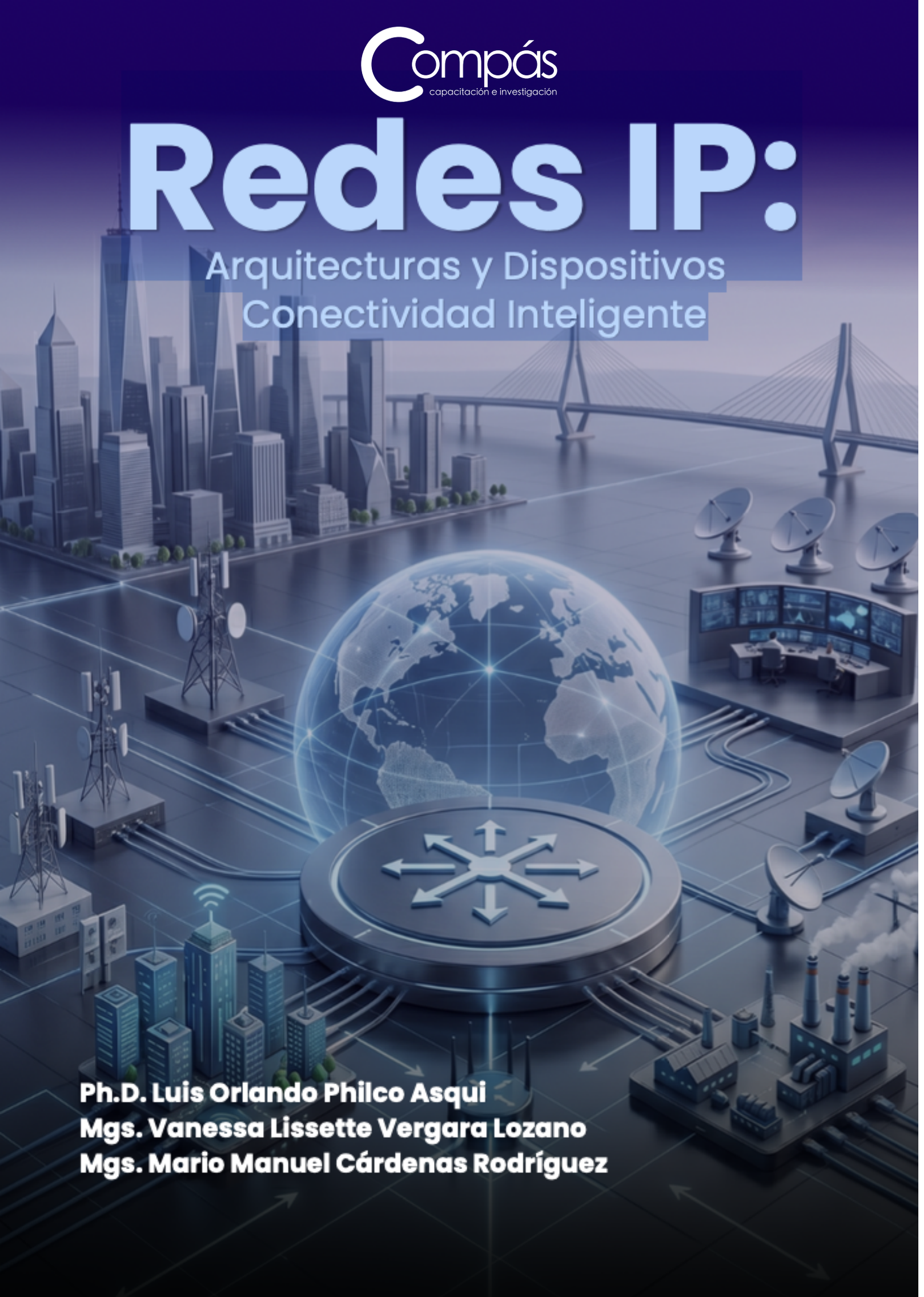
Compás
capacitación e investigación

Redes IP:

Arquitecturas y Dispositivos Conectividad Inteligente



Ph.D. Luis Orlando Philco Asqui
Mgs. Vanessa Lissette Vergara Lozano
Mgs. Mario Manuel Cárdenas Rodríguez

Redes IP: Arquitecturas y Dispositivos Conectividad Inteligente

Ph.D. Luis Orlando Philco Asqui
Mgs. Vanessa Lissette Vergara Lozano
Mgs. Mario Manuel Cárdenas Rodríguez

ISBN: 978-9942-53-199-5
Primera edición, 2026

Guayaquil, Ecuador
www.grupocompas.com
http://repositorio.grupocompas.com

Primera edición, 2026

Esta obra ha sido sometida a un proceso de evaluación bajo el sistema de arbitraje doble ciego (double-blind peer review), garantizando el anonimato tanto de los autores como de los evaluadores externos. El dictamen favorable certifica que el contenido cumple con los más altos estándares de rigor científico, calidad editorial y originalidad exigidos por la comunidad académica internacional para su indexación y reconocimiento científico.

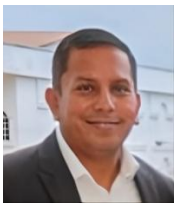
ISBN: 978-9942-53-199-5
DOI: 10.31876/9789942531995
Distribución online
Acceso abierto

Cita
Philco, L., Vergara, V., Cárdenes, M. (2026) *Redes IP: Arquitecturas y Dispositivos Conectividad Inteligente*. Editorial Grupo Compás.

Biografía de los Autores

LUIS ORLANDO PHILCO ASQUI
Ph. D de la Universidad Nacional de la Plata, Argentina.
Con experiencia en campos de Internet de las Cosas (IoT), Tecnologías de la Información y la Comunicación (TIC). Investigador en redes y telecomunicaciones con enfoque en IoT Industrial (IIoT), 5G, cómputo en el borde, redes de datos y algoritmos inteligentes distribuidos.
Docente tiempo completo en la UAE
lphilco@uagraria.edu.ec
Orcid: https://orcid.org/0000-0001-5312-3563

VANESSA LISSETTE VERGARA LOZANO
Ingeniera en Estadística e Informática y Magíster en Gestión de la Productividad y Calidad por la Escuela Superior Politécnica del Litoral (ESPOL). Con más de 15 años de experiencia en los sectores educativo, industrial y de telecomunicaciones. Proyectos de investigación aplicada orientados a la mejora de procesos y al análisis de datos. Autora y coautora de artículos científicos, con enfoque en ciencia de datos, productividad y calidad organizacional.
Docente tiempo completo en la Universidad Agraria del Ecuador (UAE).
Correo: vvergara@uagraria.edu.ec
ORCID: https://orcid.org/0000-0002-0644-5946



MARIO MANUEL CÁRDENAS RODRÍGUEZ
Máster en Docencia Superior Universitaria y Máster en Tecnologías de la Información con mención en Transformación Digital e Innovación. Diplomado en Investigación Educativa y en Derecho Digital y Legaltech. Ingeniero en Informática y Tecnólogo en Programación de Sistemas. Profesional con experiencia en Tecnologías de la Información y procesos de transformación digital, orientado a la formulación y evaluación de proyectos informáticos, de investigación educativa y agromáticos, así como a la gestión de procesos de inducción en el ámbito educativo.
Docente tiempo completo en la Universidad Agraria del Ecuador (UAE).
Correo: mcardenas@uagraria.edu.ec

INDICE GENERAL

ÍNDICE DE FIGURAS

INDICE DE TABLAS

PRÓLOGO

Las últimas décadas se han mejorado, reducido, pero sobre todo aumentado la capacidad de recursos de procesamiento y computo en los diversos dispositivos que se interconectan al internet. El paradigma del Internet de las cosas (IoT) involucran sistemas y servicios que interactúan sobre infraestructuras de comunicación cada vez más complejas. Sin embargo, la conectividad en sus inicios fue limitada a redes locales y servicios aislados y hoy se ha convertido en un pilar indispensable para la operación de empresas, industrias, ciudades y hogares. En este contexto, el presente libro constituye una contribución oportuna y rigurosa al entendimiento de las redes IP y de los mecanismos que permiten integrar arquitecturas, plataformas y dispositivos en un ecosistema de conectividad verdaderamente inteligente.

Los autores presentan una visión clara y estructurada que abarca desde las bases físicas de las redes fijas hasta las arquitecturas avanzadas de centros de datos, las tecnologías móviles de última generación y el universo creciente de dispositivos de bajo costo que conforman el Internet de las Cosas. Su enfoque es particularmente valioso porque no se limita a describir tecnologías, sino que las articula dentro de un marco conceptual donde convergen interoperabilidad, automatización, virtualización y diseño distribuido.

Este ecosistema tal como se expone en sus capítulos comprende un entorno en el que interactúan seres humanos, animales, máquinas, sensores y objetos cotidianos, todos ellos conectados mediante infraestructuras IP que permiten el intercambio continuo y eficiente de información. La interacción entre estos elementos redefine no solo la comunicación digital, sino también los modelos productivos, educativos y sociales.

La obra destaca por su equilibrio entre profundidad técnica y claridad expositiva. Cada capítulo combina fundamentos, arquitecturas y aplicaciones reales con una perspectiva actualizada sobre los desafíos y tendencias del sector. Esto convierte al libro en una referencia útil tanto para estudiantes como para docentes, ingenieros e investigadores que requieren comprender los principios que sustentan la conectividad moderna y su evolución hacia escenarios de inteligencia distribuida.

INTRODUCCIÓN

La conectividad digital y la convergencia entre redes fijas, móviles y dispositivos de bajo costo apoyados con los centros de datos de alta capacidad y si estas infraestructuras se agregan redes de sensores da conformación a un ecosistema hiperconectado capaz de operar con eficiencia, adaptabilidad y escalabilidad. Este libro ofrece una visión estructurada y profunda de esa evolución técnica, explicando cómo los distintos componentes de la red contribuyen a construir plataformas de comunicación más rápidas, seguras y autónomas.

El Capítulo 1 expone los esquemas de redes fijas y sus fundamentos tecnológicos, así desde los medios físicos como el cable de par de cobre, la fibra óptica y los estándares Ethernet ópticos que conforman la columna vertebral de las comunicaciones IP. El Capítulo 2 aborda los centros de datos como el núcleo de la conectividad moderna, donde convergen computación, almacenamiento. Se exponen las arquitecturas de alta densidad, enlaces de 10G a 400G, la virtualización de funciones (NFV), y los procesos de automatización y orquestación. El Capítulo 3 examina las redes móviles IP y las tecnologías de conectividad, movilidad, handover, Wi-Fi 6/7 e integración satelital así mismo tecnología de red celular desde 4G y 5G hasta evolución al 6G.

El Capítulo 4 presenta los dispositivos de bajo costo e IoT como Raspberry Pi, ESP32, Arduino y STM32 en comunicación con la arquitectura Edge–Fog–Cloud que pueden desplegar aplicaciones de red IP en salud, ciudades, agricultura, industria con conectividad inteligente, esto último por despliegue de esas capas cercanas al usuario para computo, almacenamiento y automatización. Por consiguiente, los autores de este libro ofrecen una base sólida para comprender fundamentos y las tendencias que darán forma a las redes hiperconectadas de los próximos años

CAPÍTULO 1: INFRAESTRUCTURAS EN REDES DE DATOS FIJAS

Se examina la evolución de las infraestructuras cableadas y centros de datos desde los componentes físicos, protocolos, modelos de comunicación y arquitecturas de servidores que conforman la base de la conectividad de datos fija. La infraestructura física de las redes LAN (Red de Área Local) y WAN (Red de Área Amplia) se compone de dispositivos de red, medios de transmisión y dispositivos finales, aunque los tipos específicos y la escala varían significativamente debido a las diferencias geográficas que cubren. Siendo las LANs locales (Ethernet, Wi-Fi) y las WANs globales, usando tecnologías como fibra óptica, satélites y enlaces arrendados para conectar múltiples LANs (ejemplo de WAN).

Las redes de datos fijos constituyen la columna vertebral de toda la comunicación digital contemporánea. A través de infraestructuras cableadas como hilos de cobre, fibra óptica o enlaces híbridos, se facilita el transporte de la información o datos con alta capacidad y velocidad este aspecto técnico permite la calidad en la funcionalidad de servicios empresariales, industriales y de nube. Estas redes no solo garantizan estabilidad, sino que además sirven como base física sobre la cual operan las arquitecturas inalámbricas, móviles y del internet de las cosas IoT que caracterizan la conectividad moderna.

En los últimos años, la transición tecnológica ha transformado el concepto de "red fija" de una estructura rígida y jerárquica hacia un entorno definido por software, programable y virtualizado. Los modelos de red tradicionales como las redes de área local (LAN. Local Area Network), metropolitana y amplia evolucionaron con la incorporación de arquitecturas SDN (Software Defined Networking) y NFV (Network Function Virtualization), permitiendo que las funciones de conmutación, ruteo y seguridad sean gestionadas desde plataformas centralizadas y flexibles.

1.1 INFRAESTRUCTURA FÍSICA Y COMPONENTES

Las infraestructuras de redes de datos y sistemas de comunicación constituyen el soporte físico y lógico que hace posible el intercambio continuo de información entre dispositivos, servidores y usuarios. Este ecosistema integra tanto elementos de hardware como enrutadores, conmutadores, paneles de parcheo, cables de cobre y fibra óptica. Su sistema operativo y aplicaciones informáticas como componentes de software y protocolos que organizan el flujo, la seguridad y la priorización de los datos.

En el ámbito de las redes de datos fijas, el medio cableado representa la espina dorsal de la conectividad. Aunque las redes inalámbricas y móviles han ganado protagonismo en las últimas dos décadas, la transmisión por cable sigue siendo insustituible en entornos donde la estabilidad, la baja latencia y la alta velocidad son requisitos críticos, como centros de datos, industrias, empresas o redes troncales de operadores. El diseño de una red cableada eficiente no solo depende del tipo de medio físico empleado, sino también de su organización estructurada y de la correcta aplicación de normas internacionales (TIA[1] 568, ISO/IEC[2] 11801, ANSI/TIA-942).

Así también de estándares como ISO/IEC 11801 que especifica el cableado genérico para instalaciones del cliente, mientras que ANSI/TIA-942 se enfoca en los requisitos de infraestructura para centros de datos. TIA/EIA (como TIA-568) proporcionan directrices para el diseño e instalación de sistemas de cableado estructurado en general. Estos estándares

[1] Asociación de Industrias de Telecomunicaciones (TIA) definen las normas para el cableado estructurado, permitiendo que infraestructuras de diferentes fabricantes sean compatibles entre sí.

[2] Organización Internacional de Normalización (ISO) y la Comisión Electrotécnica Internacional (IEC), que establecen especificaciones y buenas prácticas para tecnologías, productos y servicios (especialmente en TI y electrónica), asegurando calidad, seguridad e interoperabilidad global

garantizan compatibilidad, seguridad eléctrica y continuidad operacional entre los distintos componentes de la red.

Este preámbulo da lugar al concepto de la interoperabilidad en redes de datos, en especial la interoperabilidad técnica la cual se basa en estándares internacionales que garantizan que equipos de distintos fabricantes puedan comunicarse correctamente entre sí. Estos estándares son definidos principalmente por organizaciones como IEEE, ISO, IETF[3] y ITU-T[4].

1.1.2 INTEROPERABILIDAD TECNICA EN INFRAESTRUCTURA DE RED

La interoperabilidad técnica se refiere a la capacidad de componentes heterogéneos —equipos de red, sistemas operativos, servicios y plataformas de gestión— para intercambiar información y operar de forma coordinada bajo interfaces y protocolos comunes, sin requerir cambios estructurales en su diseño interno. En infraestructuras de redes fijas, este principio es clave para integrar soluciones multi-fabricante, habilitar la convergencia de servicios (datos, voz, video y gestión), y sostener escalabilidad y continuidad operativa mediante estándares abiertos en las distintas capas del modelo de red (física, enlace, red, transporte y aplicación).

1.1.3 PROTOCOLOS DE COMUNICACIÓN

En una red corporativa, pueden coexistir servicios de aplicación como HTTP/HTTPS, DNS, SMTP/IMAP y NTP (todos protocolos de capa de aplicación), mientras que la operación del plano de red se apoya en protocolos de control y señalización como ICMP y ARP/ND (IPv6), además de mecanismos de

---

[3] Grupo de trabajo de ingeniería de Internet, desarrollan los estándares técnicos (conocidos como RFCs) que hacen funcionar Internet

4 Unión Internacional de Telecomunicaciones (UIT) de la ONU, que crea estándares técnicos (Recomendaciones) para la conectividad global

enrutamiento y segmentación definidos por el diseño (por ejemplo, VLAN 802.1Q y, cuando aplica, protocolos de enrutamiento dinámico). Esta coexistencia se sostiene sobre un plano común IP (IPv4/IPv6) y tecnologías de acceso y agregación como Ethernet (IEEE 802.3), permitiendo integrar servicios empresariales junto con políticas de QoS (calidad de servicio) y priorización de tráfico en la misma infraestructura.

1.1.4 FORMATOS DE DATOS

Los formatos de datos definen cómo se representa la información para su transmisión e interpretación. El uso de formatos estandarizados reduce ambigüedades semánticas y facilita la integración entre plataformas. En gestión y observabilidad de redes, la interoperabilidad depende además de modelos de información para describir configuración, estado y telemetría; por ejemplo, YANG (lenguaje de modelado) junto con codificaciones como XML/JSON, o mecanismos de exportación de métricas, registros y eventos operativos. Esta estandarización permite que un NMS/Controlador consolide inventario, alarmas y desempeño de equipos multi-fabricante sin acoplarse a implementaciones propietarias.

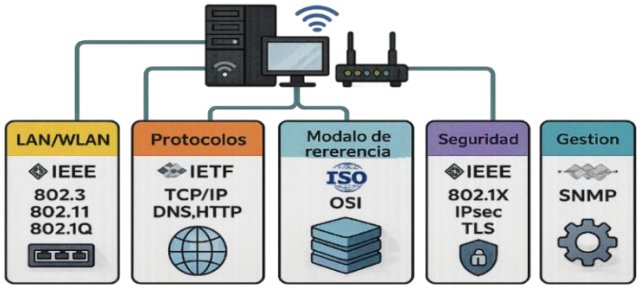
1.1.5 ESTÁNDARES TÉCNICOS

Los estándares establecen conjuntos de reglas y buenas prácticas que aseguran compatibilidad, seguridad y calidad en el intercambio de información. Así las normas ISO/IEC relacionadas con redes, seguridad de la información y gestión de servicios guían el diseño de infraestructuras interoperables, permitiendo que equipos de distintos proveedores se integren bajo criterios comunes de direccionamiento, transporte y control de acceso.

La figura 1 se sintetiza los elementos que hacen posible la interoperabilidad en una infraestructura de red basada en estándares. En el nivel de acceso, las

tecnologías LAN/WLAN definidas por IEEE 802.3 (Ethernet), IEEE 802.11 (Wi-Fi) y IEEE 802.1Q (VLAN) proporcionan los mecanismos de conectividad y segmentación del tráfico.

Figura 1. Representación de interoperabilidad mediante estandarizaciones



Sobre esta base operan los protocolos definidos por la IETF (Grupo de Trabajo de Ingeniería de Internet), donde la suite TCP/IP (Protocolo de Control de Transmisión/Protocolo de Internet) permite la comunicación extremo a extremo y servicios de aplicación como DNS (Sistema de Nombres de Dominio) y HTTP (Protocolo de Transferencia de Hipertexto) soportan la operación de sistemas y aplicaciones empresariales.

El modelo de referencia OSI (Interconexión de Sistemas Abiertos) de ISO (Organización Internacional de Normalización) actúa como marco conceptual para organizar las funciones de comunicación en capas, facilitando el diseño, la interoperabilidad y la integración de equipos de distintos fabricantes. De forma complementaria, la interoperabilidad también incorpora mecanismos de seguridad, como IEEE 802.1X (Control de acceso a la red basado en puertos) y protocolos de protección de datos como IPsec (Seguridad del Protocolo de Internet) y TLS (Seguridad de la Capa de Transporte), así como capacidades de gestión y monitoreo mediante SNMP (Protocolo Simple de Administración de Red), que permiten la supervisión, configuración y operación centralizada de la infraestructura.

Así, estos estándares conforman el fundamento técnico para redes escalables, seguras y administrables en entornos corporativos. Por ejemplo, en una red empresarial es posible integrar switches de Cisco, equipos de acceso inalámbrico de Aruba (HPE) y routers de borde de MikroTik o Juniper, todos operando sobre Ethernet (IEEE 802.3) e IP (Protocolo de Internet). A nivel de servicios, los dispositivos pueden autenticarse mediante 802.1X (Control de acceso a la red basado en puertos), ser monitoreados desde un sistema de gestión central utilizando SNMP (Protocolo Simple de Administración de Red) y transportar aplicaciones corporativas a través de HTTP/HTTPS (Protocolo de Transferencia de Hipertexto/Protocolo Seguro de Transferencia de Hipertexto). Esta interoperabilidad permite que equipos de distintos fabricantes funcionen de manera coordinada dentro de una misma infraestructura sin depender de soluciones propietarias.

1.2 MEDIOS FÍSICOS ETHERNET

Los medios físicos de transmisión constituyen la capa más baja de la arquitectura de red y son responsables de transportar señales eléctricas u ópticas entre dispositivos Hoy en día siguen siendo la columna vertebral de las redes de datos, incluso en entornos inalámbricos y tecnología de redes móviles como 5G. Los medios físicos de transmisión cableados se clasifican principalmente en par trenzado de cobre y fibra óptica, cada uno con propiedades físicas, eléctricas y ópticas distintas que determinan su uso. Como criterios de selección del medio físico cableado este puede depender de:

- Distancia del enlace
- Velocidad requerida
- Ambiente electromagnético
- Escalabilidad futura
- Presupuesto disponible

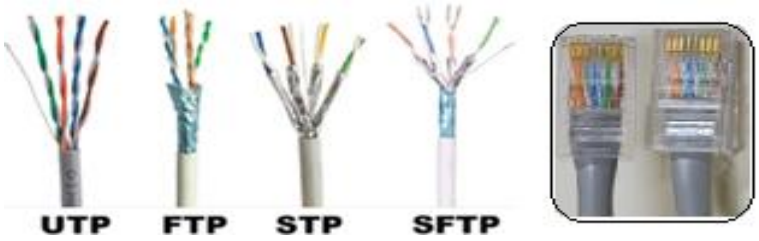
1.2.1 PAR TRENZADO DE COBRE

Es un tipo de medio de transmisión que tiene 4 pares de hilo de cobre y utilizan para transmitir señales tanto digitales como analógicas. Se detallan cuatro clases:

1. UTP (Unshielded Twisted Pair) par trenzado sin blindaje consisten en cables de cobre codificados por colores. Tiene cuatro pares de hilos dentro de su cubierta o chaqueta. Cada par está trenzado con un número diferente de giros (trensado) para ayudar a eliminar la interferencia de pares adyacentes y otros dispositivos eléctricos.
2. STP (Shielded Twisted Pair): Es un cable par trenzado blindado, contiene una envoltura de papel de aluminio adicional o una cubierta trenzada de cobre para ayudar a proteger las señales del cable de las interferencias. Este blindaje adicional evita que las interferencias electromagnéticas se filtren dentro o fuera del cable, lo que lo hace ideal para entornos con alta interferencia eléctrica.
3. FTP (Foiled Twisted Pair): Par trenzado con pantalla global, cuenta con pares que no están apantallados, pero sí tiene un blindaje de aluminio global o completo que mejora la protección o bloqueo de interferencia externa y así evitar que interrumpan las señales transmitidas a través de los cables (FS, 2021).
4. SFTP (Screened Foiled Twisted Pair): Par trenzado totalmente blindado o apantallado completo. Se basa en la construcción del cable FTP, pero en el apantallamiento global se le ha añadido una malla metálica de Baja Emisión de Humo y Cero Halógenos (LSZH, Low Smoke Zero Halogen) alrededor para aumentar el aislamiento de este cable y sobre todo producir mínimas emisiones de humo y no tóxicas al exponerse al fuego. Esto es una

opción notable en entornos donde la seguridad contra incendios es primordial.

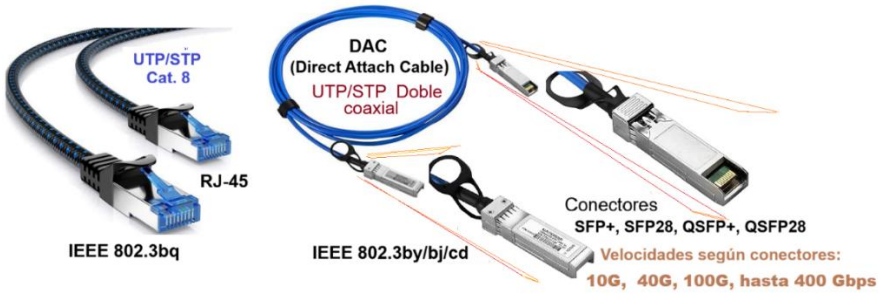
Figura 2. Distintos tipos de cables de par de cobre trenzados



Fuente: Adaptado de (Lazarus, 2021)

Se indica además que existen soluciones de cobre optimizadas para enlaces de muy corto alcance en centros de datos (estándar IEEE 802.3bq) así se tiene al cable UTP categoría 8 (Cat 8) con velocidades máximas de 40 Gbps (40GBASE-T) en distancias de hasta 30 metros. Así mismo con un tipo de cable de par tranzado, pero de solo 2 hilos de cobre (DAC, Direct Attach Cable). Denominados también como cable doble coaxial integra los transceptores directamente en los extremos del cable, eliminando la necesidad de módulos ópticos independientes. Son ampliamente utilizados para conectar puertos SFP+, SFP28, QSFP+ y QSFP28 en equipos de red modernos ya que pueden transmitir datos a velocidades de 10 Gbps, 25 Gbps, 40 Gbps, 50 Gbps y hasta 400 Gbps.

Figura 3. Cable Categoría 8 y DAC



Fuente: Adaptado de (Angolu, 2024) y (Afril, 2024)

Además, los cables DAC se pueden fabricar en varias longitudes, como 1 m, 3 m, 5 m, 7 m y 10 m, para adaptarse a diferentes configuraciones de red.

El cable DAC de la figura 3 es totalmente compatible con el conector SFP Plus MSA (acuerdo de múltiples fuentes).

Características técnicas del cable de par trenzado según su categoría, ancho de banda (MHz), velocidad máxima (Gbps) y aplicaciones, se muestran en la tabla 1.

Tabla 1. Características técnicas de las categorías de cable de par trenzado

Categoría	Ancho de banda	Velocidad máxima	Alcance típico	Aplicación
Cat 5e	100 MHz	1 Gbps	100 m	Oficinas pequeñas y medianas
Cat 6	250 MHz	1–10 Gbps	55–100 m	Redes empresariales
Cat 6A	500 MHz	10 Gbps	100 m	Centros de datos
Cat 7/7A	600–1000 MHz	10 Gbps	100 m	Instalaciones críticas, entornos industriales.
Cat 8	2000 MHz	25–40 Gbps	30 m	Enlaces troncales de alta densidad

Nota: Los cables de pares trenzados se emplean en tramos de distribución y acceso, por su flexibilidad, costo reducido y facilidad de instalación, aunque presentan limitaciones de interferencia electromagnética y pérdida de señal (atenuación) en distancias mayores a 100 metros.

En la implementación de cuartos de equipos, centros de datos y redes empresariales, se despliega el cableado estructurado siguiendo las normas EIA/TIA-568A y EIA/TIA-568B, las cuales estandarizan el orden de los pares trenzados en los conectores RJ-45 para garantizar compatibilidad, desempeño eléctrico y certificación del sistema. Estas normas permiten identificar la configuración de los cables directo, cruzado y *rollover* (consola) según el tipo de conexión requerida:

- Cable directo: conecta dispositivos de distinto tipo (PC a switch o router).
- Cable cruzado: conecta dispositivos del mismo tipo (PC a PC, switch a switch).
- Cable *rollover*: para configuración de equipos de red por el puerto de consola.

Figura 4. Cables de red y conexiones de cableado estructurado

Fuente: (Conectividad, 2022)

Hoy en día, los switches administrables y equipos de alto rendimiento incorporan la función Auto-MDI/MDIX, la cual permite detectar automáticamente el tipo de conexión, admitiendo el uso de cables directos en la mayoría de los casos. No obstante, en infraestructuras certificadas se mantiene el cumplimiento estricto de las normas EIA/TIA-568 para asegurar la calidad y trazabilidad del cableado.

1.2.2 FIBRA ÓPTICA

Es un medio de transmisión que utiliza un núcleo de vidrio o plástico para enviar información en forma de pulsos de luz a altas velocidades y a grandes distancias. Esta tecnología ofrece una mayor capacidad de transmisión y una menor latencia en comparación con el cableado tradicional. Se ha consolidado en despliegue de redes troncales y de backbone debido a su alta capacidad de transmisión, inmunidad al ruido electromagnético y bajo nivel de atenuación.

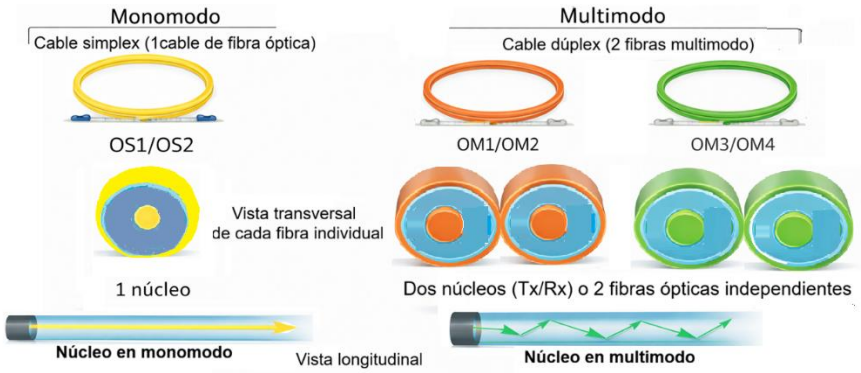
Existen dos grandes tipos de fibra óptica:

1. Fibra multimodo (OM1–OM5): para distancias cortas (hasta 2 km) en redes LAN y centros de

datos. Las distancias varían según la velocidad de la conexión y el tipo de cable (OM3, OM4, etc.).

