej. servicios de teleconsulta, tele-monitoreo, tele-fisioterapia, etc.) y sus aplicaciones en zonas urbanas-marginales o zonas rurales. En estas localidades normalmente no se cuenta con un gran ancho de banda disponible para la transmisión de datos, por ello, las investigaciones se han centrado en el desarrollo de esquemas que permitan el despliegue de servicios de e-Salud con un bajo consumo de ancho de banda (Casillas et al., 2015; González et al., 2015; Magana-Rodriguez et al., 2018; Magaña-Rodriguez et al., 2015).

Como se hizo mención anteriormente, la cantidad de médicos especialistas en México es preocupante, puesto que hay muchas zonas que no tienen acceso a una atención médica básica y los pacientes tienen que llevar a cabo largos viajes para poder recibir atención médica. Por esto, en los últimos años los

sistemas de teleconsulta han ganado mucha relevancia debido a la ayuda que pueden proporcionar a las zonas rurales (Casillas et al., 2015; González et al., 2015; Magana-Rodriguez et al., 2018).

Dentro de los sistemas de teleconsulta, la transmisión de señales de audio cardiopulmonar es deseable para una auscultación adecuada de los pacientes. No obstante, su implementación todavía no es una característica esencial en los sistemas de salud e-Salud. Para esto, los sistemas que permiten desplegar el servicio de tele-auscultación consideran que el audio cardiopulmonar se transmitirá sin comprimir. Sin embargo, en numerosos casos, la infraestructura de telecomunicaciones en las zonas rurales y urbano-marginales no es suficiente para garantizar la transmisión de señales de audio cardiopulmonar sin comprimir de calidad diagnóstica.

El conectar un estetoscopio digital a un sistema de videoconferencia no es factible, debido a que en trabajos como el de Olachea (2019), se demostró que los códecs utilizados por dichos sistemas pueden introducir espurios en la señal de audio cardiopulmonar, lo cual podría llevar a un mal diagnóstico. Por lo tanto, es necesario desarrollar esquemas que permitan transmitir audio cardiopulmonar comprimido de calidad diagnóstica y que sean factibles de desplegar en lugares con anchos de banda limitados.

En la actualidad no hay un estándar que defina la integración de auscultación digital en sistemas de teleconsulta y, pueden surgir diferentes obstáculos técnicos que deben ser superados dependiendo de las condiciones de despliegue. En este contexto, la prioridad es garantizar que no exista una alteración de la señal transmitida, lo que podría provocar un mal diagnóstico por parte del personal médico. Debido a esto, existe la creciente necesidad de desarrollar mecanismos que permitan la transmisión confiable y segura de este tipo de señales.

En respuesta a estos desafíos, en este proyecto de investigación se propone la creación de una arquitectura que habilite la transmisión segura en tiempo cuasi-real de señales de audio cardiopulmonar, haciendo uso de técnicas de compresión y encriptación, abordando así las complejidades inherentes que trae consigo este tipo de transmisiones.

1.3 Hipótesis

Es factible diseñar y validar un esquema que permita la transmisión de audio cardiopulmonar comprimido y cifrado de calidad diagnóstica en tiempo cuasi-real utilizando una arquitectura híbrida P2P y cliente-

servidor que resuelva el problema del NAT. Esto permitirá reducir el ancho de banda requerido para la transmisión de la señal sin comprometer la seguridad del flujo de información, lo cual contribuirá a habilitar servicios de teleconsulta en localidades con conectividad de datos limitada.

1.4 Objetivos

1.4.1 Objetivo general

El objetivo de este proyecto es el diseño y validación de un esquema para la transmisión de señales de audio cardiopulmonar comprimido en tiempo cuasi-real, empleando una arquitectura híbrida P2P y cliente-servidor que considere un esquema de cifrado básico para el flujo de información.

1.4.2 Objetivos específicos

- Estudiar el estado del arte relacionado con: la transmisión de audio cardiopulmonar; esquemas empleados para la transmisión de audio a través de la Internet; arquitecturas P2P y cliente-servidor; métodos de cifrado para las señales transmitidas con formato de “streaming”.
- Diseñar un sistema de transmisión de audio cardiopulmonar para aplicaciones de teleconsulta con base en una arquitectura híbrida P2P y cliente-servidor.
- Incorporar en la arquitectura híbrida (P2P y cliente-servidor) los diferentes esquemas de compresión de audio cardiopulmonar identificados en trabajos previos dentro del grupo de investigación.
- Analizar el desempeño de la arquitectura con los diferentes esquemas de compresión, considerando: la calidad de la señal de audio, la pérdida de paquetes, el retardo y la latencia que pueda presentar la señal.
- Seleccionar el esquema de compresión y características de la arquitectura que proporcionen el mejor desempeño.

- Proponer un esquema básico de cifrado para la transmisión del flujo de datos de audio cardiopulmonar que funcione en conjunto con la arquitectura y los esquemas de compresión propuestos.
- Evaluar el sistema final propuesto para su desempeño en términos de calidad de audio por medio de los profesionales de la salud.

1.5 Metodología

1. **Revisión del estado del arte.** Efectuar una búsqueda exhaustiva de literatura científica, artículos, libros y fuentes confiables relacionadas con la transmisión de audio cardiopulmonar, esquemas de transmisión de audio; arquitecturas P2P y cliente-servidor; métodos de cifrado para transmisiones de "*streaming*".
2. **Diseño del sistema de transmisión.** Realizar el diseño de la arquitectura híbrida (P2P y cliente-servidor) que permitirá la transmisión eficiente de las señales de audio.
3. **Implementación de los esquemas de compresión.** Realizar la identificación y adaptación de los diferentes esquemas de compresión de audio cardiopulmonar encontrados en trabajos previos.
4. **Análisis del desempeño.** Establecer un conjunto de métricas para evaluar el desempeño de la arquitectura, incluyendo la calidad de la señal de audio, la entrega de paquetes, así como su consumo de datos; Implementar de transmisión utilizando el esquema de compresión de MP3 y registrar los resultados obtenidos en términos de las métricas definidas; Analizar y comparar los resultados de las pruebas de forma crítica.
5. **Propuesta de esquema de cifrado.** Investigar y analizar métodos de cifrado adecuados para la transmisión segura de datos de audio cardiopulmonar en formato de "*streaming*".; Diseñar y proponer un esquema básico de cifrado que pueda ser implementado en conjunto con la arquitectura y los esquemas de compresión seleccionados.
6. **Evaluación y validación.** Llevar a cabo pruebas del sistema final propuesto, considerando la calidad de audio y la seguridad de la transmisión.; Invitar a profesionales de la salud a participar

en la evaluación del sistema y recopilar retroalimentación sobre la calidad del audio y la usabilidad en aplicaciones de teleconsulta.; Analizar y discutir los resultados obtenidos durante las pruebas y la evaluación del sistema.

1.6 Organización de la tesis

El presente trabajo de tesis se compone de 7 capítulos, los cuales se describen brevemente a continuación.

En el capítulo 2 se analiza el estado del arte en cuanto al estetoscopio electrónico, donde se examinan las partes fundamentales de este, su rendimiento acústico y las ventajas que representa su uso en el campo de la medicina. Asimismo, se investiga el estado del arte de la tele-auscultación y su integración en plataformas comerciales de telemedicina y sistemas especializados para tele-auscultación. Por otro lado, se examinan las plataformas convencionales de videoconferencia y su factibilidad de uso para la tele-auscultación. Adicionalmente, se realiza una revisión de los trabajos previos relacionados a este trabajo de investigación.

En el Capítulo 3 se realiza una revisión de los aspectos más relevantes relacionados con las arquitecturas de red y los esquemas de transmisión de señales a través de Internet, que son importantes para el desarrollo de esta tesis. Primero, se compara entre distintos modelos de red. Luego, se abordan temas clave como los mecanismos de traducción de direcciones de red y la seguridad en las redes de comunicación. Por último, se incluye un breve análisis de la tecnología WebRTC, que se toma como modelo de referencia en este trabajo de investigación.

En el Capítulo 4 se describe la arquitectura de comunicaciones diseñada para este trabajo de investigación. Se describen los distintos bloques que conforman la arquitectura y se detalla el funcionamiento de cada uno de ellos. Asimismo, se presentan las razones que llevaron a la selección de los diversos elementos que la integran.

En el Capítulo 5 se presenta la descripción de la cama de pruebas experimentales utilizada para evaluar el desempeño de la arquitectura de comunicaciones diseñada en un entorno real, así como la exposición de los resultados obtenidos. Para esto se analizan los aspectos relacionados con el consumo de datos y la entrega de paquetes en los diferentes escenarios propuestos. Adicionalmente, se presentan los resultados obtenidos a través de una evaluación con ayuda de personal de la salud.

En el Capítulo 6 se lleva a cabo el análisis y la discusión de los resultados obtenidos en los experimentos. Se interpretan los valores y comportamientos observados, comparándolos con los resultados reportados en investigaciones previas.

El Capítulo 7 expone las conclusiones del trabajo de investigación, destacando los principales aportes realizados. Además, se incluyen algunas recomendaciones y sugerencias para el desarrollo de trabajos futuros.

Capítulo 2. El estetoscopio electrónico y la tele-auscultación

En este capítulo se realizará una revisión del estado del arte del estetoscopio electrónico y su integración en sistemas de telesalud.

2.1 El estetoscopio electrónico y sus características

Desde su invención en 1816 por el médico René Laennec, el estetoscopio se convirtió en una herramienta emblemática en la medicina, al permitir a los médicos escuchar y distinguir los sonidos internos del cuerpo humano con mayor claridad (Leng et al., 2015). En sus inicios se trataba de un dispositivo mecánico, pero con el avance de la tecnología evolucionó hasta la aparición del estetoscopio electrónico. Este modelo convierte las ondas sonoras captadas en señales eléctricas, que luego son procesadas, amplificadas y transmitidas a los oídos del clínico o almacenadas para su análisis posterior, mejorando así la interpretación de los datos.

Un estetoscopio electrónico tiene cuatro componentes fundamentales: transductor/sensor, circuito de pre-amplificación, sistema de filtrado y sistema de salida. Cada uno de los componentes cumple una función específica para captar, procesar y transmitir los sonidos corporales de manera precisa. A continuación, se verá en detalle cada uno de estos componentes.

Transductor/sensor. El transductor es un dispositivo de vital importancia en los estetoscopios electrónicos, ya que permite la conversión de las vibraciones acústicas (sonido) generadas por el cuerpo humano en señales eléctricas (voltaje) (Denault et al., 2016). Existen algunas variantes de este tipo de dispositivo con diferentes características:

- **Sensores de micrófono.** El enfoque más simple consiste en colocar un micrófono detrás del diafragma del estetoscopio. Este diseño presenta dos diafragmas en serie (el del estetoscopio y el del micrófono) con una cámara de aire intermedia, lo que puede generar captación excesiva de ruido ambiental y transferencia energética ineficiente (Leng et al., 2015).
- **Sensores piezoeléctricos.** Funcionan mediante la deformación de un material cristalino que produce energía eléctrica en respuesta a presión mecánica. El diafragma del estetoscopio está mecánicamente acoplado al cristal piezoeléctrico, y su movimiento genera una señal eléctrica.

Este enfoque es utilizado en los estetoscopios Welch-Allyn Meditron y 3M Littmann (Swarup y Makaryus, 2018). Sin embargo, el mecanismo de conversión puede introducir distorsión en comparación con la detección pura del movimiento del diafragma.

- **Sensores capacitivos MEMS.** Basados en tecnología de Sistemas Micro-Electro-Mecánicos (MEMS, por sus siglas en inglés), detectan la presión acústica mediante cambios en su valor nominal de capacitancia. El diafragma central actúa como masa de prueba suspendida que puede moverse libremente y está eléctricamente aislada de una estructura fija, creando una capacitancia nominal entre ambas (Leng et al., 2015). Las ventajas incluyen tamaño reducido, producción masiva y mejor estabilidad térmica. El Thinklabs One Digital utiliza un diafragma electromagnético con superficie interna conductora que forma un transductor capacitivo (Swarup y Makaryus, 2018).
- **Sensores electromagnéticos.** Emplean el principio de inducción electromagnética para convertir las vibraciones en señales eléctricas, utilizando un diafragma con materiales conductores que se mueve dentro de un campo magnético.

Circuito de pre-amplificación. La señal eléctrica generada por el transductor suele ser débil y requiere una amplificación para mejorar la calidad del audio obtenido antes de cualquier procesamiento adicional. El circuito de pre-amplificación se encarga de realizar este proceso y además contribuye a la reducción de la interferencia, la prevención del fenómeno de *aliasing* y la optimización de la relación señal-ruido (SNR, por sus siglas en inglés).

Sistema de filtrado. El sistema de filtrado constituye un componente fundamental en el diseño del estetoscopio electrónico, ya que permite seleccionar diferentes bandas de frecuencia para optimizar la auscultación de distintos órganos, esencialmente el corazón y los pulmones. A continuación, se ejemplifican algunos de los modos de trabajo más empleados.

- **Modo campana.** Típicamente 20-200 Hz, ideal para sonidos cardíacos de baja frecuencia.
- **Modo diafragma.** Aproximadamente 100-500 Hz, óptimo para sonidos respiratorios y componentes de alta frecuencia cardíacos.
- **Modo extendido.** 20-1000 Hz, para una evaluación más completa del espectro acústico (3M™ Littmann®, 2010).

Sistema de salida. El sistema de salida constituye la interfaz crítica entre el dispositivo y el usuario, determinando la forma en que los datos acústicos procesados son entregados para su interpretación clínica. Los estetoscopios electrónicos modernos incorporan diversos componentes de salida, incluyendo auriculares binaurales optimizados para frecuencias cardiopulmonares, altavoces integrados para demostraciones educativas, y conector estándar que permite la integración con dispositivos externos. La conectividad inalámbrica mediante Bluetooth o wifi permite la transmisión de los datos auscultados a sistemas de documentación clínica, plataformas educativas o servicios de telemedicina, expandiendo significativamente la utilidad del dispositivo más allá del examen directo. Según Swarup y Makaryus (2018), estos sistemas de salida diversificados transforman el estetoscopio de una herramienta meramente auditiva a un instrumento diagnóstico ubicuo y fundamental en la práctica médica.

2.2 Rendimiento acústico en los estetoscopios

La importancia del diseño y configuración del estetoscopio, especialmente en los modelos digitales con capacidades de filtrado selectivo, responde directamente a las características acústicas de los sonidos fisiológicos que se necesitan captar. Los sonidos cardiopulmonares se manifiestan en un espectro de frecuencia amplio que abarca desde 10 Hz hasta 2000 Hz (Hassani, 2012; Marcy, 1999; Miah et al., 2017; Webster, 2014; Xu et al., 1989). Estos sonidos incluyen diversos componentes de la actividad cardíaca (ruidos S1, S2, S3 y S4); soplos; fenómenos asociados a la función valvular y flujo sanguíneo; así como ruidos respiratorios normales y sonidos adventicios (crepitantes, sibilancias, roncus) que pueden indicar condiciones patológicas específicas. Esta diversidad de componentes acústicos subraya la importancia del estetoscopio como herramienta diagnóstica esencial en la evaluación clínica integral del paciente.

Como se ha mencionado con anterioridad, existen una gran variedad de estetoscopios que pueden clasificarse en función de los niveles de rendimiento acústico, los cuales dependen de la aplicación específica para la cual será utilizado el estetoscopio. En la Figura 2 se observa una clasificación de estetoscopios conforme a dicho nivel de rendimiento acústico, el cual está dividido en cuatro categorías. El nivel 6 (amarillo) se emplea en mediciones básicas como la tensión arterial y las exploraciones rutinarias. En el nivel 7-8 (morado) entran los estetoscopios pediátricos y neonatales que son diseñados para exploraciones físicas y monitorizaciones continuas. Los estetoscopios en nivel 9-10 (rojo) buscan proporcionar una sensibilidad acústica para los entornos de cuidados intensivos. Finalmente, en el nivel 10+ (azul) entran estetoscopios que ofrecen una amplificación digital, cancelación de ruido y una conexión estable con un dispositivo móvil o una aplicación de computadora de la marca del estetoscopio.



Figura 2: Niveles de rendimiento acústico de diferentes estetoscopios comerciales, obtenido de: (3M Littmann®, 2021).

2.3 Ventajas del estetoscopio electrónico

Los estetoscopios electrónicos ofrecen algunas ventajas respecto a los estetoscopios convencionales, lo que aumenta su utilidad en la práctica clínica. A continuación, se exponen las principales ventajas asociadas a los estetoscopios electrónicos.

1. **Calidad de sonido y amplificación mejoradas.** Los estetoscopios electrónicos transforman las vibraciones acústicas del cuerpo en señales electrónicas, que pueden amplificarse, filtrarse y procesarse. Esta capacidad permite una gama más amplia de características de respuesta de frecuencia, lo que permite a los médicos oír ruidos y soplos cardíacos sutiles que pueden ser difíciles de detectar con los fonendoscopios tradicionales (Webster, 2014). La amplificación de los sonidos puede mejorar significativamente la detección de los ruidos cardíacos de baja frecuencia, que suelen ser difíciles de oír con los dispositivos acústicos estándar (Mangione et al., 2020).
2. **Reducción del ruido.** Los estetoscopios electrónicos modernos están diseñados para ser inmunes al ruido de fondo, lo que resulta especialmente beneficioso en entornos clínicos ruidosos. A menudo incorporan tecnologías de reducción del ruido que mejoran la claridad de los sonidos auscultados, lo que permite llevar a cabo evaluaciones más precisas (Webster, 2014). Esta característica es crucial para los médicos que trabajan en entornos en los que el ruido ambiental puede interferir en la auscultación.
3. **Grabación y reproducción de datos.** El poder grabar los sonidos cardiopulmonares permite a los profesionales sanitarios reproducir los sonidos para su posterior análisis, compartir las

grabaciones con otros médicos o utilizarlas con fines educativos (Webster, 2014).

- 4. Integración con la tecnología.** La integración de tecnología con los estetoscopios como lo son computadoras, laptops y dispositivos móviles facilita el procesamiento sofisticado de señales, la extracción de características e incluso su uso en aplicaciones de telemedicina (Webster, 2014).
- 5. Funciones amigables con el usuario.** La gran mayoría de estetoscopios vienen equipados con funciones simples de usar, como controles de volumen ajustables, pantallas visuales de las formas de onda del sonido y la posibilidad de cambiar entre distintos modos de sonido. Esto hace que sea más intuitivo para los médicos en entrenamiento como para los experimentados.
- 6. Beneficios educativos.** La posibilidad de visualizar y analizar sonidos contribuye significativamente a la educación y formación médica. Los estetoscopios electrónicos pueden integrarse con software educativo diseñado para ayudar a los estudiantes a reconocer diferentes sonidos cardíacos y pulmonares a través de simulaciones interactivas (Pereira et al., 2020).

2.4 Tele-auscultación

La tele-auscultación representa la intersección entre la auscultación tradicional y la telemedicina, definida como la adquisición, transmisión y análisis remoto de sonidos fisiológicos para fines diagnósticos y educativos. Este proceso implica la captura digital de sonidos cardíacos, pulmonares u otros sonidos corporales mediante un estetoscopio electrónico, seguida de su transmisión a distancia para ser evaluados por profesionales sanitarios ubicados remotamente (Lakhe et al., 2016).

Aunque la tele-auscultación es considerada un elemento fundamental en un sistema de telemedicina, no todas las plataformas de atención médica a distancia tienen soporte para este propósito. Esto se debe fundamentalmente a los retos asociados a su implementación en cuanto a la calidad y seguridad del audio transmitido.

Respecto a la seguridad, los sistemas de telemedicina deben cumplir con la Norma de Privacidad de la Ley HIPAA (Office for Civil Rights, 2025). Esta normativa establece cómo y con quién se debe compartir la información de salud protegida y define los derechos de las personas sobre estos datos médicos.

2.4.1 Integración de la tele-auscultación en plataformas de telemedicina

En la actualidad, existen varias plataformas que se dedican a ofrecer servicios de telemedicina de forma profesional, entre las que destacan Teladoc Health, Amwell, Doxy, Vsee, Zoom for healthcare y Amd Global Telemedicine. Estas soluciones han cobrado gran importancia y desarrollo en los últimos años, demostrando ser una opción eficaz y viable para la prestación de servicios médicos a distancia.

En la Tabla 1 se analizan estas plataformas en cuanto a soporte del estetoscopio para el envío de audio cardiopulmonar de calidad diagnóstica, ancho de banda requerido para su funcionamiento y cumplimiento de HIPAA con cifrado. Acorde a la información de la Tabla 1, se puede destacar que todas las plataformas de telemedicina analizadas cumplen con los criterios de HIPAA con cifrado; sin embargo, no todas tienen soporte para estetoscopio. Además, la mayoría requieren un ancho de banda elevado, pues ofrecen toda una gama de servicios de salud y presentan costos asociados a la contratación del servicio.

Tabla 1: Plataformas de telesalud existentes en el mercado y sus características.

Plataforma	Criterio HIPAA con cifrado	Soportan el uso del estetoscopio	Ancho de banda empleado
Teladoc Health (Helicon Health Ltd., s/f)	Si	Si	Desde 300 kbps
Amwell (Amwell, 2020)	Si	No	Recomienda desde 3 Mbps
Doxy.me (Doxy.me, s/f)	Si	No	0.5 Mbps hasta 2.5 Mbps
VSee (VSee, s/f)	Si	Si	0.2 Mbps hasta 2 Mbps
Zoom for Healthcare (Zoom Communications, Inc., s/f)	Si	No	0.6 Mbps hasta 3 Mbps
Agnes Connect (AMD global Telemedicine, 2020)	Si	Si	256 kbps – 2 Mbps
Zocdoc (Zocdoc, 2021)	Si	No	No especificado
Healow (Healow, s/f)	Si	No	No especificado

2.4.2 Plataformas especializadas para tele-auscultación

En la actualidad, existen también soluciones diseñadas para ofrecer únicamente servicios de tele-auscultación. En este caso se pueden mencionar los sistemas desarrollados por las empresas Littmann y EKO. Ambas plataformas ofrecen una solución para la captura de audio cardiopulmonar y su transmisión y almacenamiento en la nube para futuros análisis.

Sin embargo, estos sistemas solo pueden usarse con equipos propios de esas marcas, los cuales son considerados costosos en países de bajos recursos. Además, al igual que las plataformas analizadas en el apartado anterior, requieren de una infraestructura de telecomunicaciones de banda ancha. Por ejemplo, en el caso del sistema desarrollado por Littmann, se recomienda una conexión de 1.5 Mbps en el lado del cliente (3M™ Littmann®, 2025).

En la Tabla 2, se muestran algunas aplicaciones pertenecientes a los estetoscopios de las marcas 3M™ Littmann® y Eko Software. Donde se puede observar que sus soluciones para poder transmitir solo están habilitadas en la región de Estados Unidos, esto sucede por la infraestructura que puede soportar grandes anchos de banda. Asimismo, el poder emplear este tipo de aplicaciones no es algo barato, esto se debe a que las aplicaciones solo se pueden usar con los estetoscopios compatibles de las marcas anteriormente mencionadas.

Tabla 2: Características de aplicaciones 3M™ Littmann® y EKO.

Aplicación	Ancho de banda requerido	Precio	Marca de estetoscopios compatibles	Transmisión de audio
Littmann Learning (3M™ Littmann®, s/f)	No especificado	Gratis suscripción de \$ 40 USD •	Littmann	No
Littmann University (3M™ Littmann®, s/f)	No especificado	Modo de suscripción solo para universidades •	Littmann	Si. ■
Eko: Digital Sthetoscope (Eko Health, Inc., s/f)	No especificado	Incluido en la compra de un estetoscopio digital. •	Littmann y Eko	No
Eko Telehealth (Eko Health, Inc., s/f)	No especificado	No especificado por parte de ventas. •	Littmann y Eko	Si

Nota: • Solo habilitado en EE. UU; ■ Solo para entornos educativos.

2.4.3 Tele-auscultación en plataformas de videoconferencia convencionales

Los requerimientos de ancho de banda y los costos económicos asociados a la contratación de servicios de telemedicina y tele-auscultación profesionales han limitado su uso en países con una infraestructura de telecomunicaciones y recursos limitados, como ocurre en México. En este caso, los especialistas han optado por el uso de plataformas de videoconferencia convencionales como Google Meet, Microsoft

Teams, Zoom o WhatsApp para ofrecer servicios de tele-auscultación. En la Tabla 3 se muestran algunas de sus características principales.

Tabla 3: Características de plataformas de videoconferencia.

Plataforma	Zoom (Zoom Communications, Inc., 2023)	Microsoft Teams (MicrosoftHeidi, 2024)	Google Meet (Google LLC, 2025)	WhatsApp (WhatsApp, s/f)
Protocolos empleados	SIP-H.323 TLS 1.2 o superior	MTLS	WebRTC QUIC UDP	XMPP
		TURN		TLS
		RTP		Signal
		SRTTP		
		S2S		
Códec de audio	G.711 G.722 OPUS	SILK G.722	OPUS Lyra	AAC
				AMR-WB
				AMR-NB
				OPUS
Códec de video	H.264 H.264 de perfil alto	H.264	VP8	
		RGB24	VP9	H.264
		NV12	AV1	MPEG-4
Puertos	TCP			
	443			
	80	TCP/UDP	TCP	
	8801	50000-50059	443	
	8802	TCP	3478	TCP/UDP
	UDP	443	UDP	5222
	478	UDP	443	443
	3479	3478 – 3481	3478	
	8801 – 8810	50070 – 50089	19302 – 19309	
	20000 – 64000			
Criterio HIPAA	No	No	No	No

Estas soluciones ofrecen la ventaja de no tener costos económicos asociados, lo cual resulta beneficioso desde una perspectiva presupuestaria. No obstante, una cuestión relevante es que, en la actualidad, no están diseñadas para soportar sesiones de tele-auscultación, lo que limita su aplicabilidad en ciertos contextos clínicos. Asimismo, existen posibilidades de mejora en la administración de la seguridad, privacidad y confidencialidad de los datos del paciente.

De acuerdo con Castorena (2012), estas plataformas aún presentan requerimientos de ancho de banda

elevados, lo que puede ser un inconveniente en localidades remotas, donde las redes de datos suelen tener tasas de transferencia bajas o medianas. Asimismo, al convertir una señal cardiopulmonar analógica a digital, se necesita un muestreo y una cuantificación adecuados para evitar el deterioro de la señal. Sin embargo, estas plataformas de videoconferencia convencionales emplean códecs como el G.711 o el Opus, diseñados para la comprensión del habla normal, pero no para la transmisión de sonidos fisiológicos como los ruidos cardíacos y pulmonares, que requieren una mayor fidelidad de audio.

El uso de estas plataformas de comunicación “sin costo” puede considerarse como una solución temporal ante la falta de acceso a servicios de telemedicina más robustos y seguros, pero no es recomendable a largo plazo. Por estas razones, es necesario trabajar en el desarrollo de soluciones de tele-auscultación, que garanticen la calidad y seguridad del audio cardiopulmonar transmitido con requerimientos de ancho de banda adecuados para países de bajos recursos.

2.4.4 Trabajos previos

La ausencia de estándares internacionales que ayuden a definir la transmisión de audio cardiopulmonar mediante enlaces de comunicaciones dificulta la interoperabilidad y el desarrollo de sistemas de telemedicina eficaces. En este sentido, la calidad de los códecs empleados es crucial, ya que debe garantizar la preservación de la calidad del audio durante la transmisión.

En trabajos previos realizados (Castorena, 2012; Ramírez, 2017), se investigaron aspectos relacionados con el papel de la compresión en la calidad del audio cardiopulmonar transmitido y el ancho de banda requerido. Ambos estudios señalan que la transmisión de audio en redes con anchos de banda limitados (inferiores a 64 kbps) presenta el desafío de equilibrar la calidad del audio con la tasa de datos. Además, mencionan que problemas como la compresión excesiva del audio y la reducción en la frecuencia o resolución de muestreo representan riesgos adicionales que afectan tanto la calidad de la señal como la capacidad diagnóstica.

Por otro lado, la investigación desarrollada (Olachea, 2019) se centró en la transmisión de audio cardiopulmonar de calidad diagnóstica sobre redes de datos heterogéneas. En el trabajo se exploró la implementación de un esquema basado en el paradigma de radio en línea para llevar a cabo el proceso de tele-auscultación. La arquitectura del sistema propuesto se muestra en la Figura 3.

