

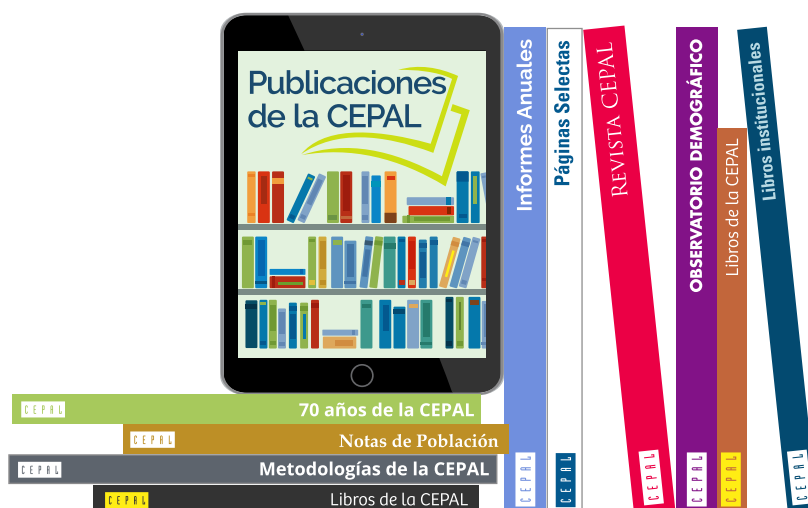
Computación en el borde

Análisis tecnológico,
económico, regulatorio
y de mercados en
América Latina y el Caribe

Omar de León



Gracias por su interés en esta publicación de la CEPAL



Si desea recibir información oportuna sobre nuestros productos editoriales y actividades, le invitamos a registrarse. Podrá definir sus áreas de interés y acceder a nuestros productos en otros formatos.

Deseo registrarme



www.cepal.org/es/publications



www.instagram.com/publicacionesdelacepal



www.facebook.com/publicacionesdelacepal



www.issuu.com/publicacionescepal/stacks



www.cepal.org/es/publicaciones/apps

Computación en el borde

**Análisis tecnológico, económico, regulatorio
y de mercados en América Latina y el Caribe**

Omar de León



Este documento fue elaborado por Omar de León, Consultor de la División de Desarrollo Productivo y Empresarial de la Comisión Económica para América Latina y el Caribe (CEPAL), y coordinado por Fernando Rojas, Asistente Superior de Asuntos Económicos de la misma División, con el apoyo de la Alianza Digital Unión Europea-América Latina y el Caribe.

Ni la Unión Europea ni ninguna persona que actúe en su nombre es responsable del uso que pueda hacerse de la información contenida en esta publicación. Los puntos de vista expresados en este estudio son del autor y no reflejan necesariamente los puntos de vista de la Unión Europea.

Las Naciones Unidas y los países que representan no son responsables por el contenido de vínculos a sitios web externos incluidos en esta publicación.

No deberá entenderse que existe adhesión de las Naciones Unidas o los países que representan a empresas, productos o servicios comerciales mencionados en esta publicación.

Las opiniones expresadas en este documento, que no ha sido sometido a revisión editorial, son de exclusiva responsabilidad del autor y pueden no coincidir con las de las Naciones Unidas o las de los países que representan.

Publicación de las Naciones Unidas
LC/TS.2024/105
Distribución: L
Copyright © Naciones Unidas, 2024
Todos los derechos reservados
Impreso en Naciones Unidas, Santiago
S.2400835[S]

Esta publicación debe citarse como: O. de León, "Computación en el borde: análisis tecnológico, económico, regulatorio y de mercados en América Latina y el Caribe", *Documentos de Proyectos* (LC/TS.2024/105), Santiago, Comisión Económica para América Latina y el Caribe (CEPAL), 2024.

La autorización para reproducir total o parcialmente esta obra debe solicitarse a la Comisión Económica para América Latina y el Caribe (CEPAL), División de Documentos y Publicaciones, publicaciones.cepal@un.org. Los Estados Miembros de las Naciones Unidas y sus instituciones gubernamentales pueden reproducir esta obra sin autorización previa. Solo se les solicita que mencionen la fuente e informen a la CEPAL de tal reproducción.