2. Fibra monomodo (OS1–OS2): puede transmitir hasta 40 km sin amplificación y hasta cientos de kilómetros con el uso de amplificadores y repetidores. Minimiza la dispersión y atenuación de la señal, lo que la hace ideal para transmisiones de datos de alta velocidad y larga distancia, como en redes de telecomunicaciones y conexiones a internet. (Villacís & Guamán, 2022).

Figura 5. Tipos de cables de Fibra Óptica



Fuente: Adaptado de (Fibconect, 2024)

La fibra monomodo se caracteriza por tener un núcleo muy estrecho, típicamente de 8 a 10 µm, lo que permite que la señal luminosa se propague en un solo modo, reduciendo la dispersión y posibilitando transmisiones de larga distancia y alta capacidad. En contraste, la fibra multimodo posee un único núcleo de mayor diámetro, generalmente de 50 µm o 62.5 µm, que permite múltiples trayectorias de la luz dentro de la fibra, facilitando la transmisión en enlaces cortos, aunque con mayores limitaciones en distancia y ancho de banda debido a la dispersión modal. (RackOnline, 2023).

Tabla 2. Características técnicas de la fibra óptica multimodo

Tipo	Núcleo / Revestimiento	Longitud de onda (nm)	Ancho de banda en función de distancia	Alcance típico	Aplicación principal
OM1	62.5/125 µm	850	200 MHz·km	275 m a 1Gbps	Instalaciones heredadas
OM2	50/125 µm	850	500 MHz·km	550 m a1Gbps	LAN estándar
OM3	50/125 µm	850	2000 MHz·km	300 m a10Gbps	Data centers
OM4	50/125 µm	850	4700 MHz·km	400 m a10Gbps 150 m a 100Gbps	Alta densidad
OM5	50/125 µm (Wideband MMF)	850–953	4700 MHz·km	150 m a 400Gbps	Transmisión por multiplexación de longitud de onda (SWDM, Short Wavelength Division Multiplexing)

1.2.3 CONECTORES PARA FIBRA OPTICA

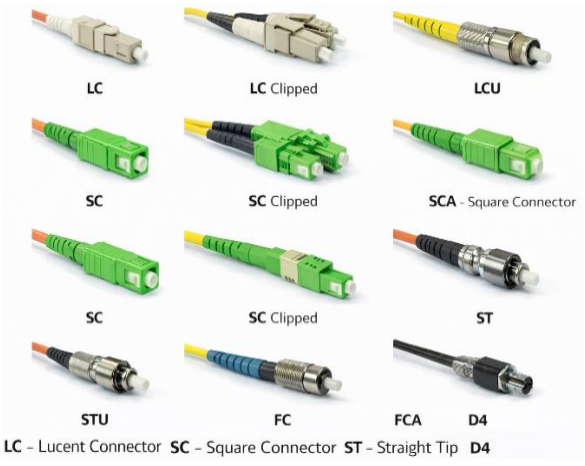
Los cables de fibra óptica necesitarán de conectores en sus extremos, que permitan una conexión y desconexión más rápidas que el empalme. Se utilizan en redes de telecomunicaciones, centros de datos y aplicaciones de internet en el hogar para la conexión entre elementos de transmisión y recepción. El conector enlaza mecánicamente el

cable de fibra, asegurando la continuidad y la eficiencia de la transmisión de datos a través de pulsos de luz (Albarrán, 2024).

Los conectores LC (Lucent Connector) y SC (Square Connector) dominan el mercado debido a su eficiencia, facilidad de uso y diseño compacto, especialmente en centros de datos modernos y aplicaciones de fibra hasta el hogar (FTTH) y otros que ya están poco utilizados como ST, FC y D4 este último de la corporación japonesa NEC (Nippon Electric Company). Se consideran conectores heredados (legacy connectors) y su uso ha disminuido considerablemente debido a la adopción de conectores más compactos y eficientes como el LC y el SC.

Un caso de conectores heredados puede ser en un campus universitario con un backbone de fibra instalado hace 15–20 años, es común encontrar un ODF (distribuidor óptico) y patch cords con conectores ST o FC en enlaces entre edificios.

Figura 6. Tipos de conectores básicos para fibra óptica



Fuente: (López, 2017)

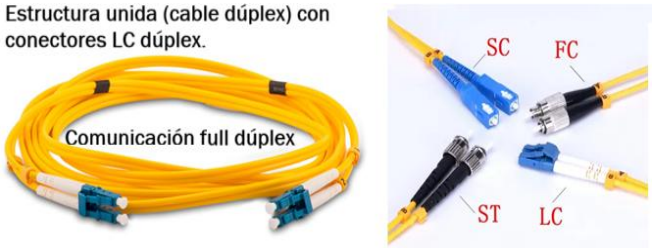
Si se moderniza el *core* con switches y módulos ópticos actuales, típicamente los nuevos enlaces y paneles se implementan con LC (y en planta

externa/FTTH con SC/APC), por lo que durante un período de transición se usan latiguillos híbridos (por ejemplo, ST–LC o FC–LC) o adaptadores en el *patch panel* para mantener servicio mientras se migra gradualmente la infraestructura.

En los conectores vigentes, existen configuraciones simplex y dúplex según la cantidad de fibras que terminan y la forma de transmisión. Un conector simplex termina una sola fibra, por lo que transporta una única portadora óptica; se utiliza típicamente cuando el enlace opera en un solo hilo por dirección mediante técnicas como BiDi (bidireccional sobre una sola fibra), o en aplicaciones específicas de instrumentación.

En redes de datos corporativas y de centros de datos, lo más común es el uso dúplex, donde se agrupan dos fibras (una para transmisión y otra para recepción), permitiendo comunicación bidireccional simultánea (full-duplex). Por esta razón, muchos conectores de uso cotidiano se presentan en formato LC dúplex o SC dúplex, facilitando el manejo de pares TX/RX en patch cords y paneles de parcheo.

Figura 7. Cable de fibra óptica full dúplex



Fuente: (Dekam, 2025)

1.2.4 CONVERSIÓN ÓPTICO-ELÉCTRICA

El proceso de conversión de señal eléctrica a haz de luz/óptica E/O y viceversa O/E es realizado por los transceptores ópticos que permiten que los equipos electrónicos como switches, routers y servidores transmitan y reciban señales a través de enlaces de fibra óptica. En un enlace óptico, la información no

viaja como señal eléctrica sino como un haz de luz modulada.

El proceso de transmisión implica tres etapas fundamentales:

- Conversión eléctrica-óptica: el transmisor, normalmente un láser (LD) o diodo emisor de luz (LED), convierte la señal eléctrica en pulsos luminosos.
- Propagación óptica: la luz viaja a través del núcleo de la fibra por reflexión interna total, minimizando la pérdida de señal (atenuación inferior a 0,2 dB/km en fibras monomodo).
- Conversión óptica-eléctrica: en el receptor, un fotodiodo (PIN o APD) convierte nuevamente la luz en señal eléctrica que puede ser interpretada por el transceptor o chipset del dispositivo.

1.2.5 CONVERTIDOR DE MEDIOS

Un convertidor de medios es un dispositivo que permite la interconversión entre distintos medios físicos de transmisión, principalmente cobre y fibra óptica, y en algunos modelos específicos entre fibra multimodo y fibra monomodo. Su uso es especialmente relevante en entornos de red donde el cableado Ethernet de cobre no puede cubrir la distancia requerida, ya sea por limitaciones físicas, interferencias electromagnéticas o restricciones de alcance (100 m), haciendo necesario el uso de fibra óptica para extender el enlace.

Figura 8, Convertidor de medios

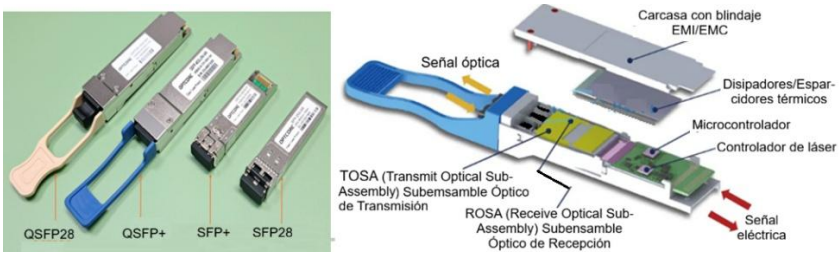
Fuente: (FiberMall, 2023)

Asimismo, los convertidores de medios desempeñan un papel clave en escenarios de acceso y agregación, ya que facilitan la conexión de la última milla de enlaces de fibra óptica hacia redes de área metropolitana (MAN) y redes exteriores, permitiendo la integración entre infraestructuras de acceso basadas en cobre y backbones ópticos de mayor capacidad y alcance.

1.2.6 TRANSCEPTORES OPTICOS

Los transceptores ópticos son módulos compactos que convierten las señales eléctricas generadas por los equipos de red (switches, routers, servidores) en señales ópticas, y viceversa. Albergan semiconductores láser y fotodiodo. Su diseño se integrar en las ranuras SFP (Small Form-factor Pluggable) de switches y equipos de red. En el mercado se encuentran módulos transceptores SFP, SFP+, QSFP, entre otros más.

Figura 9. Transceptores Ópticos



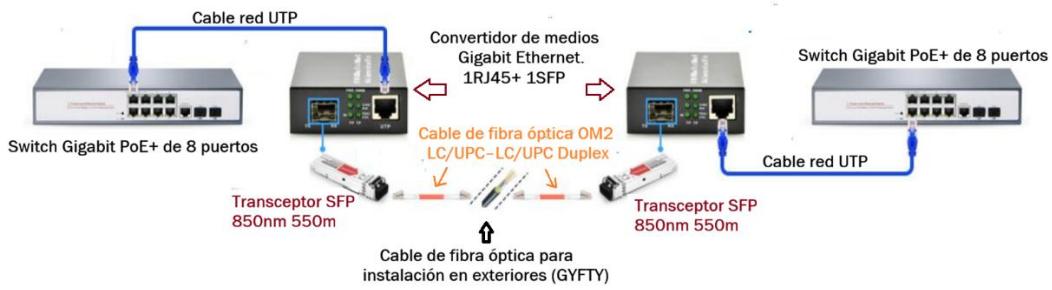
Fuente: Adaptado de (Opcore, 2024)

Los transceptores ópticos se clasifican también por su alcance óptico:

- SR (Short Reach): multimodo (850 nm, hasta 400 m).
- LR (Long Reach): monomodo (1310 nm, hasta 10 km).
- ER / ZR: monomodo (1550 nm, hasta 40–80 km).
- Estos utilizan modulación PAM-4 (Pulse Amplitude Modulation 4-level) y estándares IEEE 802.3bs/ck para Ethernet con tasas de 400 Gbps y 800 Gbps.

- Existen variantes DAC (Direct Attach Copper) y AOC (Active Optical Cable) que integran el transceptor en un solo cable para distancias cortas (≤ 7 m).

Figura 10. Conexión típica del Media Converter



Fuente: (Optico, 2025)

La figura 10 muestra un escenario típico de interconexión de redes Ethernet mediante fibra óptica, utilizado para extender la distancia de transmisión más allá del límite del cable UTP (100 m). En ambos extremos se emplean switches Gigabit, conectados localmente a través de cable UTP a un convertidor de medios Ethernet (RJ45 + SFP). Estos convertidores integran transceptores SFP multimodo de 850 nm, los cuales permiten la conversión de la señal eléctrica Ethernet a señal óptica.

La interconexión entre ambos convertidores se realiza mediante un cable de fibra óptica multimodo OM2 LC/UPC–LC/UPC dúplex, que soporta enlaces de hasta 550 m. Este tipo de arquitectura es común en campus universitarios, edificios separados, redes industriales y enlaces de acceso, donde se requiere mayor alcance, inmunidad a interferencias electromagnéticas y estabilidad, manteniendo compatibilidad con infraestructuras Ethernet existentes.

La tabla 3 detalla los formatos, entendidos como el conjunto estandarizado de características físicas, mecánicas y funcionales de los transceptores ópticos Ethernet

Tabla 3. Formatos de transceptores ópticos utilizados en redes Ethernet

Formato	Velocidades típicas	Tipo de conector óptico	Densidad y aplicación	Ejemplo de estándar IEEE
SFP (Small Form-factor Pluggable)	1 Gbps	LC duplex	LAN y acceso empresarial	IEEE 802.3z (1000BASE-LX/SX)
SFP+	10 Gbps	LC duplex	Enlaces troncales o servidores 10G	IEEE 802.3ae (10GBASE-SR/LR/ER)
SFP28	25 Gbps	LC duplex	Servidores 25G y enlaces de agregación	IEEE 802.3by (25GBASE-SR/LR)
QSFP+	40 Gbps (4×10 Gbps)	MTP/MPO	Data centers y agregación 40G	IEEE 802.3ba (40GBASE-SR4/LR4)
QSFP28	100 Gbps (4×25 Gbps)	MTP/MPO o LC	Backbones 100G, redes spine-leaf	IEEE 802.3bm / 802.3bj
QSFP56	200 Gbps (4×50 Gbps PAM-4)	MTP/MPO	Enlaces de IA o HPC	IEEE 802.3cd
QSFP-DD / OSFP	400 Gbps (8×50 Gbps PAM-4)	MTP/MPO o LC duplex	Supercentros de datos y nube	IEEE 802.3bs / 802.3ck
CFP / CFP2 / CFP4	100–400 Gbps	LC duplex	Redes troncales de telecomunicaciones	IEEE 802.3ba / 802.3bs

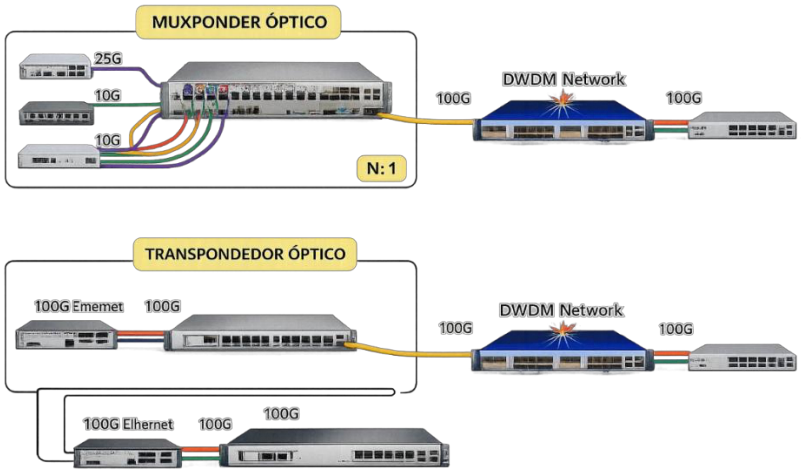
1.2.7 MUXPONDER Y TRANSPONDEDOR

En las comunicaciones por fibra óptica, los muxponders pueden combinar múltiples servicios en una única longitud de onda/enlace ascendente mediante el protocolo OTN de la UIT para mapear los servicios en el mismo enlace ascendente. Un muxponder puede maximizar la capacidad de la fibra al reducir el número de longitudes de onda

necesarias para el transporte de datos, lo que permite un uso más eficiente de las fibras y las hace ideales para el crecimiento futuro de la red. Los operadores, proveedores de servicios de internet (ISP) y otras industrias suelen adoptar muxponders, ya que ofrecen características de nivel de operador, como monitorización y gestión remotas flexibles, herramientas de diagnóstico de enlaces y monitorización bidireccional del rendimiento de las interfaces de servicio del cliente y el enlace ascendente.

Un transpondedor es un equipo (o tarjeta en un chasis) que recibe una señal "cliente" (normalmente eléctrica o óptica corta) y la convierte a una señal óptica "de línea" (line-side) adecuada para transporte en una red óptica (DWDM) (Villacís & Guamán, 2022). En otras palabras, es un tipo de tarjeta o módulo que agrega múltiples señales de baja velocidad en un único canal óptico de alta velocidad (muxponder) o convierte una señal eléctrica en una señal óptica de alta velocidad (transpondedor). Esto permite que el equipo de la red transmita hasta terabits por segundo a través de un solo par de fibras ópticas, multiplicando enormemente la capacidad de la red.

Figura 11 Conexión de Muxponder y Transpondedor



Fuente: Adaptado de (Carvajal, 2023))

Se muestran dos enfoques de transporte óptico en redes DWDM: el transpondedor óptico y el muxponder óptico. Desde el punto de vista técnico, en la parte inferior se representa un transpondedor óptico, cuya función principal es adaptar un único flujo cliente de alta velocidad (por ejemplo, 100G Ethernet) a un canal óptico DWDM de línea. Este dispositivo realiza la conversión eléctrico–óptica, aplica técnicas como FEC y modulación coherente, y transmite el flujo como una sola longitud de onda hacia la red DWDM. Este enfoque 1:1 es ideal cuando se dispone de enlaces dedicados de alta capacidad, como interconexiones directas entre centros de datos (DCI), enlaces troncales empresariales o backbone de operadores con tráfico homogéneo y elevado.

En contraste, la parte superior de la figura 10 muestra un muxponder óptico, cuya característica distintiva es la agregación de múltiples flujos cliente de menor velocidad (por ejemplo, varios enlaces de 10G y 25G) en un único canal óptico de mayor capacidad (100G). Este proceso, conocido como multiplexación o grooming, optimiza el uso del espectro óptico y reduce el número de longitudes de onda necesarias en la red DWDM. El muxponder resulta especialmente adecuado en escenarios donde existen múltiples servicios dispersos como campus universitarios, redes metropolitanas o accesos empresariales que deben concentrarse eficientemente hacia un nodo central o backbone.

En términos de casos de uso, el transpondedor se emplea cuando se requiere simplicidad, baja latencia y enlaces dedicados de alta capacidad, mientras que el muxponder se utiliza para consolidar tráfico, mejorar la eficiencia espectral, reducir costos operativos y facilitar la escalabilidad de la red. En conjunto, la figura evidencia cómo ambos dispositivos son complementarios dentro de una arquitectura de transporte óptico moderna, y cómo su selección depende del perfil de tráfico, la

topología de red y los objetivos de eficiencia y crecimiento a largo plazo.

1.3 MODULACIÓN ÓPTICA EN REDES DE DATOS

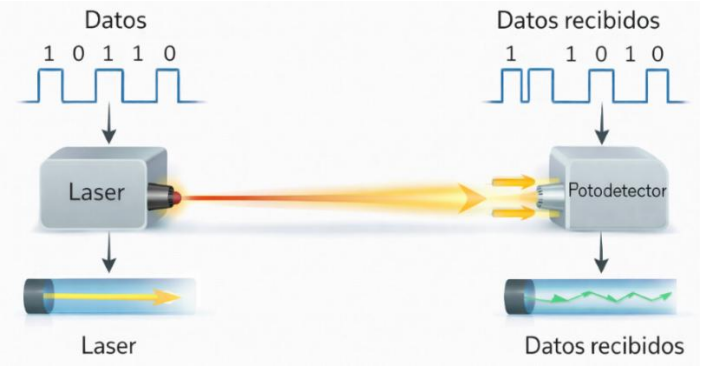
En el contexto de las redes de datos sobre fibra óptica, la portadora es una onda luminosa generada por un láser, y la modulación permite representar información digital mediante variaciones ópticas que pueden propagarse eficientemente a través del medio físico. Este proceso posibilita que los datos digitales se transformen en una señal óptica adecuada para transmisión, considerando las limitaciones y características de la fibra, como la atenuación, la dispersión y el ruido. En consecuencia, la modulación constituye un elemento clave para determinar la velocidad, el alcance y la eficiencia espectral de los enlaces ópticos modernos.

Desde esta perspectiva, la modulación no solo cumple la función básica de habilitar la transmisión de datos, sino que se convierte en un factor determinante en la evolución de las redes ópticas, especialmente en escenarios de alta capacidad como Ethernet de 100G y 400G, interconexión de centros de datos y redes DWDM.

1.3.1 MODULACIÓN DIRECTA EN ETHERNET ÓPTICO

La modulación directa ha sido el mecanismo predominante en los enlaces Ethernet ópticos de corta y media distancia, especialmente en redes LAN, campus y centros de datos. En este enfoque, la información digital se transmite mediante la activación y desactivación directa de la fuente láser, generando pulsos de luz que representan los valores binarios.

Figura 12. Modulación Directa



Fuente: Adaptado de (Adarsh, KartHa, & Santhosh, 2022)

Esta modulación funciona como un esquema de modulación en banda base de tipo No Retorno a Cero (NRZ), en el cual la información se transmite variando la amplitud de la señal óptica. Se implementa comúnmente mediante OOK (On–Off Keying), donde el bit "1" se representa por la presencia de luz y el bit "0" por la ausencia de emisión óptica.

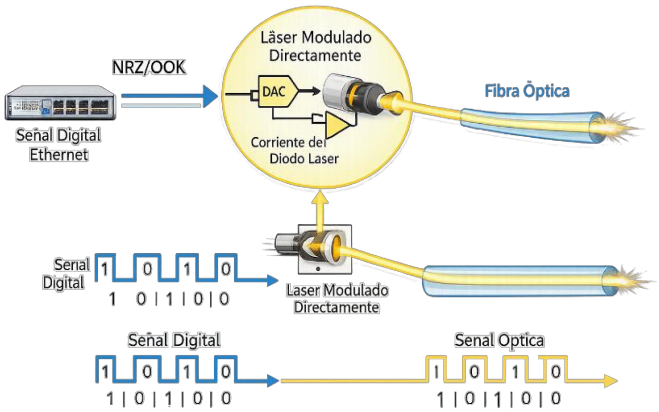
La simplicidad del proceso permite implementar transceptores ópticos de bajo costo, baja latencia y reducido consumo energético, lo que ha favorecido su adopción masiva en estándares como 1000BASE-SX/LX, 10GBASE-SR/LR y variantes de Ethernet óptico tradicional. No obstante, la modulación directa presenta limitaciones físicas claras, principalmente relacionadas con la dispersión cromática, la atenuación y la sensibilidad al ruido, lo que restringe su alcance efectivo y su escalabilidad a mayores velocidades.

Por esta razón, su uso se concentra en distancias que van desde decenas de metros hasta algunos kilómetros, donde el entorno controlado de la fibra permite mantener la integridad de la señal sin mecanismos avanzados de compensación (ETU-LINK, 2024).

1.3.2 MODULACIÓN PAM-4

El crecimiento sostenido del tráfico de datos en centros de datos, redes de agregación y backbones empresariales impulsó la necesidad de incrementar la capacidad de transmisión sin ampliar el ancho de banda óptico disponible. En este contexto surge PAM-4 (Pulse Amplitude Modulation con cuatro niveles) como una evolución directa de la modulación binaria tradicional. A diferencia de NRZ, que utiliza dos niveles de señal para representar un bit por símbolo, PAM-4 emplea cuatro niveles de amplitud, permitiendo codificar dos bits por símbolo. Este enfoque duplica la tasa de bits transmitida sobre el mismo canal óptico, sin requerir un aumento proporcional en la frecuencia de operación ni en el número de fibras. (FiberMall, 2021)

Figura 13.Modulación directa en Ethernet Óptico



Fuente: Adaptado (Albarrán, 2024)

Se aprecia en la figura 13, que la información se codifica mediante cuatro niveles distintos de amplitud óptica, permitiendo transportar 2 bits por símbolo (00, 01, 10, 11). A diferencia de NRZ/OOK, que usa solo dos niveles, PAM-4 duplica la eficiencia espectral sin aumentar el ancho de banda del canal, razón por la cual es ampliamente utilizada en Ethernet de alta velocidad (50G, 100G, 200G y 400G) sobre fibra óptica. Su principal desventaja es una menor tolerancia al ruido, lo que exige mejores condiciones de SNR, ecualización y control preciso del láser. (Villacís & Guamán, 2022).

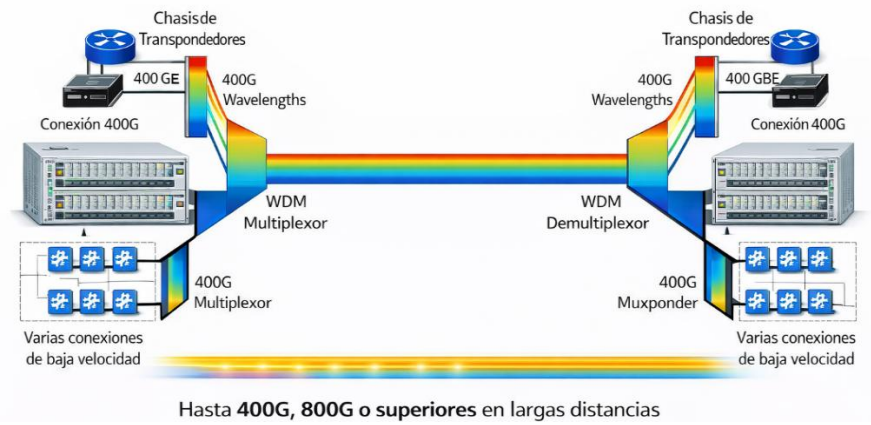
Desde la perspectiva de infraestructura, PAM-4 ha sido determinante para la adopción de Ethernet de 100G, 200G y 400G, especialmente en transceptores como QSFP28, QSFP56 y QSFP-DD, al duplicar la eficiencia espectral sin incrementar el ancho de banda. No obstante, la reducción del margen entre niveles de amplitud incrementa la sensibilidad al ruido y a las imperfecciones del canal, lo que obliga a incorporar corrección de errores avanzada (FEC) y electrónica de alta precisión (FiberMall, 2021).

En la práctica, esto supone un compromiso entre capacidad, alcance y complejidad, razón por la cual PAM-4 se emplea principalmente en enlaces de corta y media distancia, típicos de centros de datos y redes de agregación.

1.3.3 MODULACIÓN COHERENTE EN REDES DWDM

Es una tecnología avanzada que permite la transmisión de datos a velocidades extremadamente altas (como 400G, 800G o superiores) sobre distancias de miles de kilómetros. La modulación coherente representa el enfoque más avanzado en la transmisión óptica y es fundamental en redes DWDM de media y larga distancia, como backbones metropolitanos, interurbanos y enlaces de interconexión entre centros de datos. A diferencia de la modulación directa, este esquema no se limita a la amplitud de la señal, sino que modifica simultáneamente la fase, la frecuencia y la amplitud de la onda luminosa.

Figura 14. Modulación Coherente en redes DWDM



Fuente: Adaptado de (FiberMall, 2022)

En la figura 14 se visualizan múltiples flujos de datos de alta velocidad (por ejemplo, 400 GbE) son generados en chasis de transpondedores ópticos y convertidos en longitudes de onda coherentes individuales. Estas longitudes de onda se multiplexan mediante un WDM Multiplexor y se transportan simultáneamente sobre una única fibra óptica, maximizando la capacidad del enlace. En el extremo opuesto, un WDM Demultiplexor separa nuevamente las longitudes de onda, que son procesadas por transpondedores coherentes para entregar los datos a los equipos de red. Esta arquitectura permite alcanzar 400G, 800G o velocidades superiores en largas distancias, siendo típica de backbones metropolitanos, interurbanos y redes de interconexión entre centros de datos, donde se prioriza alta capacidad, eficiencia espectral y escalabilidad (Lanchi, 2020).

Este enfoque permite el uso de formatos de modulación de alto orden, como QPSK y 16-QAM, incrementando significativamente la eficiencia espectral. La modulación coherente se apoya en procesamiento digital de señal (DSP) para compensar dinámicamente efectos físicos de la fibra, como la dispersión cromática y la variación de polarización, extendiendo el alcance efectivo del enlace. Desde el punto de vista arquitectónico,

constituye la base de los transpondedores ópticos coherentes y módulos DCO empleados en sistemas DWDM modernos. Aunque implica mayor complejidad y consumo energético, su capacidad para transportar grandes volúmenes de datos sobre una sola longitud de onda la convierte en una tecnología esencial para redes ópticas de transporte de alta capacidad (Pérez, 2024).

1.4 ESTANDAR ETHERNET CON FIBRA OPTICA

El estándar de protocolo de comunicación para la LAN es el IEEE 802.3 o Ethernet, esta también es una tecnología que vincula software y/o hardware entre sí, y transmite y recibe datos a través de un medio físico de transmisión. Una red bajo este estándar conecta varias computadora, switches o conmutadores y otros dispositivos como impresoras, escáneres, etc. Una red cableada emplea cables de cobre tipo par trenzado o cables de fibra óptica. Así la IEEE 802.3 evolucionó para cubrir necesidades de red sobre fibra óptica, clasificándose principalmente por su alcance y tecnología de transmisión:

IEEE 802.3z (Short Reach, Corto alcance): Utilizan fibra multimodo (MMF) y láseres de bajo costo (VCSEL). Ejemplos comunes incluyen 10GBASE-SR, 40GBASE-SR4 y 100GBASE-SR4, (velocidades de 10, 40 y 100 giga bit por segundos) diseñados para centros de datos con alcances de hasta 300-400 metros.

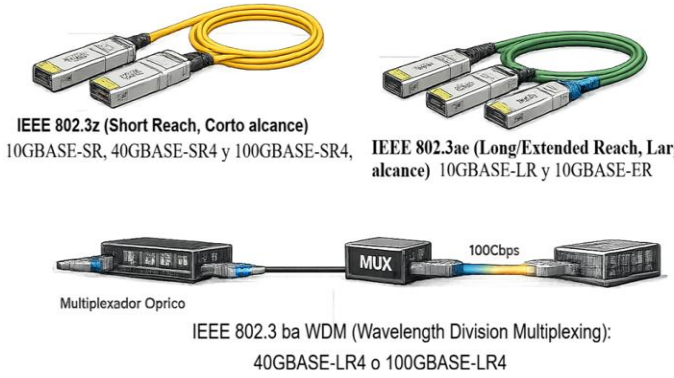
IEEE 802.3ae (Long/Extended Reach, Largo alcance): Emplean fibra monomodo (SMF) para distancias mayores. Estándares como 10GBASE-LR (10 km) y 10GBASE-ER (40 km) son pilares en redes metropolitanas y campus extensos.

IEEE 802.3 ba WDM (Wavelength Division Multiplexing): Para maximizar el ancho de banda, estándares como 40GBASE-LR4 o 100GBASE-LR4 multiplexan 4 longitudes de onda en un solo par de fibras, permitiendo altas velocidades en infraestructuras existentes. (Franco, 2023).