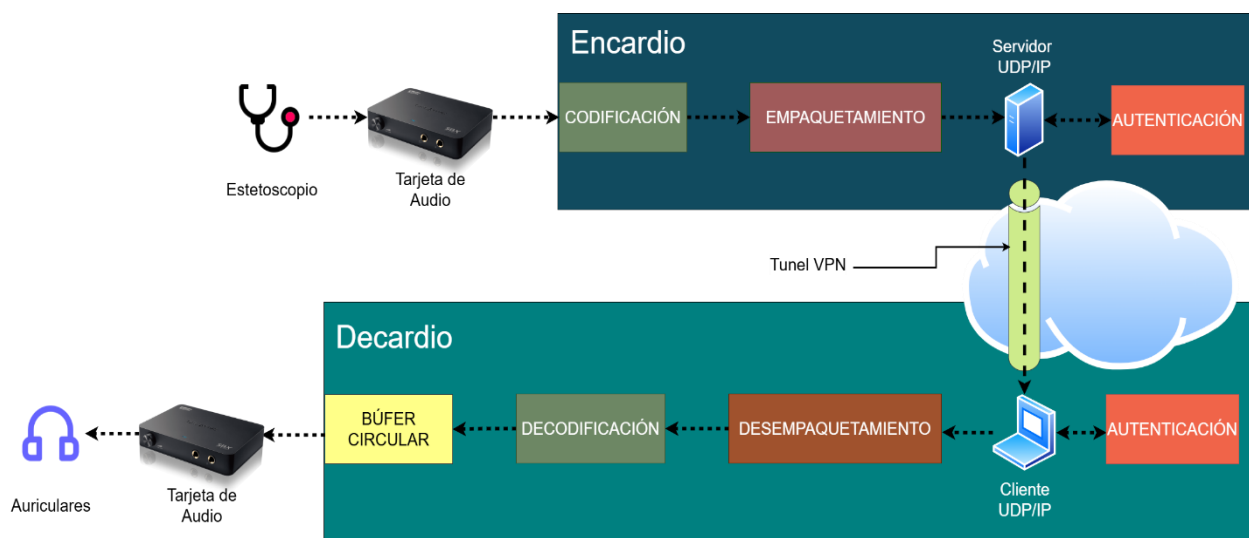


Figura 3: Esquema del sistema de transmisión de audio cardiopulmonar desarrollado en (Olachea, 2019).

El proceso inicia con la autenticación del transmisor y el receptor en el servidor UDP/IP a través de una red privada virtual (VPN, por sus siglas en inglés). Una vez validado el acceso, el transmisor comienza la captura del audio mediante un estetoscopio electrónico o digital. Posteriormente, una tarjeta de sonido **Sound Blaster X-Fi** convierte la señal de analógica a digital (en el caso del estetoscopio electrónico), tras lo cual el audio es codificado y empaquetado para su transmisión.

En el lado receptor, una vez establecida la conexión, se reciben los paquetes de audio y se realiza el proceso de desempaquetado y decodificación. Finalmente, el audio procesado se almacena en un búfer circular, debido a que previene interrupciones en la reproducción, mejorando la calidad del sonido con el apoyo de la misma tarjeta de sonido **Sound Blaster X-Fi**.

Para el proceso de codificación y decodificación se emplearon los códecs AAC y MP3 con una frecuencia de muestreo de 16 kHz y una resolución monoaural a 16 bits. Estas configuraciones se observan en la Tabla 4.

Para evaluar la calidad del audio cardiopulmonar y la efectividad del sistema de tele-auscultación, se implementaron pruebas subjetivas con personal de salud utilizando la metodología de múltiples estímulos con referencia y patrón escondidos (MUSHRA, por sus siglas en inglés). También, se desarrolló una prueba de puntuación media de opinión (MOS, por sus siglas en inglés) con el objetivo de medir la calidad de experiencia ofrecida utilizando parámetros de calidad de servicio.

Esta investigación (Olachea, 2019) ofreció resultados significativos en el ámbito de la transmisión de audio

cardiopulmonar. El estudio estableció que los códecs AAC y MP3, diseñados originalmente para la transmisión de música en directo, demostraron un rendimiento superior en la codificación y compresión de sonidos cardiopulmonares en comparación con códecs tradicionalmente utilizados en comunicaciones de voz por IP (VoIP).

Tabla 4: Configuraciones del programa con los códecs de audio.

Códec	Librería utilizada	Número de muestras para generar el gránulo	Frecuencia de muestreo	Bits por muestra
AAC	FDK-AAC	1024	16 kHz	16
MP3	LAME	576	16 kHz	16

Específicamente, códecs de voz como G722.2, G711 y EVS, presentaron deficiencias considerables al intentar reconstruir con fidelidad las señales de audio cardiopulmonar. Estas conclusiones fueron validadas mediante una evaluación realizada por profesionales del sector salud, quienes confirmaron que los códecs musicales son mejores para la transmisión de audio cardiorrespiratorio.

Adicionalmente, el estudio incluyó una evaluación exhaustiva de la calidad de experiencia desde la perspectiva de los parámetros técnicos de calidad de servicio. Los resultados indicaron que el sistema propuesto mantuvo un rendimiento óptimo ante fluctuaciones en el ancho de banda y *jitter*, cumpliendo satisfactoriamente con los estándares establecidos por la Unión Internacional de las Telecomunicaciones (ITU) para transmisiones de audio en tiempo real.

2.5 Discusión del servicio de tele-auscultación

La investigación desarrollada (Olachea, 2019) puso en evidencia la relevancia de contar con un servicio de tele-auscultación sobre redes con un ancho de banda limitado, sobre todo en poblaciones que se encuentran en zonas rurales o urbanas marginadas. Sin embargo, a pesar de los resultados positivos obtenidos, el sistema propuesto tiene algunas limitaciones relacionadas con el uso de VPNs, las cuales son necesarias para establecer la comunicación y garantizar la seguridad.

El empleo de la VPN supone dificultades para el despliegue de la plataforma en términos de administración y mantenimiento. También, aumenta la dificultad de uso por parte de los usuarios y añade sobrecarga de protocolo, lo que se traduce en un incremento del ancho de banda requerido. Además, una incorrecta

configuración de la VPN puede generar brechas de seguridad en la información de los pacientes.

Dado que la mayoría de las soluciones actuales de videoconferencia para aplicaciones médicas operan sobre infraestructuras de Internet convencionales sin necesidad de usar una VPN, es fundamental desarrollar mecanismos que faciliten la transmisión de audio cardiopulmonar en tiempo real a través de la Internet que no dependan de este tipo de tecnología.

Considerando lo anterior, es necesario continuar investigando acerca de la transmisión de audio cardiopulmonar a través de Internet, para garantizar diagnósticos precisos, en tiempo real y con los menores requisitos de ancho de banda posibles.

Capítulo 3. Conectividad en la red: La importancia de la comunicación en la era digital

En este capítulo se proporciona una revisión de los modelos y arquitecturas de red que se implementan en la actualidad para lograr una comunicación continua y segura.

3.1 Arquitectura de red

Una arquitectura de red define un conjunto de reglas y protocolos que regulan la interacción entre los dispositivos que la componen. Estas arquitecturas son fundamentales para la estandarización de las redes, y facilitan la distribución de recursos y servicios dentro de los sistemas conectados.

Existen diversas clasificaciones de arquitecturas de acuerdo con sus características, como lo son el tipo de tecnologías de transmisión que emplean (alámbricas o inalámbricas), su alcance y área de cobertura (tráfico local, metropolitano o de área amplia). Asimismo, existe la clasificación funcional, donde destacan 2 modelos principales: la arquitectura cliente-servidor y la arquitectura par-a-par (Buford et al., 2009). La primera se basa en el uso de un servidor central, mientras que la segunda permite que los nodos funcionen como clientes y servidores simultáneamente. A continuación, se ahondará en las arquitecturas cliente-servidor y par-a-par (P2P).

3.1.1 Arquitectura par-a-par (P2P)

El modelo P2P se caracteriza por ser una arquitectura descentralizada, donde los nodos pueden intercambiar la información sin tener que depender de un servidor central que controle el funcionamiento de la red. En este caso, cada nodo tiene una función equivalente y debe funcionar a la vez como cliente y como servidor, lo que permite un intercambio de datos directo entre pares.

Estas redes cuentan con mecanismos establecidos para el tratamiento de errores y el mantenimiento de la funcionalidad en caso de pérdida de nodos, lo que unido a otras características como la transferencia de archivos en forma paralela y la privacidad, convierte a estas redes en una herramienta muy útil, eficaz, robusta e inteligente para aplicaciones muy diversas. No obstante, el modelo P2P también presenta

inconvenientes en torno a su gestión, monitoreo y seguridad, de ahí la necesidad de un equilibrio entre la eficacia y el control (Amoretti y Zanichelli, 2018; Buford et al., 2009; Kurose y Ross, 2017; Zhang, 2013).

Las redes P2P se clasifican en tres tipos: (Jaideep y Battula, 2016)

- **Supernodo.** Un nodo asume el papel de supernodo, al encargarse de gestionar y facilitar la comunicación entre el resto de los nodos que forman parte de la arquitectura, así como el almacenamiento y la distribución de contenido.
- **Híbrida.** Las redes P2P híbridas reúnen características de las arquitecturas de cliente-servidor y P2P. En este caso existe un servidor que se encarga de facilitar el descubrimiento de nodos y la indexación del contenido, mientras que el intercambio de la información se realiza directamente entre los dispositivos que forman parte de la arquitectura.
- **Pura.** En una red P2P pura, todos los nodos tienen el mismo nivel de responsabilidad y capacidad. Ningún nodo se preocupa por conocer el privilegio de los demás nodos y tampoco de la existencia de algún servidor central. Cualquier par puede actuar tanto como cliente como servidor y todos los pares se envían mensajes directamente entre ellos. Este modelo tiene un nivel alto de escalabilidad y de disponibilidad, ya que la red puede seguir funcionando aun cuando se desconecten muchos de los nodos.

3.1.2 Arquitectura cliente-servidor

La arquitectura cliente-servidor es un modelo fundamental en sistemas distribuidos, basado en una división clara de roles: el cliente actúa como solicitante de servicios o recursos, mientras el servidor centraliza su provisión, gestionando desde la lógica de negocio hasta el almacenamiento de datos (Tanenbaum y Steen, 2006). Esta arquitectura mejora las responsabilidades y establece un marco para sistemas escalables.

La interacción se basa en un diálogo estructurado: el cliente envía una solicitud y espera una respuesta. Este es un mecanismo eficiente, pero dependiente de la estabilidad de la red, donde factores como la latencia pueden afectar el rendimiento. Una de las grandes ventajas que ofrece es la escalabilidad, ya que los servidores pueden adaptarse a demandas crecientes mediante la gestión de múltiples conexiones

simultáneas, lo que es ideal para entornos dinámicos. Sin embargo, la centralización de recursos exige medidas de seguridad robustas para proteger datos confidenciales, así como un mantenimiento simplificado mediante actualizaciones globales en el servidor. A pesar de los desafíos que afronta esta arquitectura, tales como la gestión de la concurrencia o la dependencia de la infraestructura de red, posee una gran capacidad para equilibrar la eficiencia, control y adaptabilidad.

La arquitectura cliente-servidor se puede clasificar en: (Coulouris et al., 2011; Schuldt, 2009)

- **Arquitectura de dos niveles.** Se compone de dos elementos (cliente y servidor), siendo el cliente el que comienza la solicitud de un servicio y el servidor el que la procesa y devuelve la respuesta al cliente.
- **Arquitectura de tres niveles.** Añade una capa intermedia entre el cliente y el servidor, normalmente una capa de negocios o de aplicación que se encarga de procesar la lógica de la aplicación, facilitando así la escalabilidad y mantenimiento de esta.
- **Arquitectura de múltiples niveles.** En este tipo de arquitectura, se añaden más capas intermedias entre el cliente y el servidor, cada una con una función especializada, lo que permite una mayor escalabilidad y gestión de la aplicación.

3.1.3 Comparativa entre la arquitectura cliente-servidor y P2P

Como se ha comentado anteriormente, las arquitecturas P2P y cliente-servidor tienen diferencias fundamentales en su estructura y funcionamiento. En algunos trabajos (Kanazawa y Takami, 2018), se habla acerca de la evaluación de las arquitecturas anteriormente mencionadas. En esta investigación se consideran parámetros como la gestión de usuarios (que abarca el control de autenticación, permisos y acceso) y la gestión de contenidos (que se enfoca en la distribución, almacenamiento y actualización de recursos). También se analiza la carga de trabajo tanto en las terminales, donde el procesamiento depende de los dispositivos finales, como en el servidor, que enfrenta una alta demanda de recursos en arquitecturas cliente-servidor o su ausencia en modelos P2P.

Otro aspecto relevante es la concentración del tráfico, que puede dirigirse a un punto central o distribuirse, afectando el número de comunicaciones y la cantidad de nodos involucrados. Además, la tolerancia a fallos

resulta fundamental para asegurar la operación continua ante fallos críticos, junto con mecanismos de prevención para evitar la integración de dispositivos falsos o maliciosos. La comparativa de los aspectos mencionados anteriormente se proporciona en la Tabla 5.

Tabla 5: Comparación entre las diferentes arquitecturas de red.

Parámetro	Cliente-servidor	P2P Híbrido	P2P Puro	P2P Supernodo
Gestión de usuarios	Fácil	Fácil	Difícil	Medio
Gestión de contenidos	Fácil	Fácil	Difícil	Difícil
Carga de trabajo (terminal)	Pequeña	Pequeña	Grande	Grande
Carga de trabajo (servidor)	Media	Pequeña	Ninguna	Ninguna
Concentración del tráfico	Media	Pequeña	Grande	Media
Número de comunicaciones	Pequeña	Media	Grande	Grande
Tolerancia a fallos	Baja	Baja	Alta	Alta
Prevención de falsificaciones	Alta	Media	Baja	Baja

Las arquitecturas cliente-servidor y P2P presentan retos diferentes de su diseño estructural puro. En lo que respecta al cliente-servidor, encontramos:

- **Vulnerabilidades en materia de seguridad** (suplantación de IP; recursos compartidos que requieren protección mediante *firewalls* y VPN).
- **Problemas de escalabilidad** (conurrencia de clientes y su configuración dinámica).
- **Dificultades operativas** (distribución de carga, sincronización, costos).

En las redes P2P, los problemas provienen fundamentalmente de la naturaleza descentralizada de esta arquitectura, por ejemplo: (Buford et al., 2009; Ford et al., 2006, 2008)

- La carencia de autoridad central (ataques Sybil y comportamiento oportunista).
- Rotación (churn) de los nodos en general (lo que compromete la estabilidad de la red).
- Falta de alineación entre redes lógicas y físicas, lo que produce problemas de eficiencia por medio del enrutamiento.
- Las características que presentan los recursos distribuidos (búsqueda débil, saturación de ancho de banda) que dificultan su gestión (Coulouris et al., 2011; Tanenbaum y Steen, 2006).

En general, el modelo cliente-servidor controla la red a costa de su resiliencia mientras que las redes P2P ceden el control a cambio de autonomía, por lo que es razonable suponer que las soluciones del futuro opten por soluciones P2P híbridas que combinen las características de los distintos tipos de red. En este contexto de complejidad arquitectónica para las redes P2P, uno de los principales retos para la conectividad de los dispositivos a través del Internet es el uso de enrutadores que implementan el mecanismo de traducción de direcciones de red (NAT) (Buford et al., 2009).

3.2 Impacto del NAT en la conectividad P2P y desafíos de la red

El surgimiento del mecanismo de “traducción de direcciones de red” (NAT) se remonta a la década de los 90’s, cuando la cantidad de dispositivos activos en Internet se incrementó sustancialmente de tal manera que comenzó a agotarse el direccionamiento IPv4 rápidamente. Este fenómeno llegó a tal punto que la Autoridad de Números Asignados de Internet (IANA, por sus siglas en inglés) decidió reservar 3 rangos de direccionamiento IP privados con el objetivo de evitar el agotamiento de direcciones (Tabla 6) y establecer las bases para la correcta implementación del NAT.

Tabla 6: Rango de direccionamientos establecidos por la IANA.

Clase	Rango	Número de direcciones
A	10.0.0.0 – 10.255.255.255	16,777,216
B	172.16.0.0 – 12.31.255.255	1,048,576
C	192.168.0.0 – 192.168.255.255	65,536

El NAT ofrece una posibilidad de conexión a Internet para los dispositivos que se encuentran en una red privada (como se observa en la Figura 4). Sin embargo, en las arquitecturas P2P el NAT complica las conexiones de los dispositivos, debido a que generalmente los nodos no cuentan con una IP y puerto de comunicación visible para poder recibir las conexiones de otros dispositivos directamente (Buford et al., 2009; Holdrege y Srisuresh, 1999; L. Zhang, 2008). Es necesario aclarar que el término “puerto de comunicación” se refiere a un elemento de nivel de transporte que permite a los dispositivos direccionar el flujo de datos a una aplicación o servicio en concreto.

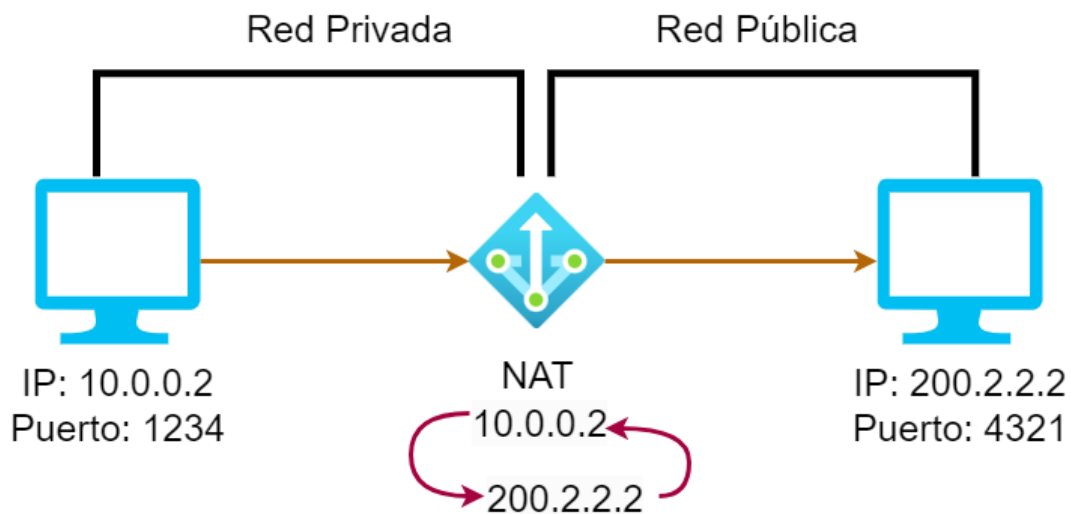


Figura 4: Funcionamiento básico del NAT.

Actualmente, existe una gran variedad de tipos de NAT, algunos de los cuales se mencionan a continuación:

- **NAT bidireccional** (*BNAT*, por sus siglas en inglés) (Chen y Jia, 2009; Smith y Hunt, 2002). El NAT bidireccional es un tipo de traducción de direcciones de red que permite la comunicación efectiva entre clientes en redes internas. En el NAT tradicional, las conexiones solo son iniciadas por los dispositivos pertenecientes a la red interna, pues son solamente ellos los que están autorizados. Entonces, el NAT bidireccional es diferente al NAT tradicional, ya que es capaz de establecer las conexiones en ambas direcciones.
- **Doble nat** (o del inglés *Twice NAT*) (Smith y Hunt, 2002; Taleb et al., 2015). El doble NAT consiste en la traducción de la dirección IP de origen, así como la de destino de los paquetes de información. Permite tener una configuración de red más flexible y es práctica en situaciones de solapamiento de los espacios de direcciones IP.

- **Multi-homed NAT** (Engelstad y Egeland, 2004; Smith y Hunt, 2002). El *multi-homed* NAT usualmente se presenta en lugares donde se cuentan con múltiples conexiones a Internet y cada conexión puede contar con un NAT distinto. Uno de los conflictos principales se encuentra en que todo el tráfico de una sesión de comunicación debe de pasar por el mismo NAT, si esto no sucede, la comunicación se interrumpirá.
- **Traducción de puertos y direcciones de red** (NAPT, de sus siglas en inglés) (Egevang y Srisuresh, 2001; Wing, 2010). El NAPT permite la traducción de direcciones IP, así como de los identificadores de transporte, como son los puertos de comunicación TCP/UDP. Esto facilita que múltiples dispositivos en una red privada compartan únicamente una dirección IP pública, utilizando diferentes puertos para lograr la identificación de las sesiones.

En la Figura 5 se muestran dos dispositivos que son parte de una red que se encuentran detrás de un NAPT, por lo cual pueden compartir una sola IP pública para salir a Internet. Asimismo, se observa que en su tabla de mapeo al salir a Internet tienen diferentes puertos de comunicación, esto con el fin de identificar sus sesiones. Este tipo de NAT es el más común en las redes actuales.

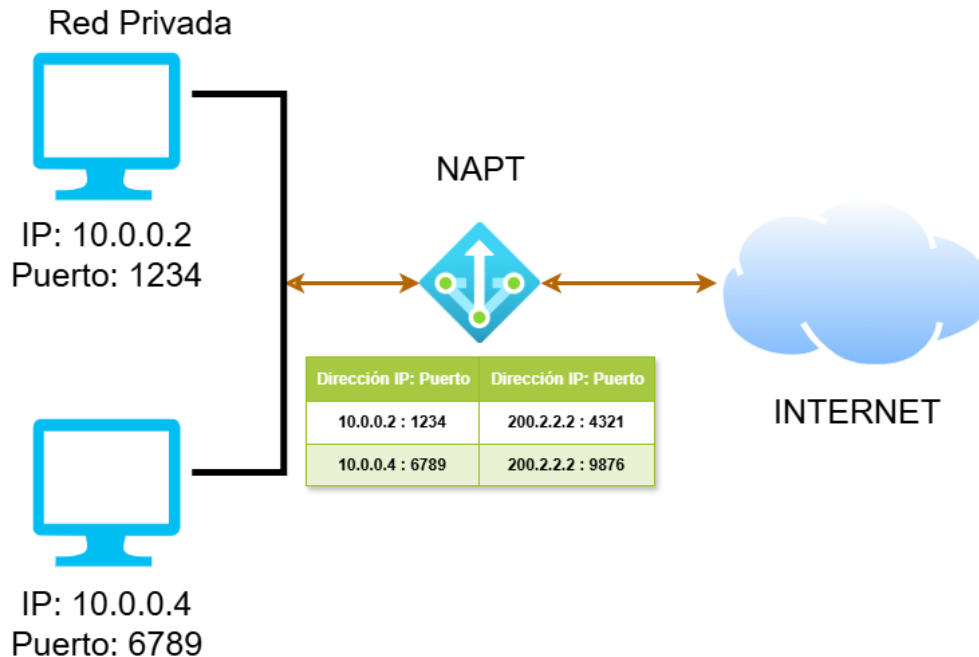


Figura 5: Funcionamiento del mecanismo de traducción de puertos y direcciones de red (NAPT).

Las cuatro principales variantes de NAPT son: (Poulin et al., 2012; Rosenberg et al., 2003; Wagner, 2015)

- **NAT de cono completo:** Un NAT de cono completo implementa que todas las peticiones provenientes de la misma tupla (dirección IP privada y puerto de comunicación) se asignan a la misma dirección IP pública. Esto permite el establecimiento de un canal de comunicación bidireccional, permitiendo que cualquier dispositivo externo pueda iniciar conexiones con el dispositivo interno mediante su dirección IP pública y su puerto, siempre que el mapeo NAT correspondiente exista en la tabla de mapeo (Figura 6-A).
- **NAT de cono restringido:** El NAT de cono restringido se comporta de manera similar al NAT de cono completo, puesto que implementa un mapeo donde cada tupla (IP y puerto privados) mantiene una correspondencia consistente con una tupla (IP y puerto público). No obstante, esta implementación añade un mecanismo de seguridad adicional, donde permite solo los paquetes entrantes de los dispositivos externos (IP pública) que hayan sido contactados previamente por un dispositivo interno, teniendo así un filtro basado en la dirección IP de origen de los paquetes externos (Figura 6-B).
- **NAT de cono con puertos restringidos:** El NAT de cono con puertos restringidos es una variante del NAT de cono restringido, donde la sutil diferencia radica en que la restricción también incluye los números de puerto. Esto sucede debido a que implementa un mapeo donde la tupla (IP y puerto privados) mantiene una conexión con una tupla (IP y puerto públicos), con un filtrado bidireccional más riguroso. Donde solo permite paquetes entrantes desde dispositivos externos que coincidan tanto en la dirección IP, como en el puerto de origen, que fueron previamente contactados por el dispositivo interno, estableciendo así un control detallado entre la sesión (Figura 6-C).
- **NAT simétrico:** El NAT simétrico implementa un mapeo dinámico; esto significa que, cuando un dispositivo interno se comunica con un dispositivo externo, se asignará un nuevo mapeo de dirección IP y puerto externos que es único para esa sesión de comunicación con el dispositivo interno. Por ello, los dispositivos externos sólo podrán mantener la comunicación mediante la sesión previamente realizada por el dispositivo interno, en caso contrario no podrán comunicarse (Figura 6-D).

Como se mencionó anteriormente, la implementación de estas variantes de NAT, aunque resuelve el problema de la escasez de direcciones IPv4, constituye uno de los principales factores que dificultan el establecimiento de redes P2P en la actualidad. Por esta razón, se han desarrollado también diferentes

técnicas para atravesar el NAT y permitir la comunicación directa entre nodos.

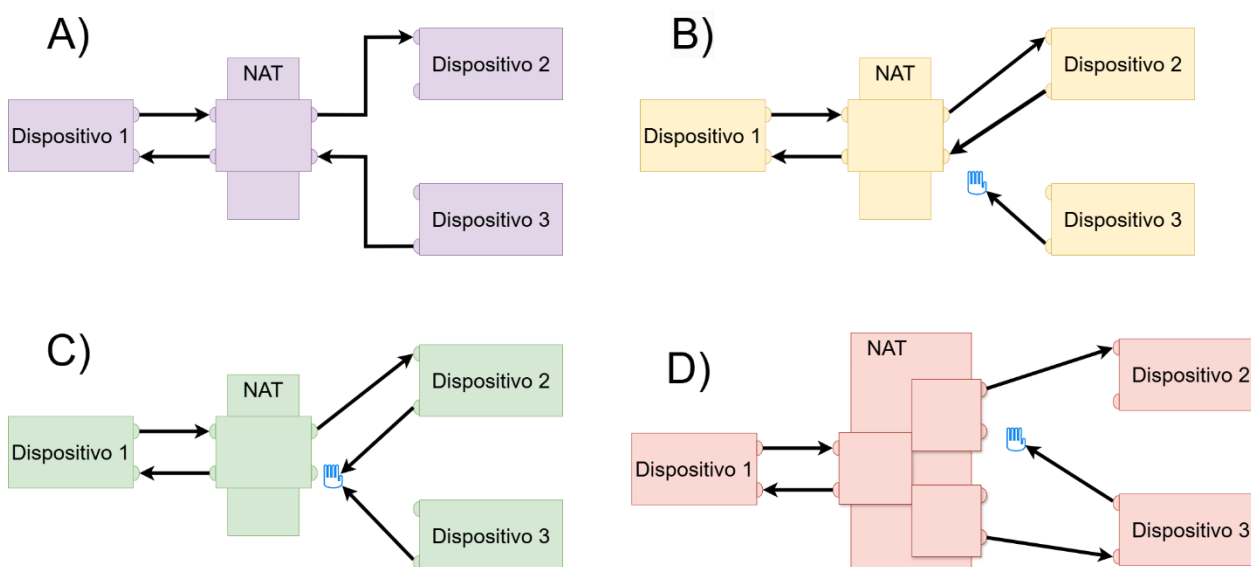


Figura 6: Descripción gráfica de los principales tipos de NAT: A) NAT de cono completo; B) NAT de cono restringido; C) NAT de cono con puertos restringidos; D) NAT simétrico.

3.3 Métodos de conectividad para atravesar el NAT

Para que los dispositivos puedan eludir el NAT se requieren algunas “acrobacias de red”, más conocidas como NAT Transversal. Estas técnicas no siempre son efectivas, puesto que dependen del tipo de NAT en el que se encuentren los dispositivos. Acorde a Rosenberg et al. (2003), una alternativa para superar las barreras del NAT es el uso de puertas de enlace de capa de aplicación (ALG, por sus siglas en inglés), las cuales funcionan como un intermediario al permitir el acceso a ciertos tipos de tráfico de aplicaciones para sobrepasar el NAT o bien un firewall. No obstante, el uso de un ALG no es muy recomendado debido a su complejidad si la red aumenta y a que el rendimiento de la red se ve afectado por la necesidad de inspeccionar y modificar el tráfico de las aplicaciones. Además, no es compatible con todos los protocolos de aplicación, y el procesamiento adicional puede ocasionar una sobrecarga de trabajo en toda la red.

En el trabajo de Ford et al. (2006), discuten acerca de estrategias donde se usan aplicaciones para señalar directamente con el NAT empleando protocolos como lo son UPnP (del inglés — Universal Plug and Play) y el MIDCOM (del inglés — Middlebox Communications). Estas implementaciones se emplean para ayudar a los dispositivos a navegar a través del NAT, pero no son recomendables por cuestiones de seguridad, esto debido a que pueden permitir a aplicaciones no autorizadas ingresar a su red privada o incluso llegar a realizar modificaciones en su tabla de NAT (Poulin et al., 2012).

Asimismo, existen mecanismos para “eludir” al NAT sin tener la necesidad de alterar el dispositivo o dejarlo sin protección alguna de forma directa, lo que los convierte en una opción factible para su implementación en arquitecturas P2P. Estos mecanismos presentan diferenciación en función del protocolo de capa de transporte para el cual han sido implementados. **En este trabajo solo se analizan las técnicas transversales del NAT que involucran el uso del Protocolo de Datagramas de Usuario** (UDP, por sus siglas en inglés). El uso de UDP está ligado al cumplimiento de los requisitos de rendimiento especificados para la arquitectura propuesta en cuanto a transmisión en tiempo real, pues se trata de un protocolo no orientado a conexión, que no realiza control de errores ni garantiza la entrega confiable, lo que reduce la latencia asociada a la comunicación.

En las investigaciones (Ford et al., 2008; Huang et al., 2019; Poulin et al., 2012; Wagner, 2015) se documentan diversos mecanismos para sobrepasar el NAT y facilitar la comunicación entre dispositivos en arquitecturas punto a punto (P2P). A continuación, se procede al análisis de estos mecanismos en su versión asociada al protocolo UDP.

Perforación UDP (UDP Hole Punching). Esta técnica es empleada para establecer conexiones de par-a-par (P2P) a través de dispositivos de NAT, específicamente mediante la implementación del protocolo UDP. Este mecanismo posibilita la comunicación directa entre dispositivos ubicados detrás de diferentes NATs, eliminando la necesidad del uso de un servidor intermediario una vez establecido el enlace inicial. En la Figura 7 se muestra el proceso de establecimiento de comunicación, donde se puede segmentar en etapas de conexión. El mecanismo inicia con la fase de registro, donde los dispositivos (cliente A y cliente B) establecen comunicación con un servidor de coordinación mediante la transmisión de mensajes UDP, permitiendo al servidor obtener las direcciones IP y puertos públicos de cada dispositivo. Subsecuentemente, en la fase de intercambio, el servidor distribuye a cada dispositivo los parámetros de conexión con su contraparte. La tercera etapa consiste en la transmisión simultánea de paquetes UDP entre los dispositivos, empleando los parámetros previamente distribuidos, donde la sincronización temporal de estos paquetes es crítica para garantizar la apertura apropiada de la sesión del NAT. Finalmente, la fase de establecimiento se completa cuando los paquetes alcanzan sus respectivos destinos, momento en el cual los dispositivos NAT de los dos clientes efectúan el mecanismo de *Hole Punching* en sus tablas de traducción de direcciones, habilitando así un canal de comunicación directa que prescinde del servidor intermediario.

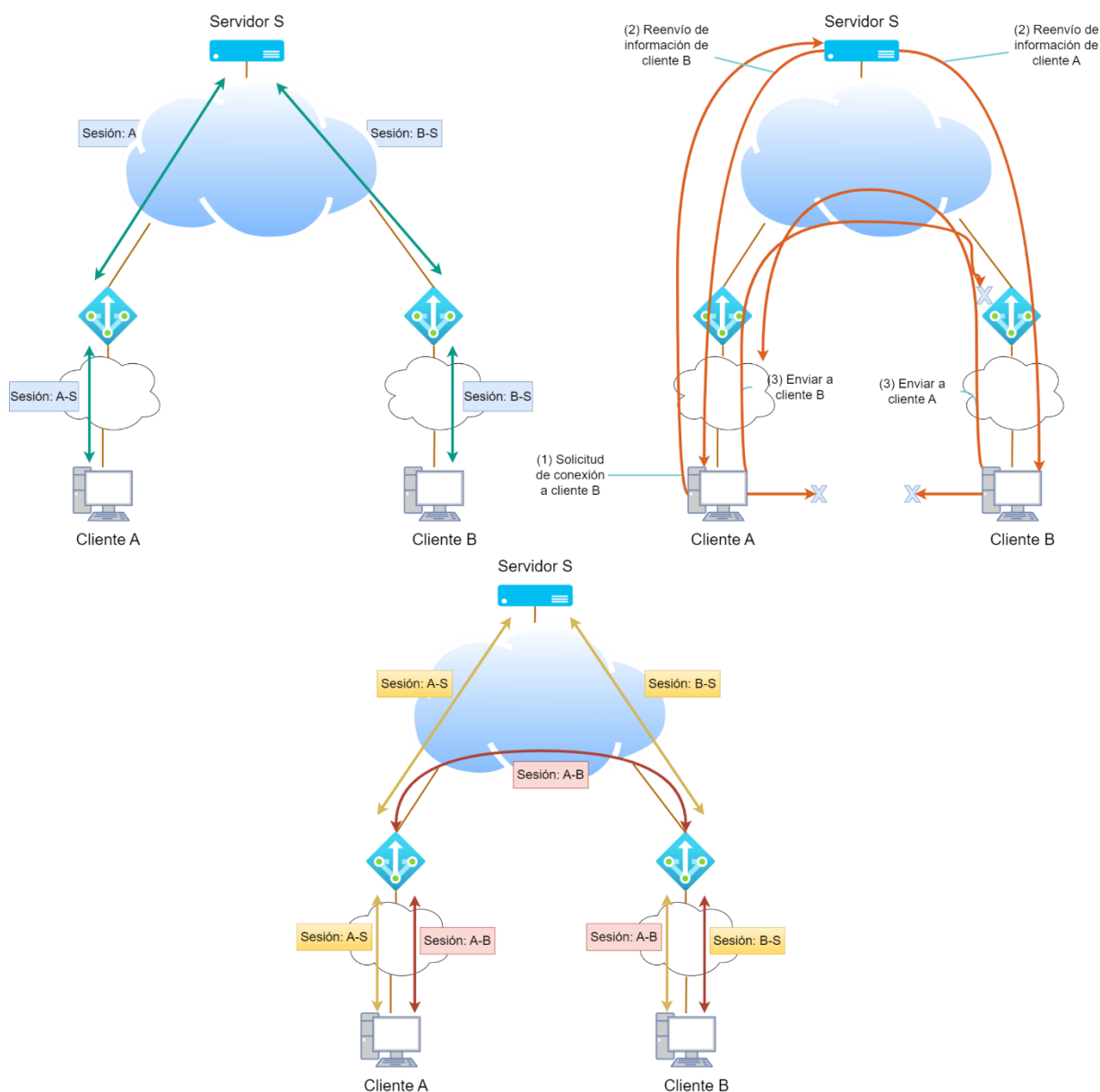


Figura 7: Diagramas que muestran la secuencia de intercambio de paquetes utilizada en el mecanismo UDP *Hole Punching*.

Connection Reversal. La conexión reversa facilita el establecimiento de enlaces directos entre un dispositivo ubicado detrás de un NAT y otro dispositivo que dispone de una dirección IP pública. Este mecanismo muestra particular eficacia en escenarios asimétricos donde la topología de red presenta un único dispositivo detrás de un dispositivo NAT, mientras que su contraparte mantiene una conectividad directa a la red global de Internet. En la Figura 8 se muestra el proceso de conexión empleando el mecanismo de conexión reversa. Primero se ejecuta una fase de registro donde el cliente con restricciones NAT (cliente A) y el cliente con IP pública (cliente B) transmiten sus parámetros de conexión a un servidor de coordinación preestablecido (servidor S). Posteriormente, cuando el cliente B requiere establecer

comunicación con el cliente A, transmite una solicitud de conexión mediante el servidor S, requiriendo que el cliente A inicie el enlace comunicacional. La fase final comprende la retransmisión de la solicitud por parte del servidor S hacia el cliente A, quien establece una conexión directa hacia el cliente B, utilizando la dirección IP pública y el puerto previamente especificado; esta conexión resulta viable debido a que el dispositivo NAT del cliente A permite el tráfico saliente generado desde su red interna.

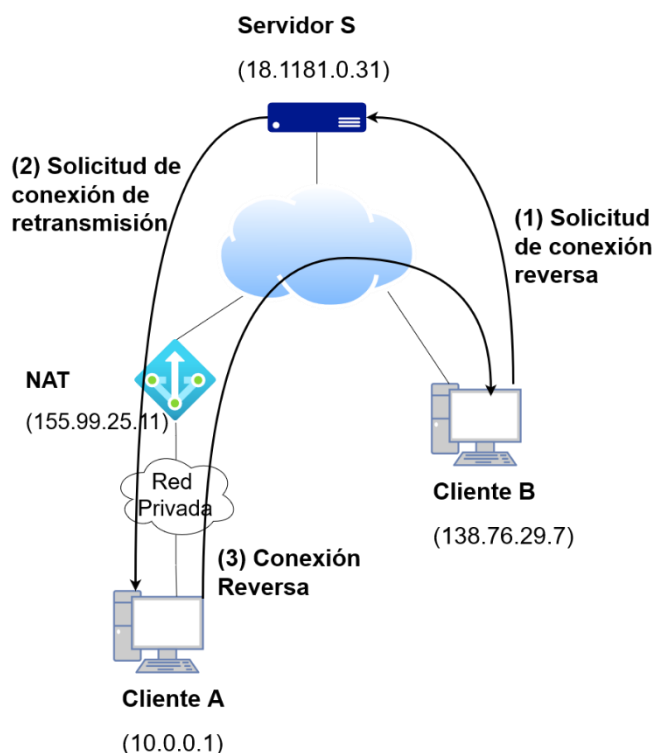


Figura 8: Diagrama del mecanismo de conexión reversa.

UDP Port Number Prediction. La predicción del número de puerto UDP es un método de comunicación con dispositivos NAT que permite adelantarse a las asignaciones de los puertos públicos para futuras comunicaciones. Esta técnica se basa en la observación del comportamiento determinístico que presentan muchos dispositivos NAT cuando realizan asignaciones de puertos, que normalmente utilizan estrategias secuenciales o algoritmos de asignación que pueden preverse. Esta característica permite que un cliente, después de haber establecido una conexión con un dispositivo NAT, pueda averiguar con certeza cuáles serán los identificadores de los puertos que se utilizarán en comunicaciones posteriores que se comiencen desde la misma dirección de red privada.

Relaying. El mecanismo *Relaying* suele ser la solución más factible al momento de conectar dos dispositivos en una arquitectura P2P, pero es una implementación poco eficiente, debido a que no se establece una comunicación directa entre los dispositivos, como se menciona en los mecanismos

anteriores. Aquí el servidor toma papel central para actuar como un intermediario para el reenvío de los paquetes entre los dos dispositivos. En la Figura 9 se muestran dos dispositivos (cliente A y cliente B) que se encuentran detrás de un NAT diferente cada uno. Ambos clientes se encuentran conectados a un servidor (servidor S). En el caso de que el cliente A decida comunicarse con el cliente B, ambos clientes deben de abrir y mantener una sesión cliente-servidor con el servidor S. El cliente A envía el mensaje hacia el servidor S, una vez que el servidor S recibe los datos del cliente A, los retransmite hacia el cliente B a través de su sesión.

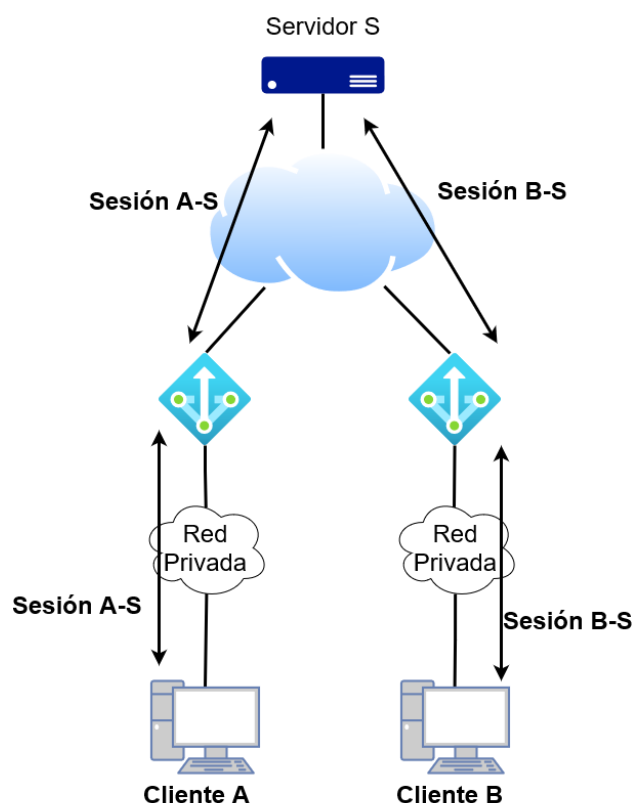


Figura 9: Diagrama del mecanismo de *Relaying*.

En la Tabla 7 se muestra una versión propia de los datos presentados por Poulin et al. (2012), en los que se agrupan las diferentes técnicas transversales del NAT y el modo en el que estas operan, en función del tipo de NAT que haya en cada uno de los dispositivos. En este apartado no se han considerado los métodos que requieren la modificación directa del comportamiento del NAT porque no han llegado a ser utilizados de un modo consensuado entre todos los dispositivos, ni tampoco la técnica UDP Port Number Prediction, ya que su efectividad depende de muchos factores externos, no del comportamiento mismo del NAT. De acuerdo con la Tabla 7, los mecanismos que mejor desempeño presentan son UDP *Hole Punching* y Relay.

Tabla 7: Funcionamiento de las técnicas transversales del NAT en relación con el tipo de NAT presente en cada extremo.

	Ninguno	Cono lleno	Cono de dirección restringida	Cono de Puerto restringido	Simétrico
Ninguno	Directo	Directo	Relay	Relay	Relay
	Relay	Relay	Con. Reversa	Con. Reversa	Con. Reversa
		Con. Reversa	<i>Hole Punching</i>	<i>Hole Punching</i>	<i>Hole Punching</i>
		<i>Hole Punching</i>			
Cono lleno	Directo	Directo	Relay	Relay	Relay
	Relay	Relay	Con. Reversa	Con. Reversa	Con. Reversa
	Con. Reversa	Con. Reversa	<i>Hole Punching</i>	<i>Hole Punching</i>	<i>Hole Punching</i>
	<i>Hole Punching</i>	<i>Hole Punching</i>	UPnP	UPnP	UPnP
Cono de Dirección restringida	Directo	Directo			
	Relay	Relay	Relay	Relay	Relay
	Con. Reversa	Con. Reversa	<i>Hole Punching</i>	<i>Hole Punching</i>	<i>Hole Punching</i>
	<i>Hole Punching</i>	<i>Hole Punching</i>	UPnP	UPnP	UPnP
Cono de puerto restringido	Directo	Directo			
	Relay	Relay	Relay	Relay	Relay
	Con. Reversa	Con. Reversa	<i>Hole Punching</i>	<i>Hole Punching</i>	UPnP
	<i>Hole Punching</i>	<i>Hole Punching</i>	UPnP	UPnP	
Simétrico	Directo	Directo			
	Relay	Relay	Relay	Relay	Relay
	Con. Reversa	Con. Reversa	<i>Hole Punching</i>	UPnP	UPnP
	<i>Hole Punching</i>	<i>Hole Punching</i>	UPnP		
	UPnP	UPnP			

3.4 Ejemplo de una arquitectura para comunicaciones híbridas – WebRTC

Como se ha mencionado anteriormente, existen mecanismos para superar las limitaciones del NAT y permitir el establecimiento de redes P2P. Una de las soluciones arquitectónicas más prometedoras para lograr la comunicación par a par, que hace uso de estas técnicas transversales del NAT, es WebRTC (del

inglés, Web Real Time Communications). Esta tecnología es una iniciativa de código abierto promovida por Google for Developers en 2013 y soportada por diferentes organizaciones y empresas. WebRTC permite a los navegadores web y aplicaciones móviles realizar comunicaciones en tiempo real mediante el intercambio directo de datos multimedia (audio, vídeo, datos) sin necesidad de complementos.

Los principales componentes de WebRTC son: (Dubois et al., 2021; Google for Developers, 2013; Grigorik, 2013)

- **MediaStream**, el cual permite acceder a la cámara y micrófono del usuario.
- **PeerConnection**, el cual administra la configuración de las llamadas y el flujo de audio y video.
- **DataChannel**, el se encarga de la transmisión bidireccional de datos entre los dispositivos.

La arquitectura de red WebRTC, mostrada en la Figura 10, presenta una estructura distribuida que requiere un mínimo de dos dispositivos terminales, complementada por componentes esenciales de infraestructura: un servidor de señalización y un servidor de utilidades para atravesar sesiones NAT (STUN, por sus siglas en inglés). Adicionalmente, puede incorporar un servidor transversal usando relés alrededor de NAT (TURN, por sus siglas en inglés) como elemento opcional para casos específicos de conectividad compleja.

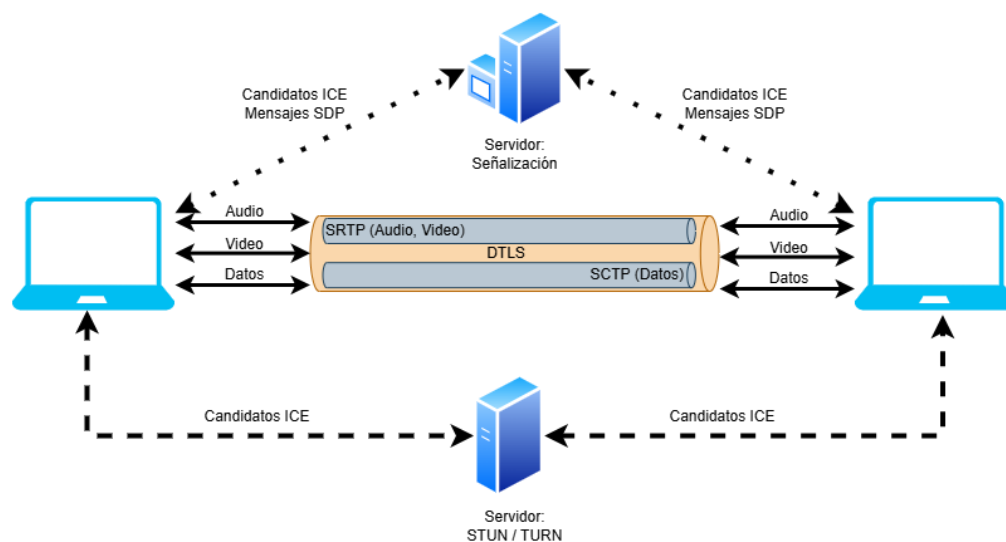


Figura 10: Modelo de comunicación empleando WebRTC.

Esta infraestructura distribuida implementa un mecanismo de autogestión para las sesiones P2P facilitando la interconexión automática entre participantes mediante el establecimiento de enlaces directos. El proceso se fundamenta en una señalización inicial que intercambia información del establecimiento de conectividad interactiva (ICE, por sus siglas en inglés), la cual incluye parámetros críticos como direcciones IP y puertos que constituyen los puntos de acceso potenciales.

En el contexto operativo, el servidor STUN cumple una función diagnóstica fundamental al permitir que los clientes identifiquen su dirección IP pública y determinen la tipología del NAT en su entorno de red. De manera complementaria, el servidor TURN opera como un intermediario de retransmisión de paquetes, proporcionando una ruta alternativa de comunicación cuando las conexiones P2P directas resultan inviables debido a restricciones impuestas por NAT o firewalls, garantizando así la continuidad de la conectividad en condiciones de red poco favorables.

El flujo de WebRTC para realizar una conexión entre dispositivos se puede observar en cinco etapas: (Manson, 2013)

1. **Conexión de usuarios.** Los dispositivos se conectan a un servidor de señalización para posteriormente realizar su registro, esto mediante el protocolo de WebSocket o bien HTTPS.
2. **Inicio de señalización.** El servidor de señalización actúa como intermediario para que los dispositivos puedan negociar la configuración de la conexión por medio del protocolo SDP (del inglés – *Session Description Protocol*).
3. **Búsqueda de candidatos.** Cada dispositivo recopila información sobre las direcciones IP y puertos disponibles para la conexión utilizando el Protocolo de Establecimiento Interactivo de Conectividad (ICE, por sus siglas en inglés) (Keränen et al., 2018). La función principal de este protocolo es examinar sistemáticamente diversas direcciones IP y puertos mediante el método *UDP Hole Punching*. Su objetivo es identificar las opciones más adecuadas para establecer canales de intercambio de datos, lo que optimiza el desempeño de la aplicación. La selección de los candidatos públicos se basa en el estándar STUN (Matthews et al., 2008), que permite detectar y describir los puntos de acceso a la red. En caso de que la comunicación directa entre dispositivos no sea posible, se recurre al protocolo TURN (Matthews et al., 2010).
4. **Negociación de las sesiones multimedia.** Los dispositivos intercambian información detallada

sobre sus características para el establecimiento de la sesión multimedia con ayuda del protocolo de descripción de sesiones (de sus siglas en inglés, *SDP*) (Begen et al., 2021), algunos de los datos compartidos son el tipo de datos, formato, resolución, tasa de bits, codécs, entre otros.

5. **Inicio de la transmisión con *RTCPeerConnection*.** Una vez completadas las etapas anteriores, los dispositivos comienzan a transmitir información, en las Figuras 10 y 11 se aprecia que el transporte de los datos multimedia se ejecuta empleando el protocolo seguro de transporte en tiempo real (de sus siglas en inglés, *SRTP*) (Carrara et al., 2004) y los canales de datos emplean el protocolo de control de transmisión de flujos (de sus siglas en inglés, *SCTP*) (Stewart et al., 2022). Dichos canales de datos son protegidos y garantizan la seguridad, gracias a que emplean el protocolo de seguridad de la capa de transporte de datagramas (de sus siglas en inglés, *DTLS*) (Rescorla et al., 2022).

Asimismo, en la Figura 11 se observa la pila de protocolos del protocolo WebRTC, donde arriba de la capa de transporte UDP tiene los protocolos de ICE, STUN y TURN para facilitar la conectividad entre dispositivos, incluso a través de NAT o de *firewalls*. Para garantizar la seguridad, implementa el protocolo DTLS en conjunto con el protocolo SRTP para cifrar, autenticar y proteger la transmisión de audio y video en tiempo real.

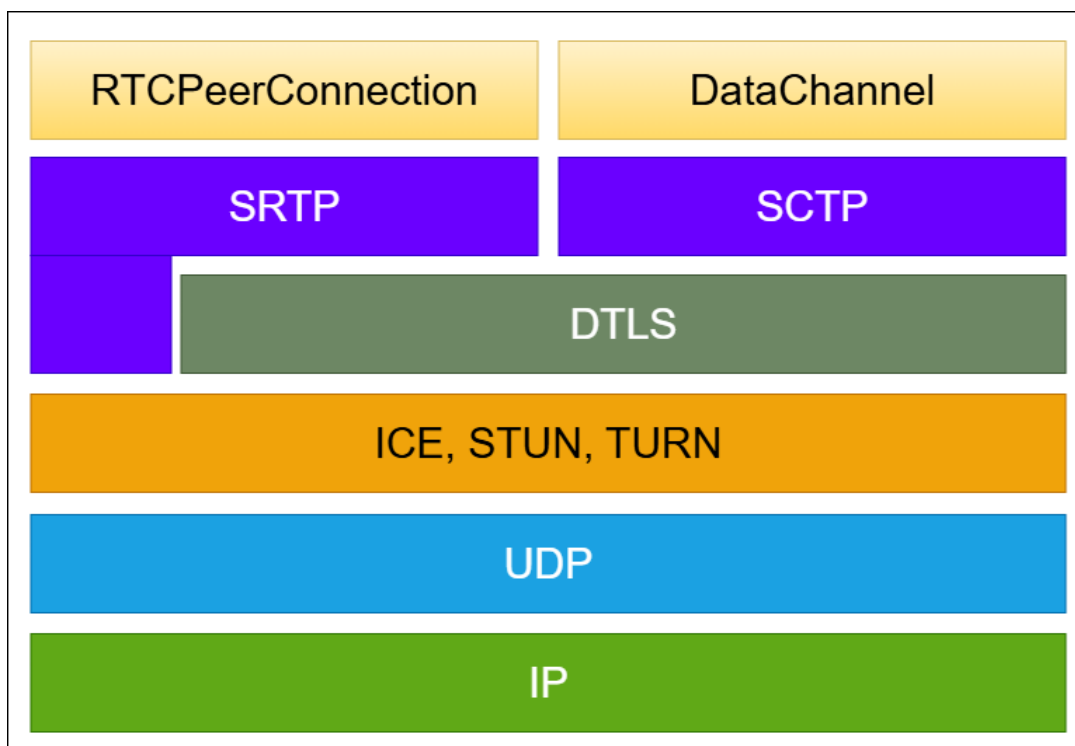


Figura 11: Pila de protocolos utilizada en WebRTC.

3.5 Componentes de seguridad en la red

Como se mencionó en el Capítulo 2, en los sistemas de telemedicina es necesario proteger la información que se transmite para que no sea accedida por actores maliciosos (p. ej. ciberdelincuentes). Por esta razón, en una arquitectura de red para la transmisión de audio cardiopulmonar es necesario incluir mecanismos de seguridad que protejan la información. Los principios fundamentales en la seguridad de la información son la confidencialidad, integridad y la alta disponibilidad (Alnahawi et al., 2024; Oppliger, 2023). Esto es conocido como la tríada CID.

- **Confidencialidad.** Es la protección de la información contra el acceso no autorizado. Esto ayuda a que únicamente las personas autorizadas tengan el acceso a la información y a la implementación de controles que eviten la divulgación accidental o intencionada de dicha información.
- **Integridad.** Se asegura que la información no sea alterada o modificada sin alguna autorización previa, esto para evitar una comprometer la confianza de los datos.
- **Disponibilidad.** Determina que la información o cualquier recurso sea accesible y utilizable en el momento que sea requerido por los usuarios autorizados.

Existen mecanismos de seguridad que se implementan para asegurar los principios fundamentales de la seguridad de la información. Dichos mecanismos operan bajo una arquitectura de seguridad multinivel que integra el cifrado de la información por medio de algoritmos criptográficos, controles de acceso mediante roles asignados, firmas digitales para validar la autenticidad, así como la integración de autenticación multifactor para la validación de la identidad de las personas. También, existen diversos protocolos de seguridad de capa de aplicación, entre los que se destacan el protocolo de Seguridad de la Capa de Transporte (TLS, de sus siglas en inglés) y el protocolo de seguridad de la capa de transporte de datagramas (DTLS, de sus siglas en inglés).

Mientras que TLS funciona sobre TCP, DTLS es una variante del protocolo TLS, adaptada para su funcionamiento con el protocolo de transporte UDP. Estos protocolos son de vital importancia en la actualidad y empleados por casi todas las aplicaciones en el mundo del Internet. De acuerdo con el apartado anterior, la tecnología WebRTC delega los aspectos de seguridad de los canales de datos al protocolo DTLS.

El funcionamiento del protocolo DTLS, al menos en su versión 1.3 (ver Figura 12), se puede segmentar en dos etapas: Inicialización (Handshake) y sesión (Record Protocol y Key derivation). En la primera etapa, se establece una comunicación segura mediante la cual los dispositivos acuerdan el algoritmo de cifrado de datos, la autenticación y la generación de las claves. Posteriormente, en la etapa de sesión, las partes negocian las suites de cifrado (Cipher Suites) que ayudan a definir los algoritmos para el cifrado de datos y la autenticación. A continuación, se lleva a cabo el intercambio de claves, en el cual se establecen las claves secretas que permiten el cifrado y el descifrado de los datos durante toda la sesión y cualquier participante puede finalizar la conversación. Una explicación a detalle se puede consultar en los trabajos de Oppliger (2023), Rescorla (2018) y Rescorla et al. (2022).