IEEE 802.3bm: Si bien el 802.3ba definió LAN-WDM, el estándar 802.3bm y los acuerdos multi-fuente (MSA) impulsaron el uso de CWDM4 (multiplexación por división de longitud de onda de 4 canales) alcanza una velocidad agregada de 100 Gbps. Así (CWDM4 MSA) es una alternativa más económica al estándar 100GBASE-LR4 en centros de datos.

IEEE 802.3ct DWDM: Para ultra largo alcance con interconexiones de velocidad a 100Gbps, es decir; 100GBASE-ZR y utiliza modulación coherente para alcanzar hasta 80 km sin amplificación óptica activa. Define la operación sobre sistemas DWDM (multiplexación por división de longitud de onda densa).

Figura 15, Estándares Ethernet sobre fibra óptica



Un proyecto a futuro son IEEE 802.3db y 802.3df que se basan en el uso de múltiples "carriles" de 100 Gbps para lograr mayores velocidades. Por ejemplo, 400G se puede lograr utilizando cuatro pares de fibra. 800 Gbps sobre fibra óptica utilizará ocho pares de fibra multimodo (hasta 100 m) o monomodo (hasta 2 km). Así también el proyecto IEEE P802.3dj, que utilizarán tecnologías de señalización de 200 Gbps por carril para lograr estas densidades y velocidades más altas (Franco, 2023).

Hay que comprender que un enlace Ethernet de alta velocidad rara vez utiliza un solo hilo de fibra para transportar toda la información. En su lugar, divide el flujo de datos en múltiples canales paralelos o "carriles" (comúnmente 4 u 8 pares de fibra) que trabajan simultáneamente. Así la señalización se

refiere a la velocidad a la que viaja la información a través de cada uno de estos carriles o canales individuales y a la tecnología o método de modulación utilizado para lograr esa velocidad.

1.5 ARQUITECTURA DE RED

Es la forma en que los servicios y dispositivos de red se estructuran juntos para responder a las necesidades de conectividad o la infraestructura de TI utilizada para proporcionar conectividad entre usuarios, objetos y aplicaciones dentro de las redes empresariales encontramos diferentes tipos de redes en función de su tamaño, ancho de banda y caudal, extensión y método de transmisión; o elementos que comunican o su alcance. La configuración de una red requiere una serie de componentes de hardware para garantizar la conectividad y el buen funcionamiento de la red como router, conmutador, puntos de acceso inalámbrico, cables de red, firewall, servidores, inversores o sistemas de energía, paneles de conexión y patchbays y armarios de red (Albarrán, 2024).

Por aquello a la hora de construirla adecuadamente se han de tener en cuenta:

Tolerancia a fallos

Es aquella que limita el número de dispositivos afectados cuando ocurren errores, ya que Internet puede fallar en ciertos momentos. Esta característica permite recuperarse rápidamente mediante redundancia y el uso de múltiples rutas entre el origen y el destino; si una ruta falla, otra asume el tráfico. Tanto la infraestructura física como los mecanismos lógicos de encaminamiento están diseñados para adaptarse a estas condiciones. Ejemplo: si se corta un enlace principal, el protocolo de enrutamiento redirige el tráfico por un enlace alterno sin interrumpir el servicio.

Escalabilidad

Es aquella que puede expandirse para dar soporte a nuevos clientes y aplicaciones sin degradar el rendimiento del servicio entregado a los usuarios existentes. Dado que cada día se incorporan nuevos usuarios y proveedores a Internet, la infraestructura debe admitir nuevas interconexiones mediante un diseño jerárquico por capas (físico y lógico), lo que facilita la gestión eficiente de recursos y el balanceo de carga. Este enfoque permite integrar crecimiento en cada capa sin causar disrupciones, aumentando capacidad de transporte y desempeño de los componentes. Además, los mecanismos de direccionamiento e identificación dentro de una internetwork (interconexión de varias redes) permiten ajustar la operación a la demanda. Ejemplo: al incorporar un nuevo segmento de red, se agrega un switch de acceso y se amplía el enrutamiento sin afectar a los usuarios ya conectados.

Seguridad

Es un parámetro fundamental al configurar una arquitectura de red, se delinea tanto en la infraestructura como en la información transportada; por ello, es necesario proteger físicamente los equipos y enlaces, y también resguardar los datos en tránsito y en acceso. Además, las redes de telecomunicaciones deben cumplir requerimientos legales y normativos para operar con estructuras seguras y confiables, garantizando confidencialidad e integridad, ya que se intercambia información empresarial y personal de carácter crítico. Esta exposición atrae a cibercriminales y obliga a aplicar controles de autenticación, segmentación y cifrado. Ejemplo: separar usuarios y servidores en VLANs y habilitar VPN para accesos remotos reduce el riesgo de intrusiones y fuga de datos.

Calidad

La calidad de servicio (QoS) es otro requisito clave, especialmente en entornos multicloud. Los servicios deben ser confiables, medibles y, en ciertos casos, garantizados, evitando que el desempeño se degrade por congestión. Esto implica aplicar mecanismos para clasificar y priorizar tráfico, administrar colas, controlar pérdidas y asignar ancho de banda según el tipo de aplicación. Ejemplo: priorizar voz y videoconferencia frente a descargas masivas reduce latencia y "jitter" en horas pico.

Gestión de red

Comprende las funciones para controlar, planificar, asignar, implementar y coordinar los recursos de la red. Además, integra y supervisa las demás capacidades (seguridad, disponibilidad, rendimiento y calidad), con el objetivo de asegurar que el acceso y los flujos de datos se administren de forma consistente y eficiente. Esto implica monitoreo, configuración, control de cambios y resolución de incidencias para mantener la operación estable. Ejemplo: un sistema NMS (Network Management System, sistema de gestión de red) genera alertas ante caída de enlaces y facilita aplicar ajustes sin interrumpir el servicio.

Figura 16. Representación de arquitecturas de red

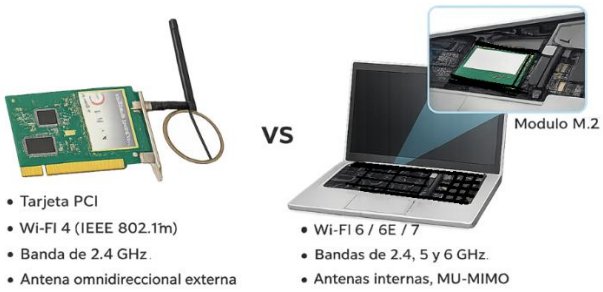
Fuente: Adaptado de (Pululart, 2020)

1.5.1 LA TARJETA DE RED

La tarjeta de red NIC (Network Interface Card) permite el envío y recepción de datos, ya sea por cable (Ethernet) o de forma inalámbrica (Wi-Fi), y posee una dirección única (MAC) para identificarse en la red, funcionando como una interfaz física y electrónica para la comunicación. Esta tarjeta era una placa de circuito impreso que se insertaba en una ranura de expansión de la placa base (como ISA, luego PCI, y más tarde PCIe). El usuario instalaba físicamente la tarjeta, conectaba el cable de red a su puerto trasero y, a menudo, instalaba controladores de software específicos para que el sistema operativo la reconociera. Ahora la NIC de equipos portátiles está en desuso por un chip LAN integrado dentro de la placa principal (mainboard) o placa base que tiene un puerto RJ45.para conectividad para equipos cableados no dedicados (PCs y laptops).

Cuando la comunicación es de forma inalámbrica se emplea la interfaz M.2 Wi-Fi. El M.2 "M punto dos" facilita la optimización del rendimiento, la densidad y la eficiencia energética, gracias a su compatibilidad con múltiples interfaces como PCI Express (PCIe), SATA y USB, lo que le permite adaptarse a diversas aplicaciones desde unidades de estado sólido (SSD) de alta velocidad hasta módulos de comunicación inalámbrica o aceleradores de hardware.

Figura 17. Tarjeta NIC vs modulo M.2



Fuente: Adaptado de (Quraishi, Tavakoli, & Ren, 2021)

En tabla 5, se muestra la comparativa entre estas dos tarjetas de red inalámbricas

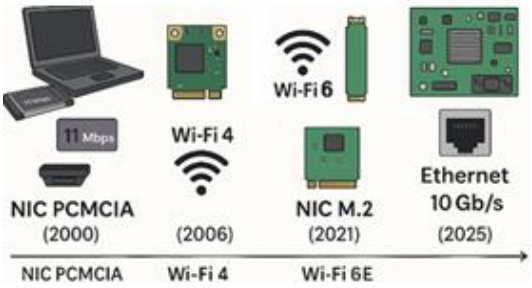
Tabla 4. Comparativa entre NIC y M.2 para red de datos inalámbrica

Aspecto	Antes (802.11b/g)	Ahora (Wi-Fi 5/6/6E/7)
Frecuencia	2,4 GHz	2,4 GHz + 5 GHz + 6 GHz
Velocidad típica	11–54 Mbps	1 Gbps (Wi-Fi 5), hasta 46 Gbps (Wi-Fi 7)
Tipo de tarjeta	PCI / PCMCIA interna	Módulos integrados (M.2, mini-PCIe) o adaptadores USB-C
Antena	Externa omnidireccional	Interna tipo MIMO (2×2, 4×4) con beamforming
Compatibilidad	Limitada a Windows XP–7	Compatible con Windows 11, Linux, macOS
Seguridad	WPA	WPA3, OWE, SAE (robusta frente a ataques)
Eficiencia	Alto consumo y latencia	OFDMA, MU-MIMO y Target Wake Time reducen consumo

NOTA: Tecnología **MIMO**: Multiple-Input Multiple-Output/Múltiple Entrada Múltiple Salida Protocolos: **WPA3**: Wi-Fi Protected Access/Acceso protegido versión 3 para Wi-Fi **OWE**: Opportunistic Wireless Encryption/Cifrado Inalámbrico Oportunista. **SAE**: Simultaneous Authentication of Equals/Autenticación Simultánea de Iguales. **OFDMA**: Orthogonal Frequency Division Multiple Access/ Acceso Múltiple por División de Frecuencia OrtogonaL

Técnicamente en laptops: la NIC conecta Wi-Fi y Bluetooth y se integran en un conector M.2 Key E, esto es una ranura física (un zócalo o slot) en la placa base para que se inserte una tarjeta de expansión, que suele ser una tarjeta inalámbrica (Wi-Fi y Bluetooth). Se indica que desde 2018 la organización a Wi-Fi Alliance introdujo la nomenclatura simple (Wi-Fi 4, 5, 6, etc.) para así facilitar su uso.

Figura 18. Evolución de la NIC



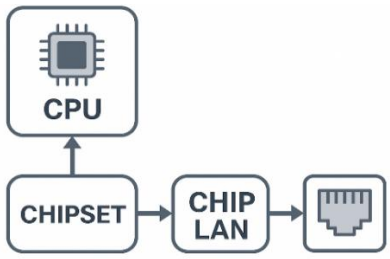
El chip de red local cableada (LAN) y el conector M.2 Wi-Fi (red inalámbrica) son métodos de conectividad diferentes pero complementarios para acceder a una red o a Internet.

- El chip LAN gestiona la conexión a través de un cable Ethernet (conexión física, generalmente más estable y rápida).
- El módulo M.2 Wi-Fi permite la conexión de forma inalámbrica (a través de ondas de radio).

Ambos sistemas pueden coexistir en un mismo dispositivo, y el usuario puede elegir cuál utilizar en función de sus necesidades.

El Chip LAN actúa como controlador de red Ethernet viene soldado (integrado) en la placa base que maneja la conexión a internet por cable (puerto RJ45). Como se aprecia en la figura 6 el chipset es parte de la tarjeta principal del dispositivo (PC, laptop, impresora, etc.) se conecta al procesador a través de interfaces como PCIe, USB o SoC, luego para conectarse por ejemplo a un switch se emplea cable de red.

Figura 19. Representación actual de la NIC integrado en chip LAN



Fuente: Autores

En redes fijas (no inalámbricas) las arquitecturas actuales (usan un PCH o Platform Controller Hub de Intel, o las soluciones de AMD), por consiguiente, el chipset actúa como un centro de control y un "puente" de comunicación. En la tabla 6 se muestran características del chipset

Tabla 5, Características del chipset

Componente	Función	Hardware de fabricante
Chipset de red LAN	Conecta hosts con la red física. Controla las velocidades y la comunicación de capa física (PHY).	Intel I219, Realtek RTL8125
Chipset de comunicación	Actúa como "puente" (PCH) para coordinar y controlar buses internos críticos como PCIe, USB y Thunderbolt, gestionando el flujo de datos entre la CPU y los periféricos.	Intel Z790, AMD X670
Controlador Wi-Fi / BT	Proporciona conectividad inalámbrica, gestionando los estándares de radiofrecuencia (802.11ax/ac/n) y Bluetooth.	Intel AX211, Qualcomm FastConnect
Chipset del switch / router	Diseñado para la conmutación y el enrutamiento de alto rendimiento en redes empresariales y centros de datos, implementando funciones avanzadas como QoS, VLANs y VXLAN.	Broadcom Trident (serie para centros de datos, hasta 400 GbE), Marvell Prestera (serie para acceso y agregación de campus)
Chipset óptico / transceptor	Módulos o cables que permiten la comunicación a través de fibra óptica o cobre de alta velocidad, convirtiendo señales eléctricas en ópticas.	SFP+ (10 Gbps), QSFP (40 Gbps o más), DAC (cables de cobre de conexión directa)

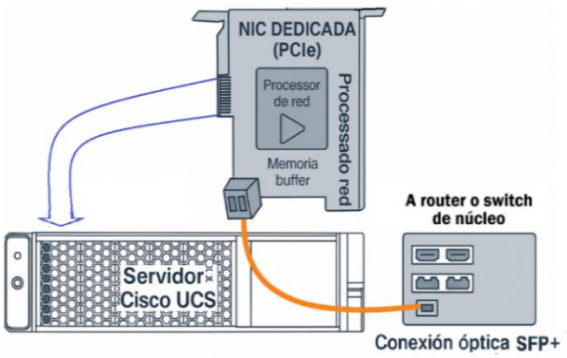
El controlador LAN integrado es un chip de red independiente soldado a la placa base que brinda conectividad Ethernet nativa; aunque no pertenece al chipset principal (PCH/AMD), se apoya en él como puente para comunicarse con la CPU y la memoria. En equipos estándar suele ofrecer un conjunto básico de funciones, normalmente sin procesador de red especializado ni buffers avanzados, lo cual es suficiente para uso doméstico y de oficina. En cambio, las NIC dedicadas están orientadas a entornos empresariales, servidores y centros de datos, donde se requieren mayor rendimiento, redundancia y capacidades avanzadas, además de la descarga de tareas del procesador principal para una operación de red más eficiente y robusta.

1.5.2 NIC COMO TARJETA DE RED DEDICADA

En entornos que requieren alto rendimiento y fiabilidad continua, como servidores, estaciones de trabajo potentes o equipos de red dedicados (por ejemplo, routers y switches empresariales), se debe instalar una NIC o un módulo físico específico para conseguir las funciones de red necesarias. Estos equipos requieren una mayor capacidad, redundancia o funciones avanzadas que no están disponibles en un chip integrado estándar.

A nivel empresarial, las infraestructuras modernas utilizan tarjetas de red dedicadas multigigabit, con velocidades de 2.5G, 10G, 25G o superiores, que proporcionan un rendimiento mucho mayor que las interfaces integradas básicas.

Figura 20. NIC de expansión para equipos dedicados



Fuente: Adaptado de (Ribeiro, 2019)

Estas NICs dedicadas suelen instalarse en ranuras de expansión PCI Express (PCIe), lo que permite reemplazarlas o actualizarlas fácilmente según las necesidades de rendimiento o el tipo de red (Ethernet, fibra óptica, etc.). A diferencia de las antiguas tarjetas PCMCIA utilizadas en portátiles, las NIC PCIe son componentes internos permanentes diseñados para un funcionamiento continuo y estable dentro del chasis del servidor, y no son extraíbles en caliente ni portátiles sino componentes internos permanentes diseñados para un funcionamiento continuo y estable dentro del chasis del servidor.

En la tabla 6 se muestran equipos con tarjetas de red NIC dedicados

Tabla 6. Equipos y sus usos que emplean NICs dedicadas

Tipo de equipo	Uso	Ejemplo de NIC dedicada
Servidor de archivos o virtualización	Maneja tráfico masivo de datos o máquinas virtuales (VMs).	Broadcom NetXtreme 10 GbE, Intel X710, Mellanox ConnectX-6 (25/40 GbE).
Firewall o Gateway empresarial	Requiere múltiples interfaces para segmentar redes.	Intel i350-T4 Quad Port, Silicom PE310G4I71.

Appliance de red (NAS, IDS/IPS, Router dedicado)	Procesa paquetes o almacenamiento en tiempo real.	QNAP LAN-10G2T, Chelsio T520-CR.
Servidor HPC (High Performance Computing)	Demanda baja latencia y altísimo ancho de banda.	NVIDIA ConnectX-7, Solarflare SFN8522 (100 GbE).
Estación de control industrial o SCADA	Conecta redes OT y IT, requiere fiabilidad y determinismo.	Intel i210, Realtek RTL8125B, módulos 2.5G.

Nota: Appliance de red es un equipo dedicado (hardware + software) diseñado para cumplir una función específica de red de forma estable y administrable, normalmente como "caja cerrada" (consola web/CLI, actualizaciones controladas y recursos propios) IDS/IPS (detección/prevención de intrusiones): un Snort sensor dedicado (o un appliance con Suricata/Snort) inspeccionando tráfico en un enlace espejo (SPAN) o en línea para detectar y/o bloquear amenazas. Ej. De router dedicado: ISR 1100) como gateway de sucursal: enrutamiento, VPN, QoS y políticas de acceso hacia Internet/MPLS.

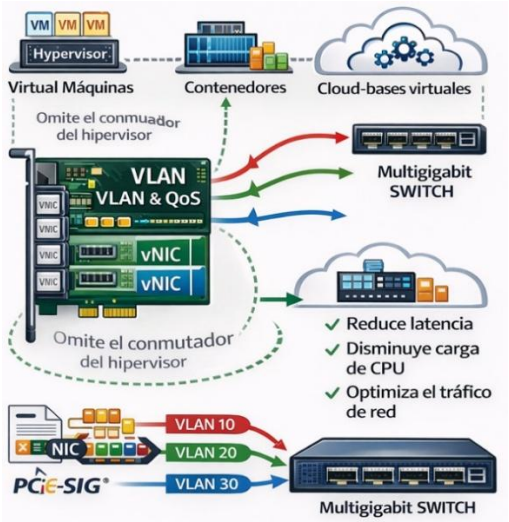
1.5.3 CARACTERÍSTICAS TÉCNICAS DE NICS DEDICADAS

Las NIC dedicadas están diseñadas para entornos de alto rendimiento, particularmente en centros de datos modernos, donde la conectividad de red debe sostener cargas de trabajo virtualizadas y tráfico intensivo con baja latencia. A diferencia de interfaces convencionales orientadas únicamente al envío y recepción de paquetes, estas NIC incorporan capacidades de aceleración y procesamiento en hardware que optimizan el plano de datos, reducen la sobrecarga del sistema operativo y mejoran la eficiencia del host.

Una característica central es el soporte SR-IOV (Single Root I/O Virtualization, Virtualización de Entrada/Salida de raíz única), especificación promovida por PCI-SIG (PCI Special Interest Group, Grupo de Interés Especial de PCI). La virtualización

SR-IOV divide una NIC física en múltiples interfaces virtuales, conocidas como vNIC (virtual NIC) implementadas mediante funciones virtuales asignables a diferentes VM o contenedores. En un escenario tradicional, el tráfico de las máquinas virtuales atraviesa un conmutador virtual gestionado por el hipervisor, lo que consume CPU y añade latencia. Con SR-IOV, cada función virtual se presenta como un dispositivo PCIe independiente y puede asociarse directamente a una VM o contenedor, reduciendo la participación del hipervisor en el encaminamiento de paquetes y logrando un rendimiento cercano al nativo. Esta aproximación mejora el aislamiento al separar recursos de memoria e interrupciones por función virtual y permite consolidar servidores sin degradar la conectividad.

Figura 21. NICs dedicadas con soporte SR-IOV, VLAN y QoS por hardware



Fuente: Adaptado de (Liu, Guo , Jiang, Song, & Zhang, 2025)

En la figura 21 se observa el efecto de SR-IOV en el plano de datos, al permitir que interfaces virtuales se asignen a máquinas virtuales con mínima intervención del hipervisor.

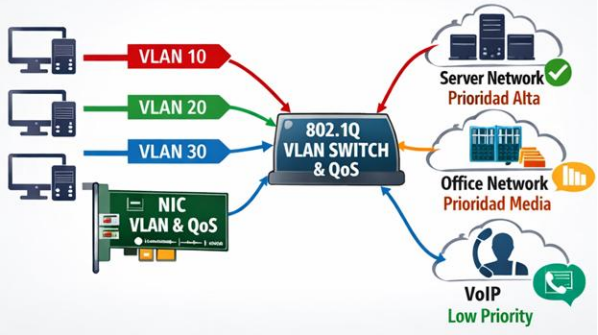
En paralelo, las NIC dedicadas integran mecanismos de descarga de procesamiento o offloading,

convirtiéndose en coprocesadores especializados para tareas de red. Entre las funciones más habituales se encuentran TSO (TCP Segmentation Offload, Descarga de segmentación TCP), donde el host entrega un bloque de datos grande y la NIC realiza la segmentación conforme al MTU (Maximum Transmission Unit, Unidad Máxima de Transmisión); Checksum Offload (Checksum Offload, Descarga de suma de comprobación), que calcula y valida sumas de verificación en transmisión y recepción; y LRO (Large Receive Offload, Descarga de recepción ampliada), que agrega segmentos entrantes para disminuir interrupciones y llamadas al stack de red del sistema operativo.

Estas optimizaciones elevan el throughput y disminuyen el consumo de CPU, aunque en equipos que realizan inspección fina de paquetes, como routers o firewalls, ciertas agregaciones pueden no ser convenientes y deben configurarse cuidadosamente. Otra capacidad relevante es la gestión de tráfico por hardware mediante VLAN según IEEE 802.1Q (Etiquetado de VLAN) y políticas de QoS aplicadas en colas internas de la tarjeta.

En un entorno empresarial típico, la segmentación puede organizarse como una VLAN para tráfico general de usuarios y navegación, otra VLAN para voz sobre IP con prioridad alta por su sensibilidad a retardo y variación de retardo, y una VLAN adicional para video, gestión o vigilancia, que también requiere tratamiento preferente y ancho de banda sostenido. Este esquema mejora la estabilidad del servicio bajo congestión, refuerza el aislamiento lógico entre dominios y facilita aplicar políticas coherentes desde el servidor hasta el switch.

Figura 22. VLAN Tagging y OoS por hardware



Fuente: Adaptado de (Gong, 2023)

La figura 22 la NIC puede aplicar segmentación lógica mediante VLAN (IEEE 802.1Q) y priorización de tráfico mediante QoS por hardware, diferenciando flujos con requisitos de latencia y criticidad distintos En conectividad física, estas NIC suelen ofrecer enlaces de alta capacidad tanto en cobre como en fibra. Para cobre, es frecuente la presencia de puertos RJ45 multigigabit, mientras que para fibra se emplean transceptores modulares como SFP/SFP+ y familias de mayor densidad para velocidades superiores.

Esta flexibilidad permite separar físicamente redes por propósito, por ejemplo una red de datos y una red de almacenamiento, y sostener aplicaciones exigentes en ancho de banda y latencia, como almacenamiento en red y cómputo de alto desempeño. Adicionalmente, algunas NIC soportan tramas jumbo, lo que reduce la sobrecarga por cabeceras en transferencias masivas, y pueden incorporar capacidades de administración como WoL (Wake-on-LAN, Activación por LAN) cuando la plataforma lo permite, facilitando operaciones remotas.

A partir de esta evolución de infraestructura física y lógica es posible integrar cómputo, almacenamiento, virtualización y automatización, aspectos que serán analizados en el siguiente capítulo.

CAPÍTULO 2:
CENTROS DE DATOS PARA REDES IP

2.1 EVOLUCIÓN DEL CENTRO DE DATOS EN LAS REDES IP

Los centros de datos han evolucionado de forma paralela al desarrollo de la informática y las redes de comunicación. En sus primeras etapas, estuvieron dominados por sistemas centralizados basados en computadoras mainframe y mecanismos de procesamiento por lotes. Posteriormente, la adopción de arquitecturas cliente–servidor, servidores en rack y virtualización marcó un punto de inflexión, dando paso a plataformas más flexibles y escalables. En la actualidad, los centros de datos constituyen el núcleo de la computación en la nube y de los servicios digitales, integrando cómputo, almacenamiento y redes bajo arquitecturas altamente automatizadas y distribuidas.

Desde una perspectiva funcional, un centro de datos moderno debe garantizar el ingreso, procesamiento, almacenamiento y distribución eficiente de grandes volúmenes de datos y servicios digitales. Los datos provienen de múltiples fuentes (usuarios, aplicaciones, sensores o sistemas externos) y pueden generarse tanto en tiempo real como de forma diferida. Estos flujos son procesados por infraestructuras de cómputo diseñadas para ejecutar aplicaciones, realizar análisis intensivos y responder de manera dinámica a las demandas de los servicios alojados.

El almacenamiento de la información se apoya en tecnologías heterogéneas que incluyen unidades de estado sólido, discos magnéticos y sistemas de almacenamiento en red, seleccionados en función de criterios de rendimiento, capacidad y disponibilidad. La interconexión entre servidores y sistemas de almacenamiento se realiza mediante infraestructuras de red de alta velocidad, basadas en conmutación Ethernet avanzada y enlaces redundantes, que aseguran baja latencia y continuidad operativa.

Un componente clave en la operación de los centros de datos es el balanceo de carga, que permite distribuir el tráfico y las cargas de trabajo entre múltiples recursos de cómputo, optimizando el uso de la infraestructura y evitando puntos únicos de congestión. De forma complementaria, las plataformas de monitoreo y gestión proporcionan visibilidad continua sobre el estado del hardware, la red y los servicios, permitiendo una administración proactiva y la detección temprana de fallas.

La seguridad constituye otro eje fundamental. Los centros de datos integran mecanismos de protección física y lógica, que incluyen control de acceso, vigilancia, firewalls, sistemas de detección y prevención de intrusiones, así como técnicas de autenticación y cifrado. Estas medidas buscan preservar la confidencialidad, integridad y disponibilidad de la información, en un contexto donde los servicios digitales son cada vez más críticos.

La escalabilidad y la resiliencia son características inherentes al diseño moderno de centros de datos. La adopción de arquitecturas modulares permite incorporar o retirar recursos de manera progresiva, mientras que la redundancia eléctrica, de enlaces y de equipamiento asegura la continuidad del servicio ante fallos previsibles o imprevistos. Esta combinación de escalabilidad y tolerancia a fallos resulta esencial para cumplir con los niveles de disponibilidad exigidos en entornos empresariales y de proveedores de servicios.

En este contexto, la red deja de ser un componente pasivo de interconexión y se convierte en el eje central de integración del centro de datos, actuando como una infraestructura de internetworking IP capaz de coordinar, optimizar y asegurar el funcionamiento conjunto de los sistemas de cómputo y almacenamiento. El centro de datos moderno deja de ser una simple instalación física para convertirse en una plataforma integrada,

donde redes, sistemas y servicios convergen. Esta convergencia sienta las bases para arquitecturas más avanzadas, caracterizadas por la virtualización, la automatización y la programabilidad, aspectos que se abordan en las secciones siguientes del capítulo.

2.1.1 INTERNETWORKING EN CENTROS DE DATOS

El internetworking en centros de datos modernos trasciende la simple conexión física entre dispositivos. Se trata de un sistema distribuido basado en protocolos IP que permite la comunicación eficiente entre múltiples dominios de red, aplicaciones y plataformas virtualizadas. En este enfoque, la red asume un rol activo en la provisión de servicios, garantizando conectividad, resiliencia y control del tráfico.

Figura 23. Centro de datos y equipos conectados

Fuente: (Lu, 2025)

La infraestructura de red se organiza a partir de la interacción entre el plano de datos, responsable del reenvío de paquetes a alta velocidad, y el plano de control, encargado de la toma de decisiones de enrutamiento, redundancia y convergencia. Esta separación conceptual permite diseñar redes escalables, tolerantes a fallos y adaptables a cambios dinámicos en la carga de trabajo.

En consecuencia, la arquitectura de red se convierte en un elemento estratégico dentro del centro de datos, ya que de su diseño dependen parámetros críticos como la latencia, el rendimiento, la

disponibilidad y la capacidad de expansión del entorno.

2.2 ARQUITECTURAS JERÁRQUICAS TRADICIONALES

Durante muchos años, los centros de datos se diseñaron siguiendo modelos jerárquicos de tres capas: acceso, agregación y núcleo. Este enfoque organizaba la red de forma piramidal y fue concebido para escenarios con tráfico predominantemente norte–sur, característicos de aplicaciones tradicionales del enfoque cliente–servidor. Este escenario presenta limitaciones significativas frente a las cargas de trabajo actuales, caracterizadas por comunicaciones este–oeste intensivas entre servidores, máquinas virtuales y servicios distribuidos.

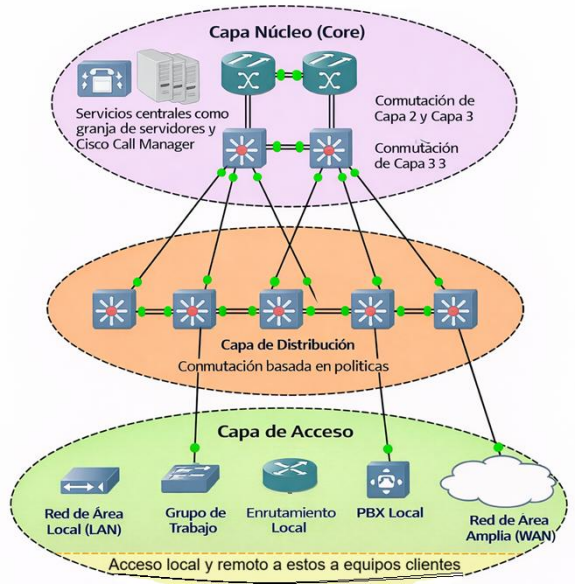
Entre las principales limitaciones se encuentran la existencia de puntos únicos de congestión, la complejidad en la expansión de la infraestructura y la dificultad para garantizar latencias predecibles a gran escala. Estas restricciones impulsaron la adopción de nuevos modelos arquitectónicos más adecuados para entornos de alta densidad y virtualización.