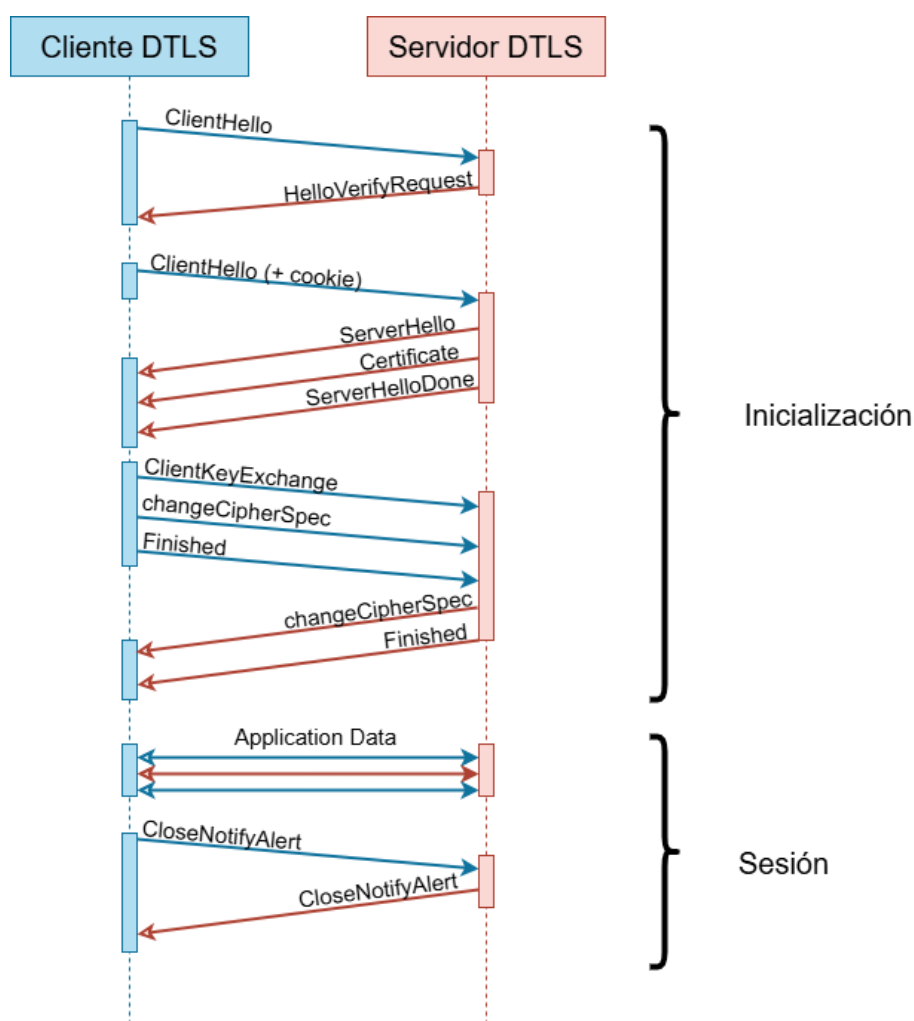


Figura 12: Diagrama de secuencia general del protocolo DTLS version 1.3.

Acorde a Paar y Pelzl (2010), una de las suites de cifrado más utilizadas en DTLS es AES en modo de encadenamiento de bloques de cifrado (CBC, por sus siglas en inglés). Este modo de operación proporciona un nivel robusto de seguridad al encadenar bloques de datos cifrados, donde cada bloque depende del

anterior. El proceso comienza con un Vector de Inicialización (IV) aleatorio de 16 bytes, que se combina mediante XOR con el primer bloque de texto plano antes del cifrado. Cada bloque subsiguiente utiliza el bloque cifrado anterior como su IV, creando así una dependencia que hace que cada mensaje cifrado sea único, incluso si el texto original es idéntico. Esta característica es fundamental para prevenir ataques de patrones y garantizar la confidencialidad de las comunicaciones en aplicaciones que requieren transmisión segura de datos sobre UDP (Figura 13).

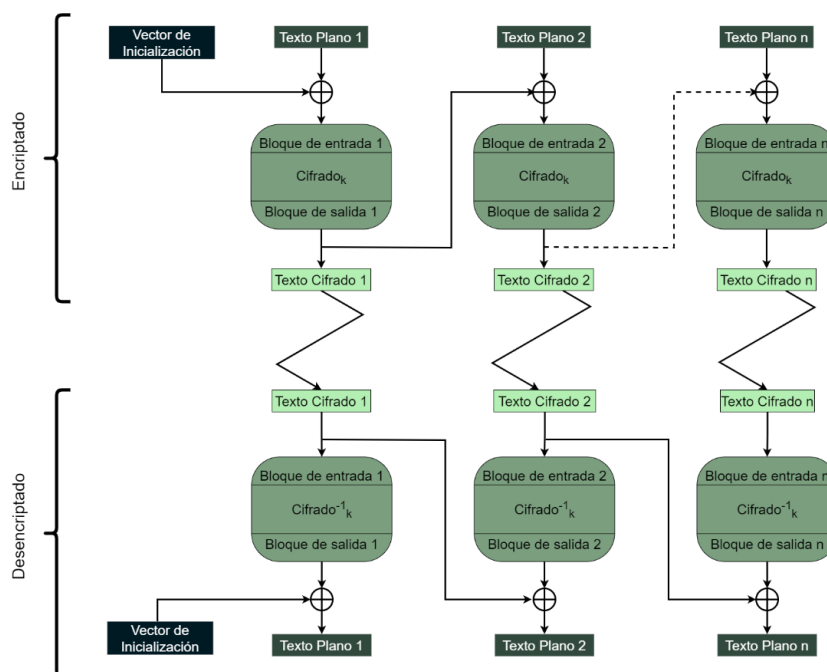


Figura 13: Diagrama que muestra el proceso de encriptación y descriptación con AES-CBC.

3.6 Conclusiones del capítulo

En este capítulo se estudiaron diferentes arquitecturas de redes de comunicaciones prevalentes en la Internet actual, desde la perspectiva de las dos configuraciones principales (cliente-servidor, P2P) y sus derivados. También se analizó la problemática que presenta el atravesar el NAT cuando se quiere establecer comunicación bajo el paradigma P2P. Derivado de esto, se estudiaron las técnicas de NAT transversal que podrían ser utilizadas para el diseño de la arquitectura para la transmisión de audio cardiopulmonar propuesta en el presente trabajo de investigación.

Adicionalmente, se analizó WebRTC, la cual es la arquitectura híbrida más utilizada actualmente con el fin de estudiar sus características. El capítulo terminó con una revisión de las técnicas de seguridad de la información utilizada en WebRTC.

Capítulo 4. Arquitectura de comunicaciones propuesta

El objetivo de este trabajo de investigación es el desarrollo de una arquitectura de comunicaciones para transmitir señales de audio cardiopulmonar comprimido en tiempo cuasi-real. El audio capturado se transmitirá de forma segura desde la ubicación donde se realice la auscultación (posiblemente zonas urbanas marginadas o zonas rurales debido a la deficiencia en la cantidad de médicos especialistas) hasta la ubicación del profesional de salud. Asimismo, se asume que la retroalimentación por parte del especialista se dará por medio de una videollamada o bien una sesión por un servicio de transmisión en línea, de tal manera el sistema desarrollado en este trabajo de investigación sea utilizado exclusivamente para el envío del audio cardiopulmonar (ver Figura 14).

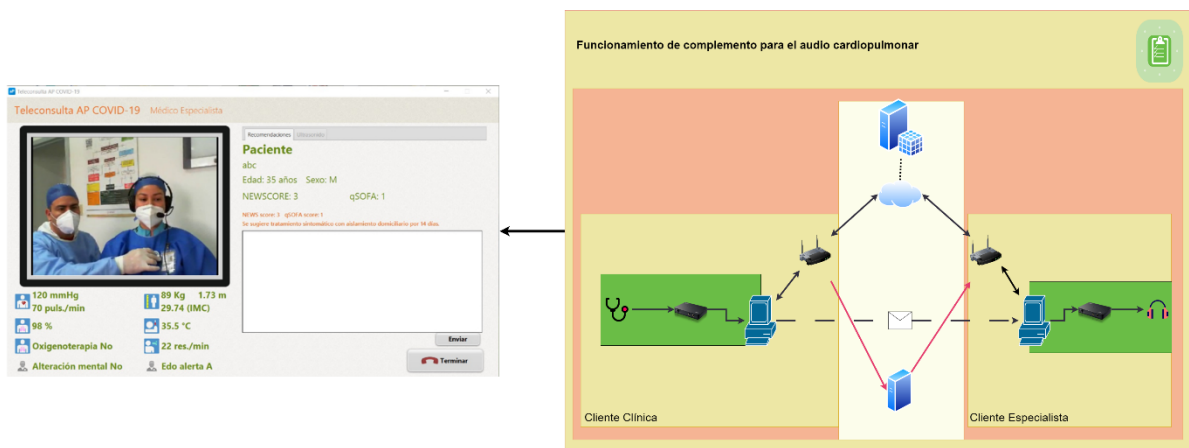


Figura 14: Integración de la arquitectura de envío de audio cardiopulmonar propuesto como módulo de un sistema de telemedicina.

Basados en la revisión del estado del arte presentada en los capítulos 2 y 3 del presente documento, en este trabajo de investigación se propone la arquitectura mostrada en la Figura 15 como la base para alcanzar los objetivos planteados en este trabajo de tesis. La arquitectura propuesta, llamada Sistema de Envío de Audio Cardiopulmonar (EAC), se puede segmentar en cuatro módulos cruciales:

1. Cliente clínica.
2. Módulo de red.
3. Cliente especialista.

4. Interfaz gráfica.

En el sistema EAC se considera que el flujo de la información sea unidireccional, desde el punto del origen del audio (i.e. cliente clínica) hasta llegar al destino del audio (i.e. cliente especialista). A partir de lo mencionado, el presente trabajo no indaga en el mecanismo de retroalimentación por parte del personal médico especialista, en cambio, se centrará en el diseño de una arquitectura que permita el establecimiento y flujo de audio cardiopulmonar entre clínica-especialista.

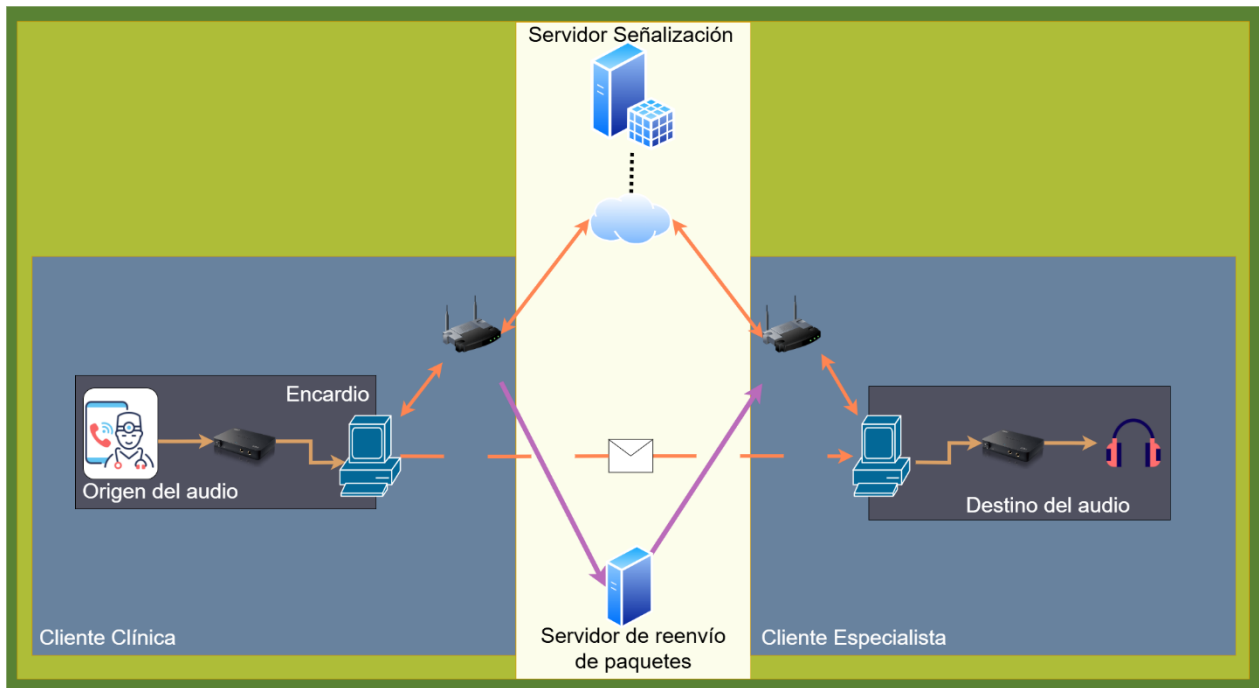


Figura 15: Diagrama de la arquitectura propuesta para la transmisión de audio cardiopulmonar a través de Internet.

Los módulos del sistema EAC están conformados a su vez por submódulos, los cuales se especifican a continuación.

El módulo de **cliente clínica** está compuesto por:

- Submódulo de captura de audio cardiopulmonar, codificación y empaquetado. Este submódulo hace uso del programa encardio, desarrollado dentro del Grupo de Investigación ARTS del CICESE como parte del trabajo de investigación reportado en (Olachea, 2019).
- Submódulo de aplicación de escritorio. Este submódulo se encarga de realizar una conexión entre el programa de encardio y el módulo de red. Se ocupa de la seguridad de la información, el re-

empaquetado del audio y su transmisión conforme al protocolo establecido.

El **módulo de red** se encarga del transporte del audio a través de Internet hasta la aplicación del cliente especialista. Este módulo está compuesto por un servidor que se encarga de:

- Implementar el mecanismo de señalización definido en la arquitectura diseñada.
- Participa en el proceso de resolución de direcciones IP y puertos.
- Cumplir con funciones de intermediario (*relay*) en el proceso de comunicación P2P.

El módulo de **cliente especialista** está conformado por:

- Submódulo de aplicación de escritorio, el cual se encarga de la conexión entre el módulo de red y el programa decardio desarrollado en (Olachea, 2019). El submódulo de aplicación de escritorio recibe los paquetes de audio provenientes del cliente clínica, retira la seguridad del paquete para desempaquetarlo y reenvía la información al programa decardio.
- Submódulo de desempaquetado, decodificación y reproducción del audio usando la aplicación decardio.

En el módulo de interfaz gráfica, como su nombre lo indica, se proporciona al usuario una interfaz para la gestión de llamada entre la clínica y el médico especialista. Para efectos de este trabajo de investigación, únicamente se contará con una pantalla que actualizará los estados de conexión acorde a la etapa de la llamada. Adicionalmente, este módulo cuenta con un contador para conocer la duración de la llamada.

A continuación, se realiza una descripción con mayor detalle acerca de los módulos anteriormente mencionados.

4.1 Módulo cliente clínica

Para la construcción del módulo cliente clínica es necesario tener en consideración diversos aspectos relacionados con la captura del audio cardiaco y pulmonar empleando un estetoscopio electrónico o

digital, por ejemplo: códec utilizado, ancho de banda requerido, y seguridad del flujo de información. En este módulo se usa la implementación del programa de encardio para la captura y conversión de la señal cardiopulmonar en una señal digital codificada. La salida de este módulo posteriormente se conecta con la aplicación de escritorio del módulo cliente clínica. Debido a que el programa encardio es una parte importante de esta implementación, en la siguiente sección se describe cómo funciona y las partes que la componen.

4.1.1 Submódulo de captura de audio cardiopulmonar, codificación y empaquetado

El submódulo de captura de audio cardiopulmonar, codificación y empaquetado se compone esencialmente de los bloques de estetoscopio, tarjeta de audio, y programa encardio, tal cual se muestra en la Figura 16. El estetoscopio y la tarjeta de audio son necesarios para realizar la auscultación y digitalizar el audio cardiopulmonar. El flujo de audio digitalizado es después alimentado al programa encardio desarrollado en (Olachea, 2019) para obtener los paquetes de datos que serán alimentados al submódulo de aplicación de escritorio. Como se comentó en la Sección 2.3, el programa encardio fue desarrollado dentro del grupo de trabajo como parte del trabajo de tesis (Olachea, 2019). Este programa se basa en el paradigma de transmisión de radio en línea para la transmisión y recepción de audio cardiopulmonar con calidad diagnóstica.

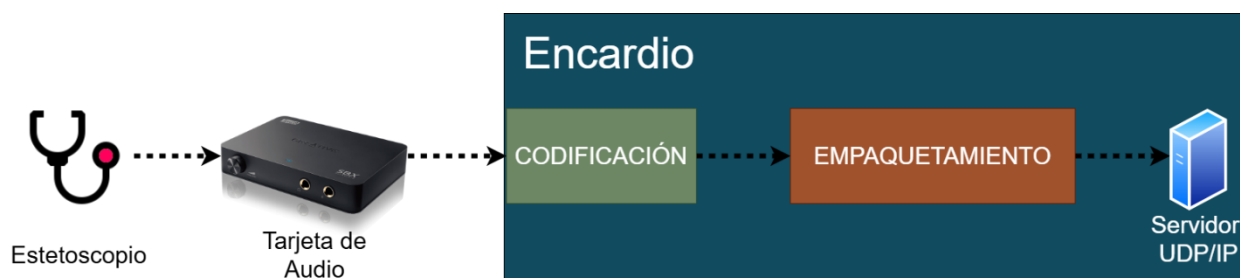


Figura 16: Diagrama de bloques del submódulo de captura de audio cardiopulmonar, codificación y empaquetado.

En las siguientes secciones se hace una descripción de los bloques del submódulo de captura de audio cardiopulmonar, codificación y empaquetado mostrados en la Figura 16.

4.1.1.1 Estetoscopio

En este trabajo de investigación, se optó por utilizar un estetoscopio electrónico (Clinical E-Scope®

stethoscope) y un estetoscopio digital (Thinklabs One), ambos de nivel 9-10. Estos dispositivos fueron seleccionados debido a sus características de filtro para escuchar los sonidos cardíacos o pulmonares, según sea el caso. Ambos instrumentos han sido previamente validados en los trabajos de Castorena (2012), Olachea (2019) y Ramírez (2017), donde se analizaron de manera detallada sus características. Las especificaciones principales de los filtros utilizados en estos dispositivos son las siguientes:

- **Estetoscopio digital – Thinklabs One.** (Thinklabs Medical LLC, 2014)
 1. **30 Hz – 500 Hz.** Con este filtro se amplifican las frecuencias bajas para los sonidos cardíacos, en especial para el tercer sonido cardíaco (S3). Es una versión amplificada del modo campana de un estetoscopio tradicional.
 2. **60 Hz – 500 Hz.** Este filtro está diseñado para la auscultación cardíaca, es una alternativa al primer filtro, esto cuando las frecuencias son demasiado intensas al momento de su uso.
 3. **80 Hz – 500 Hz.** Este filtro se utiliza para la evaluación de sonidos respiratorios y eventos valvulares cardíacos, particularmente para distinguir los componentes del segundo ruido cardíaco (S2). Al eliminar las frecuencias más bajas del espectro cardíaco, emula la función del diafragma en estetoscopios convencionales.
 4. **100 Hz – 1000 Hz.** Este filtro se especializa en la auscultación pulmonar, debido a que elimina de manera efectiva las frecuencias bajas y las vibraciones, además de extender el límite superior hasta 1000 Hz. Al proporcionar una mayor definición en los sonidos pulmonares, esto incrementa de manera moderada la captación del ruido ambiental.
 5. **20 Hz – 2000 Hz.** Este filtro constituye un modo amplio de alta sensibilidad diseñado para facilitar la grabación del audio captado por el estetoscopio.
- **Estetoscopio electrónico – Clinical E-Scope stethoscope.** (Cardionics, 2012)
 1. **20 Hz – 650 Hz.** (Sonidos cardíacos). Este filtro está diseñado para realizar la auscultación cardíaca, debido a que está calibrado para detectar de manera precisa los sonidos cardíacos fundamentales (S1, S2, S3, S4) y sus variaciones patológicas.

2. **70 Hz – 2000 Hz.** (Sonidos pulmonares). Al contar con un límite de hasta 2000 Hz, este filtro se utiliza para realizar auscultaciones pulmonares.

4.1.1.2 Tarjeta de audio

Acorde a Gunawan y Khalifa (2010), una tarjeta de audio es un componente de hardware que permite a las computadoras captar, reproducir y manipular sonidos. Se utiliza principalmente como un dispositivo de entrada y salida de audio, facilitando la grabación, síntesis y reproducción de sonidos, como música y voz. En este contexto, las tarjetas de audio externas adquieren especial relevancia debido a que generalmente ofrecen mejor calidad de sonido y menor latencia en comparación con las tarjetas de sonido integradas en las computadoras (Gunawan y Khalifa, 2010). Esta superioridad se debe a su diseño optimizado para grabación y reproducción de audio en tiempo real, proporcionando una fidelidad de estudio musical. Asimismo, suelen tener un nivel de ruido más bajo y menos distorsión armónica total (THD), lo que asegura una experiencia de sonido más clara y precisa. Por estas razones, son preferidas para la grabación profesional y la producción musical.

Con lo comentado anteriormente, en este trabajo de investigación se optó por emplear tarjetas de audio externas de la familia sound blaster, en específico, los modelos “Sound Blaster Play! 3” y “Sound BlasterX G1”. Esta elección se basa en las características técnicas de las dos tarjetas, que cumplen con los requisitos necesarios para el procesamiento de audio en este estudio. Las especificaciones detalladas se presentan en la Tabla 8.

Tabla 8: Características de las tarjetas de audio “Sound Blaster Play! 3” y “Sound BlasterX G1”.

Característica	Relación señal a ruido	Impedancia de audífonos soportada	Frecuencia de muestreo (Reproducción)	Frecuencia de muestreo (Grabación/Mic- in)
Sound Blaster Play! 3	~93 dB	Hasta 300 Ω	16-bit: 44.1, 48.0, 88.2, 96.0 kHz 24-bit: 44.1, 48.0, 96.0 kHz	16/24-bit: 44.1, 48.0kHz
Sound BlasterX G1	~93 dB	16 - 300 Ω	24-bit: 44.1, 48.0, 88.2, 96.0 kHz	24-bit: 48.0kHz

4.1.1.3 Encardio – codificación de audio

El programa encardio (Olachea, 2019), permite configurar una rutina de captura de audio en la que, después de adquirir un determinado número de muestras PCM, se inicia el proceso de compresión mediante una API de codificación. Esta API funciona como un intermediario entre el programa encardio y un conjunto de bibliotecas de códecs de audio. Una vez que se procesan las muestras, se entrega un bloque de información comprimida conocido como gránulo en el contexto de este trabajo.

La cantidad de muestras necesarias para formar un gránulo depende del códec empleado, por lo que es necesario definirlo antes de que comience el proceso de codificación por parte del programa encardio. La Tabla 4 presenta las configuraciones utilizadas para la codificación de audio en formatos MP3 y AAC. No obstante, para el desarrollo de este proyecto, solo se consideró el formato MP3.

La elección del códec MP3 se fundamenta en los estudios realizados (Castorena, 2012), quienes demostraron que las características de este códec son adecuadas para el procesamiento de audio cardiopulmonar. Esto se debe a que las frecuencias asociadas a los sonidos cardiopulmonares se encuentran por debajo de los 2000 Hz y esta es una frecuencia soportada por el códec MP3. Además, su eficiencia fue validada en los trabajos (Olachea, 2019; Ramírez, 2017), mediante la prueba de múltiples estímulos con referencia escondida (MUSHRA), con la participación de personal del área de la salud, lo que permitió confirmar que el códec MP3 es uno de los más adecuados para garantizar una calidad diagnóstica del audio cardiopulmonar.

Es necesario aclarar que el proceso de codificación normalmente toma solo unos milisegundos de tiempo, lo que es suficiente para generar un gránulo antes de que la tarjeta de sonido le envíe el siguiente conjunto de muestras de audio para su procesamiento. Una vez que se generan los suficientes gránulos (en este caso 9), se organizan por medio de paquetes para transmitirlos al médico especialista.

4.1.1.4 Encardio – empaquetado de audio

En el formato de empaquetado de audio propuesto (Olachea, 2019), una vez creado un gránulo nuevo de audio comprimido, este se guarda en un búfer transitorio que hace uso del principio FIFO (del inglés, *First In, First Out*). El búfer tiene la función primordial de agrupar un número predeterminado de gránulos consecutivos, esto con el fin de transmitirlos simultáneamente como parte principal de la carga útil de un

paquete UDP/IP que se transmitirá de forma periódica. Para conformar la carga útil del paquete UDP/IP es necesario incluir una cabecera general y cabeceras seccionales que permitan distinguir el inicio de cada gránulo. Esto es necesario para permitir la correcta decodificación del audio en el cliente especialista. La Figura 17 muestra cómo queda conformada la carga útil del paquete que contiene la información de audio.

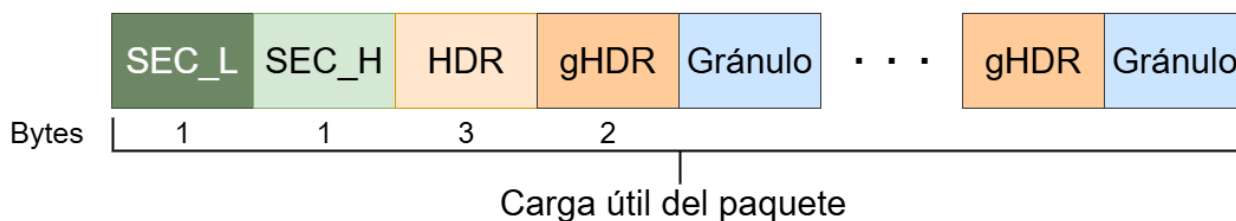


Figura 17: Formato de la carga útil para la transmisión de varios gránulos de audio codificado sobre UDP/IP.

La cantidad total de gránulos por paquete depende del tipo de codificador empleado. En el caso de MP3, con una resolución monoaural a 16 bits y una frecuencia de muestreo a 16 kHz se consideran 9 gránulos de audio por paquete (Ver Tabla 4).

4.1.1.5 Encardio – servidor UDP/IP

Para lograr la transmisión de audio cardiopulmonar, el trabajo de Olachea (2019), se basó en un paradigma de radio en línea, donde los clientes que busquen escuchar los sonidos capturados deben de conectarse a un servidor creado mediante un socket de comunicación. En este caso, el programa encardio crea un socket de comunicación UDP, al cual el cliente que tenga la aplicación decardio puede conectarse para obtener los paquetes de audio y posteriormente decodificarlos para su escucha. Para mantener la comunicación de forma segura, se empleó una red privada virtual (VPN, por sus siglas en inglés). En la Figura 18, se muestra la comunicación par-a-par empleando el servicio de VPN entre los clientes encardio y decardio.

Sin embargo, como se mencionó con anterioridad, el uso de una VPN supone dificultades para su despliegue en los entornos de conectividad actuales. Por estas razones, este trabajo de investigación se enfoca en cambiar el paradigma de conectividad para la transmisión de audio cardiopulmonar de la necesidad de utilizar una VPN, hacia utilizar un paradigma híbrido P2P/cliente-servidor. Para lograr esto, se optó por establecer una conexión local entre el programa de encardio y la aplicación de escritorio del módulo cliente clínica desarrollada en este trabajo de investigación, lo cual es descrito en el próximo apartado.

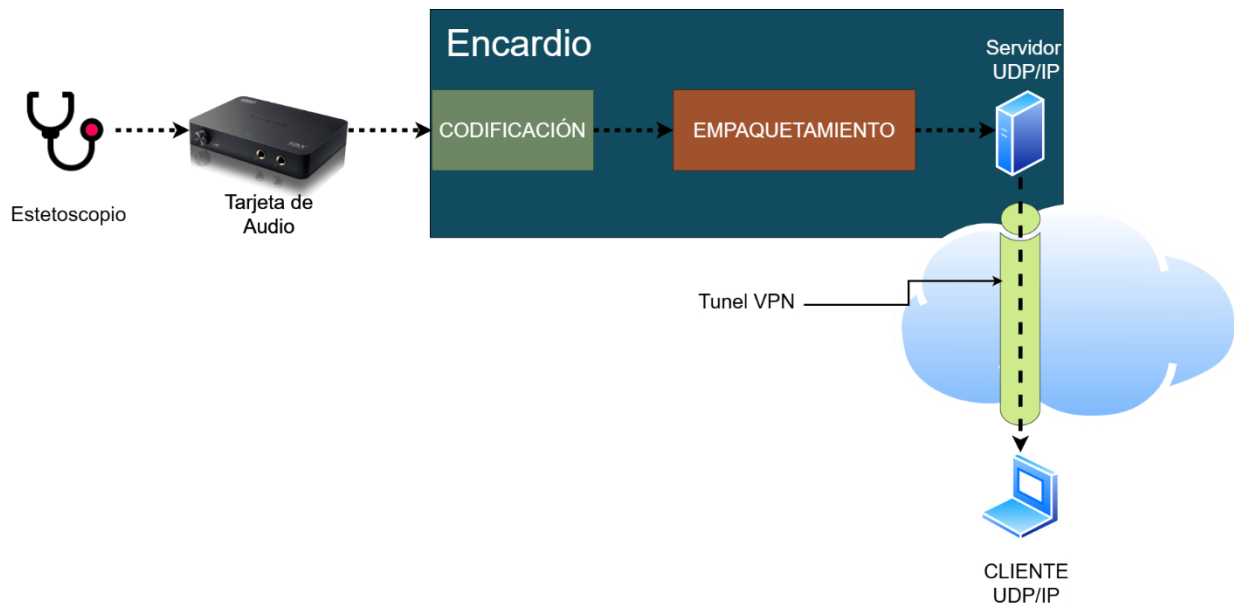


Figura 18: Diagrama a bloques del esquema desarrollado en (Olachea, 2019) para transmitir audio cardiopulmonar.

4.1.2 Submódulo de aplicación de escritorio

Para entender el proceso de conexión entre los programas de encardio y el submódulo de aplicación de escritorio del módulo cliente clínica es fundamental entender que es un socket. Un socket es una interfaz que permite la comunicación entre dos procesos, en el mismo dispositivo o en diferentes dispositivos a través de la red (Tanenbaum y Wetherall, 2012). Como se observa en la Figura 18, la aplicación de encardio crea un socket de comunicación empleando el protocolo UDP para iniciar la comunicación con su contraparte, el programa decardio. Para lograr la conexión entre las aplicaciones, se decidió usar el socket UDP con el que ya contaba la aplicación de encardio. Este se configuró para comunicarse con el submódulo de aplicación de escritorio mediante una dirección de red local y un puerto de comunicación específico (local-loop: 10007).

En el submódulo de aplicación de escritorio, se crea y configura un socket con la dirección local y el puerto de comunicación (local-loop: 10 007) para recibir los paquetes de encardio. Además, se ajusta el socket para que no bloquee la ejecución del programa de encardio. En caso de que se produzca un error durante la creación del socket, la función registra el incidente en el log del sistema y devuelve un valor de False. A continuación, se muestra el código de la función `setup_socket`, responsable de esta tarea.

```
def setup_socket(self):
    try:
        self.socket = socket.socket(socket.AF_INET, socket.SOCK_DGRAM
```

```

self.socket.bind(('127.0.0.1', self.port))
self.socket.setblocking(False)
logging.info(f"Audio receiver listening on port {self.port}")
return True
except Exception as e:
    logging.error(f"Failed to setup audio receiver socket: {e}")
    return False

```

Una vez que el flujo de comunicación se establece completamente con el cliente especialista, se lleva a cabo un proceso de empaquetamiento de los datos provenientes de la aplicación encardio. Este proceso tiene como objetivo incorporar un mecanismo de seguridad adicional (ver Figura 19). Cabe destacar que los detalles relacionados con la resolución de red, necesarios para garantizar la correcta comunicación entre las aplicaciones, se abordarán en la Sección 4.2. El funcionamiento general del empaquetamiento varía dependiendo de si la comunicación se realiza en modo punto a punto (P2P) o bajo un esquema cliente-servidor.

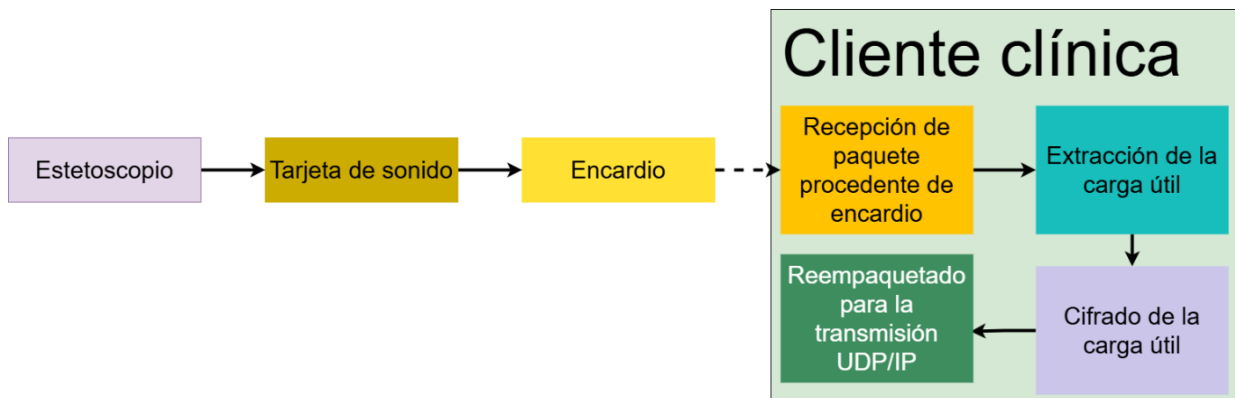


Figura 19: Diagrama a bloques del proceso de empaquetado de paquete UDP/IP por el módulo cliente clínica con destino al módulo cliente especialista.

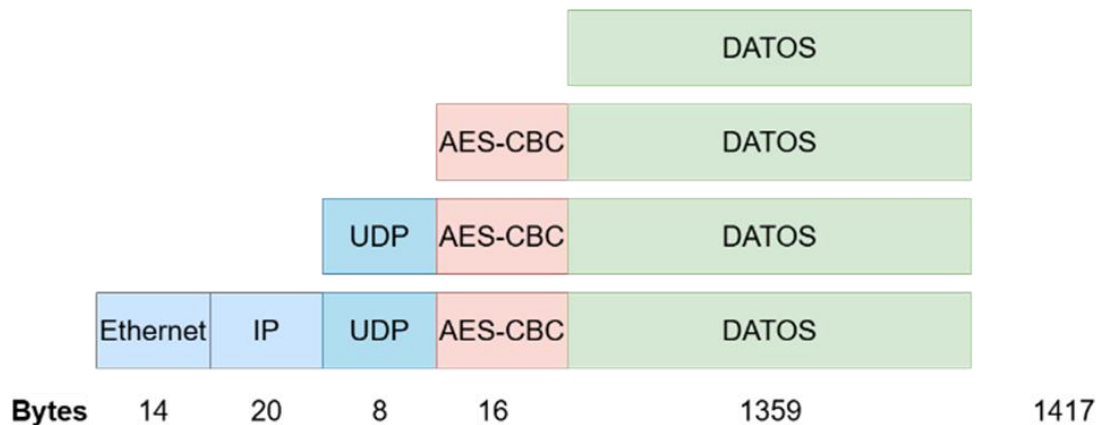


Figura 20: Formato de paquete UDP/IP por el módulo cliente clínica con destino hacia el módulo cliente especialista.

4.2 Módulo de red

En esta sección se lleva a cabo un análisis detallado de la arquitectura y del protocolo de señalización y comunicaciones propuestos en el presente trabajo de investigación. Asimismo, se explican los componentes del cliente y del servidor de señalización desarrollados.

4.2.1 Diseño de la arquitectura de señalización y comunicaciones

En el capítulo anterior, se abordó el tema de las diferentes arquitecturas de red, destacándose que la arquitectura P2P es la opción adecuada para el desarrollo de sistemas de transmisión de datos en tiempo cuasi-real. Esta arquitectura permite reducir el número de saltos de red que los paquetes deben realizar, esto comparado con la arquitectura cliente-servidor. No obstante, como se mencionó en el Capítulo 2, los sistemas comerciales que integran la auscultación a distancia suelen basarse en una arquitectura en la nube. Esto se debe principalmente a la limitada disponibilidad de direcciones IPv4 y a la lenta adopción de las tecnologías necesarias para implementar el direccionamiento IPv6, el cual resulta esencial para habilitar las comunicaciones en redes P2P.

En este sentido, la arquitectura que proporciona WebRTC representa un modelo alternativo para realizar transmisiones de punto a punto entre navegadores web en tiempo real. Esta tecnología consiste en un conjunto de protocolos de red diseñados para abordar los desafíos de las comunicaciones P2P en Internet y sirve como base para soluciones como Google Meet, Jitsi Meet, así como Zoom y Microsoft Teams en sus versiones para navegador.

En el contexto de esta investigación, la arquitectura propuesta debe ser adecuada para su implementación tanto en zonas rurales como en áreas urbano-marginales. Estas regiones presentan diversas limitaciones, que van desde la disponibilidad de personal médico hasta la infraestructura de red. Dado este escenario, no resulta viable implementar toda la lógica de los protocolos de WebRTC para la transmisión de los paquetes de datos (los cuales contienen el audio cardiopulmonar digitalizado, codificado y empaquetado) generados por el módulo de cliente clínica. Esto se debe a que el uso de WebRTC implicaría una sobrecarga innecesaria de tráfico en el sistema, lo que lo hace poco adecuado para este caso. Por esta razón, en el presente trabajo se propone la arquitectura ilustrada en la Figura 21.

Aunque está inspirada en el modelo de WebRTC, esta arquitectura ha sido específicamente diseñada para

el transporte paquetes de datos con carga útil de audio cardiopulmonar utilizando el menor ancho de banda posible.

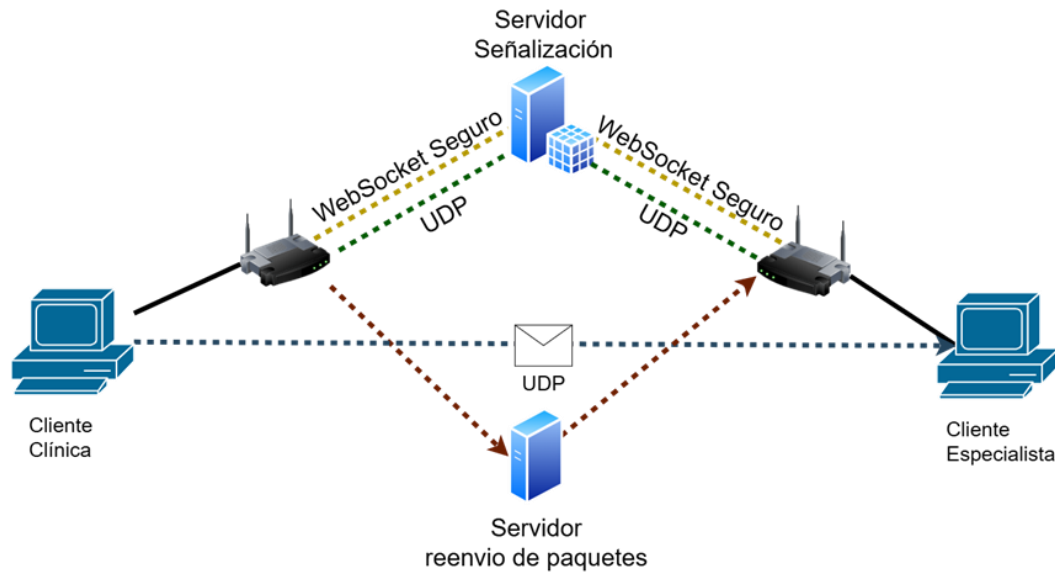


Figura 21: Arquitectura de comunicaciones propuesta en este trabajo de investigación.

La arquitectura propuesta está basada en un modelo P2P híbrido (Buford et al., 2009; Jaideep y Battula, 2016; Kanazawa y Takami, 2018), el cual incorpora un servidor de señalización que facilita la identificación y el descubrimiento de los nodos que conforman la red. Los clientes, al registrarse en el servicio, ya sea para transmitir datos (clínicas) o recibirlos (especialistas), establecen comunicación con el servidor a través de una conexión Websocket segura, siguiendo el estándar propuesto en WebRTC. Tanto los dispositivos de las clínicas como los de los especialistas realizan este procedimiento para llevar a cabo el proceso de señalización, que incluye acciones como el registro y el intercambio de claves privadas necesarias para cifrar la comunicación.

Con el objetivo de resolver los problemas asociados a la traducción de direcciones de red en el establecimiento de redes P2P, se propone el uso de las técnicas *UDP Hole Punching* como primera opción. En caso de que no se pueda establecer comunicación con *UDP Hole Punching*, entonces se utilizará el mecanismo de *Relaying* para comunicar a ambos clientes bajo el paradigma de comunicaciones cliente-servidor (Coulouris et al., 2011; Tanenbaum y Steen, 2006). Estos mecanismos se analizaron en el Capítulo 3.

El proceso de *UDP Hole Punching* utiliza un servidor de señalización que también desempeña el papel de “servidor de encuentro”. Este servidor permite realizar el intercambio seguro de paquetes UDP entre los

clientes durante este proceso. En caso de que sea posible establecer correctamente la conexión P2P entre los dispositivos, se generará un flujo de paquetes UDP en el sentido desde el cliente clínica hacia el cliente especialista. Para la transmisión se seleccionó el protocolo UDP, debido a que evita retransmisiones, lo cual a su vez minimiza la latencia y reduce la sobrecarga de encabezados comparado con el protocolo de control de transmisión (TCP).

En caso de que el proceso de conexión mediante UDP *Hole Punching* no sea exitoso, se recurre a la técnica de *Relaying*. En este escenario, el servidor actúa como intermediario, retransmitiendo los paquetes UDP generados por el cliente clínica hacia el cliente especialista, procurando minimizar la latencia.

Para garantizar la seguridad de la arquitectura, se propone el uso de encriptación AES-CBC, cuya clave privada se comparte de manera segura a través de la conexión Websocket segura previamente establecida. Este enfoque elimina la necesidad de implementar el protocolo completo de DTLS, reduciendo así la sobrecarga de datos asociada al proceso de inicialización (handshake). Los puertos seleccionados para la comunicación se encuentran dentro del rango de 1024 a 65535, ya que son los más utilizados por soluciones basadas en WebRTC (Uberti, 2012).

A continuación, se proporciona una descripción detallada de los bloques diseñados y desarrollados para la implementación de la arquitectura de señalización propuesta en este trabajo de tesis. Estos bloques son:

- Bloque de registro.
- Bloque de inicio de llamada.
- Bloque de fin de llamada.
- Bloque de mantenimiento de rutas.

Cada uno de ellos es fundamental para garantizar el funcionamiento eficiente y estructurado del sistema.

4.2.1.1 Bloque de registro

El diagrama de secuencia del bloque de registro se observa en la Figura 22, donde se muestra la

señalización inicial por parte de ambos clientes: el cliente clínica corresponde a la clínica situada en la zona rural / urbano-marginal y el cliente especialista corresponde al personal médico con especialidad.

Para la señalización, los clientes inician el proceso de *handshake* con la finalidad de establecer una conexión segura. Esto se realiza mediante el uso de un Websocket seguro (WSS en la Figura 22), para así enviar su usuario y contraseña de acceso. Con esto, el servidor de señalización puede comprobar si es correcta la información y actualizar el estado de cada cliente (estado, dirección de red y puerto). Al terminar este proceso, el servidor envía a los clientes un mensaje con la llave de cifrado usada por el servidor para poder enviar mensajes mediante el uso del canal UDP en modo cliente-servidor.

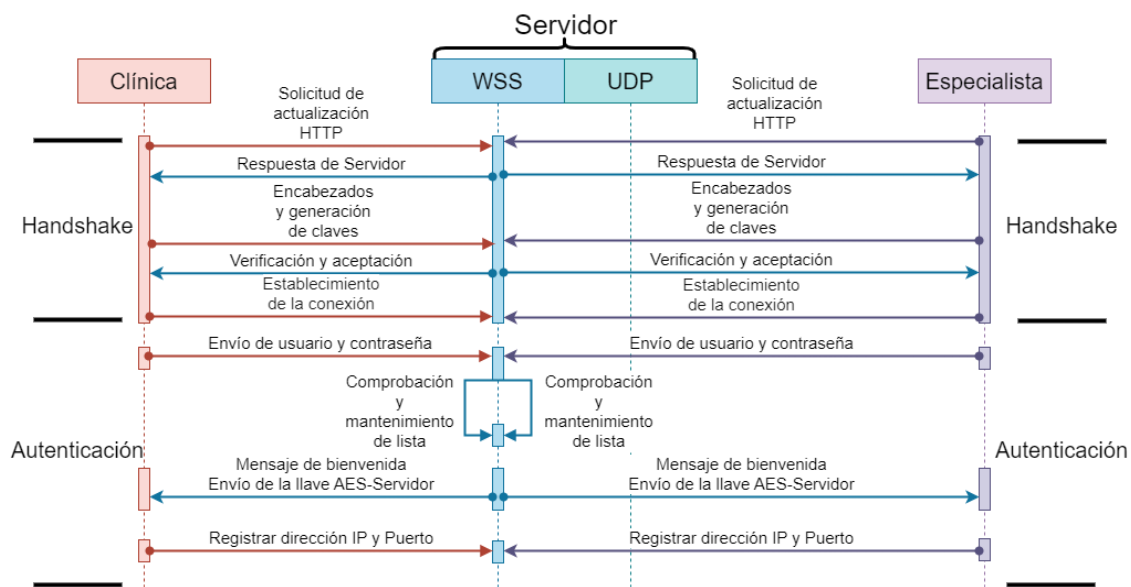


Figura 22: Diagrama de secuencia del protocolo de señalización inicial propuesto.

4.2.1.2 Bloque de inicio de llamada

El bloque de inicio de llamada se divide en dos partes. La primera parte corresponde a la solicitud de llamada y establecimiento de comunicación en modo P2P (ver Figura 23). Esta parte se ejecuta si el cliente clínica se puede contactar con el cliente especialista por medio de la implementación de la técnica de UDP *Hole Punching*. La segunda parte se ejecuta en el caso de que alguno de los dos clientes se encuentra detrás de un NAT y, por lo tanto, la técnica de *Hole Punching* no se complete de forma correcta (ver Figura 24). Note que la segunda parte se ejecuta para establecer la comunicación entre el cliente clínica y el cliente especialista a través del mecanismo de *Relaying* utilizando el paradigma de comunicación cliente-servidor.

En ambos casos, el inicio de la conexión es de la misma forma. El cliente clínica solicita la llamada (con un tiempo máximo de espera de 10 segundos) a través del servidor. En esta solicitud se realiza el envío de la clave de cifrado para la comunicación P2P. En caso de que el cliente especialista responda que quiere aceptar la llamada, el servidor de señalización intenta primero forzar a ambos clientes a abrir sus comunicaciones por medio del socket UDP, para lo cual recopila su información de dirección de red y el puerto de comunicación. Una vez que tenga esta información, el servidor de señalización reenvía la información a su contraparte para que después cada uno de los clientes envíe 1 mensaje “dummy” por segundo, con un límite máximo de 10 mensajes. Con esto se busca abrir un “agujero” en el NAT y de esta manera lograr una comunicación P2P a través de la técnica de *UDP Hole Punching*.

En caso de falla en recibir los mensajes “dummy” utilizados en la técnica *UDP Hole Punching*, el servidor manda un mensaje a los dos clientes para pasar a utilizar el mecanismo de *Relaying* para conectar a ambos clientes bajo el paradigma de comunicación cliente-servidor.

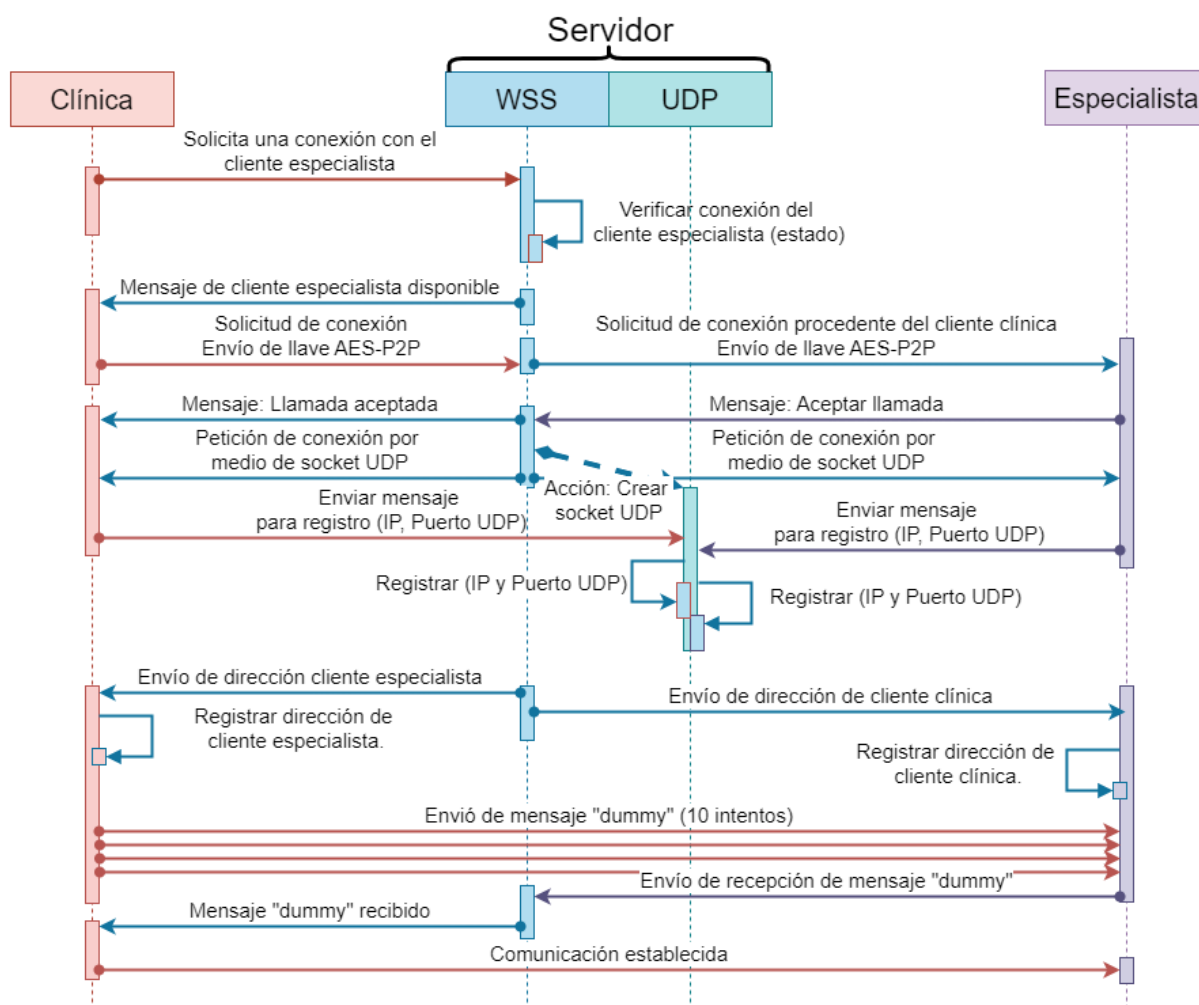


Figura 23: Diagrama de secuencia de la solicitud de llamada y establecimiento de comunicación en modo P2P.

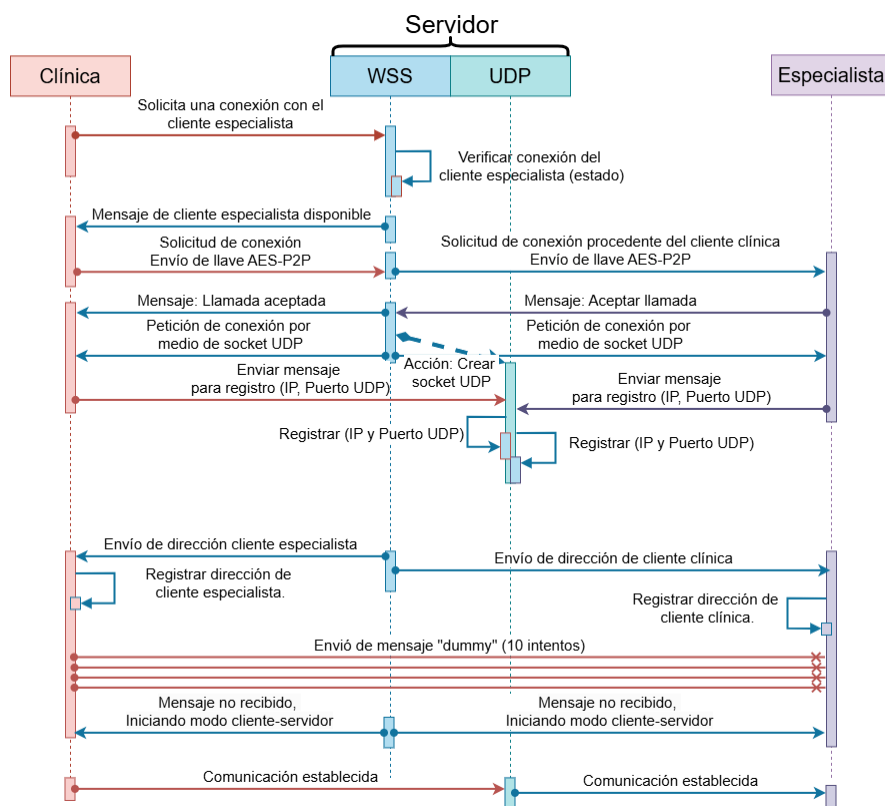


Figura 24: Diagrama de secuencia para establecer la comunicación con el mecanismo de *Relaying* bajo el paradigma de comunicación cliente-servidor.

4.2.1.3 Bloque de fin de llamada

Por medio del bloque de finalización de la llamada, cualquiera de los clientes involucrados puede finalizar la llamada. En la Figura 25, se observa que el cliente clínica manda un mensaje de finalización de llamada por medio del servidor, esto para evitar que se llegue a perder el mensaje. Una vez que el servidor manda el mensaje de finalización de llamada y este es recibido por el cliente especialista, se termina la comunicación.

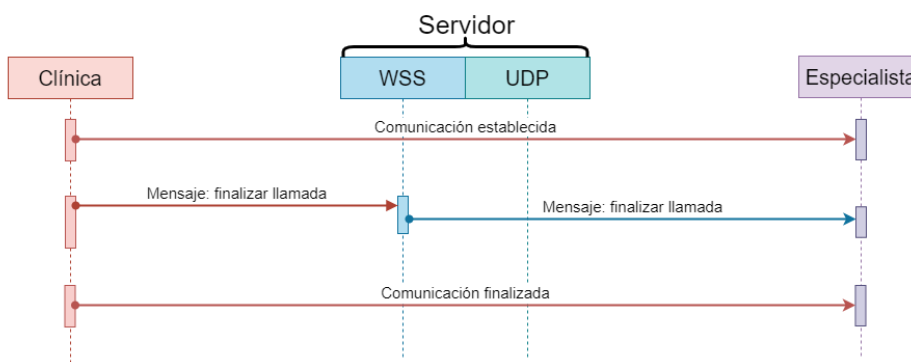


Figura 25: Diagrama de secuencia de la solicitud de finalización de llamada.

4.2.1.4 Bloque de mantenimiento de rutas

Para mantener la comunicación establecida mediante Websocket seguro y esta no se pierda por inactividad, se emplea el mecanismo denominado “Ping-Pong”. Este mecanismo consiste en el intercambio periódico de mensajes entre el cliente y el servidor para verificar que la conexión sigue activa. El servidor envía un mensaje de tipo "Ping" al cliente, y este responde con un mensaje "Pong" (Figura 26). Si el servidor no recibe una respuesta en un tiempo acordado, el servidor asume que la conexión se ha perdido y procederá a finalizarla. En el caso de la aplicación desarrollada, el tiempo límite es de 10 segundos (Price y Tino, 2009).

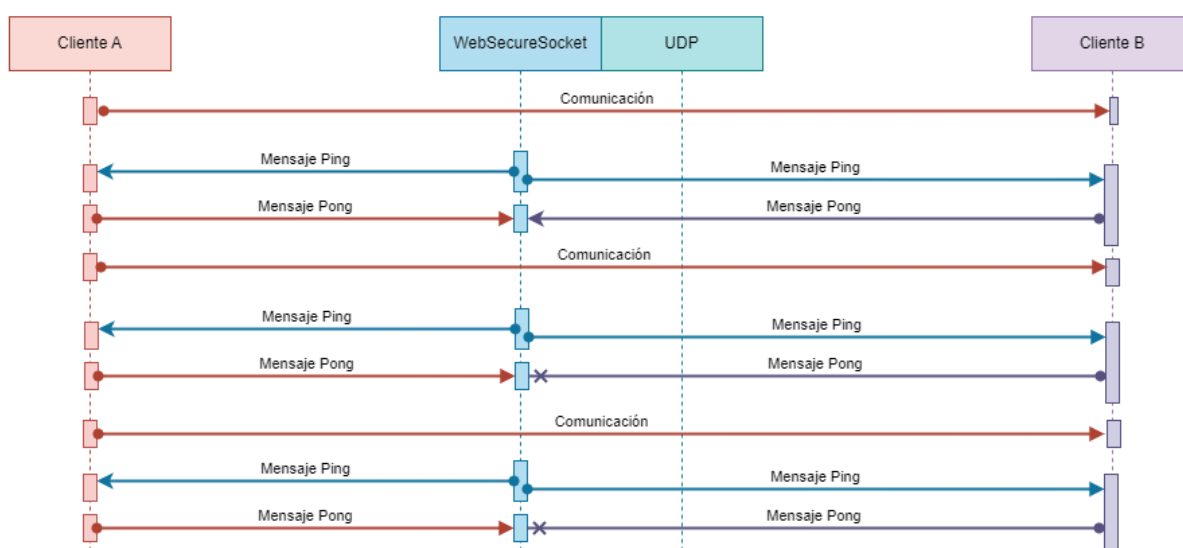


Figura 26: Diagrama de secuencia del bloque de mantenimiento de rutas, el cual emplea el mecanismo Ping-Pong.

4.3 Módulo de cliente especialista

El módulo cliente especialista consiste en 2 submódulos: aplicación de escritorio y programa decardio. Al igual que encardio, el programa decardio fue desarrollado como parte del trabajo de investigación reportado en Olachea (2019). El primer submódulo se encarga de recibir los paquetes provenientes del cliente clínica, extraer la carga útil del paquete UDP/IP, descifrar la carga útil para poder ser procesada y retransmitir la carga útil UDP/IP descriptada hacia el programa decardio. Por su parte, el segundo submódulo realiza las funciones de desempaquetamiento, decodificación y reproducción del audio cardiopulmonar.

Como en el caso del cliente clínica, se siguió una estructura modular para el desarrollo de los submódulos del cliente especialista, tal cual se muestra en la Figura 27. Cada submódulo se compone de diferentes bloques que se explican a continuación.