2.2.1 MODELO JERÁRQUICO DE TRES CAPAS

La capa de acceso proporciona conectividad directa a servidores y dispositivos finales, concentrando un gran número de enlaces de baja o media velocidad. La capa de agregación actúa como punto de consolidación, aplicando políticas de red, segmentación mediante VLANs y enrutamiento entre subredes. Finalmente, la capa núcleo se encarga del transporte de alta velocidad entre dominios de red, generalmente mediante dispositivos redundantes de gran capacidad.

Este modelo ofreció durante años simplicidad conceptual y facilidad de operación, pero su diseño asumía patrones de tráfico que ya no se corresponden con las exigencias actuales.

Figura 24. Arquitectura de tres niveles



Fuente: (Azam, 2018)

2.2.2 LIMITACIONES TÉCNICAS DEL MODELO JERÁRQUICO

El modelo jerárquico tradicional, estructurado en capas de Acceso, Distribución y Núcleo, fue concebido y optimizado para patrones de tráfico predominantemente Norte–Sur, donde la comunicación principal ocurre entre usuarios finales y servidores centralizados. Sin embargo, en entornos de centros de datos modernos, este enfoque presenta limitaciones estructurales significativas cuando se enfrenta a cargas de trabajo distribuidas y altamente interconectadas.

Uno de los principales inconvenientes radica en que el tráfico entre servidores ubicados en racks adyacentes o dentro de la misma zona lógica debe ascender innecesariamente hacia capas superiores (Distribución o Núcleo) para luego descender nuevamente, generando rutas indirectas, consumo ineficiente de ancho de banda y aumento de latencia

2.2.3 DEPENDENCIA DE SPANNING TREE PROTOCOL (STP)

Una de las debilidades más críticas del modelo jerárquico de tres capas en centros de datos es su fuerte dependencia del protocolo de árbol de expansión (STP, Spanning Tree Protocol), definido originalmente en IEEE 802.1D, como mecanismo para prevenir bucles de capa 2.

SPANNING TREE PROTOCOL (STP)

Fue diseñado para garantizar la operación libre de bucles en redes Ethernet de Capa 2, mediante la selección de un árbol lógico que desactiva enlaces redundantes. Este enfoque resultó adecuado en redes empresariales tradicionales, donde la simplicidad y la estabilidad primaban sobre el aprovechamiento total del ancho de banda.

Sin embargo, en el contexto de un centro de datos, STP introduce una ineficiencia fundamental: una parte significativa de la infraestructura física instalada permanece inactiva de forma permanente. En escenarios con alta densidad de servidores, esta subutilización limita el crecimiento efectivo de la red y dificulta la adopción de esquemas de alta disponibilidad activo–activo. Además, los tiempos de convergencia de STP ante fallos pueden resultar incompatibles con los requisitos de continuidad de servicio de aplicaciones críticas.

STP Y EL BLOQUEO DE ENLACES

En arquitecturas jerárquicas, es común desplegar enlaces redundantes entre switches con el objetivo de garantizar tolerancia a fallos. No obstante, Ethernet carece de un mecanismo nativo de prevención de bucles, por lo que STP se encarga de seleccionar un único camino activo y bloquear lógicamente los enlaces redundantes. Es decir STP obliga a "cortar" hilos o enlaces para evitar que la red colapse. Se puede indicar que STP es el enemigo del modelo "Fabric" (tejido) de centros de datos modernos que por Fabric se tiene multiples enlaces

que busca usar todos los enlaces para ser más robusta, rápida y sin bloqueos.

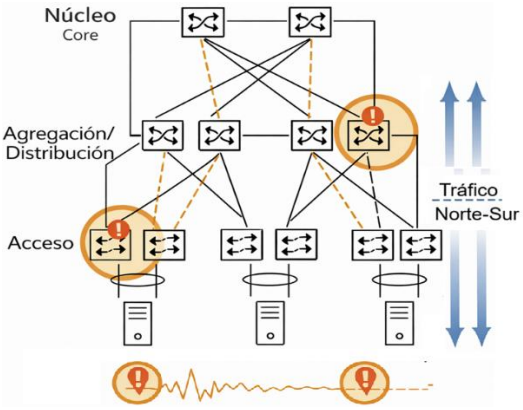
Un ejemplo operativo de STP implica qué, si un switch de acceso (por ejemplo, dentro de una LAN) dispone de dos enlaces de 10 Gbps hacia la capa de distribución, STP colocará uno de ellos en estado Blocking. Como resultado, solo 50 % del ancho de banda instalado permanece activo durante condiciones normales de operación. En términos económicos y de ingeniería, esta situación se traduce en una infrautilización permanente de la infraestructura, donde una parte significativa de la inversión en enlaces y puertos de alta capacidad permanece inactiva, a la espera de un fallo.

ACUMULACIÓN DE LATENCIA Y CUELLOS DE BOTELLA

La latencia de red es el tiempo que tardan los datos en viajar de un origen a un destino dentro de una red. Una latencia baja permite tiempos de respuesta rápidos y comunicación eficiente, mientras que una latencia alta degrada el rendimiento de las aplicaciones, puede provocar errores operativos y reduce la eficiencia general del sistema. En una arquitectura jerárquica, cada capa adicional implica procesamiento, almacenamiento temporal en colas y reenvío de paquetes. Cuando dos servidores ubicados en diferentes dominios deben comunicarse, el tráfico atraviesa múltiples dispositivos y niveles lógicos, incrementando la latencia total.

Más allá del valor absoluto de la latencia, el principal problema radica en su variabilidad esto implica rutas asimétricas, mecanismos de bloqueo de enlaces y congestión puntual generan retardos no deterministas, lo que dificulta la planificación de aplicaciones sensibles al tiempo, como bases de datos distribuidas, sistemas de análisis en tiempo real o plataformas de inteligencia artificial.

Figura 25. Latencias presentes en el tráfico Norte-Sur



Fuente: (Hpe, 2023)

Las líneas discontinuas naranjas con señales de advertencia ilustran:

Enlaces Bloqueados (STP): Las líneas discontinuas desde los switches de Acceso y Distribución muestran rutas redundantes que están lógicamente bloqueadas por el protocolo STP para prevenir bucles de red. Esto significa que, aunque físicamente hay dos caminos, solo uno está activo, lo que reduce la capacidad total utilizable y crea un punto potencial de congestión.

Punto de Congestión/Falla Único: El círculo naranja con la exclamación en la capa de Distribución/Agregación señala el cuello de botella. Todo el tráfico de los servidores (tráfico Este-Oeste) debe pasar por aquí para comunicarse con otros servidores, lo que añade latencia y variabilidad en el tiempo de respuesta.

Variabilidad de Latencia: La línea ondulada inferior con exclamaciones indica que el tiempo que tarda un paquete en ir de un servidor a otro no es constante (es variable, o jitter). Esta falta de predictibilidad impacta negativamente a las aplicaciones modernas que requieren baja latencia (como las bases de datos distribuidas o la voz sobre IP).

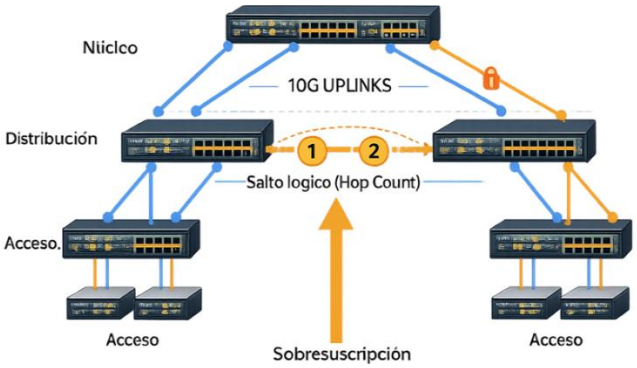
En entornos empresariales modernos, donde las aplicaciones dependen de la nube, de dispositivos IoT y de procesamiento de datos en tiempo real, la latencia se convierte en un factor crítico,

especialmente para cargas de trabajo sensibles al tiempo como computación de alto rendimiento (HPC), análisis en streaming y sistemas distribuidos.

INCREMENTO DEL NÚMERO DE SALTOS

En una arquitectura jerárquica tradicional, cada paquete de datos debe atravesar múltiples dispositivos intermedios antes de alcanzar su destino final, lo que incrementa el número de saltos dentro del dominio de conmutación y enrutamiento. Cada salto introduce latencia adicional asociada al procesamiento de cabeceras, la búsqueda en tablas de reenvío y la toma de decisiones de encaminamiento, incluso cuando estas operaciones se realizan en hardware especializado.

Figura 26, Representación de saltos



Fuente: Adaptado de (Network Baseline, 2026)

Las líneas azules continuas representan los enlaces físicos activos por los cuales circula el tráfico de datos seleccionado por los mecanismos de control de la red (por ejemplo, SPT en arquitecturas jerárquicas tradicionales) y reflejan el flujo real de tráfico predominante de tipo Norte–Sur entre las capas de acceso, distribución y núcleo. Por su parte, la línea naranja curva y segmentada indica la presencia de un salto lógico (hop count) dentro de la capa de distribución. Dicho salto corresponde a los procesos internos que ocurren en los switches de esta capa, tales como el análisis de cabeceras, la consulta a tablas de reenvío y la toma de decisiones de conmutación o enrutamiento, los cuales introducen

latencia adicional sin implicar cables o enlaces adicionales.

SOBRESUSCRIPCIÓN DE ENLACES (OVERSUBSCRIPTION)

La sobre-suscripción describe la relación entre la capacidad total de los enlaces de acceso y la capacidad disponible en los enlaces ascendentes hacia las capas de agregación y núcleo. En otras palabras, la sobre-suscripción determina cuántos usuarios o dispositivos comparten un recurso de red limitado. Así esta relación se vuelve más estrecha a medida que el tráfico asciende del acceso hacia el núcleo. Este aspecto técnico de sobresuscripción se traduce en congestión, incremento de latencia y pérdida de rendimiento, afectando de forma directa la experiencia de usuario y el cumplimiento de acuerdos de nivel de servicio.

El ratio o tasa de sobre-suscripción se obtiene dividiendo el ancho de banda total descendente (hacia los hosts) entre el ancho de banda total ascendente (hacia el nivel superior). Véase este escenario típico de sobre-suscripción:

- Un switch de acceso con 48 puertos de 1 Gbps (48 Gbps agregados).
- Uno o dos enlaces ascendentes de 10 Gbps hacia la capa de distribución.

Esto genera ratios de sobresuscripción elevados (por ejemplo, 4.8:1), donde la capacidad de entrada supera ampliamente la capacidad de salida. Bajo cargas simultáneas, esta condición provocará una congestión persistente, la pérdida de paquetes y por lo tanto muchas retransmisiones a nivel de transporte.

El impacto es especialmente severo en aplicaciones de almacenamiento en red (SAN), bases de datos distribuidas y sistemas de análisis en tiempo real, donde el rendimiento depende de flujos constantes y baja latencia.

Sin embargo, existen valores de ratio de sobre-suscripción aceptables en donde hay dos casos; uno con usuarios finales en entornos de oficina (PCs, teléfonos IP, impresoras) y otro con un servidor de base de datos o un nodo de virtualización que demanda tráfico constante y ráfagas masivas ambos se conectan al switch de acceso de la red.

- Acceso a Distribución: Se consideran aceptables ratios dependiendo de si el entorno es una oficina estándar o un centro de datos de hasta 20:1 (para navegación web, correo, etc.) o 3:1 (rendimiento crítico en centro de datos),
- Distribución a Núcleo: Las ratios suelen ser mucho más bajos, idealmente entre 4:1 y 2:1, para evitar cuellos de botella en el backbone.

Cuando no se controla ratios de sobre-suscripción elevados se conforma:

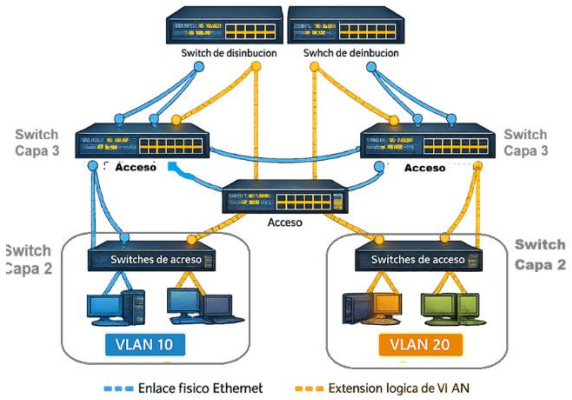
- Saturación de Buffers: Cuando los enlaces ascendentes se saturan, los paquetes se acumulan en la memoria temporal (buffers) de los switches. Una vez llenos, el switch descarta paquetes, forzando retransmisiones y aumentando la latencia.
- Degradación en Aplicaciones de Tiempo Real: El "jitter" (variación de latencia) causado por la congestión afecta gravemente a la voz sobre IP (VoIP) y videollamadas, provocando cortes o pérdida de sincronía.
- Inestabilidad en la Nube: En nubes públicas, donde las máquinas virtuales (VM) comparten hardware de forma dinámica, una sobresuscripción mal gestionada puede degradar el rendimiento de cargas de trabajo críticas e interactivas

MOVILIDAD DE CARGAS Y VIRTUALIZACIÓN

Las arquitecturas jerárquicas tradicionales fueron concebidas para entornos esencialmente estáticos, lo que contrasta con los modelos actuales de virtualización, en los cuales las cargas de trabajo y los servicios se desplazan dinámicamente entre nodos de cómputo. Este cambio de paradigma implica una transformación profunda en los patrones de tráfico, en la gestión de la conectividad y en los requisitos de la infraestructura de red, que deja de ser un componente pasivo para convertirse en un elemento crítico para la continuidad y el rendimiento de los servicios.

En los centros de datos modernos, las máquinas virtuales y los contenedores se migran dinámicamente mediante mecanismos de live migration o vMotion, y las cargas de trabajo se redistribuyen de forma continua con el objetivo de optimizar el consumo energético, la disipación térmica y la utilización de los recursos computacionales. Estas operaciones generan un volumen significativo de tráfico este–oeste entre nodos de cómputo, asociado a la replicación de datos, la sincronización de estados y la interacción constante entre servicios distribuidos, hasta representar la mayor proporción del tráfico interno del centro de datos.

Figura 27, Representación de extensión de VLANS



Fuente; Adaptado de (Abid, 2025)

En la arquitectura mostrada, las VLANs se crean y gestionan en los switches de Capa 2 mediante puertos de acceso y troncales, permitiendo la segmentación lógica del tráfico. La extensión de estas VLANs a través de enlaces troncales hacia los switches de Capa 3 posibilita la comunicación entre VLANs, pero también amplía el dominio de broadcast y la superficie de fallo, introduciendo restricciones estructurales a la movilidad de servicios. En este contexto, el diseño jerárquico introduce restricciones estructurales a la movilidad de servicios.

La combinación de enlaces redundantes inactivos debido al uso de STP, elevados niveles de sobresuscripción en los enlaces de agregación y latencias variables derivadas del tránsito por múltiples capas limita la continuidad de las comunicaciones durante la migración de cargas y reduce la eficiencia global en el uso del ancho de banda. Como resultado, la previsibilidad del rendimiento extremo a extremo se ve comprometida, afectando negativamente a aplicaciones distribuidas y a entornos virtualizados de alta densidad.

Para mitigar la interrupción de conectividad durante la movilidad de cargas, las redes jerárquicas recurren con frecuencia a la extensión de VLANs a través de múltiples capas y dominios físicos, con el fin de preservar la identidad lógica de los servicios. No obstante, esta estrategia introduce vulnerabilidades técnicas adicionales, entre ellas la expansión excesiva del dominio de broadcast, el aumento de la superficie de fallo y la propagación de tormentas de broadcast o errores de configuración a gran escala.

La pérdida de aislamiento lógico asociada a VLANs extendidas convierte a la red en un sistema más frágil y complejo de operar, donde un fallo localizado puede afectar a amplias zonas del centro de datos. En conjunto, estas limitaciones evidencian que las arquitecturas jerárquicas tradicionales no están

alineadas con los requerimientos de movilidad, escalabilidad y resiliencia de los entornos virtualizados modernos, lo que ha motivado la adopción de arquitecturas más planas y altamente interconectadas, diseñadas específicamente para soportar tráfico este–oeste intensivo y una movilidad eficiente de servicios.

2.3 ARQUITECTURAS MODERNAS DE CENTROS DE DATOS

A diferencia de las arquitecturas tradicionales, donde predominaba el tráfico Norte–Sur (cliente–servidor), las aplicaciones virtualizadas, los contenedores y las arquitecturas de microservicios están generando un alto volumen de tráfico Este–Oeste (East–West) entre servidores. Este comportamiento exige redes con baja latencia, alta capacidad de conmutación y múltiples caminos simultáneos entre cualquier par de nodos.

En este contexto, la escalabilidad deja de ser vertical (equipos más grandes y costosos) y pasa a ser horizontal, permitiendo añadir nuevos dispositivos de forma incremental sin rediseñar la infraestructura existente. Asimismo, los mecanismos tradicionales basados en protocolos de Capa 2 y en la prevención de bucles mediante STP resultan ineficientes, ya que bloquean enlaces redundantes y limitan el aprovechamiento del ancho de banda disponible.

Las arquitecturas modernas adoptan, por tanto, un enfoque basado en Capa 3 distribuida, donde todos los enlaces activos pueden ser utilizados simultáneamente mediante técnicas de balanceo multipath, como ECMP (Equal-Cost Multi-Path). Este modelo elimina dependencias de un único camino, mejora la resiliencia y permite un uso eficiente de los recursos de red.

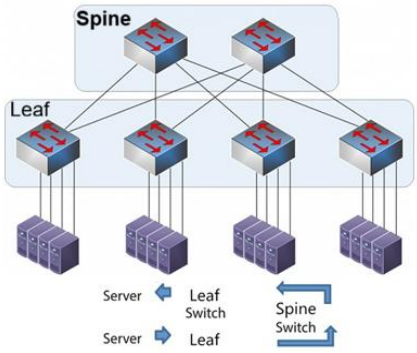
Sobre estos principios de conectividad uniforme, redundancia activa y enrutamiento distribuido, emerge la arquitectura Spine–Leaf, (garantiza una ruta de solo dos saltos entre switches) considerada el

fundamento de las actuales implementaciones de IP Fabric (muchos enlaces posibles) en centros de datos modernos.

2.3.1 ARQUITECTURA SPINE–LEAF E IP FABRICS

Es un diseño de red de dos capas que sustituye a la antigua jerarquía de tres capas o niveles. La arquitectura Spine–Leaf organiza los switches en dos niveles claramente definidos: los switches leaf, que conectan servidores y almacenamiento. Aquí los switches o nodos spine, actúan como núcleo de alta capacidad.

Figura 28. Arquitectura Spine-Leaf



Fuente: (User Datacenter, 2022)

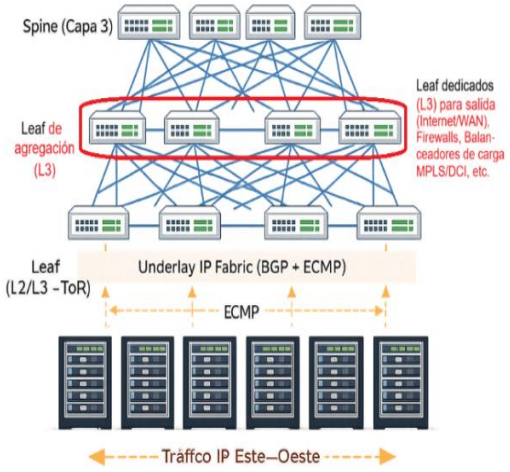
Los switches del núcleo de la red (capa 3 o enrutamiento) están interconectados en una malla completa con todos los switches Leaf inferiores. Proporciona múltiples caminos activos entre hojas siendo su función principal es reenviar paquetes a alta velocidad entre los diferentes switches Leaf. Garantiza una latencia baja y con característica uniforme y predecible esto es ideal para aplicaciones criticas o sensibles.

2.3.2 CONECTIVIDAD DE MALLA COMPLETA

El principio fundamental del modelo Fabric en arquitecturas Spine–Leaf es la conectividad de malla completa entre los niveles Leaf y Spine. Aunque en escenarios reales esta arquitectura puede incluir una capa intermedia que representa:

- Leaf adicional (doble capa de Leaf/Leaf–Leaf o Aggregation Leaf)
- Leaf dedicados (Service Leaf/)
- Clos con más etapas (multistage Clos)

Figura 29. Arquitectura Spine–Leaf con doble capa Leaf



Fuente: Adaptado de (Siemon, 2025)

La arquitectura puede extenderse con nodos o switches Leaf adicionales dedicados a funciones de salida, seguridad o interconexión. Cada switch Leaf de agregación (Capa 3 o Top-of-Rack) establece enlaces directos hacia todos los switches Spine, conformando una topología de interconexión total entre ambos planos. Esta disposición crea múltiples rutas simultáneamente disponibles para el transporte de tráfico IP entre los diferentes racks de servidores. Como resultado, la red elimina puntos únicos de congestión, mejora la tolerancia a fallos y permite el uso eficiente del ancho de banda agregado del sistema.

Así mismo la figura 29 ilustra conectividad de malla completa y su relación con la comunicación predominante de tipo Este–Oeste entre los servidores del centro de datos. El funcionamiento de este plano de conectividad o transporte (underlay) se basa en enrutamiento distribuido, esto puede darse con la combinación de BGP como protocolo de control y ECMP como mecanismo de balanceo multipath.

BGP El protocolo Border Gateway Protocol (BGP) se emplea para el intercambio de información de enrutamiento entre los dispositivos del Fabric, permitiendo la construcción de un plano de control distribuido y altamente escalable. En entornos de centros de datos, es común el uso de eBGP entre nodos Leaf y Spine para simplificar el dominio de fallas y facilitar la expansión de la infraestructura. BGP opera sobre TCP (puerto 179) y selecciona rutas mediante atributos, lo que permite implementar políticas de encaminamiento flexibles y eficiente

ECMP Equal-Cost Multi-Path (ECMP) permite que el tráfico IP, representado en la figura por las rutas paralelas, se distribuya de forma equilibrada entre múltiples caminos de igual costo hacia un mismo destino. Mediante algoritmos de hashing por flujo, ECMP optimiza la utilización del ancho de banda disponible y evita la subutilización de enlaces. Además, ante la falla de un enlace o dispositivo, el tráfico se redirige automáticamente por las rutas restantes, proporcionando alta disponibilidad y rápida convergencia.

Este enfoque permite que todos los enlaces del Fabric permanezcan activos y participen en el reenvío de tráfico, eliminando la necesidad de protocolos de prevención de bucles como STP y reduciendo la dependencia de dominios extensos de Capa 2. Como resultado, la arquitectura soporta de manera eficiente los patrones de tráfico este-oeste predominantes en los centros de datos modernos, caracterizados por comunicaciones intensivas entre servidores.

2.3.3 VIRTUALIZACIÓN DE RED SOBRE IP FABRIC

Las arquitecturas modernas de centros de datos no solo optimizan la conectividad física mediante IP Fabrics, sino que también incorporan mecanismos de virtualización de red que permiten abstraer los servicios lógicos de la infraestructura subyacente.

La figura 29 muestra la separación entre la infraestructura física del centro de datos, basada en un Fabric con enrutamiento distribuido y balanceo ECMP, y una capa superior de virtualización implementada mediante tecnologías de encapsulación. En este modelo, la red física proporciona conectividad IP de alta disponibilidad (underlay), mientras que la segmentación lógica, el aislamiento de cargas y la extensión de dominios de red se implementan en una capa superpuesta (overlay).

El uso de VXLAN permite crear múltiples redes virtuales identificadas mediante VNI (VXLAN Network Identifier), facilitando la segmentación a gran escala y el soporte de entornos multi-tenant. Esta arquitectura permite desacoplar la operación de la red lógica de las restricciones físicas, facilitando la movilidad de cargas de trabajo, la integración con entornos de nube y la extensión de servicios entre centros de datos.

Este enfoque constituye la base para los mecanismos avanzados de control y señalización implementados mediante BGP y EVPN, los cuales se analizan a continuación.

2.3.4 UNDERLAY VS OVERLAY

La arquitectura moderna de centros de datos separa la red en dos planos: Underlay (físico) y Overlay (lógico). El underlay proporciona conectividad IP básica y transporte físico (cables, switches, routers), mientras que el overlay crea redes virtuales sobre el underlay, permitiendo movilidad de máquinas virtuales, segmentación avanzada y aislamiento multitenant sin depender de la topología física subyacente.

Underlay (Red Física)

Tiene estas características:

- Función: Infraestructura base, transporte de paquetes IP, enrutamiento físico.

- Componentes: Switches físicos (Spine/Leaf), routers y cables de fibra óptica
- Protocolos de enrutamiento: Generalmente Capa 3 (OSPF, IS-IS, eBGP) para una alta disponibilidad.
- Objetivo: Transparencia y conectividad de alta velocidad entre VTEPs (VXLAN Tunnel Endpoints).
- Métrica de éxito: El "Next Hop" (Próximo salto).

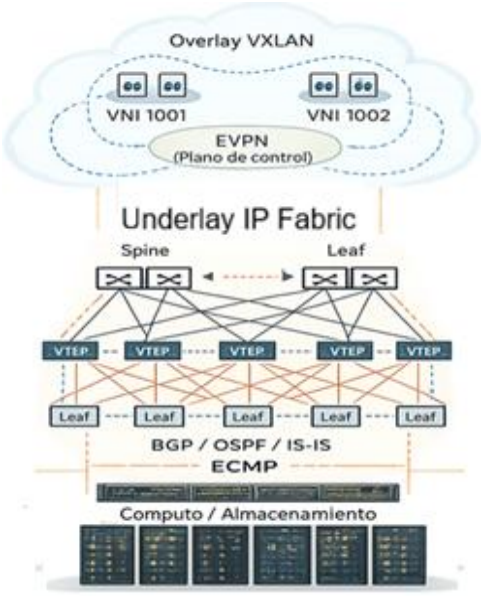
Overlay (Red Virtual)

Tiene las características siguientes:

- Función: Lógica, segmentación, aislamiento, movilidad de cargas.
- Componentes: Túneles virtuales VTEPs (VXLAN Tunnel EndPoint) son puntos finales de túnel y controladores de software sobre el underlay.
- Protocolos: En plano de control como protocolos de enrutamiento BGP-EVPN para gestionar la información de accesibilidad y protocolo de encapsulación genérica de virtualización de red (GENEVE, Generic Network Virtualization Encapsulation).
- Objetivo: Segmentación (Multitenancy), movilidad de máquinas virtuales y servicios de seguridad.
- Métrica de éxito: Flexibilidad y aislamiento de tráfico.

Esta separación permite desacoplar la operación de la red lógica de la infraestructura física, facilitando la automatización, la escalabilidad y la operación en entornos dinámicos.

Figura 30. Virtualización y separación entre Underlay IP Fabric y Overlay VXLAN



Fuente: Adapatado de (HPE-Juniper, 2023)

En arquitecturas VXLAN, la funcionalidad de VTEP se implementa en los switches Leaf, los cuales encapsulan y desencapsulan el tráfico de los segmentos virtuales (VNI) para su transporte sobre el underlay IP Fabric. Esta separación permite desacoplar la topología física de la estructura lógica de la red.

2.3.5 ALCANCES OPERATIVAS EN ARQUITECTURAS FABRIC

INTEGRACIÓN DE LOS PLANOS DE RED

Las arquitecturas modernas de centros de datos se basan en la operación coordinada de múltiples planos funcionales que permiten desacoplar la infraestructura física de los servicios de red.

En este modelo:

- El plano físico está compuesto por una topología Spine–Leaf que proporciona conectividad de alta capacidad y baja latencia.
- El plano de transporte (Underlay) implementa el enrutamiento IP distribuido mediante

protocolos como BGP, OSPF o IS-IS, junto con balanceo de carga ECMP.
- El plano lógico (Overlay) utiliza tecnologías de encapsulación como VXLAN y mecanismos de control como BGP EVPN para proporcionar segmentación, aislamiento y movilidad de cargas.

Esta separación funcional permite que los servicios de red evolucionen sin depender de cambios en la infraestructura física.

BENEFICIOS OPERATIVOS DEL MODELO FABRIC

La adopción de arquitecturas basadas en IP Fabric ofrece ventajas significativas frente a los diseños tradicionales:

- Eliminación de dominios extensos de Capa 2 y reducción del uso de STP
- Utilización activa de todos los enlaces disponibles mediante ECMP
- Alta disponibilidad y rápida convergencia ante fallos
- Escalabilidad horizontal mediante la adición de nodos Leaf o Spine
- Reducción de la complejidad operativa en entornos de gran escala
- Soporte eficiente para patrones de tráfico este-oeste

Estas características permiten optimizar el rendimiento y mejorar la resiliencia de la infraestructura.

ALINEACIÓN CON ENTORNOS DE VIRTUALIZACIÓN Y NUBE

Las arquitecturas Fabric están diseñadas para soportar los requerimientos de los entornos de cómputo modernos, incluyendo:

- Infraestructuras de virtualización masiva
- Plataformas basadas en contenedores y microservicios
- Entornos multi-tenant

- Integración con infraestructuras híbridas y multi-cloud
- Movilidad dinámica de cargas de trabajo

El uso de Overlay mediante VXLAN permite extender redes lógicas entre centros de datos o hacia entornos de nube pública sin modificar la topología física.

LA AUTOMATIZACIÓN EN UN DATA CENTER

La automatización de red consiste en definir reglas, plantilla o *template* (molde de configuración que ya trae la estructura correcta de comandos/parámetros, pero deja ciertos valores como variables para rellenarlos automáticamente) y políticas (decisiones de diseño/seguridad/calidad que la red debe cumplir) para que la infraestructura se configure y mantenga de forma repetible y consistente, reduciendo al mínimo la intervención del administrador a hacer los cambios de forma manual. El objetivo es una operación estandarizada: mismos parámetros, mismos perfiles, mismos resultados, incluso cuando el número de dispositivos en la red crece.