4.3.1 Submódulo de aplicación de escritorio

La aplicación de escritorio del cliente especialista crea y configura un socket UDP no bloqueante para la comunicación continua con el cliente clínica, ya sea en modo P2P o cliente-servidor (descritos en la Sección 4.2.1). Esto permite que el programa se ejecute sin interrupciones mientras se reciben los paquetes del cliente clínica.

Una vez establecido el flujo de comunicación con el cliente clínica, hay un bloque que permite extraer la carga útil del paquete UDP/IP enviado por el cliente clínica (bloque lila en la Figura 27). Después, la carga útil se alimenta al bloque de descifrado, el cual se encarga de descifrar la carga útil cifrada con el algoritmo AES-CBC de 128 bits. La llave para poder descifrar la información se proporciona a través del protocolo de señalización previamente explicado en la Sección 4.2.

Una vez descifrada la carga útil UDP/IP, el bloque de reenvío de carga útil (bloque verde en la Figura 27) se encarga de mandar los datos hacia el programa decardio. Para esto se abre un socket UDP con dirección de red local del dispositivo (local-loop) y un puerto de comunicación específico (10 007). El programa decardio es configurado para que espere los datos en esta combinación de dirección local/puerto para procesarlos y reproducir el audio cardiopulmonar en el dispositivo que hospeda al módulo cliente especialista.

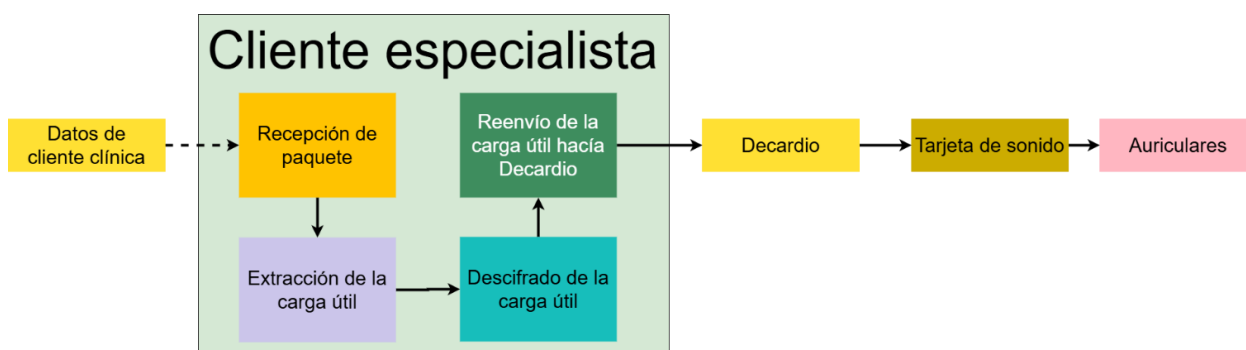


Figura 27: Diagrama a bloques del proceso del submódulo del cliente especialista para su reenvío hacia la aplicación decardio.

A continuación, se da una descripción de los bloques que conforman al programa decardio y cómo es su funcionamiento.

4.3.2 Submódulo de desempaquetamiento, decodificación, y reproducción del audio

El submódulo de desempaquetamiento, decodificación y reproducción de audio se compone de manera esencial de los bloques: programa decardio, tarjeta de audio y auriculares. El programa decardio realiza el desempaquetamiento, decodificación de los datos descifrados proporcionados por el submódulo de aplicación de escritorio.

En las siguientes subsecciones se hace una descripción de los bloques del submódulo de desempaquetamiento, decodificación y reproducción del audio, mostrados en la Figura 28.

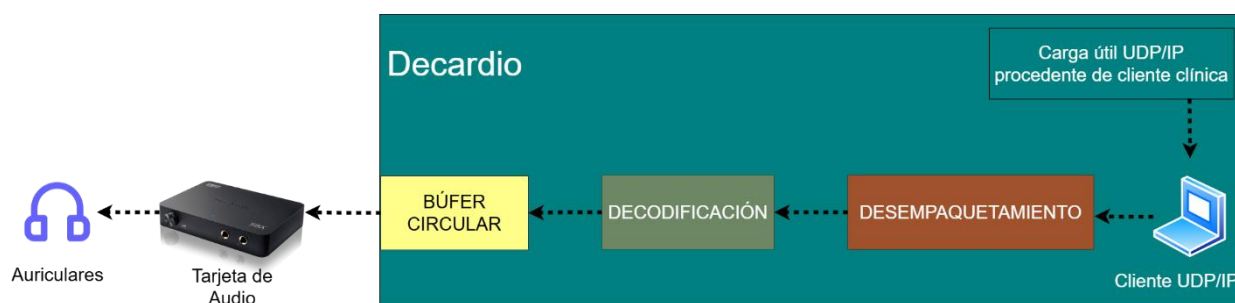


Figura 28: Diagrama de bloques del sistema decardio.

4.3.2.1 Decardio – Cliente UDP/IP

Para realizar la recepción de la recepción de audio cardiopulmonar contenido en la carga útil descifrada del paquete UDP/IP procesado por la aplicación de escritorio, el programa decardio espera la recepción de dicha carga útil a través de paquetes UDP que son enviados a una dirección local y puerto previamente configurados. Es decir, la aplicación decardio utiliza el socket de comunicación UDP abierto por la aplicación de escritorio para recibir los paquetes de audio y después decodificarlos para su escucha (ver Figura 18).

4.3.2.2 Decardio – Desempaquetamiento

En el formato de desempaquetado de audio propuesto por (Olachea, 2019), se tiene que al recibir el primer paquete de datos en el programa decardio se comienza el proceso de preparación de las rutinas encargadas de la recuperación y decodificación del audio. Del encabezado general de la información contenida en la carga útil se extraen parámetros clave, como la información del códec, la tasa de transmisión, la frecuencia de muestreo y el número de secuencia de los paquetes. Esta información se emplea para la inicialización de la API de captura/reproducción de audio, donde utiliza la frecuencia de muestreo en el encabezado del paquete recibido. Cabe recordar que cada paquete de carga útil contiene varios gránulos de audio cardíaco codificado separados por un encabezado (ver Figura 20 en la Sección 4.1.1.4).

4.3.2.3 Decardio – Decodificación y búfer circular

La decodificación de los gránulos de audio recibidos en la carga útil se realiza después de haber inicializado la biblioteca correspondiente al códec utilizado (LAME para formato MP3 y FDK-AAC para formato AAC). Cuando se detecta en la carga útil que el número de secuencia no corresponde al esperado, dicha carga útil no se procesa. En su lugar, el sistema cuantifica los gránulos contenidos en el paquete e integra en el búfer de reproducción un vector de silencio con una duración equivalente al número de gránulos. La reproducción del audio se produce de forma asíncrona con la recepción de paquetes de datos; de esta forma se pueden almacenar los paquetes de audio y reducir los efectos que el jitter pueda ocasionar en la reproducción del sonido. Asimismo, se estableció el tamaño del búfer circular de forma experimental.

4.3.2.4 Salida de audio

Normalmente, se podría pensar que cualquier par de audífonos o bien auriculares bastan para poder escuchar el audio cardiopulmonar en calidad diagnóstica, pero esto no es así. Los audífonos y auriculares convencionales no están diseñados para reproducir con precisión los sonidos que son de interés clínico, por lo que el uso de equipos diseñados específicamente para el diagnóstico y monitoreo de sonidos biológicos es esencial, sobre todo para las frecuencias que albergan los sonidos cardiopulmonares.

Por ello, se decidió emplear los auriculares Thinklabs Earbud Headphones - Deep Bass, los cuales fueron

específicamente diseñados para realizar auscultaciones (Thinklabs Earbud Headphones - Deep Bass, s.f.).

4.4 Módulo de interfaz gráfica

El diseño y las aplicaciones correspondientes a los clientes clínica y especialista como el servidor de señalización fueron desarrollados en el entorno de Visual Studio Code, empleando el lenguaje de programación Python. Para la interfaz gráfica se utilizó la biblioteca PyQt5, ya que facilita la conexión entre Python y Qt. Qt es un conjunto de herramientas multiplataforma orientado a objetos, ampliamente utilizado para desarrollar programas que emplean interfaces gráficas de usuario (Qt Group, s/f).

Para hacer más efectiva la interacción entre las clínicas y los especialistas, se tomó la decisión de realizar un diseño amigable para los clientes (clínica y especialista). Como se observa en la Figura 29-A, el cliente clínica es el único que puede comenzar la llamada. Esto se decidió así por razones de seguridad referentes con las normas HIPAA. En su contraparte, en la Figura 29-B, se muestra el diseño del cliente especialista, donde se puede observar que puede aceptar o rechazar la petición de llamada.

Ambos clientes pueden finalizar la llamada en el momento que decidan, además cuentan con un reloj que inicia al mismo tiempo que la llamada y una etiqueta que indica la etapa de la llamada.

El desarrollo del módulo de interfaz gráfica completa el desarrollo del sistema propuesto, ya que permite la interacción directa de los usuarios con este y facilitando su validación en entornos reales, en colaboración con el personal de la salud. Este módulo no solo asegura que las funcionalidades del sistema sean accesibles e intuitivas, sino que también permite recopilar retroalimentación valiosa del personal médico, lo que contribuye a identificar posibles mejoras y a mejorar su desempeño.

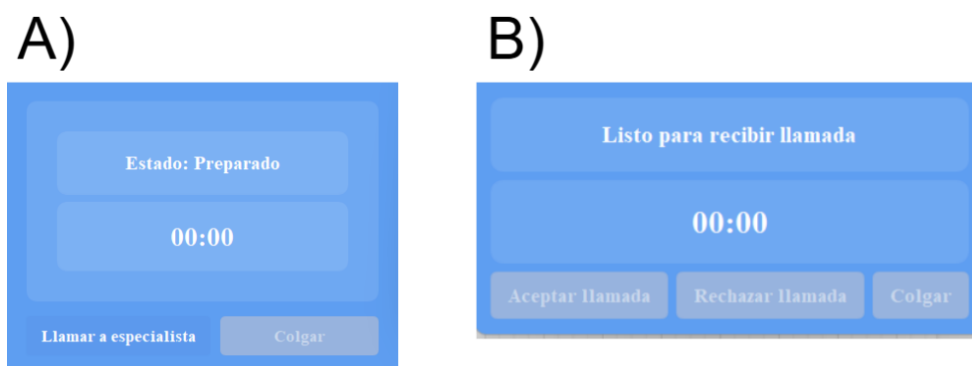


Figura 29: Diseño de la interfaz gráfica. A) Cliente clínica; B) Cliente especialista

Capítulo 5. Implementación de la cama de pruebas y resultados experimentales

En este capítulo se presentan los resultados de una serie de experimentos realizados para evaluar el comportamiento de la arquitectura diseñada en cuanto al protocolo de señalización, el consumo de datos y las principales métricas de red en 3 escenarios diferentes. Además, se realiza una comparativa de las plataformas convencionales y el sistema propuesto (sistema de Envío de Audio Cardiopulmonar (EAC)) por medio de un cuestionario subjetivo inspirado en la prueba de múltiples estímulos con referencia escondida (MUSHRA, de sus siglas en inglés), con ayuda de personal de la salud en el estado de Durango, México.

5.1 Descripción de la cama de pruebas

Para lograr una correcta implementación de las pruebas, se realizó la instalación del servidor de señalización desarrollado en una computadora modelo Dell XPS con un sistema operativo Ubuntu en su versión 20.04.6 LTS. La computadora cuenta con un procesador Intel Core-i7 y 8 GB de memoria RAM. La aplicación del cliente clínica se ejecutó en una laptop de la marca MSI, con un sistema operativo Windows 10 Enterprise en su versión 22H2. Este dispositivo cuenta con un procesador Intel Core-i7 y 64 GB de memoria RAM. La aplicación de cliente especialista se instaló en una laptop de la marca Huawei con un procesador Intel Core-i7 y 16 GB de memoria RAM y con Windows 11 Home en su versión 23H2 como su sistema base.

En la Figura 30, se muestran los escenarios propuestos para evaluar la arquitectura. Los 3 escenarios cuentan con un servidor de señalización, que obtuvo una dirección de red pública, esto para aplicar de manera correcta la técnica de *UDP Hole Punching*. En estas figuras el candado cerrado significa que los clientes se encuentran detrás de un NAT.

En la Figura 30-A se muestra el primer escenario de evaluación, donde los clientes clínica y especialista se encuentran en una red abierta. Es decir, ninguno de los clientes se encuentra detrás de un NAT o bien cuenta con una dirección de red IP pública. El cliente clínica fue conectado a la red WiFi llamada “CICESE” desplegada a nivel campus en el Centro de Investigación y de Educación Superior de Ensenada, Baja California (CICESE). Por su parte, el cliente especialista se conectó a la red llamada “AirporBI” localizada en las instalaciones del laboratorio de redes inalámbricas y sistemas embebidos del CICESE. Ambos clientes

cuentan con una dirección IP pública y están detrás del mismo firewall de red.

En el segundo escenario (Figura 30-B) el cliente clínica se mantiene con la misma conectividad que en el escenario 1, mientras que el cliente especialista se conecta a la Internet a través de una red privada. Es decir, el cliente especialista se encuentra detrás de un NAT. Para la implementación del servicio de conectividad se utilizó un dispositivo de banda ancha móvil como medio de acceso a Internet. Específicamente, se seleccionó el equipo modelo E5577 del fabricante HUAWEI, cuyas características técnicas incluyen compatibilidad con redes wifi de 2.4 GHz y 5 GHz. Para los propósitos de las pruebas realizadas, se empleó una tarjeta SIM operada por la Comisión Federal de Electricidad (CFE).

En el tercer escenario (Figura 30-C) se mantuvo la conectividad del cliente especialista del escenario 2, mientras que el cliente clínica se colocó en una red privada detrás de un NAT. Para ello, la conexión del cliente clínica fue realizada en una casa-habitación que cuenta con un proveedor de servicios de Internet. El proveedor de servicios para este escenario fue TotalPlay, el cual utiliza un CGNAT (del inglés— Carrier-Grade NAT) en su red. Este tipo de NAT suele ser conveniente para grandes empresas, debido a que permite la traducción de direcciones a gran escala y permite que distintos dispositivos compartan una única dirección de red pública de forma global (CISCO Systems, 2019).



Figura 30: Escenarios para evaluar el sistema propuesto. A) Escenario 1. B) Escenario 2. C) Escenario 3.

Para la captura y análisis de las pruebas de consumo de datos y desempeño de red en los escenarios propuestos se empleó la herramienta de monitoreo de red WireShark (Sharpe et al., 2023). WireShark es un analizador de protocolos de red de código abierto ampliamente reconocido en el ámbito de la ciberseguridad y las telecomunicaciones. Esta aplicación permite examinar detalladamente el tráfico de red en tiempo real, capturando y decodificando los paquetes de datos que circulan a través de una interfaz de red. Por otro lado, el análisis estadístico del consumo de datos se llevó a cabo utilizando el módulo pandas del lenguaje de programación Python, a partir de un archivo con extensión csv generado mediante

el módulo estadístico "Gráficas de E/S" de la herramienta Wireshark.

5.2 Pruebas de establecimiento de sesión y transmisión de datos

La verificación del correcto funcionamiento del protocolo de señalización constituye el primer paso en la evaluación integral del sistema. Para ello, en cada escenario se realizó la ejecución de cada una de las aplicaciones para iniciar la conexión entre el cliente clínica y el cliente especialista.

En los tres escenarios considerados, el primer paso consiste en el intento de establecer la sesión por parte de ambos clientes. Para esto, ambos clientes se registran en el servidor de sesión y luego comienzan a intercambiar mensajes con el servidor a través de Websockets seguros. El intercambio de mensajes se mantiene por 10 segundos hasta que la llamada es aceptada por ambos clientes. Si en 10 segundos la llamada no es aceptada, es necesario reiniciar el intento de establecimiento de sesión. Una vez aceptada la llamada por ambas partes, se intenta realizar el proceso de *UDP Hole Punching*.

En la Figura 31, se muestra el proceso de comunicación entre un cliente y el servidor empleando la herramienta de WireShark. Se observa la periodicidad del mecanismo de "ping-pong" implementado en websockets, visto anteriormente en el apartado 4.2. Este mecanismo mantiene la conexión activa entre el servidor de señalización y los clientes, mediante la transmisión de mensajes periódicos.

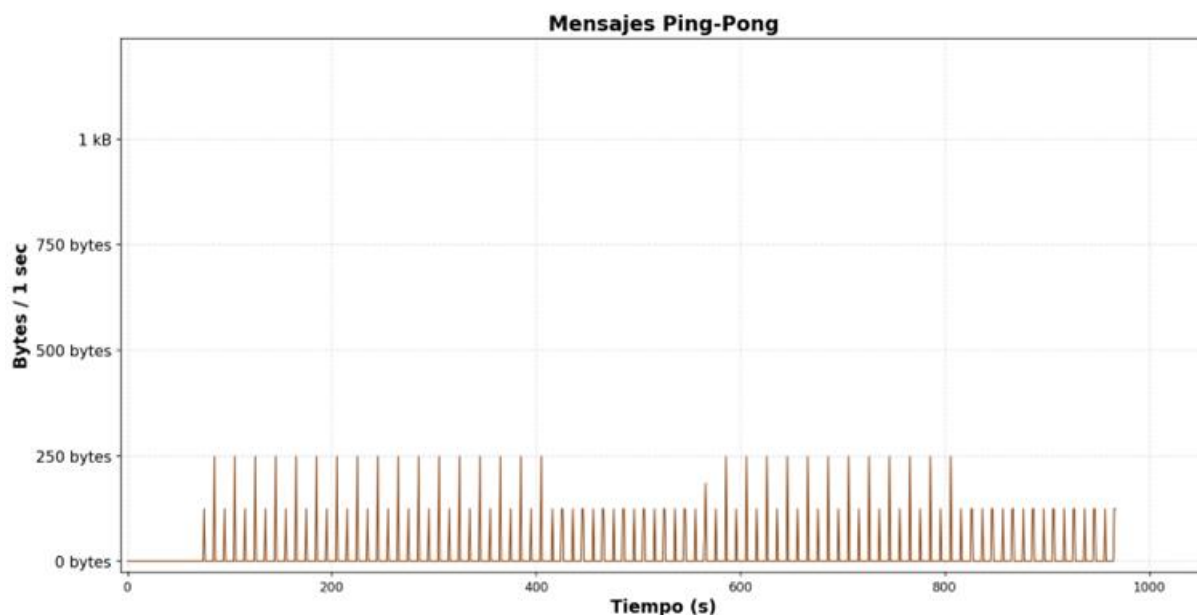


Figura 31: Mecanismo de ping-pong entre cliente-servidor de señalización.

5.2.1 Resultados de pruebas para el escenario 1

En este escenario, debido a que ambos clientes cuentan con una dirección IP pública y están detrás del mismo firewall, el proceso de *UDP Hole Punching* debería de ser exitoso. Para comprobar esto, en la Figura 32 se muestra el registro de los mensajes del cliente clínica. En esta Figura se puede ver que el cliente clínica envía 2 mensajes UDP “dummy” con la intención de poder abrir un agujero en el NAT. Después se puede comprobar que el cliente clínica recibió el mensaje de comprobación del mensaje “dummy”. Lo cual indica que se estableció de manera correcta la transmisión de par-a-par. En la Figura 33, se muestran los mensajes del registro provenientes del cliente especialista, donde se puede observar que el cliente recibe una nueva llamada (aceptar o rechazar) y en este caso la acepta. Luego de esto, envía un mensaje de verificación de “dummy” al cliente clínica mediante el servidor para indicar el establecimiento de una comunicación en el modo P2P.

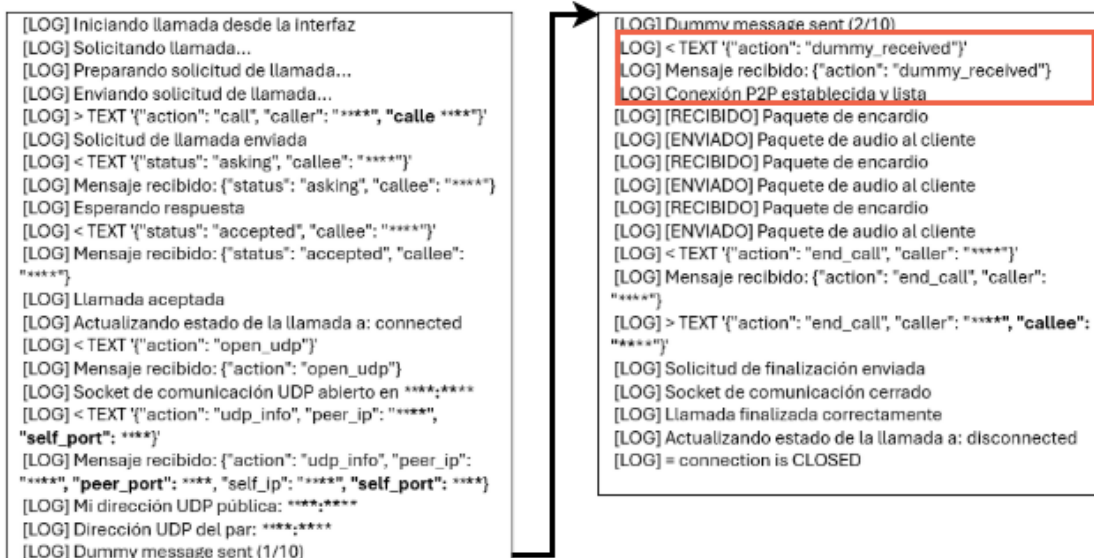


Figura 32: Registro de mensajes del cliente clínica. (Escenario 1)

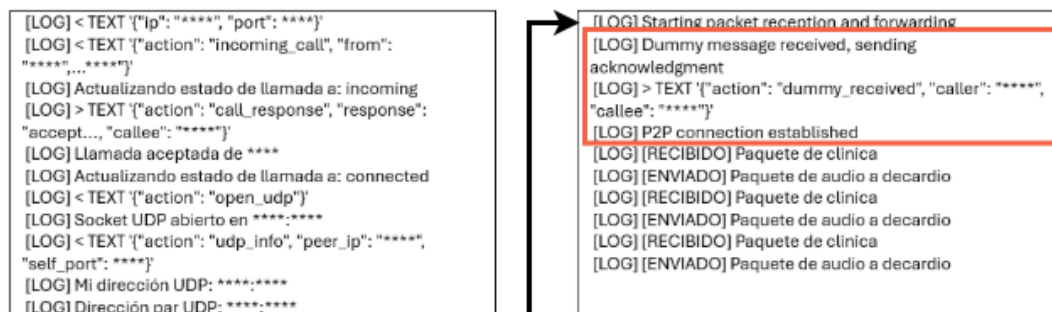


Figura 33: Registro de mensajes del cliente especialista. (Escenario 1)

5.2.2 Resultados de pruebas para el escenario 2

En la Figura 34 se muestran los mensajes del registro durante la etapa del establecimiento de la comunicación para el escenario 2. Al igual que en el escenario anterior, una vez aceptada la llamada por ambas partes, se intenta realizar el proceso de *UDP Hole Punching*. En la Figura 34, se observa que el cliente clínica envía 10 mensajes UDP “dummy” con la intención de poder abrir un agujero en el NAT en caso de que sea posible. No obstante, en la Figura 35 se observa que estos mensajes no llegan a su destino y el cliente especialista manda una notificación al servidor para comenzar la transmisión en modo cliente-servidor.

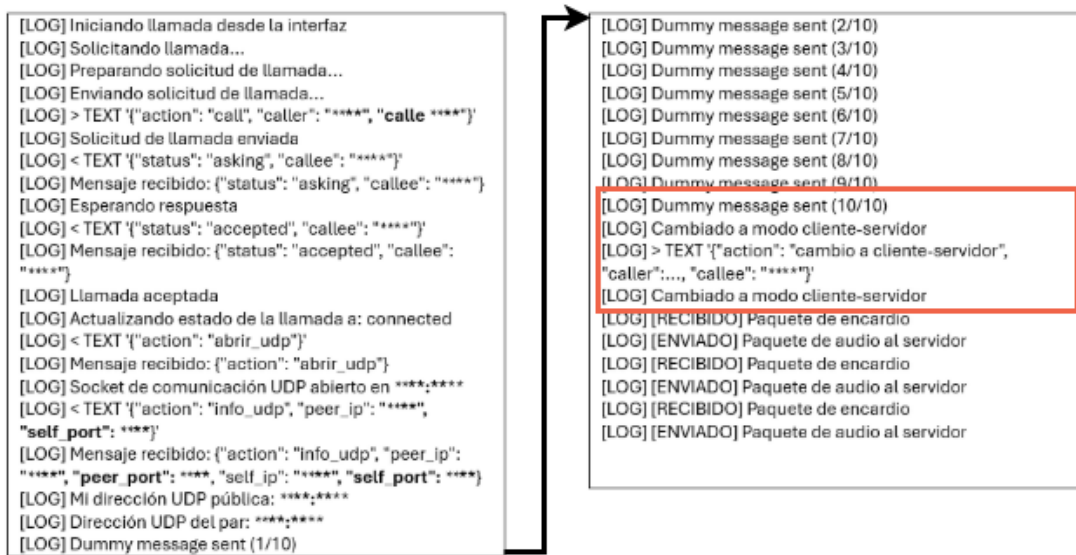


Figura 34: Registro de mensajes del cliente clínica. (Escenario 2)

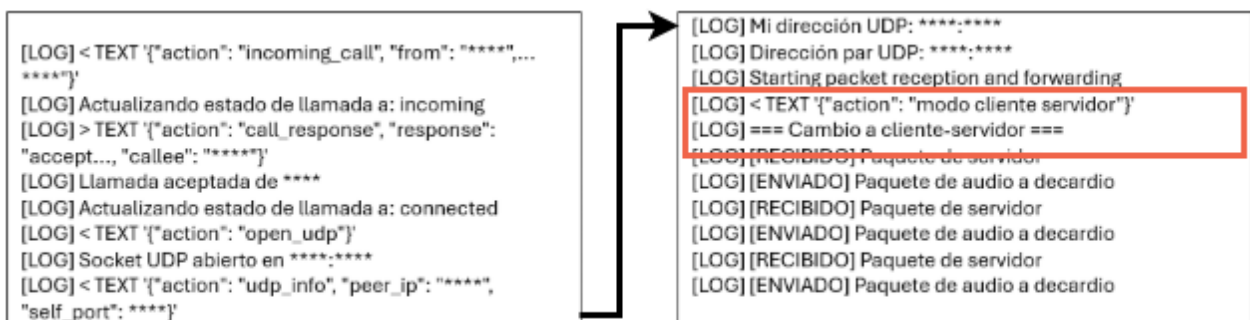


Figura 35: Registro de mensajes del cliente especialista. (Escenario 2)

5.2.3 Resultados de pruebas para el escenario 3

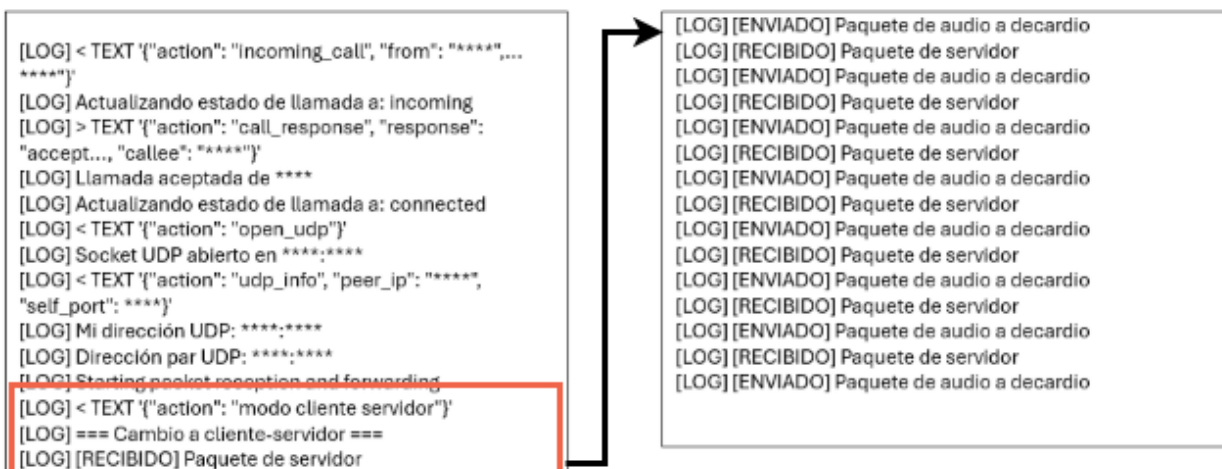
Igual que en el escenario anterior una vez aceptada la llamada por ambos clientes, se intenta realizar el proceso de *UDP Hole Punching*. Para este caso, el cliente clínica envía 10 mensajes UDP “dummy” con la intención de abrir un agujero en el NAT, en caso de que sea posible (ver Figura 36). Por su parte, en la Figura 37 se muestra que el cliente especialista no recibe un mensaje “dummy”, por lo cual manda una notificación al servidor para comenzar la transmisión en modo cliente-servidor.



```
[LOG] Iniciando llamada desde la interfaz
[LOG] Enviando solicitud de llamada...
[LOG] > TEXT '{"action": "call", "caller": "****", "calle...****"}'
[LOG] Solicitud de llamada enviada
[LOG] < TEXT '{"status": "asking", "callee": "****"}'
[LOG] Mensaje recibido: {"status": "asking", "callee": "****"}
[LOG] Esperando respuesta
[LOG] < TEXT '{"status": "accepted", "callee": "****"}'
[LOG] Mensaje recibido: {"status": "accepted", "callee": "****"}
[LOG] Llamada aceptada
[LOG] Actualizando estado de la llamada a: connected
[LOG] < TEXT '{"action": "abrir puertos"}'
[LOG] Mensaje recibido: {"action": "abrir puertos"}
[LOG] Socket de comunicación UDP abierto en ****:****
[LOG] < TEXT '{"action": "udp_info", "peer_ip": "****", "self_port": ****}'
[LOG] Mensaje recibido: {"action": "udp_info", "peer_ip": "****", "self_port": ****, "self_ip": "****", "self_port": ****}
[LOG] Mi dirección UDP pública: ****:****
[LOG] Dirección UDP del par: ****:****

[LOG] Dummy message sent (1/10)
[LOG] Dummy message sent (2/10)
[LOG] Dummy message sent (3/10)
[LOG] Dummy message sent (4/10)
[LOG] Dummy message sent (5/10)
[LOG] Dummy message sent (6/10)
[LOG] Dummy message sent (7/10)
[LOG] Dummy message sent (8/10)
[LOG] Dummy message sent (9/10)
[LOG] Dummy message sent (10/10)
[LOG] === Cambio al modo cliente-servidor ===
[LOG] > TEXT '{"action": "switch_to_client_server", "caller": "...", "callee": "****"}'
[LOG] < TEXT '{"action": "use_client_server_mode"}'
[LOG] Mensaje recibido: {"action": "modo cliente-servidor"}
[LOG] === Cambio al modo cliente-servidor ===
[LOG] [RECIBIDO] Paquete de encardio
[LOG] [ENVIADO] Paquete de audio al servidor
[LOG] [RECIBIDO] Paquete de encardio
[LOG] [ENVIADO] Paquete de audio al servidor
[LOG] [RECIBIDO] Paquete de encardio
```

Figura 36: Registro de mensajes del cliente clínica. (Escenario 3)



```
[LOG] < TEXT '{"action": "incoming_call", "from": "****", "to": "****"}'
[LOG] Actualizando estado de llamada a: incoming
[LOG] > TEXT '{"action": "call_response", "response": "accept...", "callee": "****"}'
[LOG] Llamada aceptada de ****
[LOG] Actualizando estado de llamada a: connected
[LOG] < TEXT '{"action": "open_udp"}'
[LOG] Socket UDP abierto en ****:****
[LOG] < TEXT '{"action": "udp_info", "peer_ip": "****", "self_port": ****}'
[LOG] Mi dirección UDP: ****:****
[LOG] Dirección par UDP: ****:****
[LOG] Starting packet reception and forwarding
[LOG] < TEXT '{"action": "modo cliente servidor"}'
[LOG] === Cambio a cliente-servidor ===
[LOG] [RECIBIDO] Paquete de servidor

[LOG] [ENVIADO] Paquete de audio a decardio
[LOG] [RECIBIDO] Paquete de servidor
[LOG] [ENVIADO] Paquete de audio a decardio
[LOG] [RECIBIDO] Paquete de servidor
[LOG] [ENVIADO] Paquete de audio a decardio
[LOG] [RECIBIDO] Paquete de servidor
[LOG] [ENVIADO] Paquete de audio a decardio
[LOG] [RECIBIDO] Paquete de servidor
[LOG] [ENVIADO] Paquete de audio a decardio
[LOG] [RECIBIDO] Paquete de servidor
[LOG] [ENVIADO] Paquete de audio a decardio
[LOG] [RECIBIDO] Paquete de servidor
[LOG] [ENVIADO] Paquete de audio a decardio
[LOG] [RECIBIDO] Paquete de servidor
[LOG] [ENVIADO] Paquete de audio a decardio
```

Figura 37: Registro de mensajes del cliente especialista. (Escenario 3)

En la Figura 38, se observa una captura de pantalla del sistema en funcionamiento. Esto engloba los tres escenarios contemplados, debido a que en cada caso se logró establecer una comunicación exitosa entre los clientes para transmitir de manera correcta el audio cardiopulmonar.



Figura 38: Captura de pantalla de una llamada entre los clientes.

5.3 Análisis del consumo de datos y la tasa de entrega de paquetes (PDR) en los escenarios propuestos

Durante la realización de las pruebas en los escenarios anteriormente descritos, se analizaron el consumo de datos y la tasa de entrega de paquetes (PDR, de sus siglas del inglés). El objetivo de este análisis es evaluar el rendimiento de la red en términos de eficacia en la transmisión de datos y la capacidad del sistema para entregar los paquetes de manera eficiente en cada uno de los escenarios propuestos.

Para la prueba de consumo de datos, se decidió que cada escenario se repitiera tres veces. La duración de la prueba se fijó en 904 segundos para el escenario 1 con conexión P2P. La duración de las pruebas para los otros escenarios fue de 877 segundos para el escenario 2 y de 878 segundos para el escenario 3. El tráfico de red generado por los dos clientes fue capturado y analizado con WireShark, el cual fue configurado acorde al protocolo, direccionamiento de red y puertos de comunicación establecidos en el sistema.

5.3.1 Consumo de datos

En la Tabla 9 se observan los resultados obtenidos del promedio del consumo de datos para cada escenario propuesto. El único escenario donde se pudo transmitir de par-a-par fue el escenario 1, esto debido a que ambos clientes se encontraban dentro de una red sin la presencia de NAT. El cliente clínica presentó un consumo de datos de 5 541 856 bytes con una tasa de transmisión promedio de 63.05 kbps, mientras que el cliente especialista tuvo un consumo de datos de 5 541 808 bytes y una tasa de recepción promedio de 63.05 kbps. Esto indica una distribución equitativa del ancho de banda entre ambos clientes.

En el escenario 2, el sistema EAC entró en el modo de transmisión cliente-servidor. En este escenario, el cliente clínica presentó un consumo de datos de 5 383 760 bytes con una tasa de transmisión promedio de 63.94 kbps, mientras que el cliente especialista tuvo un consumo de datos de 5 363 840 bytes y una tasa promedio de 64.73 kbps. Este escenario muestra un aumento en las tasas de transmisión promedio en comparación con el escenario 1.

En el escenario 3, bajo el modo de transmisión cliente-servidor, el cliente clínica registró un consumo de datos de 5 342 592 bytes y una tasa de transmisión promedio de 64.67 kbps, mientras que el cliente especialista tuvo un consumo de datos de 5 342 592 bytes y una tasa promedio de 65.73 kbps.

En general, los resultados obtenidos reflejan un comportamiento consistente en las tasas de transmisión promedio entre los clientes en ambos modos de transmisión, con ligeras variaciones que podrían ser atribuidas a las características específicas de cada escenario y experimento. Cabe hacer notar que todas las tasas de transmisión están muy por debajo de los 128 kbps requeridos por aplicaciones como TeleSteth de 3M o Eko Sensora de eko.

Tabla 9: Consumo de datos obtenido por cada escenario propuesto.

Escenario	Modo de transmisión	Tiempo total (Seg)	Cliente clínica		Cliente especialista	
			Consumo de datos (bytes)	Tasa de transmisión promedio (kbps)	Consumo de datos (bytes)	Tasa de recepción promedio (kbps)
1	P2P	904	5 541 856	63.05	5 541 808	63.05
2	C/S	877	5 383 760	63.94	5 363 840	64.73
3	C/S	878	5 342 592	64.67	5 342 592	65.73

5.3.2 Tasa de entrega de paquetes (PDR)

La tasa de entrega de paquetes (PDR) es una métrica fundamental para evaluar la eficacia y fiabilidad de la transmisión de datos en redes, especialmente en entornos inalámbricos y de baja potencia. El PDR se define como la proporción entre el número de paquetes recibidos correctamente y el total de paquetes transmitidos (Ecuación 1), lo que permite cuantificar la calidad del enlace y la eficiencia de los protocolos de enrutamiento en condiciones adversas (Tripathi et al., 2012).

$$PDR(\%) = \frac{(Paq. Recibidos)}{(Paq. Enviados)} * 100 \quad (1)$$

En la Tabla 10 se evalúa el PDR en los escenarios propuestos. En el escenario 1, con el modo de transmisión P2P se observa que en la transmisión no hubo pérdidas de paquetes. Los paquetes enviados por parte del cliente se recibieron por parte del especialista en su totalidad, teniendo así un 100 % de entrega de paquetes. En el escenario 2, con el modo de transmisión cliente-servidor, se observa una mayor pérdida de paquetes respecto de los demás escenarios, puesto que cuenta con una pérdida de 19 paquetes, dando así un 99.53% de PDR. Esto se podría atribuir al proveedor de la red empleada y el incremento de la complejidad del entorno de red.

En el escenario 3 se obtuvo un PDR del 99.97% dado que solo se perdió un paquete en este escenario. El hecho de tener un mejor PDR en este escenario en comparación al escenario 2 puede atribuirse a las condiciones particulares de la red en el momento de realizar los experimentos.

Tabla 10: Relación de entrega de paquetes.

Escenario	Modo de transmisión	Cliente clínica	Cliente especialista	PDR (%)
		Paquetes enviados	Paquetes recibidos	
1	P2P	2782	2782	100
2	C/S	4076	4057	99.53
3	C/S	4054	4053	99.97

5.4 Evaluación del sistema propuesto por especialistas de la salud

Para llevar a cabo la evaluación del sistema propuesto, se contó con la participación de especialistas de la salud, quienes aportaron su experiencia y conocimientos. Su colaboración permitió analizar el

funcionamiento del sistema y compararlo con otras plataformas de videoconferencia utilizadas (p. ej. Zoom, Google Meet, Microsoft Teams) en su práctica profesional. Para la comparación se realizó un formulario basado en la prueba MUSHRA (Anexo B), que compara de manera directa las plataformas convencionales como lo son Zoom, Google Meet, Microsoft Teams, respecto al sistema EAC por medio de la percepción acústica de las personas.

Para la comparación del sistema, se emplearon las tecnologías que se observan en la Figura 39. Se usaron dos tipos de estetoscopios: uno electrónico y otro digital, con el objetivo de analizar su desempeño subjetivo y evaluar la calidad del audio cardiopulmonar transmitido a través del sistema propuesto. Asimismo, cada cliente cuenta con su propia tarjeta de audio externa y el uso de una banda ancha móvil para tener una segunda red privada.



Figura 39: Herramientas utilizadas para la evaluación de los sistemas convencionales.

Las pruebas se llevaron a cabo en el estado de Durango, realizadas en consultorios médicos y en casa-habitación, con el objetivo de reducir los ruidos ambientales. Para ello, se contó con la colaboración de 10

médicos generales. Los resultados generales obtenidos a partir del formulario propuesto se presentan en la Figura 40. En esta Figura se observa la evaluación de cuatro aspectos clave: la claridad del audio, el ruido inducido durante la transmisión, los retrasos en la transmisión, las interrupciones en la comunicación y el volumen percibido del audio cardiopulmonar.

En la Figura 40 se aprecia que el sistema EAC tiene una clara ventaja en la transmisión de audio cardiopulmonar respecto a sistemas de videoconferencia como Zoom, y Teams, debido a las características de los códecs de voz empleados. En cuanto a la plataforma Google Meet presentó un desempeño más competitivo, debido al uso de códec de voz adaptativo (OPUS). Sin embargo, los médicos señalaron que introducía un leve ruido que podría confundirse con una sibilancia, lo cual puede llevar a un incorrecto diagnóstico médico.

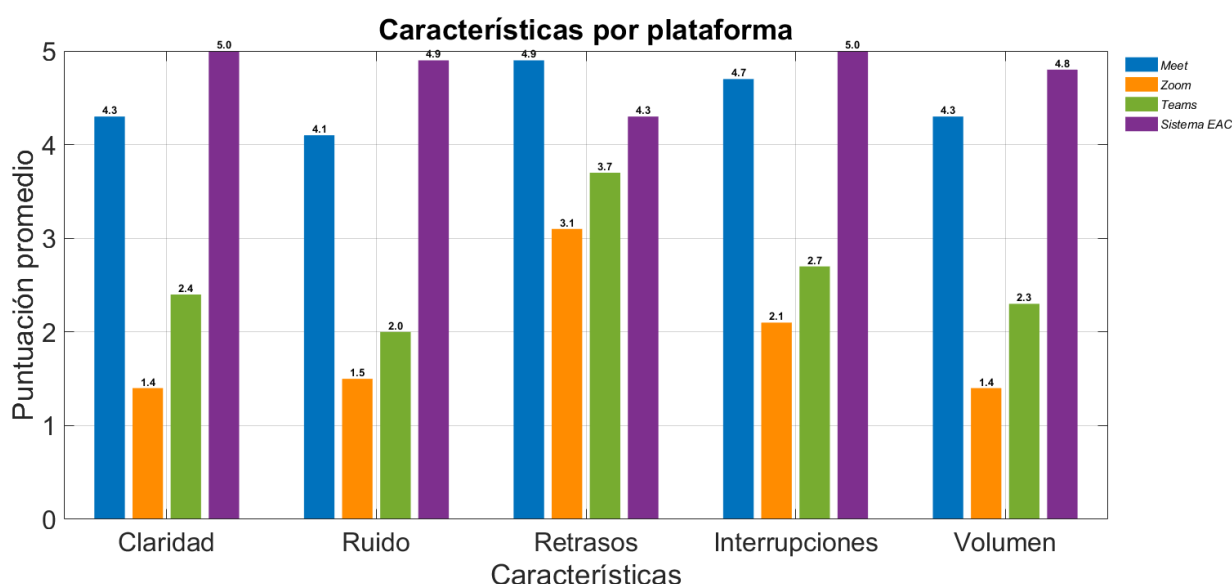


Figura 40: Características de las plataformas evaluadas.

En la Figura 41 se presenta la puntuación promedio de la calidad diagnóstica del audio cardiopulmonar, donde 5 representa una calidad diagnóstica excelente y 0 una calidad diagnóstica deficiente. Se observa que el sistema EAC obtuvo un promedio de 4.9/5, lo cual se considera una calidad diagnóstica sobresaliente. Le sigue la plataforma Google Meet con un promedio de 4.4/5, mientras que Teams y Zoom obtuvieron puntuaciones de 2/5 y 1/5, respectivamente.

Con base en los resultados obtenidos, todas las personas encuestadas manifestaron su preferencia por el Sistema EAC para la transferencia de audio cardiopulmonar. Esto se evidencia claramente en la Figura 42, donde el Sistema EAC ocupa el primer lugar, seguido por Google Meet en segundo lugar, Microsoft Teams

en tercer lugar y Zoom en último lugar. Los especialistas en salud seleccionaron el Sistema EAC como su opción preferida por su mayor precisión en el análisis y la confianza que genera al momento de escuchar el audio transmitido, lo que facilita la emisión de un diagnóstico médico adecuado.

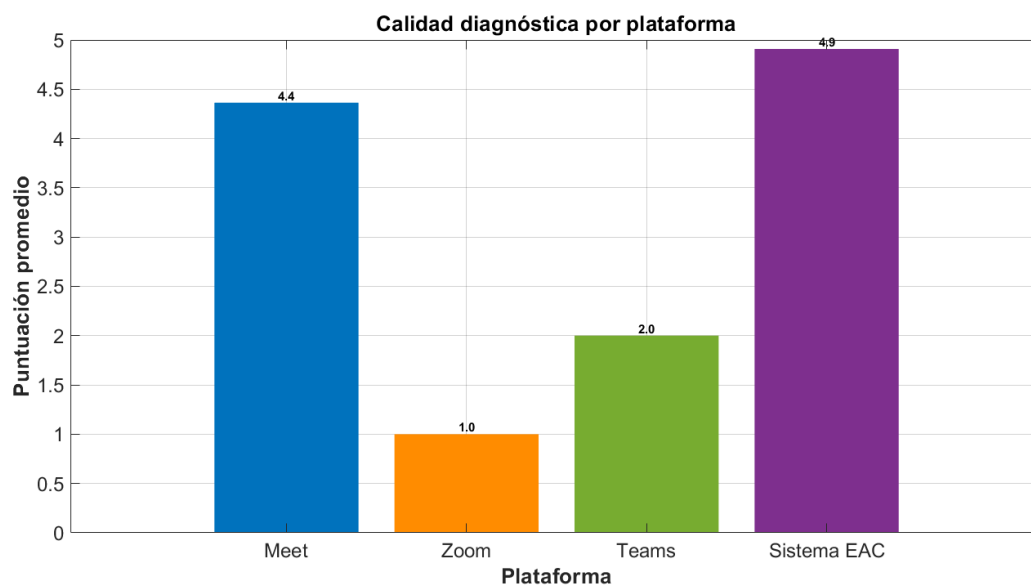


Figura 41: Calidad diagnóstica por plataforma.

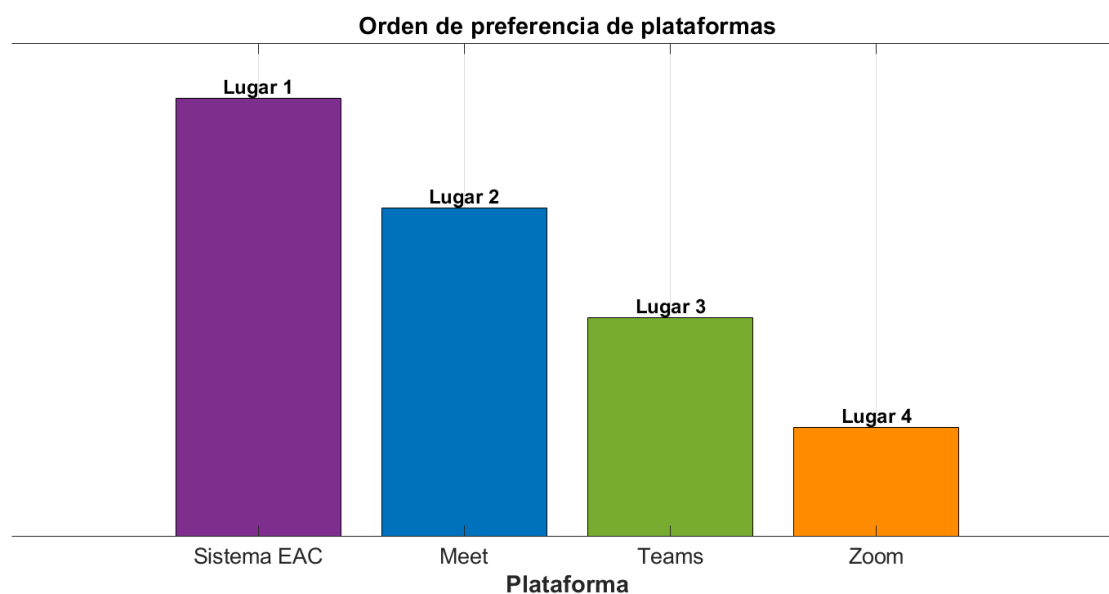


Figura 42: Preferencias de plataformas para el envío de audio cardiopulmonar.

Capítulo 6. Discusión de resultados

En este capítulo se lleva a cabo un análisis detallado de los resultados presentados en el capítulo anterior. Se resumen y destacan los aspectos más significativos relacionados con el rendimiento de la arquitectura propuesta, basándose en las mediciones previamente documentadas.

6.1 En relación con el establecimiento de la sesión

Los tres escenarios propuestos en la cama de pruebas fueron planteados con el propósito de verificar el correcto funcionamiento de las técnicas transversales de NAT implementadas en la arquitectura:

- El establecimiento de una conexión par-a-par (P2P) mediante la técnica de *UDP Hole Punching*.
- Uso de la técnica de retransmisión (*Relaying*) en el caso de que no sea posible el establecimiento de la comunicación de Par-a-Par (P2P) implementando la técnica de *UDP Hole Punching*.

Los resultados presentados en la Sección 5.2 muestran que las técnicas descritas funcionan correctamente dentro de la arquitectura propuesta. En el escenario 1, se comprobó que la técnica de *UDP Hole Punching* logra establecer una conexión entre los clientes involucrados en la tele-auscultación. Por su parte, en el escenario 2 como en el escenario 3, se demostró que la técnica de *Relaying* para la retransmisión de los mensajes implementando el servidor funciona de manera automática y efectiva cuando la técnica transversal de *UDP Hole Punching* no consigue establecer una comunicación de Par-a-Par (P2P).

En el segundo escenario evaluado, y de acuerdo con la Tabla 7, debería de haber sido factible el establecimiento de la comunicación Par-a-Par (P2P) entre los clientes. Sin embargo, en la realización de las pruebas, no fue posible lograr una conexión de Par-a-Par (P2P) debido a la presencia de un firewall que bloqueó los puertos utilizados. No obstante, este comportamiento permitió validar el correcto funcionamiento de la arquitectura en situaciones donde no es posible establecer una comunicación P2P. En este caso, se comprobó que los clientes lograron comunicarse de manera automática mediante el mecanismo de *Relaying*, una vez que el protocolo determinó que no era posible abrir un agujero en el NAT.

6.2 En relación con el consumo de datos y la tasa de entrega de paquetes

Un descubrimiento de este trabajo de investigación se debe a la significativa reducción en el consumo de datos que ofrece la arquitectura propuesta en comparación con las plataformas que ofrece la empresa Littmann. Debido a que recomiendan tener un ancho de banda de 1.5 Mbps por cada cliente en la plataforma de TeleSteth y para tener el servidor instalado, recomiendan tener un ancho de banda de 100 Mbps (3MTM Littmann®, 2025). Asimismo, para los estetoscopios de la empresa 3M™ Littmann® se debe de contar con un ancho de banda mínimo de 128 kbps para la transmisión del audio cardiopulmonar (3M™ Littmann®, 2010). Por lo que el sistema propuesto es superior en este sentido, ya que tiene una tasa de transferencia de datos de 65.73 kbps, que es 1.9 veces menor a la transmisión realizada con Littmann.

Respecto a la tasa de entrega de paquetes, en general, se obtuvo un alto nivel de eficiencia en la transmisión de datos, con valores de PDR superiores al 99.5% en todos los escenarios evaluados. El modo de transmisión punto a punto (P2P) destacó por alcanzar un PDR del 100%, lo que indica una transmisión perfecta sin pérdida de paquetes. Por otro lado, el modo cliente-servidor presentó ligeras pérdidas de paquetes, con PDRs que oscilaron entre el 99.5338% y el 99.9753%, lo que refleja una red confiable pero susceptible a pequeñas variaciones en la calidad de la transmisión. Estos resultados sugieren que, aunque ambos métodos son altamente eficientes, el entorno y las condiciones de transmisión pueden influir en el rendimiento general.

6.3 En relación con la evaluación realizada por el personal de la salud

El apoyo del personal de la salud ha sido clave para demostrar que el sistema EAC ofrece una calidad diagnóstica del audio cardiopulmonar en comparación con las plataformas convencionales. Las plataformas convencionales emplean códecs de voz que están desarrollados para la comunicación humana, lo que significa que comprimen el audio y eliminan frecuencias que no son esenciales para el habla, pero que sí son cruciales en un entorno médico. El contar con un códec que soporte las frecuencias cardiopulmonares y mantenga la calidad de estas tiene un impacto en el campo de la telemedicina, debido a que permitiría a los profesionales de la salud identificar los sonidos sutiles, como lo son los soplos en el corazón o irregularidades respiratorias.

El sistema EAC emplea los códecs identificados en (Olachea, 2019) como adecuados para la transmisión de audio cardiopulmonar. Sin embargo, fue necesario evaluar que al incorporarlos en la arquitectura la

calidad diagnóstica del audio cardiopulmonar transmitido se mantuviera. Los resultados reportados en la Sección 5.4 permitieron validar esto, ya que la evaluación por parte de los médicos involucrados en el estudio señaló que el sistema contaba con calidad diagnóstica. Además, los médicos señalaron que, comparado con otros sistemas, el sistema EAC proporcionaba mayor claridad, menores niveles de ruido, menores interrupciones al momento del envío de audio cardíaco y mayor volumen respecto a la referencia de audio cardiopulmonar local.

Capítulo 7. Conclusiones, aportaciones y trabajo futuro

En este capítulo se presentan las conclusiones derivadas del presente trabajo de investigación. Asimismo, se presentan las principales contribuciones de este trabajo, junto con las posibles direcciones para futuras investigaciones.

El objetivo principal de esta investigación fue proponer un sistema de transmisión de audio cardiopulmonar con calidad diagnóstica, diseñado para atravesar redes de Internet y garantizar la llegada del audio al destino. Durante el desarrollo del trabajo, se efectuaron diversas pruebas con el propósito de evaluar el funcionamiento de la arquitectura híbrida propuesta en comparación con estudios previos. En una primera etapa, se verificó el correcto funcionamiento de la arquitectura, así como su desempeño técnico. Posteriormente, se analizó el rendimiento del códec MP3 a través de la opinión de profesionales de la salud, quienes participaron en una serie de pruebas y completaron un formulario. Este ejercicio tuvo como finalidad recopilar las percepciones de los médicos sobre la facilidad de uso y la efectividad de esta arquitectura para la transmisión de audio con calidad diagnóstica en el contexto de la tele-auscultación.

A continuación, se presenta un resumen de los resultados obtenidos en las pruebas, así como de los comentarios de los médicos que colaboraron en el análisis del sistema.

7.1 Conclusiones

A partir de los resultados presentados en este trabajo de investigación, podemos enumerar las siguientes conclusiones:

Se concluye que la arquitectura híbrida propuesta, basada en un modelo P2P y el protocolo UDP, es viable para la transmisión de tipo streaming de señales cardiopulmonares. La implementación de técnicas como *UDP Hole Punching* y *Relaying* permitió superar las limitaciones del NAT, logrando establecer conexiones exitosas entre clientes en diversos escenarios típicos de despliegue en Internet.

Se incorporaron los programas de audio cardiopulmonar desarrollados en el trabajo de investigación de Olachea (2019), a la arquitectura propuesta. Esto facilitó la implementación del sistema y permitió la evolución directa de la línea de investigación dentro del grupo.

La arquitectura propuesta incorpora el cifrado AES-CBC de 128 bits, lo que refuerza la confidencialidad de los datos transmitidos al proteger la información cardiopulmonar contra accesos no autorizados. Este método de cifrado asegura que los datos sean ilegibles para terceros durante la transmisión.

El uso del códec MP3 en la arquitectura permitió mantener una tasa de transmisión promedio inferior a 50 kbps, lo que representa una mejora significativa en comparación con la tasa mínima requerida de 128 kbps para aplicaciones similares.

Con el apoyo y la validación de los médicos que colaboraron en el desarrollo de esta tesis, se concluye que la propuesta de sistema de transmisión de audio cardiopulmonar es adecuada, al menos en su versión prototipo, para su implementación en entornos de telemedicina.

Lo anterior demuestra que la compresión MP3 es una solución eficiente para reducir el consumo de datos sin comprometer la calidad diagnóstica del audio cardiopulmonar.

En general, la arquitectura mostró una notable eficiencia en su consumo de datos, requiriendo apenas 65.73 kbps de ancho de banda mientras mantenía una transmisión estable y de calidad. Esto no solo facilita su implementación en redes con limitaciones de capacidad, sino que también la hace más accesible para escenarios donde el consumo de datos es un factor crítico, como en áreas rurales o con infraestructura limitada.

7.2 Aportaciones

Las principales aportaciones logradas con el desarrollo de este trabajo de investigación se mencionan a continuación.

- Se demostró que es posible realizar una arquitectura híbrida P2P / cliente-servidor para la transmisión de audio cardiopulmonar con calidad diagnóstica.
- Se demostró que la calidad diagnóstica de los códecs reportados en (Olachea,2019) se mantiene al ser utilizados en una arquitectura sin VPN y con una comunicación P2P.

- Se demostró que la arquitectura propuesta es superior para la transmisión de audio cardiopulmonar que las soluciones comerciales de videoconferencias.
- Se desarrolló una arquitectura híbrida P2P / cliente-servidor que puede ser utilizada para la transmisión en vivo de otro tipo de señales

7.3 Trabajo futuro

Aunque los resultados de la investigación han sido positivos, aún hay aspectos que pueden mejorarse, los cuales se plantean a continuación como parte del trabajo futuro.

- Evaluación de la arquitectura con parámetros de red como Jitter y latencia.
- Integración de un modelo de IA que ayude a detectar si hay algún problema de salud con el sonido cardiopulmonar.
- Implementar más códecs para la realización de las pruebas, como es el caso del códec AAC.
- Añadir ayuda para las zonas que se deben de auscultar acorde a las indicaciones del especialista.
- Implementación de un algoritmo cifrado autenticado con datos asociados (de sus siglas en inglés - AEAD), como AES-GCM, para garantizar la confidencialidad, integridad y autenticidad de cada mensaje.