Algunas automatizaciones típicas son:

Provisionamiento L2/L3 (Capa 2/Capa 3): creación masiva de VLAN (Red de Área Local Virtual) y VRF (Instancia de Reenvío y Enrutamiento Virtual), interfaces SVI (Interfaz Virtual de Switch), gateways, direccionamiento, rutas estáticas o parámetros de enrutamiento (BGP (Protocolo de Puerta de Enlace Fronteriza) y EVPN (Red Privada Virtual Ethernet)).

Perfiles de puertos: aplicar configuraciones uniformes en decenas/centenas de puertos (modo access/trunk), asignación de VLAN, port-security (seguridad de puerto), storm-control (control de tormentas) y QoS

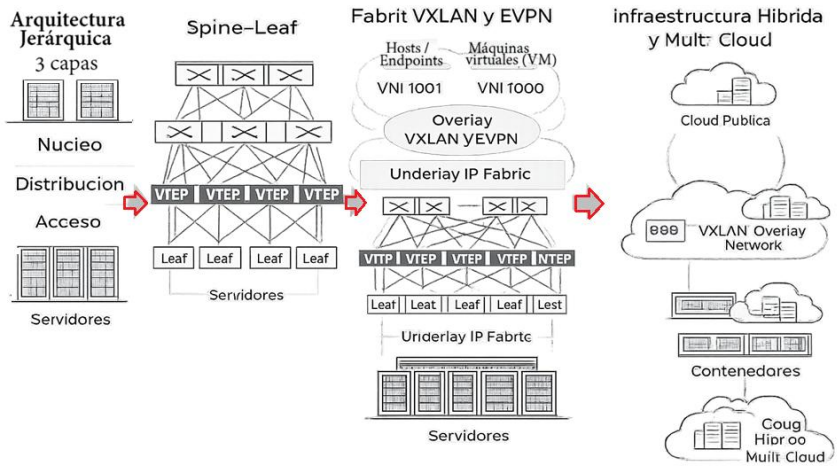
Políticas y seguridad: reglas repetibles de ACL (Lista de Control de Acceso), microsegmentación (según plataforma) y políticas de firewall (cuando aplica).

Cambios operativos: backups de configuración, auditorías (compliance), validación de estado, despliegue de cambios por lotes y rollback (reversión de cambios).

Servicios bajo demanda: Pueden ser el "alta de una nueva red para una aplicación" (segmento/VLAN/VRF más políticas y conectividad).

En este contexto, la red deja de ser un componente estático y se convierte en una plataforma dinámica orientada a servicios.

Figura 31, Evolución de las arquitecturas de centros de datos



Fuente; (AscentOptics, 2025)

Las arquitecturas Spine–Leaf basadas en IP Fabric, junto con mecanismos de virtualización mediante VXLAN y control distribuido mediante BGP EVPN, representan el modelo de referencia para el diseño de redes en centros de datos modernos. Este enfoque proporciona la orquestación

ORQUESTACIÓN

En la práctica, Spine–Leaf más IP Fabric es la base típica sobre la que se montan automatización y orquestación en data centers modernos, especialmente cuando además se usa overlay (por ejemplo VXLAN/EVPN) y controladores (plataformas de software) que actúan como plano de control y gestión centralizada de la red del centro de datos.

Así, un orquestador es la plataforma que traduce una solicitud de servicio en acciones coordinadas sobre varios dominios de la infraestructura. En lugar de configurar dispositivos o recursos aislados, el orquestador trabaja con un modelo de "estado deseado" y ejecuta flujos para aprovisionar lo necesario en cómputo, almacenamiento, red y seguridad, dejando el servicio listo y operativo.

Por ejemplo, al desplegar una nueva aplicación, el orquestador puede solicitar redes y segmentación, asignar recursos de VM (Virtual Machine, Máquina Virtual) o contenedores, aplicar políticas de acceso, conectar la aplicación a balanceo y activar monitoreo, asegurando consistencia entre lo que se pidió y lo que realmente quedó desplegado. En arquitecturas Spine–Leaf con IP Fabric, el orquestador suele apoyarse en controladores y APIs para realizar cambios de forma masiva y verificable.

La diferencia entre controlador y orquestador es que el controlador administra la red como un sistema dentro del data center, aplicando configuraciones y políticas en el fabric o en el overlay y verificando que la conectividad y la segmentación se mantengan estables, mientras que el orquestador administra el servicio completo, coordinando red, seguridad, cómputo y almacenamiento para desplegarlo y sostenerlo durante su ciclo de vida.

Un ejemplo puntual es el despliegue de una nueva aplicación para un laboratorio: el controlador de red crea y distribuye el segmento correspondiente en el **F**abric, configura el comportamiento del overlay, y asegura que los switches leaf apliquen la política de conectividad de forma consistente; el orquestador, en cambio, además de solicitar ese segmento, aprovisiona la VM o los contenedores donde correrá la aplicación, asigna recursos de almacenamiento, aplica reglas de seguridad y deja activo el monitoreo, de modo que el servicio quede operativo de extremo a extremo sin configurar componente por componente.

2.3.6 INTEGRACIÓN DE CENTRO DE DATOS CON EDGE COMPUTING

En el escenario de la hiperconectividad el centro de datos se integra de forma estrecha con plataformas de Edge Computing (computación en el borde), donde el cómputo y el almacenamiento se distribuyen cerca de las fuentes de los "datos". Así, las decisiones de diseño ya no se realizan por separado entre red, servidores y almacenamiento, sino considerando el comportamiento conjunto de aplicaciones, flujos de datos y requerimientos de procesamiento.

Bajo esta integración, el centro de datos y el Edge soportan infraestructuras virtualizadas, contenedores y plataformas hiperconvergentes, incluyendo HCI (Hyper-Converged Infrastructure, Infraestructura Hiperconvergente), así como esquemas de almacenamiento distribuido y movilidad dinámica de cargas de trabajo. El resultado es un modelo donde una misma aplicación puede ejecutarse y escalar entre diferentes ubicaciones, con etapas de filtrado, agregación y analítica que reducen el volumen de datos transportado hacia la nube.

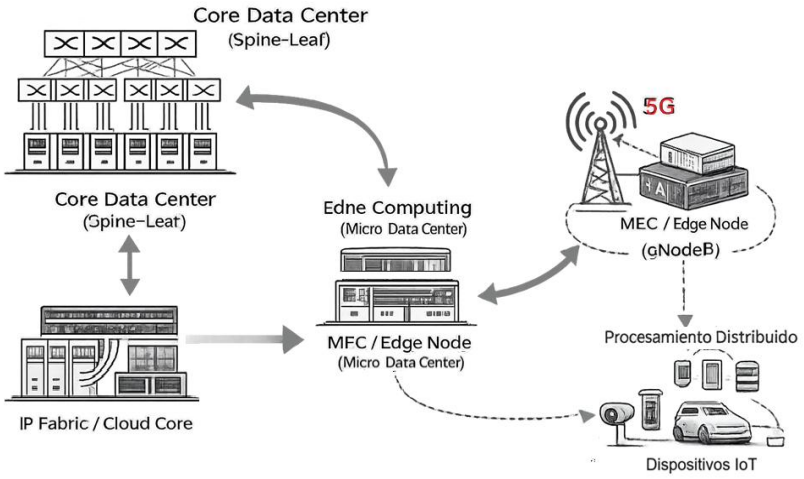
Esta capacidad es crítica en escenarios de alta densidad de dispositivos, donde la generación masiva de información exige respuestas rápidas y decisiones en tiempo real, especialmente cuando se ejecutan algoritmos de inteligencia artificial próximos al origen de los datos. En términos de red, la conectividad IP cumple el rol de columna vertebral que habilita el funcionamiento coordinado entre cómputo y almacenamiento distribuidos.

Las arquitecturas Spine–Leaf y el modelo de IP Fabric aportan escalabilidad y baja latencia interna al centro de datos, mientras que la virtualización de red mediante Overlay (capa superpuesta) permite aislamiento de servicios y segmentación lógica consistente entre el núcleo y el borde. A su vez, el Underlay (capa subyacente) mantiene el transporte

eficiente y confiable de los datos. Esta separación facilita que la red se comporte como un recurso programable, apto para automatización y orquestación, donde los servicios pueden aprovisionarse y adaptarse con rapidez frente a cambios en la demanda.

La adopción de Edge también se vuelve esencial en redes móviles de nueva generación, particularmente con 5G (Quinta Generación), donde se emplea MEC (Multi-access Edge Computing, Computación en el Borde de Acceso Múltiple) para desplegar servicios cerca del usuario y reducir latencia en aplicaciones críticas. En conjunto, la integración Centro de Datos–Edge exige arquitecturas planas basadas en enrutamiento distribuido, soporte para operación híbrida y multi-nube, y capacidad de administrar topologías geográficamente distribuidas..

Figura 32. Computing Edge hacia redes móviles e IoT



Fuente: Adaptado de (Meng, 2022)

La figura 32 representa un escenario donde un Core Data Center (centro de datos central) con arquitectura Spine–Leaf e IP Fabric se integra con nodos de Edge Computing, incluyendo micro centros de datos y plataformas MEC asociadas a redes 5G (Quinta Generación). En este contexto, la orquestación es el mecanismo que coordina de forma unificada los recursos distribuidos de cómputo

y almacenamiento para desplegar servicios cerca del origen de los datos, sin perder control central.

Esto permite decidir dónde ejecutar una carga de trabajo según latencia, capacidad disponible o congestión de red, asignar recursos a VM o contenedores, y mantener políticas coherentes de conectividad y seguridad entre el núcleo y el borde. Al mismo tiempo, la orquestación soporta el ciclo de vida del servicio mediante escalamiento, recuperación ante fallas y actualizaciones, lo que resulta crítico cuando los dispositivos IoT generan tráfico continuo y las aplicaciones requieren respuesta en tiempo real.

CAPÍTULO 3:
REDES MÓVILES IP
Y CONECTIVIDAD AVANZADA

En escenarios de movilidad y ubicuidad, la conectividad inalámbrica es un factor estructural de las redes móviles, tanto en WLAN corporativas de corto alcance como en entornos de mayor cobertura MAN/WAN. En la práctica, incluso en redes celulares como 5G, el usuario puede recibir un servicio aparentemente fijo (misma experiencia de acceso, aplicaciones persistentes y políticas de seguridad) aunque su punto de conexión cambie continuamente. Por lo tanto, las redes de datos móviles se caracterizan por la capacidad de una arquitectura IP de mantener continuidad de servicio.

Esa continuidad de servicio mientras los nodos (o dispositivos, artefactos con una dirección IP) cambian de punto de conexión depende, en primer lugar, del traspaso (handover[5]/roaming[6]): para que el usuario no perciba el cambio, la transición debe anticiparse antes de que el enlace actual caiga por debajo de umbrales operativos. En escenarios heterogéneos el reto se amplifica al handover vertical, donde el terminal puede pasar entre tecnologías distintas (Wi-Fi a 5G, de 5G a conectividad satelital) sin interrumpir una sesión de VoIP (Voz sobre IP) o un túnel de una red privada virtual (VPN, Virtual, Private Network).

La decisión de traspaso ya no depende únicamente de la potencia de señal, sino de métricas multidimensionales como latencia, jitter, pérdida de paquetes IP, congestión y carga del dominio de destino. En segundo lugar, el plano de control actúa como el "cerebro" del estado: resguarda identidad, sesión y políticas (QoS, seguridad y priorización), de modo que, si un vehículo de emergencia se desplaza, su prioridad se mantenga al moverse entre

[5] Proceso donde un dispositivo cambia de un punto de acceso (AP o celda) a otro durante una sesión activa.

[6] Mecanismo de continuidad del servicio cuando el usuario/dispositivo cambia de dominio de red (administrativo o tecnológico) y debe seguir autenticado, autorizado y enrutable sin perder la sesión.

celdas o dominios de cobertura; para ello, se separa identidad de ubicación mediante mecanismos de movilidad o gestión de sesión que sostienen la cobertura de recepción de señal.

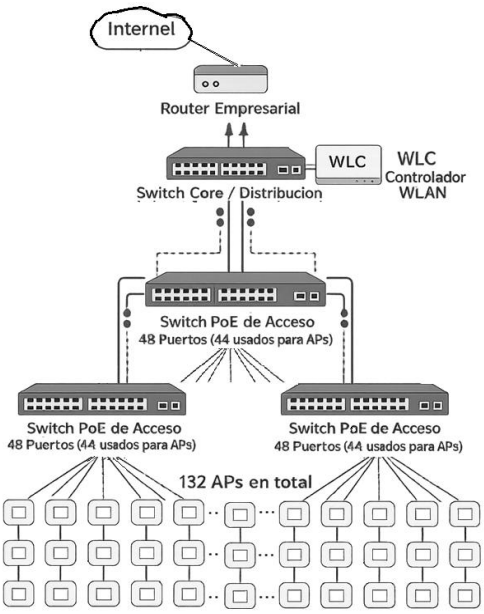
Así mismo, el plano de usuario transporta tráfico con tratamiento diferenciado: en movilidad debe ser flexible para sostener desempeño durante el traspaso y, cuando el contexto lo exige (misión crítica o industria), aplicar segmentación lógica de recursos o slicing[7] para asegurar que telemetría y flujos críticos tengan prioridad sobre tráfico no esencial incluso durante eventos de cambio de acceso.

3.1 RED CORPORATIVA CON MOVILIDAD

Las redes corporativas de datos con dispusimos móviles se apoyan con la conectividad inalámbrica del estándar IEEE 802.11 o Wi-Fi que mantengan la continuidad de sesión, dirección y políticas mientras el usuario se desplaza por un edificio o campus. En lugar de tratar cada punto de acceso (AP) como una red independiente, la LAN inalámbrica (WLAN) se gestiona como un sistema controlador (WLC, Wireless LAN Controller) o por una arquitectura de Cloud/SD-WLAN (Software-Defined Wireless Local Area Network, administrado en la nube) que centraliza autenticación, autorización, segmentación y calidad de servicio.

[7] El slicing crea redes virtuales con distintas políticas y QoS sobre la misma infraestructura.

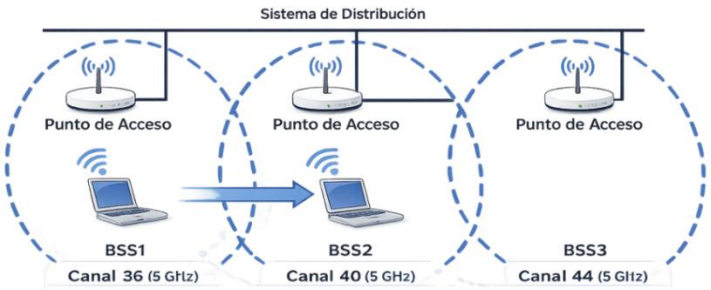
Figura 33. Arquitectura de control/gestión de un WLC



Fuente; Autores

Por ejemplo, en un despliegue corporativo de alta densidad (aeropuertos, grandes hospitales, etc.) un switch PoE de 48 puertos puede conectar ≈44–46 APs, reservando 1–2 puertos para uplinks hacia el switch core/distribución (y/o redundancia) y algunos puertos de reserva. El WLC, conectado al switch core, centraliza la configuración (SSIDs, VLANs, QoS) y la coordinación del roaming entre APs, incluyendo la planificación de canales y potencias para reducir interferencia co-canal. Así, cuando el usuario se desplaza y cambia de AP y por tanto puede cambiar de BSS/BSSID y de canal radioeléctrico la infraestructura procura que el traspaso sea transparente para la capa IP, manteniendo sesión y políticas sin que el usuario perciba un cambio de identidad de red.

Figura 34. Conjunto de servicios extendidos (ESS) y soporte a la movilidad.

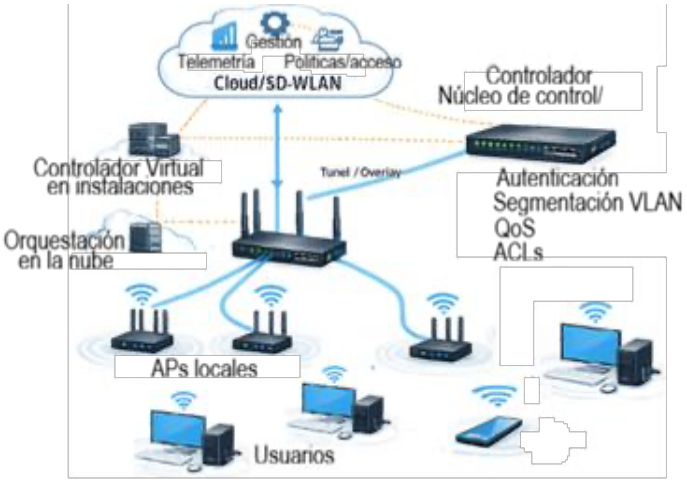


Fuente: Adaptado de (Salazar, 2016)

La WLAN corporativa puede operar como un dominio lógico único donde el usuario conserva su identidad y perfil de acceso (credenciales/rol) aunque cambie de BSSID durante el roaming. En ese traspaso, la infraestructura mantiene políticas consistentes —segmentación por rol, ACL y parámetros de QoS— para que el tráfico se trate de forma uniforme antes, durante y después del cambio de AP.

En arquitecturas con controlador, el tráfico del cliente puede encapsularse hacia un núcleo lógico de la WLAN mediante CAPWAP, permitiendo que el controlador centralice autenticación, asignación de VLAN/segmento, QoS y reglas de acceso. En una Cloud/SD-WLAN, esta lógica se moderniza: el controlador puede ser virtual y distribuido (on-premises o nube), y la gestión basada en políticas/intención sincroniza el estado entre APs para habilitar roaming rápido y coherencia operativa

Figura 35. Arquitectura Cloud/SD-WLAN



Fuente: Autores

En una Cloud/SD-WLAN, esa lógica se extiende y moderniza: el "controlador" puede ser virtual (software) y estar distribuido (on-premises o en la nube), con un plano de gestión que define intención y políticas (roles, microsegmentación, perfiles de aplicación, prioridades) y un plano de control que sincroniza el estado entre APs para habilitar roaming rápido y coherencia operativa.

En lugar de depender de configuraciones aisladas por AP, la red se gestiona como un dominio coherente, donde el usuario recibe el mismo perfil y las mismas restricciones al moverse por el campus. Además, estos diseños suelen incorporar telemetría continua (calidad de enlace, interferencia, carga por AP, experiencia de aplicación) y automatización para ajustar parámetros radio y decisiones de asociación, reduciendo degradaciones durante el traspaso y mejorando la estabilidad del servicio.

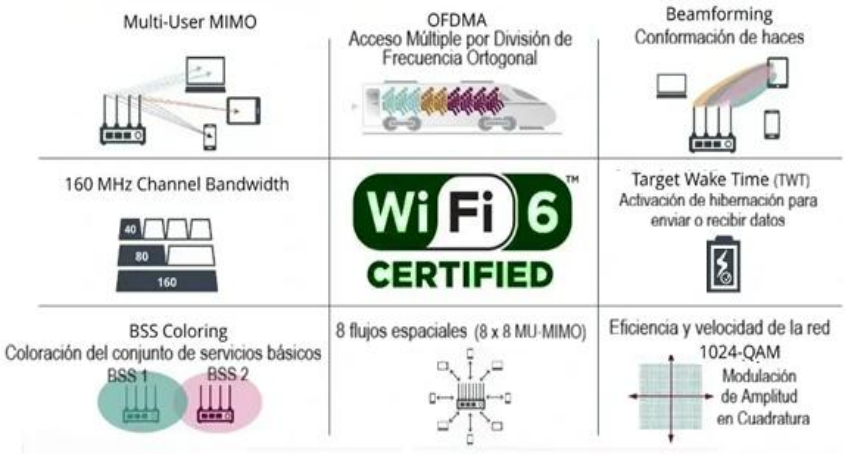
Bajo ambas aproximaciones, el uso de tunelización/overlay ya sea entre AP y controlador o entre AP y una capa de agregación, mantiene consistente el camino de reenvío hacia la red corporativa y simplifica la continuidad de políticas: el cambio físico de AP y de BSSID se gestiona dentro del dominio WLAN sin convertirse en un cambio "visible" para las aplicaciones IP. En consecuencia, el usuario

mantiene su conectividad operativa con mínima interrupción y con políticas persistentes (segmentación, control de acceso y QoS) a lo largo del movimiento.

3.2 WI-FI 6

Las WLANs basadas en IEEE 802.11 operan sobre el espectro electromagnético no licenciado (bandas de frecuencias libres), De la misma manera Wi-Fi 6 (IEEE 802.11ax), fue diseñado para mejorar la eficiencia y el rendimiento en entornos con múltiples dispositivos conectados simultáneamente. Véase sus tecnologías clave integradas.

Figura 36, Características técnicas de Wi-Fi 6 y 6E



Fuente: (Network.Switch, 2025)

El acceso a la conectividad puede distribuirse en tres bandas de frecuencias: 2.4 GHz, 5 GHz y 6 GHz, lo que amplía la capacidad de canalización y mejora la reutilización espacial. La banda de 2.4 GHz ofrece alcance favorable, pero suele estar más congestionada; 5 GHz aporta mayor disponibilidad de canales y mejor capacidad; y 6 GHz añade espectro adicional que facilita diseños con más canales reutilizables y menor presión de dispositivos legados. En este marco, Wi-Fi 6E habilita el uso de 6 GHz manteniendo la base técnica de 802.11ax, mientras que Wi-Fi 7 explota ese espectro con mayor flexibilidad de operación. Para un campus de alta densidad, el objetivo no es maximizar la velocidad

pico de una estación, sino sostener capacidad agregada y baja variabilidad de desempeño cuando miles de dispositivos comparten el medio.

ANCHO DE CANAL

Las tres bandas de uso libre para Wi-Fi (2.4, 5 y 6 GHz) se organizan en canales base de 20 MHz. En Wi-Fi 6 un canal de 20 MHz ofrece de 100 a 150 Mbps por flujo espacial en condiciones ideales (depende de la modulación y codificación, número de flujos, calidad de la señal SNR[8] y eficiencia). El ancho de canal determina cuánta porción de espectro se usa para transmitir: a mayor ancho, mayor throughput teórico, pero también menor cantidad de canales disponibles y más probabilidad de interferencia en entornos con muchos APs cercanos. Al respecto cuando es mayor la modulación por amplitud de cuadratura (64-QAM, 256-QAM, 1024-QAM) se logran altas velocidades de datos.

Figura 37. Ancho de canal para Wi-Fi

Fuente: (Llorach, 2024)

Agrupando canales de 20 MHz se puede crear anchos de 40, 80 o 160 MHz, sin embargo, cuando un AP usa 160 MHz: "se monta" sobre muchos canales que ya están ocupados. sufre colisiones. el SNR baja. la modulación (1024-QAM) ya no funciona. el AP

[8] La SNR en comunicaciones inalámbricas no depende solo del ruido térmico, sino de la suma de interferencias, ruido del receptor, distorsión por multipath y ruido impulsivo. Un alto SNR permite usar modulaciones QAM elevadas y mejorar la velocidad y estabilidad del enlace. Si el ruido aumenta, el SNR baja

retrocede a 64-QAM o 16-QAM. Se detallan características de anchos de canal.

- 40 MHz (2×20): Incrementa el throughput teórico, pero reduce a la mitad el número de canales. En 5 GHz puede funcionar bien si el entorno no está saturado; en 2.4 GHz no se recomienda porque provoca solapamiento severo y empeora la coexistencia.
- 80 MHz (4×20): Permite altas tasas, pero es más sensible al ruido y a la congestión. En escenarios densos, el rendimiento real puede caer por colisiones y retransmisiones. Es útil para tráfico pesado (p. ej., streaming local, copias a NAS).
- 160 MHz (8×20): Maximiza el throughput teórico, pero exige mucho espectro limpio. En 5 GHz suele ser difícil sostenerlo por disponibilidad limitada y coexistencia; es más viable en 6 GHz (Wi-Fi 6E), donde hay más espectro y más canales no solapados.

La elección del ancho de canal es un compromiso: canales más anchos incrementan el throughput potencial por enlace, pero reducen el número de canales reutilizables en el sitio y pueden incrementar la contención; canales más estrechos facilitan reutilización espacial y suelen mejorar la capacidad total del sistema cuando la densidad de APs y clientes es alta. Por ello, la planificación de canales, potencias y solapes de cobertura se convierte en una pieza estructural de la movilidad corporativa, ya que un traspaso entre APs ocurre dentro de un entorno donde el medio puede estar saturado y el rendimiento depende de cuán bien se administre esa competencia.

EFICIENCIA EN ALTA DENSIDAD

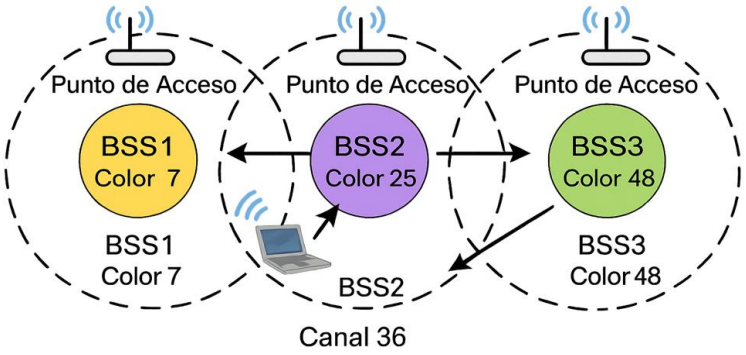
En este contexto, Wi-Fi 6 (IEEE 802.11ax) logra eficiencia en alta densidad, incorporando operación multiusuario con:

- OFDMA que permite que una sola transmisión asigne recursos a múltiples estaciones dentro del mismo intervalo, reduciendo overhead y mejorando el comportamiento bajo congestión.
- MU-MIMO que incrementa la capacidad agregada al permitir flujos espaciales simultáneos hacia múltiples estaciones, sujeto a condiciones de canal y a la capacidad del AP y los clientes. Estas técnicas no "eliminan" la contención, pero la administran con mayor granularidad, lo que se traduce en mejoras prácticas en escenarios reales de campus.

BSS COLORING Y REUSO ESPACIAL

Wi-Fi 6 introduce mecanismos de reutilización espacial como BSS Coloring, que permiten diferenciar celdas superpuestas (OBSS, Overlapping Basic Service Set) mediante un identificador ("color") asociado al BSS. Con esta información, estaciones y APs pueden tomar decisiones más finas sobre el acceso al medio cuando detectan energía en el canal: en lugar de tratar toda señal como si perteneciera al mismo dominio, el sistema puede distinguir entre transmisiones intra-BSS e inter-BSS y ajustar umbrales de decisión para habilitar mayor simultaneidad cuando el OBSS se recibe con potencia baja. El beneficio buscado es reducir contención y retransmisiones en entornos con múltiples APs cercanos, elevando la capacidad efectiva del sitio; no obstante, su ganancia depende de una planificación RF consistente, ya que configuraciones excesivamente agresivas pueden incrementar interferencia y degradar latencia y jitter en aplicaciones sensibles.

Figura 38.Representación de BSS Coloring



Fuente: (Natkaniec & Bieryt, 2023)

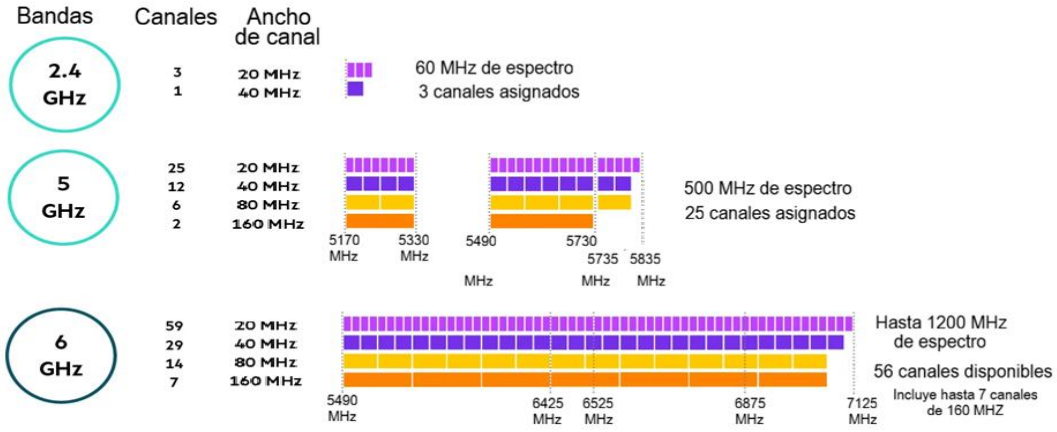
La figura 36 ilustra el principio de BSS Coloring introducido en Wi-Fi 6 (802.11ax) para mejorar la reutilización espacial cuando existen celdas Wi-Fi superpuestas (OBSS): cada círculo punteado representa el área de cobertura de un Punto de Acceso (un BSS distinto), pero todos operan en el mismo canal (ej., canal 36), lo que en un escenario clásico incrementa la contención porque las estaciones interpretan cualquier señal como ocupación del medio y aplazan transmisiones.

Los "colores" mostrados no son colores físicos ni algo visible en el aire; son un identificador lógico (BSS color) que se inserta en la señalización 802.11ax para que los clientes distingan si la energía detectada proviene de su propio BSS o de otro BSS cercano, habilitando decisiones de acceso al canal más finas (p. ej., permitir transmisión concurrente cuando el OBSS se recibe débilmente) y reduciendo retransmisiones. A nivel de implementación, esto se habilita como una función del sistema Wi-Fi (AP/controlador WLAN) mediante configuración y políticas RF; no es "un software de colores" separado ni una opción estética del router/AP, sino un mecanismo del estándar para identificar BSS y optimizar el uso del canal en alta densidad.

3.3 WI-FI 6E

WiFi 6E permite el acceso a la nueva banda de 6 GHz, lo que se traduce en 1200 MHz de espectro limpio en EE. UU. y gran parte del resto del mundo, y casi 500 MHz de espectro limpio en la Union Europea. Esto representa la mayor asignación de espectro sin licencia y casi triplica el espectro disponible para wifi. La «E» de Wi-Fi 6E significa Extended (ampliada) para ofrecer mayor capacidad, canales más amplios y menos interferencias. La banda de 6 GHz ofrece suficiente espacio para múltiples canales de 160 MHz, lo cual es difícil en la banda de 5 GHz.

Figura 39. Franjas de banda en Wi-Fi 6



Fuente: (HPE, 2025)

Además, el Wi-Fi 6E introduce ventajas técnicas que mejoran drásticamente la experiencia del usuario, especialmente en latencia y capacidad pura:

- Espectro "Puro" y sin DFS: A diferencia de la banda de 5 GHz, donde muchos canales están sujetos a DFS (Selección Dinámica de Frecuencia) para no interferir con radares climáticos o militares, la banda de 6 GHz es de uso exclusivo para Wi-Fi. Esto elimina las desconexiones temporales o saltos de canal que ocurren cuando el router detecta una señal de radar.
- Latencia Ultra-Baja: Al no haber dispositivos antiguos (Wi-Fi 4 o 5) compitiendo por el aire en los 6 GHz, la latencia puede ser tan baja

como 2 milisegundos. Esto es crítico para aplicaciones de tiempo real como Realidad Virtual (VR), Realidad Aumentada (AR) y gaming competitivo.

- Canales de 160 MHz masivos: En la banda de 5 GHz solo caben dos canales de 160 MHz sin solaparse. En 6 GHz, dependiendo de la región, puedes tener hasta 7 canales de 160 MHz totalmente independientes, permitiendo velocidades multi-gigabit de forma simultánea para varios usuarios.
- Seguridad Obligatoria: Para que un dispositivo se certifique como Wi-Fi 6E, es obligatorio el uso de WPA3, el protocolo de cifrado más avanzado, lo que garantiza una red más robusta contra ataques de fuerza bruta.
- Limitación de Alcance: Como contrapartida técnica, la frecuencia de 6 GHz tiene una menor capacidad de penetración en paredes y obstáculos que la de 2.4 o 5 GHz. Por ello, se recomienda para uso en la misma habitación o espacios abiertos para maximizar el rendimiento

3.4 WI-FI 7

Wi-Fi 7 (IEEE 802.11be) refuerza el rol de la WLAN como infraestructura de movilidad avanzada al extender capacidades PHY/MAC de muy alto rendimiento y mayor flexibilidad operativa, incluyendo operación multi-enlace y técnicas para aprovechar de forma más eficiente canales anchos y/o parcialmente interferidos.

Figura 40. Características técnicas de Wi-Fi 7

Fuente (Nayarasi, 2024)

EL Wi-Fi 7 implementa lla certificación "Wi-Fi Agile Multiband" (conocida como MBO, Multiband Operation) orientada a que red y cliente compartan información para tomar mejores decisiones de roaming/banda, mejorar balanceo de carga y uso de recursos; en controladoras empresariales. Como se ha indicado se implementa como MBO y se integra con políticas para reducir "*sticky clients*" (clientes pegados a un AP) y empujar transiciones más inteligentes. En despliegues corporativo, MBO no reemplaza las funciones clásicas de movilidad, sino que opera como una capa de coordinación entre cliente e infraestructura para que el roaming y la selección de banda sean guiados por políticas. Sobre la base de mecanismos de roaming rápido (continuidad de sesión mediante reutilización de claves) y de funciones de asistencia al traspaso (descubrimiento de APs vecinos y transición de BSS),

Así, MBO añade señalización inteligente basada en telemetría de congestión (ocupación del medio, utilización de canal y condición por banda), permitiendo que la WLAN mitigue *sticky clients*, distribuya estaciones hacia celdas/bandas con mayor capacidad efectiva y mejore la eficiencia de *airtime* (porcentaje de tiempo que el canal está ocupado) esto quiere decir cuánta información útil

se logra transportar por cada unidad de tiempo de canal consumido.

No obstante, en un entorno corporativo moderno, la movilidad trasciende el dominio de la red de acceso: la continuidad de los servicios de voz, video y aplicaciones críticas depende de una arquitectura de transporte y conectividad WAN capaz de integrar sedes, usuarios remotos y recursos en la nube sin fisuras. En este escenario, las comunicaciones satelitales se consolidan como una extensión estratégica de la infraestructura IP

3.5 COMUNICACION SATELITAL PARA REDES MÓVILES

En escenarios donde la infraestructura terrestre de fibra o microondas es inexistente o vulnerable, las comunicaciones satelitales se integran como un componente crítico de la arquitectura. Esta tecnología no solo garantiza la conectividad en ubicaciones remotas, sino que actúa como un plano de transporte resiliente que extiende el alcance del core hacia el borde, asegurando que la movilidad y el acceso a servicios en la nube se mantengan operativos sin importar la geografía. Sin embargo, aunque tanto el acceso terrestre como el satelital terminan integrándose a una arquitectura IP y a servicios cloud, no son equivalentes desde el punto de vista operacional.

Así el satélite introduce su propio medio de acceso, mayor latencia y con frecuencia más jitter (por gestión del enlace, asignación de recursos radio y, según la constelación, handovers), además de un perfil de disponibilidad/degradación distinta donde influyen condiciones de propagación y visibilidad. En su aspecto técnico de capacidad, la WAN terrestre suele ser más predecible por provisión local, mientras que en satélite la capacidad efectiva es compartida y puede variar por carga y estado del enlace. Por consiguiente, la QoS se controla en planos diferentes: en la LAN inalámbrica se optimiza principalmente la eficiencia de tiempo de aire (airtime efficiency) y la

priorización busca proteger el rendimiento de voz y video (minimizando latencia, variación de retardo y pérdidas bajo contención).

En cambio, en la WAN satelital la estabilidad depende más de mecanismos de control de tráfico como el moldeado de tráfico (shaping) y la limitación/policía de tráfico (policing), además de una gestión cuidadosa de colas en el equipo del cliente (CPE, customer premises equipment) y en el gateway del operador. En términos operativos, el shaping "suaviza" la tasa de envío ajustándola a un perfil sostenido para evitar ráfagas que llenen buffers, mientras que el policing hace cumplir un límite descartando o remarcando paquetes cuando se excede la tasa contratada; ambos se combinan con políticas extremo a extremo para evitar colas excesivas (bufferbloat) que elevan la latencia y amplifican el jitter y así mantener una experiencia de usuario consistente (quality of experience, QoE) más allá del perímetro de la WLAN.

Figura 41. Conexión satelital a estaciones terrenas

Fuente: Autores

Desde la perspectiva de red móvil IP, el satélite se integra de dos formas principales.

1. Como backhaul satelital: un sitio celular (eNodeB/gNodeB) puede ofrecer servicio 4G/5G localmente, pero transportar el tráfico hacia el núcleo del operador a través de un enlace satelital cuando no hay fibra ni microondas terrestres viables.

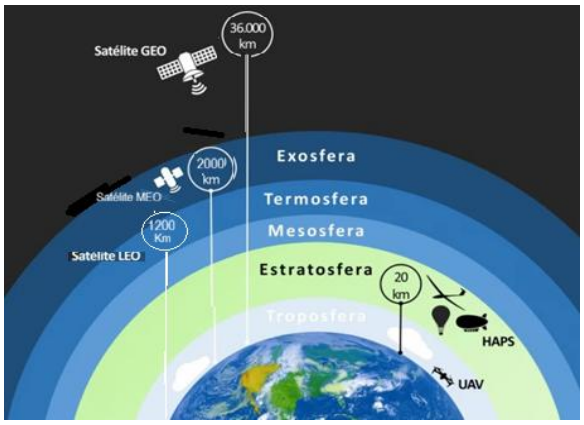
2. Como acceso satelital directo: usuarios, terminales VSAT o dispositivos IoT se conectan al satélite y, mediante el telepuerto, ingresan a redes IP y servicios. Ambas formas sostienen el mismo objetivo: garantizar disponibilidad y alcance de servicios móviles, incluso fuera del *footprint* (huella) terrestre.

Esta convergencia se vuelve más relevante con 5G porque 5G no es solo aumento de velocidad: es una red orientada a servicios (*service-based*) con capacidades de QoS, segmentación lógica (*slicing*), edge/MEC y automatización. Bajo esta lógica, el satélite deja de ser únicamente un "enlace alternativo" y se incorpora como parte de una conectividad multitransporte, donde el núcleo de red puede aplicar políticas, priorización y control de sesiones independientemente de si el transporte es fibra, microondas o satélite. El resultado es un ecosistema en el que Wi-Fi, 4G/5G y satélite trabajan como capas complementarias para habilitar servicios actuales: conectividad rural, continuidad operativa, acceso en movilidad (marítimo/aéreo), IoT masivo y comunicaciones críticas.

3.5.1 ORBITAS TERRESTRES

Existen muchos tipos de órbitas en que los satélites pueden ser colocados en el espacio y el tipo de órbita sea baja, media o geoestacionaria (LEO, MEO, GEO) condicionan métricas esenciales para redes móviles, como latencia, variación de retardo y dinámica de handover.

Figura 42. Orbitas terrestres



Fuente: (InfoEspacial, 2025)

Mientras sistemas GEO ofrecen amplia cobertura con retardos mayores, constelaciones LEO reducen latencia y habilitan experiencias más cercanas a redes terrestres, aunque exigen gestión de movilidad más intensa (handover frecuente entre satélites) y planificación de enlace. En consecuencia, el satélite se alinea naturalmente con los conceptos de red móvil IP que se desarrollarán a continuación: movilidad, continuidad de sesión, calidad de servicio y evolución tecnológica hacia 6G.

En escenarios de telemetría y automatización (agricultura, monitoreo ambiental, logística), el satélite habilita un flujo de datos IP desde estaciones remotas hacia centros de datos, donde los servicios móviles y la analítica pueden operar de forma integrada. Esto conecta con la visión de 5G/6G: una red que no solo transporta datos, sino que integra conectividad con plataformas digitales, edge y servicios inteligentes, manteniendo operación incluso en entornos con infraestructura terrestre limitada.

Así mismo se pueden desplegar arquitecturas de satélites de alto rendimiento (HTS) y muy alto rendimiento (VHTS) revolucionan la conectividad al ofrecer capacidades de datos 10-100 veces mayores que los satélites tradicionales, superando los 100 Gbips. Utilizan múltiples haces puntuales

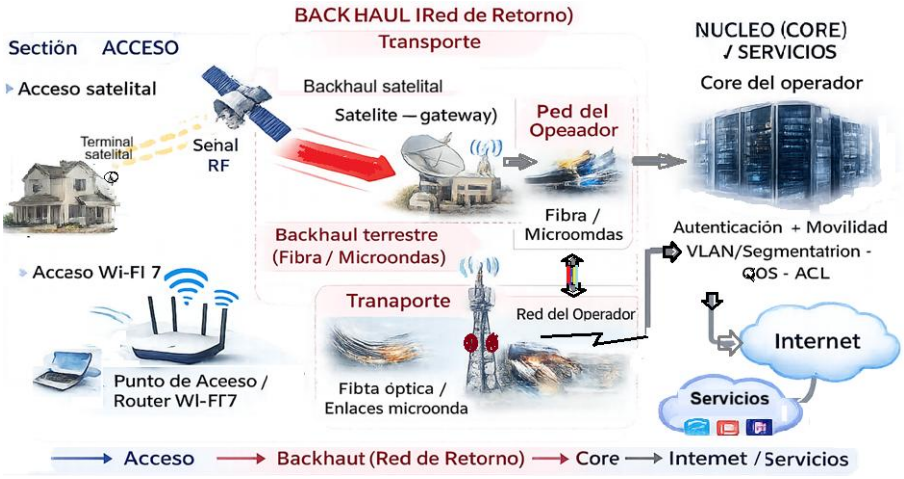
estrechos, reutilización de frecuencias y bandas Ka/Ku para reducir el coste por bit.

Los satélites HTS pueden operar tanto en órbita geoestacionaria (GEO), siendo la mayoría, como en órbitas no geoestacionarias (NGSO) es decir están en orbitas LEO o MEO.

3.5.2 FUNCIONAMIENTO DEL ENLACE SATELITAL

Un enlace satelital se basa en dos tramos: uplink (tierra a satélite) y downlink (satélite a tierra). El satélite opera como repetidor/regenerador mediante transpondedores (recepción, filtrado, amplificación y retransmisión en otra frecuencia), permitiendo transportar tráfico IP entre estaciones terrenas (telepuertos) y terminales remotos.

Figura 43. Backhaul satelital



Fuente: Autores

La estación terrena se comporta como un gateway hacia Internet o hacia la red del operador; por ello, desde el punto de vista de arquitectura, el satélite puede modelarse como un medio WAN con requisitos de línea de vista y antenas direccionales. Aquí el transporte (llamado Backhaul) conecta la estación terrena con el resto del mundo. Además, se encarga de mover los datos de miles de usuarios simultáneamente usando torres con conexión de fibra óptica o enlaces de radio de alta capacidad. Este enfoque permite habilitar conectividad de

banda ancha en zonas sin fibra, y soportar servicios IP con cobertura extensa, manteniendo el satélite como una capa de transporte o acceso complementaria para redes Wi-Fi y móviles.

3.5.3 ESCENARIOS DE USO

Se detallan algunos casos de uso:

(a) Telemetría y estaciones remotas (IoT/monitoreo): En estaciones de campo (ambiental, agrícola, industrial), el nodo de adquisición concentra datos de sensores, aplica preprocesamiento/compresión y los envía por un módem satelital a un centro de datos. Este patrón reduce costos de transmisión y consumo energético, y garantiza conectividad en ubicaciones aisladas.

(b) Comunicaciones más allá de línea de vista (BLoS) en movilidad aérea/marítima: En aeronaves no tripuladas o aeronaves en general, el satélite habilita el control y la telemetría BLoS (Beyond Line of Sight), superando las limitaciones de enlaces terrestres LoS. En paralelo, enlaces terrestres pueden coexistir para servicios locales o para segmentos donde existe cobertura.

El diagrama ilustra cómo un satélite LEO puede actuar como backhaul/transporte o ruta alternativa para que múltiples celdas/nodos terrestres mantengan conectividad con el núcleo IP, algo útil en áreas remotas, fallas de infraestructura o para aumentar resiliencia y cobertura.

FACTORES PARA LA RED MÓVIL IP

Los aspectos satelitales que más influyen en redes IP y que sirven como puente a 4G/5G/6G son:

- Latencia y jitter: dependen de órbita y de la ruta hacia el gateway; afectan servicios interactivos y control.
- Disponibilidad del enlace: degradación por lluvia (especialmente Ku/Ka), alineación de antena, obstrucciones.

- Direccionalidad y LoS: requieren planificación física y radioeléctrica.
- Continuidad del servicio y movilidad: en LEO, la continuidad implica handovers frecuentes (concepto análogo a movilidad celular).
- Rol como transporte/backhaul: integra sitios remotos a la red IP cuando no hay fibra.

Con estas consideraciones, el satélite es una extensión del transporte/acceso en redes IP: aporta cobertura y continuidad donde la infraestructura terrestre es limitada, y al mismo tiempo introduce retos (latencia, disponibilidad, movilidad y handover) que son centrales en el diseño de redes móviles modernas. A continuación, se aborda las redes celulares desde 4G hasta 5G y sus proyecciones hacia 6G, enfocándose en la RAN, el transporte y el núcleo de red como plataforma de servicios.

3.6 RED CELULAR

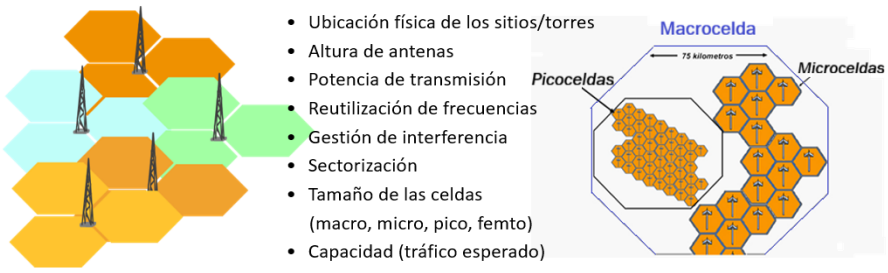
Una red celular es una infraestructura de telecomunicaciones inalámbrica basada en la subdivisión de una zona geográfica en áreas menores denominadas celdas. A diferencia de los sistemas de radio de alta potencia que usan un único transmisor para cubrir grandes distancias, la red celular utiliza múltiples estaciones base de baja potencia, permitiendo que las frecuencias de radio se reutilicen en diferentes áreas sin interferencias mutuas.

Su funcionamiento técnico se basa en la conmutación de circuitos(2G y 3G) o paquetes (Desde 4G es red "todo IP") y en la interconexión de tres elementos clave: el equipo de usuario (UE), la estación base BTS, Base Transciver Station) que puede tener diversa denominación dependiendo su generación de tecnología (/NodeB/eNodeB/gNodeB) y el núcleo de la red (Core), que gestiona la movilidad y el enrutamiento.

3.6.1 PLANIFICACIÓN DE CELDAS

Desde una perspectiva técnica, la planificación de celdas es parte de tres pilares; reutilización de frecuencias, jerarquía de capas o tamaños de sus celdas y sectorización para cobertura. Así desde el 2G hasta el 5G se emplea la planificación de celdas para optimizar el servicio.

Figura 44. Planificación de celdas



Fuente Autores

Una celda es el área geográfica cubierta por una estación base que proporciona servicio radioeléctrico a los dispositivos móviles. Para garantizar la continuidad del servicio durante el desplazamiento del usuario, las redes implementan mecanismos de handover (traspaso), que permiten la transferencia automática de la conexión entre celdas adyacentes sin interrupciones perceptibles. Este principio, vigente desde las primeras generaciones, sigue siendo esencial en 4G (LTE, Long Term Evolution) y 5G, donde la movilidad y la calidad de experiencia dependen de una gestión eficiente de la cobertura y la capacidad.

El diseño de una red celular implica múltiples variables de ingeniería orientadas a optimizar el rendimiento global. Entre los aspectos más relevantes se encuentran la ubicación física de los sitios, la altura de la radiobase/estación base y características de las antenas, la potencia de transmisión y el dimensionamiento del tamaño de las celdas según el entorno y la densidad de tráfico.

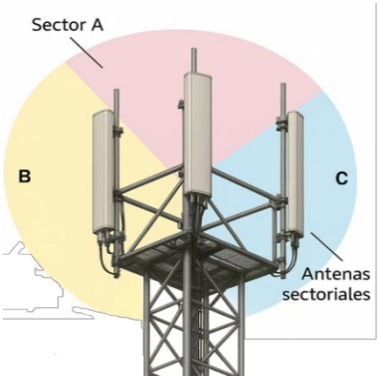
En la reutilización de frecuencias, que permite emplear el mismo espectro en celdas no adyacentes

para maximizar la capacidad del sistema. Sin embargo, este mecanismo requiere una adecuada gestión de interferencias, especialmente la interferencia co-canal (CCI), que puede degradar el rendimiento si no se controla mediante técnicas de planificación, control de potencia y coordinación entre estaciones base.

La reutilización de frecuencias es el principio que habilita alta capacidad en redes celulares (2G hasta 4G), y la sectorización (junto con la planificación radio) controla la interferencia inherente a dicha reutilización para equilibrar cobertura y eficiencia espectral. En 5G este fundamento se mantiene, pero se refuerza con técnicas como Massive MIMO y beamforming, densificación con small cells y mecanismos de gestión/coordinación de interferencias, lo que permite operar con reutilización más agresiva (reuse ≈ 1) y mejor desempeño, aun cuando exista solapamiento físico entre celdas.

La Sectorización y Tilting: implica que en lugar de una antena omnidireccional, se usan antenas direccionales llamadas también antenas sectoriales para dividir la celda en sectores (típicamente 3 sectores de 120°). Además, se aplica el Electrical/Mechanical Tilt (inclinación de la antena) para confinar la energía dentro de los límites geográficos de la celda y evitar el "over-shooting" (que la señal llegue más lejos de lo planeado e interfiera con otras). En la sectorización las estaciones base suelen utilizar antenas sectoriales (dividen la celda en 3 o 6 sectores de 120° o 60°).

Figura 45. Antenas sectoriales



Fuente: Autores

La arquitectura desde 4G a 5G adopta un enfoque heterogéneo, combinando macroceldas para cobertura amplia con micro, pico y femtoceldas para incrementar capacidad en zonas de alta demanda. Asimismo, la sectorización permite dividir la cobertura de una estación base en múltiples sectores direccionales, mejorando la eficiencia espectral.

3.6.2 RED DE 4 GENERACIÓN

LTE, (Long Term Evolution, Evolución a Largo Plazo), es un estándar de comunicación inalámbrica de cuarta generación (4G) diseñado para ofrecer mayor velocidad, menor latencia y mayor capacidad de red que el 3G. Fue la primera generación en ser All-IP ya no utiliza los nodos de conmutación de circuitos (empleados hasta en 3G). Incluso la voz se convirtió en paquetes de datos mediante VoLTE (Voice over LTE). Se detallan aspectos fundamentales de LTE o 4G:

1. Esquemas de Acceso Múltiple (Capa Física)

A diferencia de 3G (W-CDMA), LTE usa dos esquemas distintos según la dirección del tráfico para optimizar el consumo de batería y la velocidad:

- Downlink (Bajada): OFDMA (Orthogonal Frequency Division Multiple Access). Divide el canal en miles de subportadoras ortogonales,

permitiendo una alta eficiencia espectral y resistencia al ruido.

- Uplink (Subida): SC-FDMA (Single Carrier FDMA). Se eligió porque tiene una menor relación de potencia pico a promedio (PAPR), lo que evita que el amplificador del móvil consuma demasiada energía, extendiendo la vida de la batería.

2. Modos de Duplexación (FDD vs TDD)

LTE es sumamente flexible con el espectro y puede operar de dos formas:

- FDD (Frequency Division Duplex): Usa dos bandas de frecuencia separadas (una para subir y otra para bajar) simultáneamente.
- TDD (Time Division Duplex): Usa una sola banda de frecuencia y alterna el tiempo entre subida y bajada. Es ideal para tráfico asimétrico (cuando descargas mucho más de lo que subes).

3. Arquitectura de Red

La red se simplificó eliminando nodos intermedios para reducir la latencia (que puede ser inferior a 5ms). Asi se detallan sus componentes esenciales:

Estación Base inteligente eNodeB (evolved Node B): Es la red de acceso radio de LTE es la parte "visible" de la infraestructura, A diferencia de generaciones anteriores de estaciones base, el eNodeB toma decisiones de gestión de recursos de radio por sí solo y asume funciones que antes estaban en el núcleo de la red, como:

- Gestión de recursos de radio (RRM): Asigna las frecuencias y la potencia a cada móvil en milisegundos.
- Compresión de cabeceras IP: Para que los datos ocupen menos espacio en el aire.
- Cifrado: La seguridad de la conexión se gestiona directamente en la torre

Interfaz X2: Permite que un eNodeB se comunique directamente con los eNodeB vecinos esto es vital para el Handover. Así cuando un dispositivo se mueve, las torres se ponen de acuerdo entre ellas para "pasarse" tu conexión sin tener que preguntar al núcleo de la red (EPC), de esta forma se eliminan micro-cortes.

Interfaz S1: Es la conexión al core o núcleo de red (EPC). Se divide en dos:

- S1-MME: Para el plano de control (señalización y mensajes de red).
- S1-U: Para el plano de usuario (tus datos, fotos, videos)

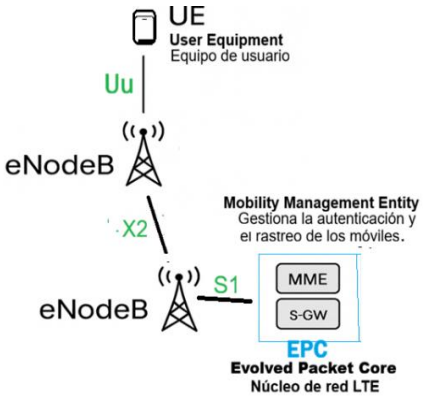
EPC (Evolved Packet Core): El núcleo de la red. Sus componentes clave son:

MME (Mobility Management Entity): El "cerebro" que gestiona la autenticación y el rastreo de los móviles.

S-GW (Serving Gateway): Enrutador que actúa como punto de anclaje para los datos cuando el equipo de usuario se mueve entre celdas.

P-GW (Packet Data Network Gateway): Es el gateway de borde que conecta la red móvil con redes externas (Internet, VPN corporativas, IMS, nubes públicas). No solo "da salida", sino que concentra funciones críticas del plano de usuario (user plane) y de control asociadas a datos:

Figura 46. Red celular LTE



Fuente: Autores

4. Capacidad y Ancho de Banda

- Canales flexibles: LTE no tiene un ancho de banda fijo; puede operar en bloques de 1.4, 3, 5, 10, 15 y hasta 20 MHz.
- MIMO (Multiple Input Multiple Output): Usa múltiples antenas tanto en la torre como en el dispositivo para transmitir varios flujos de datos a la vez sobre la misma frecuencia, multiplicando la velocidad

3.7 ESPECTRO DE RED CELULAR 4G Y 5G EN ECUADOR

En Ecuador, existen 18,4 millones de líneas activas (2026) manteniéndose tres operadores con red celular: Claro, Tigo (antes Movistar) y la empresa estatal CNT (Corporación Nacional de Telecomunicaciones). En paralelo, la habilitación de espectro para 5G se acelera a partir de decisiones regulatorias y de operación comercial inicial, con CNT como primer operador en activar 5G de forma pública y con proyección de expansión.

Desde la perspectiva de cobertura, es importante distinguir "líneas" vs "cobertura poblacional": el Plan Sectorial del MINTEL[9] reporta 94,58% de cobertura poblacional en 2G/3G y 78,49% en 4G (con base en reportes oficiales). Esto explica por qué 4G domina en adopción/uso urbano, pero aún convive con capas heredadas en zonas donde 4G no es universal o donde persisten terminales legacy (MINTEL, 2025). LTE se sostiene en una distribución típica de bandas: bandas bajas para cobertura y penetración en interiores (indoor), y bandas medias para capacidad urbana. La asignación regulatoria de espectro para Servicio Móvil Avanzado (SMA[10]) permite identificar

[9] Ministerio de Telecomunicaciones y de la Sociedad de la Información es la encargada de diseñar políticas y planes para el desarrollo de las TIC, el gobierno electrónico y la conectividad.

[10] En Ecuador es el servicio móvil de banda ancha tipo 3G/4G/5G, Es decir una categoría de servicio regulatorio sólo de Ecuador (local) es diferente a AWS.

qué bloques bajo 6 GHz tiene cada operador y su equivalencia con bandas móviles 3GPP[11] (4G/5G).

Figura 47. Franjas de bandas de frecuencias para 4G en Ecuador

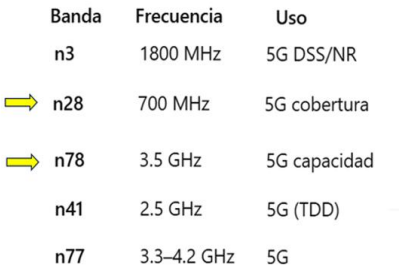
Banda	Frecuencia	Uso
B1	2100 MHz	3G/4G
B2	1900 MHz	3G/4G
B3	1800 MHz	4G
B4	AWS 1700/2100	4G
B5	850 MHz	2G/3G
B7	2600 MHz	4G capacidad
B8	900 MHz	2G/3G/4G
B28	700 MHz	4G cobertura

- **Claro:** LTE en **AWS 1700/2100 (B4)**
- **CNT:** LTE en **AWS (B4)** y **700 (B28)**
- **Movistar/Tigo:** LTE en **1900 (B2)**

Fuente: Autores

Para CNT EP se dispone de espectro en 700 MHz, 1900 MHz y (AWS[12], Advanced Wireless Services) 1700/2100 MHz (ancho total asignado: 100 MHz). En términos de identificadores 3GPP, estas franjas se asocian a B28/n28 (cobertura), B2/n2 (capacidad y migración progresiva desde bandas heredadas) y B4/n66 (capacidad urbana).

Figura 48, Bandas de frecuencias más usadas (sub 6GHz)



Banda	Frecuencia	Uso
n3	1800 MHz	5G DSS/NR
n28	700 MHz	5G cobertura
n78	3.5 GHz	5G capacidad
n41	2.5 GHz	5G (TDD)
n77	3.3–4.2 GHz	5G

Fuente: Autores

[11] 3rd Generation Partnership Project. Proyecto de Asociación de Tercera Generación es una colaboración internacional entre 7 organizaciones de desarrollo de estándares de telecomunicaciones (ARIB, ATIS, CCSA, ETSI, TSG, UIT y TTA) que colaboran para crear y mantener las especificaciones técnicas de las redes móviles

[12] Servicios Inalámbricos Avanzados: Es el nombre de banda/espectro para el rango 1700/2100 MHz (uplink/downlink), muy usado para LTE/5G en varios países. No es un "servicio" regulatorio local, sino un rótulo técnico internacional.

En Ecuador, el despliegue de 5G en bandas por debajo de 6 GHz (sub-6 GHz) se concentra en dos objetivos operativos: capacidad en zonas urbanas y cobertura en áreas extensas e interiores (indoor).

Banda n78 (3,5 GHz) como eje de capacidad

La banda n78 (3,5 GHz) se utiliza como franja principal para capacidad, debido a que soporta canales más anchos y mejor desempeño por celda en zonas de alta demanda (puntos de congestión (hotspots)). En documentación oficial se reportan autorizaciones y ajustes de espectro para operación móvil en el rango 3400–3500 MHz, vinculados al inicio comercial y ordenamiento de bloques. Este tipo de despliegue exige densificación (densification) respecto a bandas bajas, porque el alcance efectivo por sitio es menor.

Banda n28 (700 MHz) como capa de cobertura

La banda n28 (700 MHz) se proyecta como capa de cobertura, por su mejor propagación y penetración en interiores (indoor), permitiendo ampliar huella geográfica con menos sitios. A nivel de arquitectura radio, su rol es complementar la capacidad urbana de 3,5 GHz con una capa de alcance amplio para continuidad de servicio.