Literatura citada

- 3M™ Littmann®. (s/f). *Student and educator resources*. Recuperado el 6 de marzo de 2025, de <https://www.littmann.com/en-us/home/education/>
- 3M™ Littmann®. (2010). *Scope-to-Scope Software User Manual TS1000C*. 3M™ Littmann®. <https://multimedia.3m.com/mws/media/688129O/package-insert-for-3m-littmann-scope-to-scope-software-kit.pdf>
- 3M™ Littmann®. (2021). *Elija el fonendoscopio adecuado para ayudarle a dar lo mejor de usted*. 3M™ Littmann®. <https://multimedia.3m.com/mws/media/2009147O/littmann-comparison-chart-2021-es-fv.pdf>
- 3M™ Littmann®. (2025). *Frequently Asked Questions About Littmann Stethoscopes*. https://littmann.3m.com.sa/3M/en_SA/littmann-sa/my-stethoscope/faq/
- Agnisarman, S., Madathil, K. C., Smith, K., Ashok, A., Welch, B. M., & McElligott, J. (2017). Lessons learned from the usability assessment of home-based telemedicine systems. *Applied ergonomics*, 58, 424–434. <https://doi.org/10.1016/j.apergo.2016.08.003>
- Alnahawi, N., Müller, J., Oupický, J., & Wiesmaier, A. (2024). A Comprehensive Survey on Post-Quantum TLS. *IACR Communications in Cryptology*, 1(2), Article 2. <https://doi.org/10.62056/ahee0iuc>
- AMD global Telemedicine. (2020). *Agnes Connect: A Collaborative Patient Assessment Tool for Enterprise-Wide Telehealth Applications*. https://www.solutionzinc.com/hubfs/AMD_AGNES_Connect_Datasheet.pdf
- Amoretti, M., & Zanichelli, F. (2018). P2P-PL: A pattern language to design efficient and robust peer-to-peer systems. *Peer-to-Peer Networking and Applications*, 11(3), 518–547. <https://doi.org/10.1007/s12083-017-0551-y>
- Amwell. (2020). *2020 Amwell Manual Contents*. https://americanwell-atlas.s3.amazonaws.com/media/filer_public/38/d9/38d9bf64-e9d8-4d2b-8c51-e5f090c1875a/2020_amwell_manual.pdf
- Amwell. (2023). *Amwell is Again the Most Popular Telehealth App of the Year*. <https://business.amwell.com/about-us/news/press-releases/2016/amwell-is-again-the-most-popular-telehealth-app-of-the-year>
- Bedingham, W., Oster, C. D., Rogers, D. J., & Schmidt, T. P. (2016). *Electronic stethoscope system for telemedicine applications* (United States Patent No. US9480400B2). <https://patents.google.com/patent/US9480400B2/en>
- Begen, A. C., Kyzivat, P., Perkins, C., & Handley, M. J. (2021). *SDP: Session Description Protocol* (Request for Comments No. RFC 8866). Internet Engineering Task Force. <https://doi.org/10.17487/RFC8866>
- Buford, J. F., Yu, H., & Keong Lua, E. (2009). *P2P Networking and Applications*. Morgan Kaufmann. <https://www.perlego.com/book/1813735>

- Carrara, E., Norrman, K., McGrew, D., Naslund, M., & Baugher, M. (2004). *The Secure Real-time Transport Protocol (SRTP)* (Request for Comments No. RFC 3711). Internet Engineering Task Force. <https://doi.org/10.17487/RFC3711>
- Casillas, M., Villarreal-Reyes, S., González, A. L., Martinez, E., & Perez-Ramos, A. (2015). Design Guidelines for Wireless Sensor Network Architectures in mHealth Mobile Patient Monitoring Scenarios. En *Mobile Health: A Technology Road Map* (pp. 401–428). Springer. <https://doi.org/10.1007/978-3-319-12817-7>
- Castorena, L. (2012). *Análisis y adaptación de codecs para la transmisión de audio cardíaco sobre redes inalámbricas con tasas de datos bajas/medias* [Tesis de Maestría, Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California]. Repositorio institucional. <http://cicese.repositorioinstitucional.mx/jspui/handle/1007/2481>
- Chen, Y.-C., & Jia, W.-K. (2009). Challenge and solutions of NAT traversal for ubiquitous and pervasive applications on the Internet. *Journal of Systems and Software*, 82(10), 1620–1626. <https://doi.org/10.1016/j.jss.2009.03.078>
- CISCO Systems. (2019). *Carrier Grade Network Address Translation*. CISCO. https://www.cisco.com/c/en/us/td/docs/ios-xml/ios/ipaddr_nat/configuration/xr-3s/nat-xr-3s-book/iadnat-cgn.pdf
- Coulouris, G., Dollimore, J., Kindberg, T., & Blair, G. (2011). *Distributed systems: Concepts and design* (5th ed.). Pearson Education, Inc.
- Cruz, F. R. G., Magwili, G. V., Collado, M. R. L., Rosario, K. M., Aguila, R. M. L., & Egargue, J. C. C. (2019). Electronic Stethoscope with Noise Cancellation for Telehealth Peripheral. *2019 3rd International Conference on Electrical, Telecommunication and Computer Engineering (ELTICOM)*, 155–159. <https://doi.org/10.1109/ELTICOM47379.2019.8943913>
- Denault, A. Y., Couture, P., Vegas, A., Buithieu, J., & Tardif, J.-C. (2016). *Transesophageal Echocardiography Multimedia Manual* (2nd ed.). CRC Press. <https://www.perlego.com/book/4364747>
- Doxy.me. (s/f). *The Simple, Free, and Secure Telemedicine Platform*. Doxy.Me. Recuperado el 14 de marzo de 2025, de <https://doxy.me/en/>
- Dubois, S., Mogren, C., & Zhukov, A. (2021). *WebRTC for the curious*. WebRTCforcurious.com. <https://webrtcforthe curious.com/docs/webrtc-for-the-curious.pdf>
- Egevang, K. B., & Srisuresh, P. (2001). *Traditional IP Network Address Translator (Traditional NAT)* (Request for Comments No. RFC 3022). Internet Engineering Task Force. <https://doi.org/10.17487/RFC3022>
- Eko Health, Inc. (s/f). *Compare Plans*. Eko Health. Recuperado el 6 de marzo de 2025, de <https://www.ekohealth.com/pages/compare-plans>
- Engelstad, P., & Egeland, G. (2004). NAT-Based Internet Connectivity for On-Demand Ad Hoc Networks. En R. Battiti, M. Conti, & R. L. Cigno (Eds.), *Wireless On-Demand Network Systems* (pp. 344–358). Springer. https://doi.org/10.1007/978-3-540-24614-5_25

- Ford, B., Kegel, D., & Srisuresh, P. (2008). *State of peer-to-peer (P2P) communication across network address translators (NATs)* (Request for comments No. RFC 5128). IETF. <https://doi.org/10.17487/RFC5128>
- Ford, B., Srisuresh, P., & Kegel, D. (2006). *Peer-to-Peer Communication Across Network Address Translators* (No. arXiv:cs/0603074). arXiv. <https://doi.org/10.48550/arXiv.cs/0603074>
- González, A. L., Galaviz-Mosqueda, A., Villarreal-Reyes, S., Magana, R., Rivera, R., Casillas, M., & Villasenor, L. (2015). Current Situation and Challenges for mHealth in the Latin America Region. En *mHealth Multidisciplinary Verticals* (pp. 413–432). CRC Press. <https://www.perlego.com/book/1599394>
- Google for Developers (Director). (2013, mayo 20). *Real-time communication with WebRTC: Google I/O 2013* [Video recording]. <https://www.youtube.com/watch?v=p2HzZkd2A40>
- Google LLC. (2025). *Requisitos del hardware de Meet*. <https://support.google.com/a/answer/4541234?hl=es#zippy=%2Crequisitos-de-ancho-de-banda>
- Grigorik, I. (2013). *High Performance Browser Networking*. O'Reilly Media, Inc.
- Gunawan, T. S., & Khalifa, O. O. (2010). PC sound card based instrumentation and control. *International Conference on Computer and Communication Engineering (ICCCE'10)*, 1–4. <https://doi.org/10.1109/ICCCE.2010.5556784>
- Hassani, A. H. E. (2012). *eHealth and Remote Monitoring*. IntechOpen. <https://www.perlego.com/book/2017379>
- Healow. (s/f). *healow—Book An Appointment*. Recuperado el 14 de marzo de 2025, de <https://healow.com/apps/jsp/webview/index.jsp>
- Heinze-Martin, G., Olmedo-Canchola, V. H., Bazán-Miranda, G., Bernard-Fuentes, N. A., & Guízar-Sánchez, D. P. (2018). Los médicos especialistas en México. *Gaceta Médica de México*, 154(3). <https://doi.org/10.24875/GMM.18003770>
- Helicon Health Ltd. (s/f). *Teladoc Health Cloud Video Consultation Solution (Telemedicine, Telehealth)—Digital Marketplace*. Recuperado el 5 de marzo de 2025, de <https://www.applytosupply.digitalmarketplace.service.gov.uk/g-cloud/services/714847543639094>
- Holdrege, M., & Srisuresh, P. (1999). *IP Network Address Translator (NAT) Terminology and Considerations* (Request for Comments No. RFC 2663). Internet Engineering Task Force. <https://doi.org/10.17487/RFC2663>
- Huang, F., Yu, L., Shen, T., & Hu, S. (2019). The P2P solution research and design based on NAT traversing technology. *2019 IEEE 3rd Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, 1347–1351. <https://doi.org/10.1109/IMCEC46724.2019.8984136>
- Jaideep, G., & Battula, B. P. (2016). Survey on the present state-of-the-art of P2P networks, Their Security Issues and Counter Measures. *International Journal of Applied Engineering Research*, 11(1), 616. <https://doi.org/10.37622/IJAER/11.1.2016.616-620>

- Kanazawa, K., & Takami, K. (2018). Standby P2P Network to Substitute for Client-Server Network in Event of Interruption of Communication with the Server. *TENCON 2018 - 2018 IEEE Region 10 Conference*, 0931–0936. <https://doi.org/10.1109/TENCON.2018.8650292>
- Keränen, A., Holmberg, C., & Rosenberg, J. (2018). *Interactive connectivity establishment (ICE): A protocol for network address translator (NAT) traversal* (Request for Comments No. RFC 8445). Internet Engineering Task Force. <https://doi.org/10.17487/RFC8445>
- Khemapech, I., Sansrimahachai, W., & Toachooddee, M. (2019). Telemedicine – Meaning, Challenges and Opportunities. *Siriraj Medical Journal*, 71(3), Article 3. <https://doi.org/10.33192/Smj.2019.38>
- Kurose, J. F., & Ross, K. W. (2017). *Computer networking: A top-down approach* (7th ed.). Pearson Education, Inc.
- Lakhe, A., Sodhi, I., Warriar, J., & Sinha, V. (2016). Development of digital stethoscope for telemedicine. *Journal of Medical Engineering & Technology*, 40(1), 20–24. <https://doi.org/10.3109/03091902.2015.1116633>
- Leng, S., Tan, R. S., Chai, K. T. C., Wang, C., Ghista, D., & Zhong, L. (2015). The electronic stethoscope. *BioMedical Engineering OnLine*, 14(1), 66. <https://doi.org/10.1186/s12938-015-0056-y>
- Loreto, S., & Romano, S. P. (2014). *Real-Time Communication with WebRTC*. O'Reilly Media, Inc.
- Magana-Rodriguez, R., Villarreal-Reyes, S., Galaviz-Mosqueda, A., Rivera-Rodriguez, R., & Conte-Galvan, R. (2018). An Adaptive Cross-Layer Admission Control Mechanism for Telemedicine Services over the IEEE 802.22/WRAN Standard. *IEICE Transactions on Communications*, E101.B(4), 1029–1044. <https://doi.org/10.1587/transcom.2017EBP3157>
- Magaña-Rodriguez, R., Villarreal-Reyes, S., Galaviz-Mosqueda, A., Rivera-Rodriguez, R., & Conte-Galvan, R. (2015). Telemedicine Services over Rural Broadband Wireless Access Technologies: IEEE 802.22/WRAN and IEEE 802.16 WiMAX. En *Mobile Health: A Technology Road Map* (pp. 743–769). Springer. <https://doi.org/10.1007/978-3-319-12817-7>
- Mangione, S., Sullivan, P., & Wagner, M. S. (2020). *Physical Diagnosis Secrets E-Book: Physical Diagnosis Secrets E-Book* (3a ed.). Elsevier. <https://www.perlego.com/book/4279284>
- Manson, R. (2013). *Getting Started with WebRTC*. Packt Publishing. <https://www.perlego.com/book/390145>
- Marcy, T. (1999). *Medical management of pulmonary diseases*. CRC Press. <https://www.perlego.com/book/4577459>
- Matthews, P., Rosenberg, J., & Mahy, R. (2010). *Traversal Using Relays around NAT (TURN): Relay Extensions to Session Traversal Utilities for NAT (STUN)* (Request for Comments No. RFC 5766). Internet Engineering Task Force. <https://doi.org/10.17487/RFC5766>
- Matthews, P., Rosenberg, J., Wing, D., & Mahy, R. (2008). *Session Traversal Utilities for NAT (STUN)* (Request for Comments No. RFC 5389). Internet Engineering Task Force. <https://doi.org/10.17487/RFC5389>

- Miah, S. J., Hasan, J., & Gammack, J. G. (2017). On-Cloud Healthcare Clinic: An e-health consultancy approach for remote communities in a developing country. *Telematics and Informatics*, 34(1), 311–322. <https://doi.org/10.1016/j.tele.2016.05.008>
- MicrosoftHeidi. (2024, junio 25). *Preparar la red de la organización para Microsoft Teams—Microsoft Teams*. <https://learn.microsoft.com/es-es/microsoftteams/prepare-network>
- OECD. (2023). *Health at a Glance 2023*. OECD Publishing. <https://doi.org/10.1787/7a7afb35-en>
- Office for Civil Rights. (2025). *Summary of the HIPAA Privacy Rule*. U.S. Department of Health & Human Services. <https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html>
- Olachea, A. (2019). *Diseño de mecanismos “Crosslayer” para la transmisión de señales cardio-pulmonares sobre redes de datos heterogéneas*. [Tesis de Maestría, Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California]. Disponible en PDF.
- Oppliger, R. (2023). *SSL and TLS: Theory and practice* (3rd ed.). Artech house.
- Paar, C., & Pelzl, J. (2010). *Understanding cryptography: A textbook for students and practitioners*. Springer. <http://link.springer.com/book/10.1007/978-3-642-04101-3>
- Pereira, D., Amélia-Ferreira, M., Cruz-Correia, R., & Coimbra, M. (2020). Teaching Cardiopulmonary Auscultation to Medical Students using a Virtual Patient Simulation Technology. *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 6032–6035. <https://doi.org/10.1109/EMBC44109.2020.9175920>
- Pettingil, B. F., & Tewes, F. R. (2020). Post COVID-19 Will Telemedicine replace Traditional Medicine? *Archives of Health Science*, 1–5. [https://doi.org/10.31829/2641-7456/ahs2020-4\(1\)-137](https://doi.org/10.31829/2641-7456/ahs2020-4(1)-137)
- Poulin, M.-A., Maldague, L. R., Daigle, A., & Gagnon, F. (2012). *NAT traversal in peer-to-peer architecture*. https://www.sce.carleton.ca/~fgagnon/Publications/NATTraversal_TR_SCE-12-04.pdf
- Price, R., & Tino, P. (2009). Still alive: Extending keep-alive intervals in P2P overlay networks. *2009 5th International Conference on Collaborative Computing: Networking, Applications and Worksharing*, 1–10. <https://doi.org/10.4108/ICST.COLLABORATECOM2009.8268>
- Qt Group. (s/f). *Qt / Tools for Each Stage of Software Development Lifecycle*. Recuperado el 13 de marzo de 2025, de <https://www.qt.io>
- Ramírez, F. (2017). *Diseño de un protocolo para la transmisión de audio cardio-pulmonar sobre redes heterogéneas con tasas de datos bajas/medias* [Tesis de Maestría, Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California]. Repositorio institucional. <http://cicese.repositorioinstitucional.mx/jspui/handle/1007/1273>
- Rescorla, E. (2018). *The Transport Layer Security (TLS) Protocol Version 1.3* (Request for Comments No. RFC 8446). Internet Engineering Task Force. <https://doi.org/10.17487/RFC8446>
- Rescorla, E., Tschofenig, H., & Modadugu, N. (2022). *The Datagram Transport Layer Security (DTLS) Protocol Version 1.3* (Request for Comments No. RFC 9147). Internet Engineering Task Force. <https://doi.org/10.17487/RFC9147>

- Rosenberg, J., Huitema, C., Mahy, R., & Weinberger, J. (2003). *STUN - Simple Traversal of User Datagram Protocol (UDP) Through Network Address Translators (NATs)* (Request for Comments No. RFC 3489). Internet Engineering Task Force. <https://doi.org/10.17487/RFC3489>
- Salud, C. N. de E. T. en. (2024). *Conceptos en Telesalud*. gob.mx. <http://www.gob.mx/salud|cenetec/acciones-y-programas/que-es-la-telesalud-y-la-telemedicina>
- Schuldt, H. (2009). Multi-Tier Architecture. En L. LIU & M. T. ÖZSU (Eds.), *Encyclopedia of Database Systems* (pp. 1862–1865). Springer US. https://doi.org/10.1007/978-0-387-39940-9_652
- Seah, J. J., Zhao, J., Wang, D. Y., & Lee, H. P. (2023). Review on the Advancements of Stethoscope Types in Chest Auscultation. *Diagnostics*, 13(9), Article 9. <https://doi.org/10.3390/diagnostics13091545>
- Sharpe, R., Warnicke, E., & Lamping, U. (2023). *Wireshark User's Guide: Version 4.5.0*. 373.
- Smith, M., & Hunt, R. (2002). Network security using NAT and NAPT. *Proceedings 10th IEEE International Conference on Networks (ICON 2002). Towards Network Superiority (Cat. No.02EX588)*, 355–360. <https://doi.org/10.1109/ICON.2002.1033337>
- Stewart, R. R., Tüxen, M., & Nielsen, K. (2022). *Stream Control Transmission Protocol* (Request for Comments No. RFC 9260). Internet Engineering Task Force. <https://doi.org/10.17487/RFC9260>
- Swarup, S., & Makaryus, A. N. (2018). Digital stethoscope: Technology update. *Medical Devices: Evidence and Research*, 11, 29–36. <https://doi.org/10.2147/MDER.S135882>
- Taleb, T., Hadjadj-Aoul, Y., & Samdanis, K. (2015). Efficient Solutions for Enhancing Data Traffic Management in 3GPP Networks. *IEEE Systems Journal*, 9(2), 519–528. *IEEE Systems Journal*. <https://doi.org/10.1109/JSYST.2013.2283957>
- Tanenbaum, A. S., & Steen, M. V. (2006). *Distributed systems: Principles and paradigms* (2nd ed.). Pearson Education. Inc.
- Tanenbaum, A. S., & Wetherall, D. J. (2012). *Redes de Computadora* (5th ed.). Prentice Hall. https://gc.scalahed.com/recursos/files/r161r/w25733w/redes_de_computadoras-freelibros-org.pdf
- Telehealth: What Is It, How to Prepare, Is It Covered?* (2020, agosto 26). National Institute on Aging. <https://www.nia.nih.gov/health/medical-care-and-appointments/telehealth-what-it-how-prepare-it-covered>
- Thinklabs Earbud Headphones—Deep Bass*. (s/f). Thinklabs. Recuperado el 6 de marzo de 2025, de <https://store.thinklabs.com/products/thinklabs-earbud-headphones-super-bass>
- Thinklabs Medical LLC. (2014). *Thinklabs One User's Manual*. Thinklabs Medical LLC.
- Tripathi, J., Oliveira, J. C. de, & Vasseur, J. P. (2012). *Performance evaluation of the routing protocol for low-power and lossy networks (RPL)* (Request for Comments No. RFC 6687). Internet Engineering Task Force. <https://doi.org/10.17487/RFC6687>
- Uberti, J. (2012). *WebRTC: Real-time Audio/Video and P2P in HTML5*. <https://docs.google.com/presentation/d/16vD-RApRli9W76z36hKe-B3jjS1R-J9qX4lpBq2JrHg/edit#slide=id.p19>

- Uscher-Pines, L., & Mehrotra, A. (2014). Analysis of Teladoc use seems to indicate expanded access to care for patients without prior connection to a provider. *Health Affairs (Project Hope)*, 33(2), 258–264. <https://doi.org/10.1377/hlthaff.2013.0989>
- VSee. (s/f). *VSee | Most Trusted HIPAA Compliant Telemedicine Solution*. Recuperado el 14 de marzo de 2025, de <https://vsee.com/>
- Wagner, J. (2015). *UDP hole punching in TomP2P for NAT traversal* [University of Zurich]. <https://files.ifi.uzh.ch/CSG/staff/bocek/extern/theses/BA-Jonas-Wagner.pdf>
- Ward, M. M., Jaana, M., & Natafji, N. (2015). Systematic review of telemedicine applications in emergency rooms. *International Journal of Medical Informatics*, 84(9), 601–616. <https://doi.org/10.1016/j.ijmedinf.2015.05.009>
- Webster, J. G. (2014). *The Physiological measurement handbook*. CRC Press. <https://www.perlego.com/book/1599879>
- WhatsApp. (s/f). *Better calling across desktop and mobile*. WhatsApp.Com. Recuperado el 21 de enero de 2025, de <https://blog.whatsapp.com/better-calling-across-desktop-and-mobile>
- Wilson, J. J. (1928). Rene Theophile Hyacinthe Laennec. *Journal of the National Medical Association*, 20(4), 174–177.
- Wing, D. (2010). Network Address Translation: Extending the Internet address space. *IEEE Internet Computing*, 14(4), 66–70. IEEE Internet Computing. <https://doi.org/10.1109/MIC.2010.96>
- Xu, J., Chen, Q., Zhang, Y., & Liu, S. (1989). Spectrum analysis of lung sounds. *Images of the Twenty-First Century. Proceedings of the Annual International Engineering in Medicine and Biology Society*, 1676–1677 vol.5. <https://doi.org/10.1109/IEMBS.1989.96401>
- Zhang, L. (2008). A retrospective view of network address translation. *IEEE Network*, 22(5), 8–12. IEEE Network. <https://doi.org/10.1109/MNET.2008.4626226>
- Zhang, Y. (2013). *An analysis on the application of P2P in network*. 841–844. <https://doi.org/10.2991/isccca.2013.212>
- Zocdoc. (2021). *Zocdoc | Find a Doctor Near You | Book Doctors Online*. <https://www.zocdoc.com/>
- Zoom Communications, Inc. (s/f). *Comprehensive Telehealth Platform for Providers*. Zoom. Recuperado el 14 de marzo de 2025, de <https://www.zoom.com/en/industry/healthcare/>
- Zoom Communications, Inc. (2023). *Zoom system requirements: Windows, macOS, Linux*. https://support.zoom.com/hc/en/article?id=zm_kb&sysparm_article=KB0060748

Anexos

Anexo A: Glosario

- **Socket.** Un socket es un punto de comunicación entre dos programas que se ejecutan en una red. Es la interfaz que permite la transmisión de datos entre un cliente y un servidor.
- **UDP.** El protocolo de datagrama de usuario (UDP) es un protocolo de comunicación que permite el envío de datos sin establecer una conexión previa, lo que lo hace rápido pero menos confiable.
- **TCP.** El protocolo de control de transmisión (TCP) es un protocolo orientado a la conexión que garantiza la entrega confiable de datos entre dos dispositivos en una red.
- **P2P.** El modelo Par-a-Par (P2P) es una arquitectura de red donde los nodos actúan como clientes y servidores simultáneamente, compartiendo recursos directamente entre ellos sin necesidad de un servidor central.
- **NAT.** La traducción de direcciones de red (NAT) es un proceso que permite a múltiples dispositivos en una red local compartir una única dirección IP pública para acceder a Internet.
- **HTTPS.** El protocolo de transferencia de hipertexto seguro (HTTPS, por sus siglas en inglés) es una versión segura de HTTP que utiliza cifrado SSL/TLS para proteger la comunicación entre el cliente y el servidor.
- **WebSocket.** WebSocket es un protocolo de comunicación que permite una conexión bidireccional y persistente entre un cliente y un servidor, ideal para aplicaciones en tiempo real.
- **DTLS.** El Protocolo de Seguridad de Datagramas de Transporte (DTLS, por sus siglas en inglés) proporciona seguridad para aplicaciones que utilizan UDP, ofreciendo cifrado y autenticación.
- **WS seguro.** WebSocket seguro (WS seguro) es una versión de WebSocket que utiliza cifrado TLS para garantizar la seguridad de las comunicaciones bidireccionales.

- **MUSHRA.** La evaluación de la Escala de Calificación de Sujeto Humano para Audio Multidimensional (MUSHRA) es un método para evaluar la calidad de audio percibida en experimentos controlados.
- **MOS.** El Puntaje de Opinión Media (MOS, de sus siglas en inglés) es una métrica utilizada para evaluar la calidad percibida de audio o video en comunicaciones, basada en la opinión de usuarios.
- **TLS.** Es un protocolo criptográfico que proporciona seguridad en las comunicaciones a través de redes. TLS garantiza la confidencialidad, integridad y autenticación de los datos transmitidos, y es ampliamente utilizado en aplicaciones como HTTPS, correo electrónico y VPNs.
- **Códec.** Un códec es un dispositivo o software que se utiliza para codificar y decodificar datos digitales, principalmente en formatos de audio, video o multimedia. Los codecs son esenciales para comprimir archivos grandes, facilitando su almacenamiento, transmisión y reproducción, mientras se busca mantener la mayor calidad posible.
- **Tele-auscultación.** Es una técnica médica que permite realizar la auscultación a distancia, mediante el uso de tecnologías de la comunicación.

Anexo B: Formulario comparativo de las plataformas de videoconferencia convencionales con el sistema EAC.

Este formulario fue diseñado con el propósito de recopilar información sobre la evaluación del sistema propuesto (Sistema EAC) en comparación con los sistemas de videoconferencia (Meet, Zoom, Teams) para el envío de audio cardiopulmonar, con la colaboración de personal de la salud.

- **Nombre del evaluador**

- **Ambiente de pruebas**
 - Silencioso
 - Ruido moderado
 - Ruidoso

- **¿Qué tan claro percibe los sonidos cardíacos y pulmonares?**
 1. Malo
 2. Pobre
 3. Regular
 4. Bueno
 5. Excelente

- **¿Qué tan claro considera que es el sonido transmitido en comparación con el sonido original del estetoscopio?**
 1. Malo
 2. Pobre
 3. Regular
 4. Bueno

5. Excelente

- **¿Cuánto ruido o artefactos percibe en el audio transmitido?**

1. Malo (mucho ruido)

2. Pobre (ruido notable)

3. Regular (ruido controlado)

4. Bueno (casi sin ruido)

5. Excelente (sin ruido)

- **¿Nota retrasos considerables en la transmisión del sonido?**

1. Malo (mucho retraso)

2. Pobre (retraso notable)

3. Regular (retraso regular)

4. Bueno (retraso mínimo)

5. Excelente (sin retraso)

- **¿Percibe interrupciones o cortes durante la transmisión?**

1. Malo (muchas interrupciones)

2. Pobre (interrupciones frecuentes)

3. Regular (interrupciones ocasionales)

4. Bueno (interrupciones mínimas)

5. Excelente (sin interrupciones)

- **¿Qué tan bien se distinguen los cambios de volumen en los sonidos cardíacos y pulmonares, como entre sonidos suaves y fuertes durante la auscultación?**

1. Malo
 2. Pobre
 3. Regular
 4. Bueno
 5. Excelente
- **¿Considera que el audio transmitido es de calidad diagnóstica?**
 1. Malo (No permite llevar a cabo un diagnóstico)
 2. Pobre (Permite un diagnóstico limitado, con dudas)
 3. Regular (Suficiente para diagnósticos básicos)
 4. Bueno (Adecuado para diagnósticos confiables)
 5. Excelente (Excelente calidad para diagnósticos confiables)
 - **¿Qué plataforma prefieres para analizar los sonidos cardíacos y pulmonares?**
 - Zoom
 - Google Meet
 - Sistema EAC (envío de audio cardiopulmonar)
 - Microsoft Teams
 - **¿Por qué prefieres la plataforma anterior?**
 - Facilidad de uso
 - Interfaz intuitiva
 - Mayor precisión en el análisis
 - Mayor confianza en los resultados

- Otro
- **Ordena las plataformas de mejor a peor según tu preferencia:**
 - 1. Zoom
 - 2. Google Meet
 - 3. Sistema EAC (envío de audio cardiopulmonar)
 - 4. Microsoft Teams
- **¿Qué mejorarías en las plataformas para el envío de audio cardiopulmonar?**
- **¿Recomendaría alguna de las plataformas para este propósito médico?**
 - Zoom
 - Google Meet
 - Sistema EAC (envío de audio cardiopulmonar)
 - Microsoft Teams