3.7.1 REORDENAMIENTO Y REUTILIZACIÓN DE ESPECTRO 6 GHZ

El refarming (reordenamiento y reutilización de espectro) es el mecanismo operativo para acelerar 5G sin depender únicamente de nuevas bandas: consiste en migrar espectro usado por tecnologías antiguas hacia LTE/NR, priorizando bandas con alto valor de cobertura/capacidad. La banda 5925–7125 MHz (6 GHz) se mantiene como punto estratégico por su ancho disponible para capacidad, con discusión regulatoria sobre su uso como espectro no licenciado para Wi-Fi 6E/7 o como espectro para sistemas móviles IMT/5G (International Mobile

Telecommunications, telecomunicaciones móviles internacionales), según decisiones nacionales.

3.8 RED CELULAR DE QUINTA GENERACIÓN

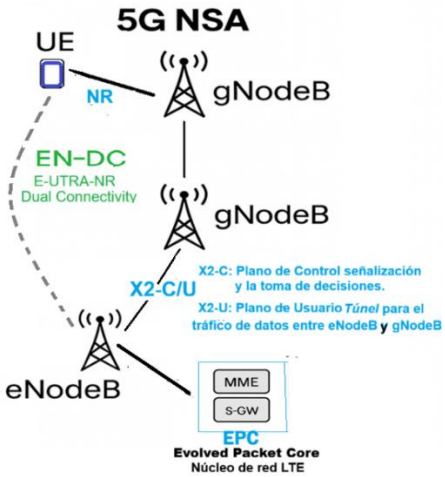
La tecnología 5G New Radio (NR) representa el salto más significativo en redes de telecomunicaciones móviles, diseñado no solo para conectar personas, sino para digitalizar industrias enteras a través de una velocidad y latencia sin precedentes. Sin embargo, su despliegue global no ha sido uniforme, sino que se ha desarrollado bajo dos esquemas técnicos distintos. Actualmente, existen dos arquitecturas principales para implementar esta red celular.

1. 5G NSA (Non-Standalone, no autónoma): despliegue 5G anclado a 4G; el acceso NR (New Radio, nueva radio) se apoya en LTE/EPC para el plano de control y parte de la movilidad, por lo que 5G funciona como una extensión de capacidad sobre la infraestructura 4G existente.
2. 5G SA (Standalone, autónoma): despliegue 5G nativo; el acceso NR se conecta directamente al 5GC (5G Core, núcleo 5G) y la red opera sin anclaje 4G, habilitando capacidades extremo a extremo como menor latencia operativa, políticas más finas y segmentación de red (network slicing, segmentación de red).

En Ecuador, el despliegue inicial de 5G se alinea con un esquema NSA (Non-Standalone, no autónoma) por razones de costo y tiempo de implementación, reutilizando la infraestructura 4G existente. Dado que el país ya inició despliegues 5G comerciales desde 2026, es razonable plantear que, en un horizonte cercano de aproximadamente tres a cinco años, puedan aparecer implementaciones 5G SA (Standalone, autónoma) de forma progresiva, comenzando en zonas urbanas y entornos de alta demanda. La característica técnica central de NSA es el anclaje en LTE mediante EN-DC (E-UTRA-NR Dual Connectivity, conectividad dual LTE-NR).

En este esquema, el UE (User Equipment, equipo de usuario) mantiene una conexión simultánea con un eNodeB (eNB, estación base LTE) y un gNodeB (gNB, estación base 5G). El eNB actúa como nodo maestro para la coordinación, mientras el gNB aporta recursos NR para incrementar capacidad. El resultado operativo es que el usuario percibe mejoras de throughput y eficiencia radio cuando NR está disponible, sin que la red dependa de un núcleo 5G completo para funcionar.

Figura 49. Arquitectura de red celular 5G NSA



Fuente: Autores

Desde la perspectiva de arquitectura, la diferencia práctica entre NSA y SA se entiende separando plano de control (control plane, plano de control) y plano de usuario (user plane, plano de usuario). En NSA, el plano de control se mantiene asociado al ecosistema LTE y al EPC (Evolved Packet Core, núcleo de paquetes evolucionado). El EPC contiene funciones clásicas como MME (Mobility Management Entity, entidad de gestión de movilidad) para control y señalización, y pasarelas de datos como S-GW (Serving Gateway, pasarela de servicio) y P-GW (Packet Data Network Gateway, pasarela hacia red de datos) para el tránsito IP. Con esto, la sesión del usuario, la autenticación y buena parte de la movilidad se resuelven con procedimientos heredados de 4G, mientras NR se usa como "acelerador" de capacidad.

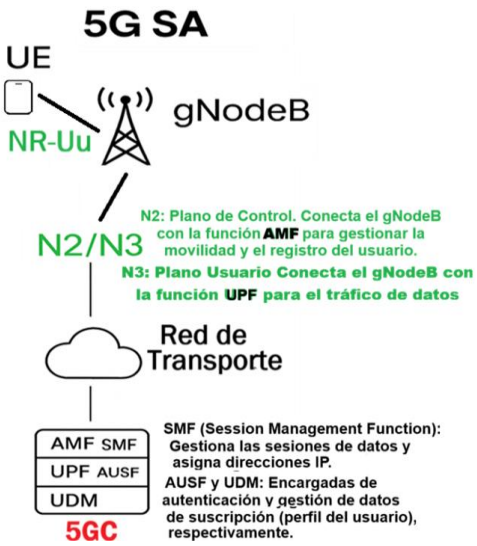
En la práctica, el comportamiento del tráfico en NSA puede resumirse así: el usuario se registra y se gestiona con el anclaje LTE (control), y cuando el terminal y la celda permiten EN-DC, el tráfico de datos se distribuye para aprovechar NR (usuario). Esta separación explica por qué NSA es atractivo para un despliegue inicial: el operador puede habilitar 5G por sectores o sitios concretos (por ejemplo, en áreas con alta demanda), manteniendo continuidad de servicio sobre LTE cuando NR no está disponible o cuando el usuario está fuera de cobertura 5G.

Este diseño también impone límites técnicos frente a una red 5G totalmente autónoma. Como el núcleo sigue siendo EPC, ciertas capacidades avanzadas de 5G quedan condicionadas o no se entregan de forma completa extremo a extremo. Por ejemplo, funciones como segmentación de red (network slicing, segmentación de red) y algunas optimizaciones profundas de latencia requieren típicamente un núcleo 5G nativo. Además, el desempeño final no depende solo del acceso NR: si el transporte (transport, red de transporte) y la

agregación del operador no están dimensionados (capacidad y latencia), el beneficio radio se "aplana" y el cuello de botella se desplaza a backhaul y core.

La 5G SA (Standalone, autónoma), es la tecnología independiente/autónoma la 5G "pura" tecnológicamente hablando donde el gNB se conecta a un 5GC (5G Core, núcleo 5G) y el control deja de depender de LTE. En SA aparecen funciones como AMF (Access and Mobility Management Function, función de acceso y movilidad) y SMF (Session Management Function, función de gestión de sesión) en el plano de control, y UPF (User Plane Function, función de plano de usuario) para el tráfico de datos. Esta separación permite una operación 5G nativa con políticas más flexibles y capacidades avanzadas, pero requiere migración del core, ajustes operativos, y una madurez mayor en el ecosistema de red.

Figura 50. Arquitectura de red celular 5G SA



Fuente: Autores

Las implementaciones 5G SA en Ecuador podrían desplegarse desde 20230 de forma progresiva, comenzando en zonas urbanas y entornos de alta demanda. Esta transición a SA dependerá principalmente de la adopción operacional de un

5GC (5G Core; núcleo 5G), del dimensionamiento de la red de transporte y de la disponibilidad de espectro y sitios suficientes para soportar servicios avanzados con latencia controlada, manteniendo continuidad de voz y movilidad.

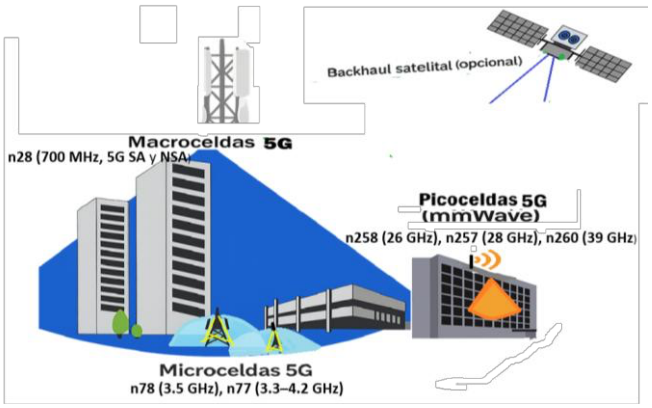
3.8.1 ARQUITECTURA HETEROGÉNEA DE CELDAS EN 5G

La arquitectura heterogénea de celdas en 5G, conocida como HetNets (Heterogeneous Networks), es un modelo de despliegue donde múltiples capas de celdas de diferentes tamaños, potencias y radios de cobertura coexisten para maximizar la capacidad de la red. A diferencia de las generaciones anteriores centradas en grandes torres, el 5G utiliza este enfoque "multi-capa" para manejar el enorme tráfico de datos y la densidad de dispositivos.

Esta arquitectura se organiza principalmente en cuatro niveles de celdas que interactúan entre sí:

- Macrocélulas (Macrocells): Son las torres tradicionales de gran escala. Proporcionan la capa de cobertura base en exteriores, con alcances que van desde 1 hasta 30 km. Operan usualmente en frecuencias bajas y medias para asegurar movilidad continua.
- Microcélulas (Microcells): Celdas de menor potencia que las macro, diseñadas para cubrir áreas urbanas densas o "zonas calientes" donde la señal macro es insuficiente.
- Picocélulas (Picocells): Estaciones base muy pequeñas (del tamaño de una mochila o un router) desplegadas en lugares públicos con gran afluencia como estadios, centros comerciales o aeropuertos.
- Femtocélulas (Femtocells): Nodos de muy corto alcance (menos de 10-50 metros) diseñados para uso doméstico o pequeñas empresas, conectados típicamente a través de la banda ancha del usuario para mejorar la cobertura en interiores.

Figura 51.Tamaño de celdas en 5G



Fuente: Autores

En la figura 51 se muestra una macrocelda 5G (cobertura en color azul) operando en n28 (700 MHz), típicamente usadas para cobertura amplia y mejor penetración en interiores; esta banda puede soportar despliegues tanto NSA como SA según el núcleo de red disponible. También se aprecian alguna microceldas 5G (cobertura en color celeste) en bandas medias como n78 (3,5 GHz) y n77 (3,3–4,2 GHz), cuyo objetivo es incrementar capacidad urbana y densidad de usuarios, requiriendo más sitios que 700 MHz por su menor alcance.

En la parte derecha de la figura 51 se ilustran picoceldas 5G (cobertura en color naranja) en ondas milimétricas (mmWave, ondas milimétricas), con ejemplos n258 (26 GHz), n257 (28 GHz) y n260 (39 GHz), orientadas a capacidad localizada en áreas de altísima demanda (hotspots) mediante celdas pequeñas y enlaces direccionales. Y en parte superior derecha, se representa la opcion de backhaul satelital (satellite backhaul, red de retorno satelital) indica un enlace de transporte opcional para conectar sitios 5G en ubicaciones donde no existe fibra, actuando como alternativa o respaldo del transporte terrestre para integrar estas celdas al núcleo de red.

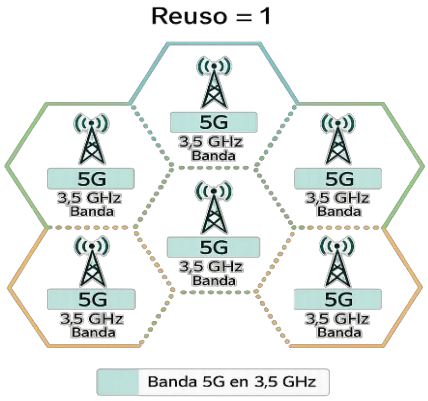
3.8.2 REUTILIZACIÓN DE FRECUENCIA DE FACTOR 1

En 5G sí aplica el concepto de reutilización de frecuencias; lo que cambia es que, en muchos despliegues, se busca operar con reutilización 1 (reuse-1, reutilización 1), es decir, celdas vecinas usan la misma portadora y el mismo bloque de espectro. Esto no elimina la interferencia intercelda, pero maximiza la eficiencia espectral y la capacidad agregada del sistema, siempre que la red gestione bien la interferencia y la asignación de recursos (5G Américas, 2024).

Desde el punto de vista físico y de control radio, la viabilidad de reutilización 1 en 5G se apoya en tres pilares.

1. OFDMA: Permite dividir el ancho de banda en subportadoras y agruparlas en bloques de recursos (RB, resource blocks; bloques de recursos), asignando a cada usuario una fracción tiempo–frecuencia específica.
2. Scheduling (planificación) del gNodeB: Asigna dinámicamente RBs en función de calidad de canal, demanda y QoS; con ello, la red puede evitar asignar a un usuario RBs que estén altamente interferidos en ese instante, priorizar usuarios con mejores condiciones radio o aplicar políticas de equidad.
3. Beamforming (formación de haces) y Massive MIMO (multiple-input multiple-output masivo; MIMO masivo), que concentran energía hacia el usuario y disminuyen la radiación no deseada hacia celdas vecinas, mejorando la separación espacial incluso compartiendo la misma frecuencia.

Figura 52. Representación de Reúso de frecuencia Factor 1



Fuente: Autores

En la figura 52 se muestran los siguientes aspectos técnicos

- Identidad de Banda: Todas las celdas (tanto la central como las periféricas) operan en la misma banda de 3,5 GHz.
- Ausencia de Fragmentación: A diferencia de modelos de reuso 3 o 7 (donde las celdas vecinas deben usar frecuencias diferentes para evitar interferencias), aquí el 100% del espectro disponible está activo en cada una de las antenas.
- Eficiencia Espectral: Esto visualiza la máxima eficiencia posible, ya que no hay "desperdicio" de frecuencias para crear zonas de guarda entre celdas

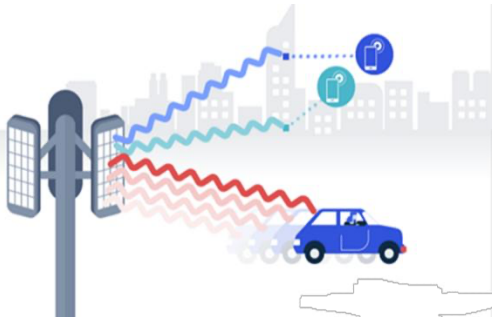
En consecuencia, la reutilización 1 en 5G se interpreta como una estrategia de máxima reutilización soportada por un control fino de recursos: los usuarios comparten la misma banda, pero no compiten “a ciegas”, sino que se separan por asignaciones coordinadas en tiempo–frecuencia y por direccionalidad espacial. Aun así, el desempeño depende de que el despliegue esté correctamente sincronizado y dimensionado (potencias, tilt, densificación, backhaul) y de mecanismos de mitigación de interferencia entre

celdas, porque en bandas medias sub-6 GHz la interferencia intercelda sigue siendo un factor dominante cuando hay alta carga y solapamiento de cobertura. (5G Américas, 2024).

3.8.3 MASSIVE MIMO

Para que este esquema de Reúso de Frecuencia Factor 1, funcione en la realidad sin que la interferencia bloquee a los usuarios en los bordes de las celdas (hexágonos en figura 51), el sistema aplica Massive MIMO y Beamforming para que la energía no se disperse por toda la celda, sino que viaje en haces dirigidos hacia los dispositivos. En 5G con Massive-MIMO se crean decenas de beams a distintos ángulos, cada uno actuando como un "mini-sector". Los beams se activan solo cuando hay usuarios o se apagan cuando no hay tráfico.

Figura 53. Massive MIMO

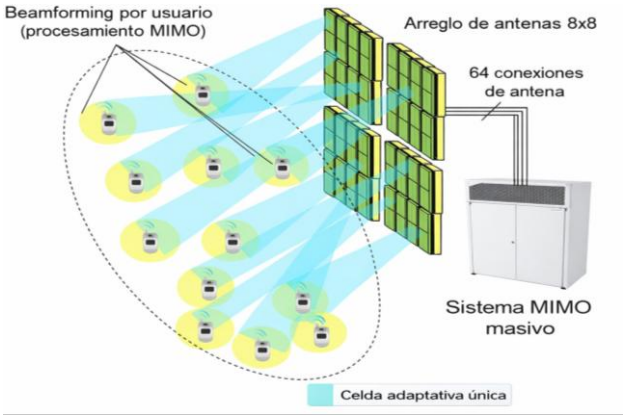


Fuente: Autores

Los sectores no dependen de paneles físicos, sino de beams dinámicos generados electrónicamente. Véase el ejemplo de una antena 64T64R Massive-MIMO, esta puede generar:

- 12 beams de control
- 32 beams de datos
- Separadamente en uplink y downlink
- Todos gestionados por IA del gNodeB
- Esto reemplaza la antigua sectorización 3×120°.

Figura 54 Antena 64T64R MIMO Masivo



Fuente: Autores

3.9 REDES MÓVILES Y COMPUTACIÓN DISTRIBUIDA

La evolución hacia las redes de quinta generación ha transformado la infraestructura de telecomunicaciones de un sistema estático y centralizado a un ecosistema dinámico basado en redes móviles inteligentes y computación distribuida. Esta convergencia es fundamental para gestionar la complejidad técnica del reuso de frecuencias factor 1 y la densificación de celdas, permitiendo que la red deje de ser un simple canal de transmisión para convertirse en una plataforma capaz de procesar datos y tomar decisiones en tiempo real cerca del usuario. Bajo este paradigma, la inteligencia se desplaza hacia el extremo de la red, facilitando la gestión autónoma de recursos y la ejecución de servicios de ultra-baja latencia que definen la era del 5G.

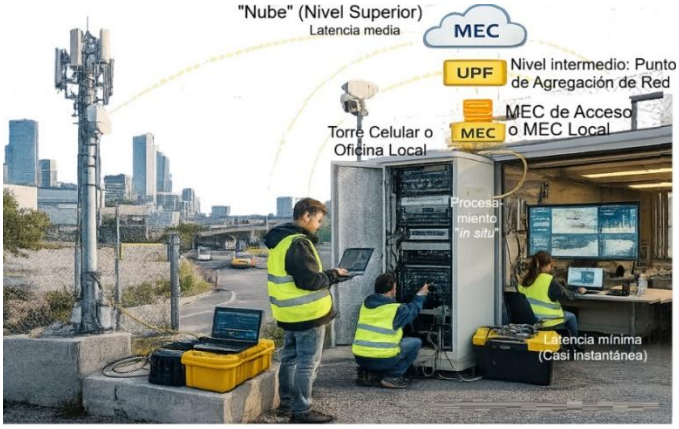
3.9.1 COMPUTACIÓN EN EL BORDE PARA REDES MÓVILES

La computación en el borde multiacceso (MEC, multi-access edge computing; computación en el borde multiacceso) extiende capacidades de procesamiento y almacenamiento hacia puntos cercanos al usuario, reduciendo la dependencia de un centro de datos remoto. El objetivo técnico es disminuir latencia y variabilidad de retardo, descargar tráfico del core y habilitar aplicaciones que no toleran ida y vuelta hacia la nube central.

MEC se integra con 5G principalmente a través del plano de usuario y su ubicación dentro de la red de transporte o en nodos regionales.

Un elemento central es el despliegue de la función de plano de usuario (UPF, user plane function; función de plano de usuario) en el borde. Al ubicar la UPF más cerca del acceso radio, los flujos de datos pueden "salir" hacia servicios locales sin recorrer todo el núcleo. Esto reduce latencia, minimiza congestión en rutas troncales y permite aplicar políticas de encaminamiento y seguridad por servicio. En redes 5G, esta aproximación facilita separar tráfico corporativo, video analítica o telemetría industrial y procesarlo localmente, manteniendo control de QoS extremo a extremo.

Figura 55. MEC



Fuente: Autores

En aplicaciones sensibles al tiempo, MEC se utiliza para analítica de video, realidad aumentada, control industrial asistido, monitoreo en tiempo real y servicios de misión crítica. El rendimiento depende tanto del acceso radio como del transporte y del dimensionamiento del borde: si el enlace de retorno o la capacidad de cómputo en el borde no están alineados, la ganancia se reduce. Por ello, MEC se diseña como parte de la arquitectura de red y no como un servidor aislado.

3.9.2 AUTOMATIZACIÓN E INTELIGENCIA EN REDES MÓVILES

La automatización en redes móviles busca reducir la intervención manual y sostener niveles estables de calidad de servicio y calidad de experiencia (QoE, quality of experience) en escenarios altamente dinámicos (movilidad, variaciones de carga, interferencia, fallas). En la práctica, esto se implementa mediante funciones de auto-gestión y optimización que actúan sobre el acceso radio, el transporte y el núcleo.

Las redes autoorganizadas (SON, Self-Organizing Networks) agrupan capacidades de auto-configuración, auto-optimización y auto-reparación. En el despliegue, SON reduce tiempos de puesta en servicio al automatizar parámetros de vecindad y cobertura; durante la operación, ajusta parámetros de potencia, balanceo de carga y selección de celdas para evitar congestión localizada; y ante eventos de falla, aplica acciones de compensación (por ejemplo, ampliar cobertura temporal de celdas adyacentes). En 4G/5G, estas funciones se relacionan con la reducción de incidencias operativas y con la estabilidad de movilidad en entornos densos.

La analítica en tiempo real incorpora telemetría de red para observar carga por celda, calidad radio, consumo de recursos y patrones de movilidad. A partir de esta telemetría, se puede realizar predicción de tráfico (traffic prediction, predicción de tráfico) con modelos estadísticos o de aprendizaje automático, anticipando picos por zona y franja horaria. En operación, esto permite dimensionar recursos antes del evento (por ejemplo, ajustar umbrales de handover, priorizar tráfico o redistribuir usuarios) en lugar de reaccionar cuando la degradación ya se produjo.

La optimización de movilidad se centra en el proceso de handover (traspaso) y en la asignación de

recursos radio. En redes densas, la movilidad mal parametrizada produce "ping-pong" (cambios repetidos de celda), cortes de sesión o incrementos de latencia. Por ello, se emplean técnicas de aprendizaje automático (ML, machine learning; aprendizaje automático) para ajustar umbrales y temporizadores en función de contexto (velocidad del usuario, calidad de señal, carga de la celda, tipo de servicio). En paralelo, el control de recursos busca maximizar eficiencia espectral sin deteriorar flujos sensibles al tiempo (voz, video interactivo) mediante políticas de programación y priorización coherentes con QoS/QoE.

En el plano de operación, la red incorpora mecanismos de detección de fallos y anomalías basados en correlación de eventos y análisis de series temporales. Se detectan comportamientos como degradación súbita de cobertura, incremento anormal de retransmisiones, caída de capacidad, celdas fuera de servicio o enlaces de transporte congestionados. Estas detecciones activan flujos de corrección automática, reduciendo el tiempo medio de reparación (MTTR, mean time to repair; tiempo medio de reparación) y mejorando disponibilidad.

Figura 56 Automatización de redes móviles

Fuente: Autores

En este enfoque, las políticas se expresan como objetivos de alto nivel y la red ejecuta ajustes sin

intervención constante. Cuando se combina con redes basadas en intención (IBN, intent-based networking; redes basadas en intención), el operador declara metas operacionales (ejemplo, "priorizar servicios críticos en horas pico" o "mantener latencia bajo cierto umbral en una zona"), y el sistema traduce esas metas en configuraciones y acciones sobre el RAN, transporte y core, validando continuamente el cumplimiento.

Esta gestión basada en la intención (IBN) en un despliegue 5G SA donde se mueve el control desde "configurar dispositivos" hacia "declarar objetivos medibles" y dejar que una plataforma traduzca automáticamente esa intención en diseño, provisión y validación de red, cerrando el ciclo con telemetría

3.9.3 CASO ECUADOR Y DESPLIEGUE DE 5G SA

Para masificar 5G, el principal cuello de botella no suele estar en el acceso radio (RAN), sino en la red de transporte (Transport Network): diseño de metro/agregación, selección de sitios de red (nodos), definición de rutas ópticas (ductos, tramos, empalmes) y coherencia integral con capacidad, latencia y resiliencia. En Ecuador con Guayaquil, Quito y Cuenca como principales ciudades de despliegue inicial para 5G NSA y posterior evolución a 5G SA, la tarea técnica no se reduce a "instalar fibra" ni a ejecutar actividades aisladas (tendido por ruta, instalación de nodos, configuración de equipos[13] o cierre de anillos[14]), sino a convertir requerimientos de servicio en un diseño ejecutable de extremo a extremo.

---

[13] Routers (PE/Core), switches de agregación metro, transporte óptico DWDM/ROADM/OTN, equipos de sincronización PTP/SyncE, demarcación NTU/NID, seguridad (firewall) y plataformas de control/automatización (SDN/NMS), más servidores edge/MEC cuando se acerca UPF para 5G SA

[14] Completar una topología redundante y operable, incorporando mecanismos de protección (ERPS/MPLS FRR/OTN), diversidad física, dimensionamiento de capacidad, políticas de QoS, transporte de sincronización PTP/SyncE, y OAM para detección y recuperación rápida ante fallas

En ese contexto, un objetivo razonable para el clúster metropolitano de Guayaquil y el corredor Guayaquil–Quito–Cuenca es definir una red metro de agregación que interconecte gNodeB hacia una UPF[15] regional/edge con latencia E2E ≤ 15 ms para tráfico URLLC[16] y ≤ 25 ms para eMBB[17], garantizando disponibilidad ≥ 99.95% mediante diversidad física real (doble ruta por sitio sin compartir ducto ni trazado crítico) y una capacidad mínima de 10–25 Gbps por sitio, escalable por crecimiento de demanda.

La implementación puede apoyarse en transporte IP/MPLS o Segment Routing (SR-MPLS/SRv6) con ingeniería de tráfico y SLA por clase, complementado con mecanismos de sincronización bajo un presupuesto de error compatible con operación 5G. A nivel de arquitectura, se requiere decidir la ubicación óptima de nodos metro (agregación/PE) y POPs[18] regionales (mini-datacenter edge), priorizando reutilización de infraestructura, mitigación de zonas de riesgo y cumplimiento de restricciones municipales. Finalmente, el despliegue se operacionaliza con un plan por oleadas (deployment in waves), que estandariza fases geográficas y técnicas (habilitación de POPs, interconexión de sitios, activación de QoS/sincronización, pruebas OAM y aceptación), reduciendo riesgo y acelerando el time-to-market (velocidad de respuesta del

[15] User Plane Function: Componente del Core 5G encargado de procesar y transportar los datos del usuario.

[16] uRLLC(Ultra-Reliable Low-Latency Communication) Servicio clave de la tecnología 5G, diseñado para ofrecer una latencia extremadamente baja (aprox.1ms) y una confiabilidad ultra alta (99.999%)

[17] eMBB (Banda Ancha Móvil Mejorada) Ofrece velocidades de datos extremadamente rápidas (hasta 20 Gbps), mayor capacidad de red y menor latencia para streaming 8K, realidad virtual (VR) y realidad aumentada (AR) en dispositivos móviles.

[18] Points of Presence o Puntos de Presencia son centros de datos o nodos de red distribuidos en diferentes provincias o zonas geográficas (fuera del núcleo central) para acercar la inteligencia de la red al usuario final.

operador ante la demanda del mercado) sin sacrificar calidad.

Este enfoque deja preparada la base para el siguiente salto evolutivo: con 6G la presión sobre el transporte será aún mayor (más densificación, más edge, requisitos más estrictos de latencia y confiabilidad), por lo que avanzar desde hoy hacia modelos basados en intención y automatización multi-dominio es una condición habilitante para escalar la red hacia la próxima generación.

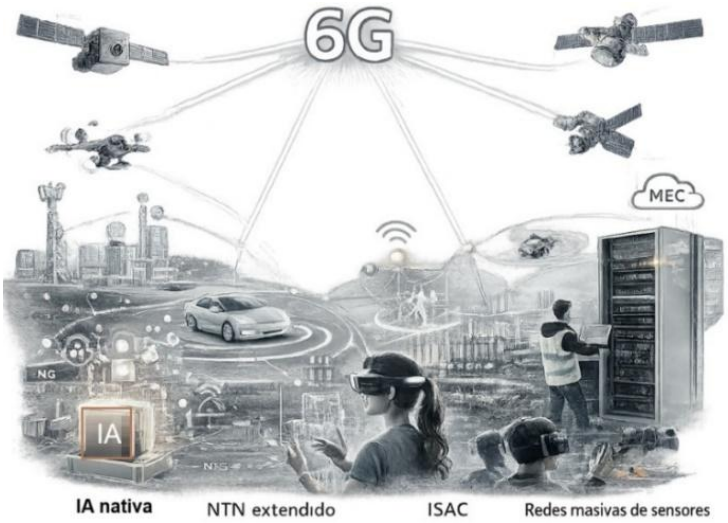
3.10 EVOLUCIÓN HACIA RED 6G

La red 6G se espera su despliegue comercial alrededor de 2030, esta tecnología la red incorpora inteligencia de forma nativa y amplía el ecosistema de conectividad. El concepto de red nativa con IA (AI-native, nativa con IA) implica que la optimización, la seguridad y la operación se diseñan desde el inicio para funcionar con automatización avanzada y decisiones guiadas por datos, manteniendo objetivos de rendimiento y resiliencia. Aparte de este aspecto de IA nativa, 6G también tienes :

- Cobertura 3D Global: Red de muchos satélites (órbita baja o LEO), drones y aviones integrados en la red para dar señal en el mar, aire y desiertos.
- Gemelos Digitales (Digital Twins): Representaciones virtuales exactas de ciudades o personas que se actualizan en tiempo real gracias a la latencia de microsegundos (1000 veces más rápida que el 5G).
- Internet de los Sentidos y Hologramas: Imágenes que muestran comunicación por hologramas 3D de alta fidelidad o interfaces cerebro-computadora.
- Comunicación y sensado integrados (ISAC, integrated sensing and communication; comunicación y sensado integrados), donde

el sistema radio no solo transmite datos, sino que extrae información del entorno (presencia, movimiento, características del canal) se convierte en una especie de radar inteligente para habilitar aplicaciones de seguridad, logística y control.

Figura 57. Escenarios de red celular 6G



Fuente: Autores

CAPÍTULO 4: DISPOSITIVOS IOT DE CONECTIVIDAD INTELIGENTE

En los entornos IoT actuales la conectividad inteligente en plataformas del IoT no se limita a la transmisión de datos, sino a su capacidad de integrarse en infraestructuras basadas en el protocolo IP. El enfoque de conectividad inteligente es precisamente cómo hardware de dispositivos IoT (ESP32[19], STM32[20], etc.) utilizan protocolos específicos para integrarse de forma eficiente en la pirámide Edge–Fog–Cloud.

4.1 DISPOSITIVOS IOT DE BAJO COSTO

Los dispositivos IoT de bajo costo constituyen la capa de adquisición de datos y control en sistemas ciberfísicos, permitiendo la interacción directa con el entorno mediante sensores y actuadores. Su adopción masiva se debe a su bajo costo, bajo consumo energético y facilidad de integración con redes IP. Así algunas plataformas o tarjeta controladoras como Raspberry Pi, ESP32, Arduino y STM32 se utilizan ampliamente en aplicaciones de monitoreo, automatización y control distribuido dentro de arquitecturas Edge–Fog–Cloud.

Raspberry Pi es un computador de placa única (SBC, Single Board Computer) basado en procesadores ARM, diseñado para ejecutar sistemas operativos completos como Linux. A diferencia de los microcontroladores, ofrece mayor capacidad de procesamiento, almacenamiento y conectividad (Ethernet, Wi-Fi, Bluetooth). Esto lo convierte en una plataforma adecuada para funciones de edge computing (computación en el borde), tales como agregación de datos, ejecución de servicios MQTT, procesamiento local, pasarelas IoT (IoT Gateway) y aplicaciones de analítica ligera.

---

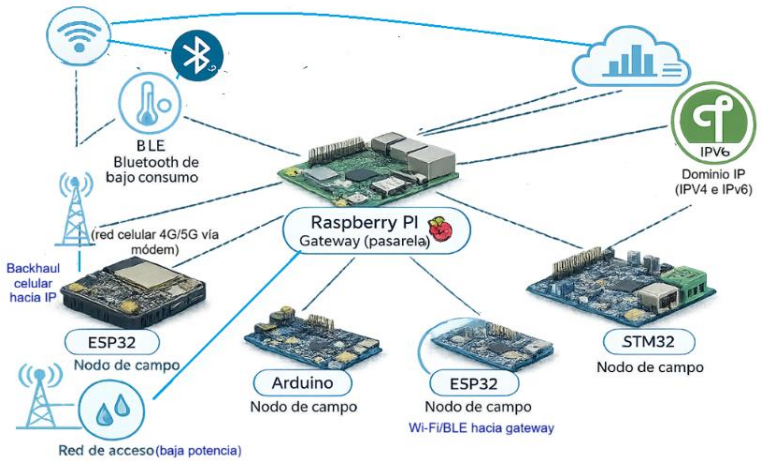
[19] Espressif Systems es la empresa fabricante del microcontrolador de 32 bits. Tiene una arquitectura de doble núcleo y mayor potencia.

[20] STMicroelectronics. Es la empresa fabricante del microcontrolador de 32 bits (núcleos ARM Cortex-M) es el estándar en la industria profesional (médica, automotriz, robótica) por su robustez

El ESP32 es un microcontrolador de bajo consumo con conectividad integrada Wi-Fi y Bluetooth (BLE, Bluetooth Low Energy), optimizado para aplicaciones IoT distribuidas. Su arquitectura permite la lectura de sensores, control de actuadores y transmisión directa de datos hacia plataformas en la nube o servidores locales mediante protocolos ligeros como MQTT o HTTP. Gracias a sus modos de ahorro de energía, es adecuado para nodos alimentados por batería en escenarios de monitoreo remoto.

Las plataformas Arduino están orientadas a la adquisición y control en tiempo real mediante microcontroladores de arquitectura simple y programación directa sobre hardware. Su principal ventaja es la facilidad de desarrollo y la amplia disponibilidad de módulos (sensores, actuadores, interfaces de comunicación). En entornos IoT, Arduino suele operar como nodo de campo, encargado de la captura de variables físicas y el envío de información hacia un gateway o dispositivo de mayor capacidad de procesamiento.

Figura 58. Dispositivos IoT e integración al dominio IP



Fuente: Autores

Por su parte, los microcontroladores STM32, basados en núcleos ARM[21] Cortex-M, ofrecen mayor rendimiento y eficiencia energética que plataformas de desarrollo básicas, por lo que se emplean con frecuencia en sistemas embebidos de operación continua y requerimientos temporales más estrictos. En IoT industrial, su valor radica en la disponibilidad de interfaces embebidas para integración con sensores, actuadores y equipos de control, tales como UART (Universal Asynchronous Receiver-Transmitter: transmisión serie asíncrona), SPI (Serial Peripheral Interface: interfaz serie de periféricos), I2C (Inter-Integrated Circuit: bus integrado) y CAN (Controller Area Network: red de área de controladores), además de conectividad IP cuando se incorpora Ethernet o módulos inalámbricos según el diseño.

En conjunto, estas plataformas habilitan nodos distribuidos capaces de adquirir variables físicas, ejecutar control local y entregar datos a un gateway o a un nodo de borde con mayor capacidad de cómputo. Esta base de hardware facilita la construcción de infraestructuras IoT escalables en arquitecturas Edge–Fog–Cloud, donde la heterogeneidad se gestiona mediante pasarelas y servicios de integración. Los mecanismos de comunicación orientados a eficiencia (protocolos de aplicación y tecnologías de acceso) se desarrollan en el apartado siguiente, como parte de la conectividad inteligente extremo a extremo.

4.2 PROTOCOLOS DE RED PARA CONECTIVIDAD OT

La selección de plataformas de hardware en IoT no debe entenderse de forma aislada, sino como el soporte físico que habilita la conectividad inteligente. La "inteligencia" operativa reside en la

[21] Advanced RISC Machine: Es una arquitectura de procesadores basada en RISC (Reduced Instruction Set Computer o Computadora de Conjunto de Instrucciones Reducido), que prioriza la eficiencia energética y la ejecución rápida de instrucciones sencillas.

capacidad de transformar variables del mundo físico en datos integrables a infraestructuras basadas en IP, ya sea mediante conectividad IP directa (Wi-Fi/Ethernet) o mediante pasarelas (gateways) que acoplan redes de baja potencia al dominio IP. Así, el hardware fija el presupuesto de cómputo y energía, mientras que el modelo de comunicación determina la eficiencia extremo a extremo (sobrecarga, latencia, confiabilidad y escalabilidad).

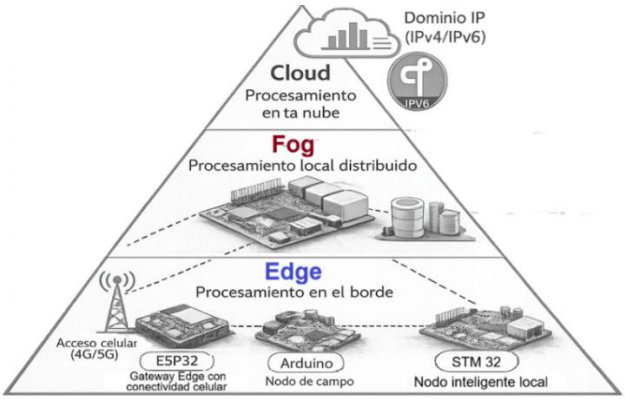
Plataformas como ESP32 y Raspberry Pi materializan esta conectividad por sus roles complementarios dentro de la pirámide Edge–Fog–Cloud. El ESP32 opera típicamente como nodo embebido de campo para adquisición y control, mientras que la Raspberry Pi se usa como gateway (pasarela)/edge node (nodo de borde) para agregación de datos, preprocesamiento y exposición de servicios hacia redes IP.

En entornos heterogéneos, muchos nodos de ultra-bajo consumo (por ejemplo, Arduino o STM32 según configuración) no participan directamente en IP y dependen de enlaces locales y/o módulos de comunicación; en estos casos, el gateway asume la función de concentración y acoplamiento al dominio IP, permitiendo interoperabilidad con servicios distribuidos en el borde, la capa intermedia y la nube. Esta separación de responsabilidades optimiza consumo en el extremo y centraliza la complejidad de integración en el borde.

Dentro de este conjunto de tecnologías de baja potencia, una opción especialmente relevante por su alineación con redes IP es Thread. Thread es una alternativa moderna basada en IEEE 802.15.4 con soporte nativo de IPv6 (Internet Protocol version 6, protocolo de Internet versión 6) en la red de sensores, lo que simplifica la integración con infraestructuras IP y facilita el encaminamiento a través de gateways hacia servicios Edge–Cloud, reduciendo la

necesidad de traducciones complejas en redes IoT de corto alcance.

Figura 59 Pirámide Edge-Fog-Cloud



Fuente: Autores

La Tabla 8 resume la sinergia hardware-conectividad indicando el rol de cada plataforma, sus interfaces y el tipo de conectividad resultante dentro de una infraestructura IoT basada en IP.

Tabla 7 Comparación de plataformas y protocolos de red

Plataforma	Rol en la Arquitectura	Protocolos/Interfaces	Tipo de Conectividad
Raspberry Pi	Gateway (pasarela) / Edge node (nodo de borde)	HTTP/REST (servicio web), MQTT broke (intermediario), CoAP proxy (pasarela CoAP), AMQP (mensajerío empresarial, opcional)	IP nativa (Ethernet/Wi-Fi)
ESP32	Nodo de campo inteligente	MQTT (telemetría) CoAP (servicios ligeros), HTTP/REST (configuración/AP), BLE (Bluetooth Low Energy: Bluetooth de bajo consumo)	IP nativa por Wi-Fi + no-IP por BLE (requiere pasarela para integración IP en BLE)
Arduino	Sensor/actuador base	UART/Serial (serie) I2C (Inter-Integrated Circuit bus integrado), SP (Serial Peripheral	No-IP por defecto; IP posible con módulo

		Interface: interfaz serie periférica) hacia gateway; Ethernet/Wi-Fi (po módulo)	(depende del hardware)
STM32	Nodo industrial/crítico (tiempo real)	CAN, Modbus (protocolo industrial), RS-485 (enlace industrial) 6LoWPAN (IPv6 sobre 802.15.4), Thread (malla IPv6) cuando aplica	Híbrida: industrial no-IP en campo + IP (IPv6) de baja potencia cuando usa 6LoWPAN/Thread

Nota: Un escenario típico de conectividad IP en IoT consiste en nodos de campo ESP32/STM32 que adquieren variables físicas y se comunican con un gateway Raspberry Pi (Ethernet/Wi-Fi). La Raspberry Pi agrega y traduce datos de dispositivos no-IP (por ejemplo, Arduino vía UART/I2C/SPI) y publica la telemetría hacia servicios Edge–Fog–Cloud mediante protocolos de aplicación (por ejemplo, MQTT/HTTP) sobre el dominio IP (IPv4/IPv6).

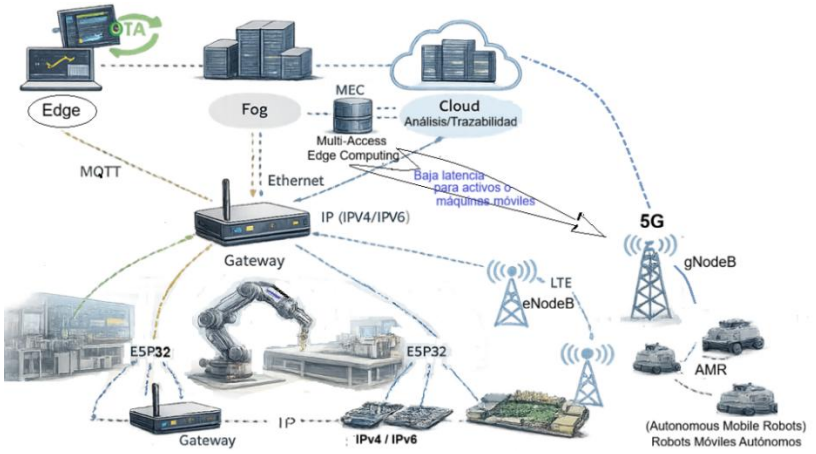
4.3 Heterogeneidad como convergencia al dominio IP

En un despliegue IoT real, la heterogeneidad no proviene únicamente de usar distintas plataformas (Arduino, ESP32, STM32 o Raspberry Pi), sino de la coexistencia de múltiples capacidades de red, enlaces de acceso e interfaces de campo dentro de un mismo sistema. Esta diversidad es esperada: algunos nodos están optimizados para sensado y control con consumo mínimo, otros para agregación y gestión, y no todos pueden ejecutar una pila IP completa. Por ello, el objetivo arquitectónico no es "homogeneizar" el hardware, sino converger funcionalmente al dominio IP (IPv4/IPv6) para lograr interoperabilidad y administración coherente.

La convergencia ocurre cuando el sistema define un plano común de integración basado en IP, donde las aplicaciones y servicios consumen datos mediante mecanismos estandarizados, independientemente del enlace local utilizado en el extremo. En este

enfoque, los nodos no-IP o de conectividad limitada se comunican mediante interfaces locales (por ejemplo, UART, I2C, SPI o enlaces de corto alcance), mientras que un nodo con mayor capacidad, típicamente un gateway (pasarela), concentra la telemetría, normaliza formatos y expone servicios sobre IP hacia Edge–Fog–Cloud. De esta forma, tecnologías distintas pueden coexistir sin romper la arquitectura: el extremo se mantiene eficiente y el dominio IP actúa como capa de unificación.

Figura 60. Convergencia al dominio IP + MEC + movilidad



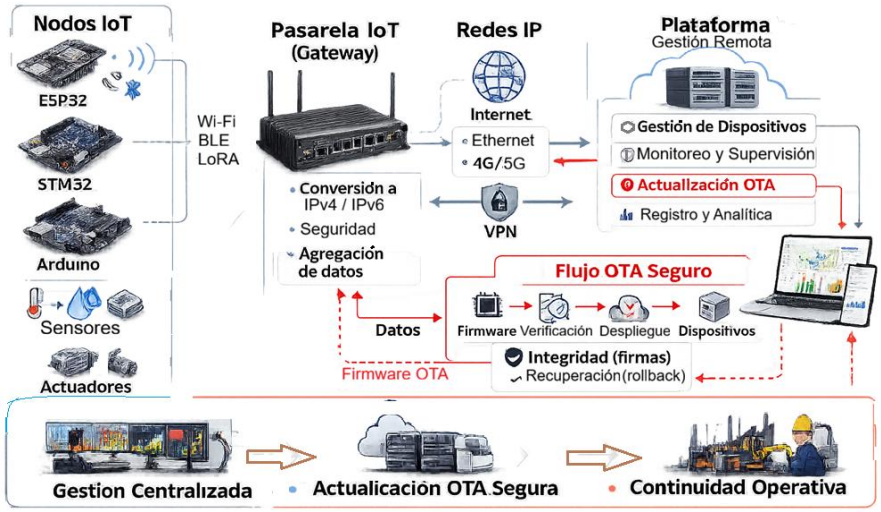
Fuente: Autores

La figura 60 resume la heterogeneidad de enlaces (Wi-Fi, BLE, LTE/5G) y su convergencia al dominio IP (IPv4/IPv6) mediante gateway y servicios Edge–Fog–Cloud, incorporando MEC (Multi-access Edge Computing: computación en el borde multiacceso) para baja latencia en movilidad industrial. La red heterogénea se convierte en conectividad inteligente cuando la convergencia al dominio IP (IPv4/IPv6) habilita integración e interoperabilidad, escalabilidad y operación remota. Esto permite funciones como supervisión centralizada, reconfiguración y despliegue de actualizaciones remotas OTA (Over The Air: actualización remota inalámbrica) según las necesidades del sistema.

4.4 CASO: PASARELA IOT Y GESTIÓN REMOTA SOBRE IP

En este contexto, la pasarela IoT (gateway) actúa como punto de convergencia al dominio IP (IPv4/IPv6), integrando nodos de campo heterogéneos y exponiendo servicios de supervisión y control a niveles superiores (Edge–Fog–Cloud). A diferencia de una red puramente local, el uso de IP proporciona un plano común para administración remota, integración con plataformas de analítica y orquestación de cambios operativos sin intervención física permanente en el extremo.

Figura 61, Arquitectura de gestión remota



Fuente: Autores

Esta arquitectura permite la gestión centralizada de dispositivos distribuidos geográficamente mediante conectividad IP, donde la pasarela IoT actúa como intermediario entre los nodos de campo y las plataformas de supervisión. En operación normal, los dispositivos transmiten telemetría hacia el gateway, el cual realiza agregación, filtrado y encapsulación sobre IPv4/IPv6 antes de enviar la información a través de redes WAN (Ethernet, VPN o 4G/5G) hacia el sistema de gestión. Desde este nivel, los operadores pueden supervisar el estado de los equipos, ajustar parámetros de operación y ejecutar acciones de control sin requerir acceso físico al sitio.

Un elemento crítico del esquema es el mecanismo de actualización OTA (Over-the-Air), que permite distribuir nuevas versiones de firmware o configuraciones de manera remota. Este proceso se realiza de forma controlada, incorporando verificación de integridad mediante firmas digitales o hashes, despliegue escalonado por grupos de dispositivos y mecanismos de recuperación (rollback) en caso de fallo, garantizando la continuidad del servicio y reduciendo el riesgo operativo.

4.4.1 CASOS DE USO TÍPICOS DE ESTA ARQUITECTURA

La misma estructura de pasarela IoT y gestión sobre IP más OTA (Over-the-Air) se aplica en múltiples sectores. En industria y manufactura, permite actualizar remotamente controladores y nodos de monitoreo, supervisar líneas de producción y ajustar parámetros operativos sin detener el proceso. En energía y servicios públicos, soporta medidores inteligentes y estaciones remotas, habilitando la actualización de firmware en dispositivos de campo y la gestión centralizada de infraestructura distribuida. En agricultura de precisión, integra sensores de suelo y variables climáticas a través de un gateway, facilitando el ajuste remoto de umbrales y algoritmos de control, además de reducir desplazamientos para mantenimiento.

En ciudades inteligentes, se utiliza en alumbrado público, estaciones ambientales y sistemas de semaforización, donde la administración desde centros de control y la actualización masiva de dispositivos son críticas. En logística y activos distribuidos, habilita configuración, telemetría, diagnóstico y actualizaciones remotas de equipos desplegados en flotas o ubicaciones alejadas.

Así la integración de pasarelas IoT con capacidades de gestión remota transforma la operación de sistemas distribuidos, al habilitar supervisión continua, mantenimiento proactivo y el despliegue seguro de actualizaciones OTA. Este enfoque reduce los costos

operativos, minimiza los tiempos de intervención en campo y establece una conectividad y control escalable, alineada con el modelo arquitectónico Edge–Fog–Cloud.

4.5 CONCLUSIÓN DE LOS AUTORES

Desde la perspectiva de las redes IP, la convergencia entre pasarelas IoT, gestión remota y mecanismos de actualización OTA (Over the Air) consolida un modelo operativo para sistemas distribuidos, capaz de sostener supervisión continua, mantenimiento proactivo y despliegues controlados sin intervención física permanente en el extremo. En conjunto, estas capacidades reducen costos operativos, disminuyen tiempos de intervención en campo y fortalecen la continuidad del servicio mediante verificación de integridad y recuperación ante fallos (rollback), elementos críticos en infraestructuras reales.

En este marco, el modelo Edge–Fog–Cloud se consolida como referencia de diseño para integrar adquisición de datos, control, analítica y administración centralizada sobre un plano común basado en IP (IPv4/IPv6), facilitando interoperabilidad, crecimiento por etapas y una operación más eficiente.

REFERENCIAS

5G Américas. (2024). *Estatus de las bandas de espectro radioeléctrico sub-1 GHz en América Latina.* Obtenido de https://www.5gamericas.org/wp-content/uploads/2024/10/WP-espectro-bandas-bajas-ES-2024.pdf

Abid, A. (2025). *Virtual LAN*. Obtenido de https://n9.cl/7crcxw

Adarsh, U., KartHa, V., & Santhosh, C. (2022). *Spectroscopy: A promising tool for plastic waste management*. Obtenido de TrAC Trends in Analytical Chemistry: https://www.sciencedirect.com/science/article/abs/pii/S0165993622000176

Afril. (2024). *Cable DAC SFP+ de 10 G*. Obtenido de https://www.afril.net/product-p-495305.html

Albarrán, C. (2024). *Qué es la fibra óptica y para qué sirve*. Obtenido de https://www.redestelecom.es/especiales/fibra-optica-que-es-como-funciona-y-mejores-aplicaciones/

Angolu. (2024). *CAT 8 Patch Cable (40 Gbit/s, 2000 MHz)*. Obtenido de https://www.angolube.com/producto-p-713105.html

AscentOptics. (2025). *La tecnología y las perspectivas de aplicación de los módulos ópticos de 800G*. Obtenido de https://ascentoptics.com/blog/es/the-technology-and-application-prospects-of-800g-optical-modules/

Azam, W. (2018). *Modelo de red jerárquica de 3 capas de Cisco*. Obtenido de https://w7cloud.com/cisco-3-layer-hierarchical-network-model-core-distribution-access/

Carvajal, E. (2023). *Transceptor SFP+ 10GBase-SR para Cisc*. Obtenido de https://es.scribd.com/document/668569086/DATA-FIBRA-Transceptor-multimodo-SR-SFP

Conectividad. (2022). *¿CONOCES LOS ESTÁNDARES DE CABLEADO ESTRUCTURADO?* Obtenido de https://conectividad.com.gt/cableado_estructurado/

Dekam. (2025). *Cables de fibra óptica simplex vs. dúplex: Guía completa para compradores*. Obtenido de https://dekam-fiber.com/es/simplex-vs-duplex-fiber-optic-cable/

ETU-LINK. (2024). *Introducción a la modulación DML y EML para láseres de módulos ópticos*. Obtenido de https://es.etulinktechnology.com/blog/introduction-to-dml-and-eml-modulation-for-optical-module-lasers_b817

Fibconect. (2024). *Fibra monomodo versus fibra multimodo*. Obtenido de https://fibconet.com/es/single-mode-fiber-vs-multi-mode-fiber/

FiberMall. (2021). *¿Qué es la señal PAM4 y sus desafíos de prueba de caracterización?* Obtenido de https://www.fibermall.com/es/blog/what-is-pam4-signal-and-its-challenges.htm?srsltid=AfmBOooRr5A3ahz8U4El4cRkOmhqjsNgp0PBoPCPEvuGgoTbuBTfB1_w

FiberMall. (2022). *Combine el transceptor QSFP-DD con DWDM Coherent*. Obtenido de https://www.fibermall.com/es/blog/400g-dwdm-combine-qsfp-dd-with-dwdm-coherent.htm

FiberMall. (2023). *Dos tipos de convertidores de medios de fibra*. Obtenido de https://n9.cl/ryd23

Franco, J. (2023). *Estudio del desempeño de la transmisión de datos de una red de 100gbps DSP*. Obtenido de https://dspace.ups.edu.ec/bitstream/123456789/25064/1/UPS-GT004409.pdf

FS. (2021). *Cable par trenzado UTP, FTP y SFTP*. Obtenido de https://www.fs.com/es/blog/utp-or-stp-cables-for-10gbase-t-network-4970.html

Gong, S. (2023). *Diferencia entre VLAN y Subred, Cuántas VLAN Usar* . Obtenido de https://www.router-switch.com/faq/vlan-vs-subnet-difference-and-best-practices.html

Hpe. (2023). *¿Qué es la arquitectura spine-leaf?* Obtenido de https://www.hpe.com/lamerica/es/what-is/spine-leaf-architecture.html

HPE. (2025). *¿Qué es el Wi-Fi 6E?* Obtenido de https://www.hpe.com/mx/es/what-is/wi-fi-6e.html

HPE-Juniper. (2023). *Guía de arquitectura de estructura EVPN-VXLAN para centros de datos*. Obtenido de https://www.juniper.net/documentation/mx/es/software/nce/sg-005-data-center-fabric/topics/topic-map/data-center-interconnect-ipvpn.html

InfoEspacial. (2025). *LEO, MEO y GEO: pros y contras de tres órbitas terrestres*. Obtenido de https://www.infoespacial.com/texto-diario/mostrar/5387314/leo-meo-geo-pros-contras-tres-orbitas-terrestres

Lanchi, L. (2020). *Diseño de una red de transporte DWDM con longitudes de onda*. Obtenido de de 400Gbps para la ciudad de Guayaquil

Lazarus. (2021). *¿Qué Cable de Red usar para Cada Situación?* Obtenido de https://n9.cl/cpe4s

Liu, Y., Guo , J., Jiang, B., Song, Y., & Zhang, P. (2025). *FastIOV: Inicio rápido de la virtualización de E/S*

de red de paso para contenedores seguros. Obtenido de https://dl.acm.org/doi/10.1145/3689031.3696066#:~:text=Abstract-,Abstract,and%20asynchronous%20VF%20driver%20initialization.

Llorach, J. (2024). *Qué es el ancho del canal wifi y por qué no necesitas un router con 160MHz*. Obtenido de https://bandaancha.eu/articulos/que-ancho-canal-wifi-que-no-necesitas-10386

López, C. (2017). *Cuatro tipos comunes de conectores de fibra óptica*. Obtenido de https://n9.cl/fp37kk

Lu, L. (2025). *¿Qué es una unidad de rack*. Obtenido de https://www.lianjer.com/es/unidades-de-rack-ru/

Meng, L. (2022). *Análisis del papel de la tecnología de comunicación 5G basada en redes de sensores inteligentes en la construcción y diseño del Internet de las cosas en zonas francas*. Obtenido de https://n9.cl/y7gux

MINTEL. (2025). *Plan Sectorial de las Telecomunicaciones y de la Sociedad de la Información 2024-2025*. Obtenido de https://n9.cl/sd0bd

Natkaniec, M., & Bieryt, N. (2023). *Un análisis de las redes inalámbricas mixtas IEEE 802.11ax en la banda de 5 GHz*. Obtenido de https://n9.cl/1dcwz

Nayarasi. (2024). *Descripción general de Wi-Fi 7*. Obtenido de https://mrncciew.com/2024/12/14/wi-fi-7-overview/

Network Baseline. (2026). *Arquitectura de 3 capas*. Obtenido de https://www.instagram.com/p/DTsrwfzkalS/

Network.Switch. (2025). *Puntos de acceso Cisco Wi-Fi 6/6E*. Obtenido de https://n9.cl/lq104

Opcore. (2024). *SFP+, SFP28, QSFP+, QSFP28, QSFP56, QSFP-DD, QSFP112 vs OSFP, ¿cuáles son las diferencias?* Obtenido de https://n9.cl/zpqbw

Optico. (2025). *1000 Media Converter*. Obtenido de Productos: https://www.opticomfiber.com/fiber-media-converter/media-converter.html

Pérez, I. (2024). *DWDM: Sistemas de Instalación, Pruebas y Mantenimiento*. Obtenido de Tesis U, de SANTANDER : https://noesis.uis.edu.co/server/api/core/bitstreams/f043d145-5b42-4b16-959a-fc6b2bd99436/content

Pululart. (2020). *9 métricas imprescindibles en una tienda online*. Obtenido de https://pululart.com/blog/metricas-para-tienda-online/

Quraishi, H., Tavakoli, E., & Ren, F. (2021). *Un estudio de arquitecturas de sistemas y técnicas para la virtualización de FPGA*. Obtenido de EEE Transactions on Parallel & Distributed Systems, vol. 32, n.º 09: doi: 10.1109/TPDS.2021.3063670.

RackOnline. (2023). *¿Cuáles son las características de la fibra monomodo?* Obtenido de https://n9.cl/87nqt

Ribeiro, L. (2019). *Acesso e Upgrade da SIU via WinFiol*. Obtenido de Scribd: https://es.scribd.com/document/421898451/Procedimento-Lte-Siu

Salazar, J. (2016). *Redes Inalámbricas*. Obtenido de Techpedia: https://upcommons.upc.edu/server/api/core/bitstreams/b685c311-e3aa-4898-9ec1-f35d955468ec/content

Siemon. (2025). *Arquitecturas de Red Emergentes para Centros de Datos de IA y sus*

Aplicaciones. Obtenido de https://files.siemon.com/es/guide/guide-siemon-emerging-ai-data-center-network-architectures-and-applications-es.pdf

User Datacenter. (2022). *Arquitectura Leaf Spine vs Arquitectura Tradicional*. Obtenido de https://blogs.salleurl.edu/es/arquitectura-leaf-spine-vs-arquitectura-tradicional

Villacís, G., & Guamán, G. (2022). *1. Implementación de bastidor de fibra óptica en el laboratorio de telecomunicaciones de la Universidad Estatal Península de Santa Elena*. Obtenido de https://n9.cl/v5t2u

GLOSARIO

API	Application Programming Interface
BSSID	Basic Service Set Identifier
BLE	Bluetooth Low Energy
COAP	Constrained Application Protocol
DWDM	Dense Wavelength Division Multiplexing
ECMP	Equal-Cost Multi-Path
EPC	Evolved Packet Core):
HPC	High-performance computing
HTTPS	Hypertext Transfer Protocol Secur
IS-IS	Intermediate System to Intermediate System
ISM	Industrial, Scientific, and Medical
JSON	JavaScript Object Notation
MEC	Multi-access Edge Computing
MIMO	Multiple Input Multiple Output
NFV	Network Function Virtualization
NAS	Network Attached Storage
NIC	Network Interface Controller
OFDMA	Orthogonal Frequency Division Multiple Access
OTA	Over The Air
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
SDN	Software Defined Networking
SFP	Small Form-factor Pluggable)

SFP28 28	Small Form-Factor Pluggable
SNMP Protoco	Simple Network Management
SNR	Signal to Noise Rate
STP	Spanning Tree Protocol
QSFP Pluggable	Quad Small Form-factor
STP	Shielded Twisted Pair
UTP	Unshielded Twisted Pair
TLS	Transport Layer Security
VPN	Virtual Private Network
VXLAN Network	Virtual Extensible Local Area
WLAN	Wireless Local Area Network
WAN	Wide Area Network
WLC	Wireless LAN Controller

ISBN: 978-9942-53-199-5

Compás
capacitación e investigación