Índice

Introducción	5
I. Descripción de la computación en el borde (EC)	7
A. Descripción de la estructura conceptual y física, y funciones de la EC	7
1. Consideraciones generales sobre la EC	7
2. Edge Computing vs. Cloud Computing	10
3. Tecnologías emergentes relacionadas a la EC	11
4. Resumen de los beneficios que impulsan la EC	15
B. Principales características de las acciones del ETSI con relación a la MEC	16
C. Evolución de la EC y de los proveedores de las tecnologías directamente relacionadas	19
D. Soberanía de los datos y la EC	22
E. Sostenibilidad de la computación en la nube y en el borde	24
F. Resumen gráfico de la estructura de la EC	27
II. Evolución de los mercados relacionados a la EC	29
A. Descripción de los submercados y su papel en la EC	29
B. Los CSP como usuarios de la MEC y proveedores de servicios sobre ella	33
C. Organización territorial de la EC en los hiperescaladores	38
D. Distintos aspectos sobre las proyecciones para el futuro	41
III. Aspectos económicos de la EC	45
A. CAPEX y OPEX en las diferentes modalidades de nube vs. servicios prestados y calidad	45
B. Estimaciones de costos de Edge vs. Cloud	47
C. Gestión de costos de los servicios en la nube	52
D. Resumen comparativo del impacto de la EC y su relación con la CC	54

IV. Tendencias en el despliegue previsto de la EC en Latinoamérica, el Caribe y en otras regiones.....	57
A. América Latina.....	57
1. Aspectos generales de la EC en la región	57
2. Despliegues previstos o en fase de desarrollo en LAC	61
B. Otras regiones.....	66
V. Impacto en países y en verticales seleccionados	71
A. Impactos en el sector productivo en general.....	71
B. Impactos en verticales seleccionados	74
C. Impactos en las Pequeñas y Medianas Empresas. Mejoramiento de la eficiencia en la PyMEs y los cambios tecnológicos y de gestión.....	78
D. Conjunción de la EC, la Internet de las Cosas y la 5G en el desarrollo productivo.....	82
VI. Requerimientos de política pública para el despliegue y adopción de la EC	85
A. Alianza Europea para los datos industriales, Edge y Cloud.....	85
1. Anuncio de la Alianza Europea para los datos industriales, Edge y Cloud	86
2. Declaración de la Alianza Europea para los Datos Industriales, Edge y Cloud	87
3. Hoja de ruta de la tecnología industrial europea para la próxima generación Cloud-Edge	89
B. Documento de trabajo de los operadores de red de telecomunicaciones de Europa.....	90
C. Aplicabilidad en la región	92
Anexo A1.....	95

Cuadros

Cuadro 1	Comparación cualitativa de EC vs. CC.....	54
Cuadro 2	Índice de preparación para las tecnologías de vanguardia en América Latina y el Caribe en 2022, por país, publicado en 2023	58
Cuadro 3	Proporción de las empresas de los integrantes de la muestra por cantidad de empleados	66
Cuadro 4	Ingresos de la MEC y su crecimiento, por vertical.....	69
Cuadro 5	Prestaciones y usos de los diferentes tipos de EC.....	72
Cuadro 6	Estado de avance de la EC en diferentes verticales (2023)	75
Cuadro 7	Las PyMEs saben que existe una desconexión entre ellas y las empresas, brecha de comprensión de la relación entre ellas	79

Diagramas

Diagrama 1	Tipos básicos de estructura de la EC	28
Diagrama 2	Evolución prevista de las fases de EC	32
Diagrama 3	Evolución de los CSP hacia la nube	38
Diagrama 4	Capas de informática en la nube	48
Diagrama 5	Comparación de costos TCO	50
Diagrama 6	Aplicaciones adecuadas para la MEC	70

Introducción

Existe una fuerte tendencia hacia la adopción de la Computación en el Borde (EC) en sus versiones privada y pública. Se habilita así un conjunto de ventajas o requerimientos como son el muy bajo retardo de respuesta, la reducción importante de costos de transporte de grandes datos al procesarlos en forma cercana o enviarlos preprocesados, la reducción del área de exposición de los datos confidenciales y sensibles, la mejor resiliencia de su red, el cumplimiento con las disposiciones nacionales de residencia de datos, entre otros, que justifican plenamente los relativamente mayores costos que surgen de la reducción de las economías de escala de la nube tradicional.

Con esta estructura de procesamiento de muy bajo retardo y alta capacidad se habilitan las tecnologías emergentes para las cuales la Computación en el Borde es el núcleo de su desarrollo y operación, y de la habilitación de nuevos casos de uso, incluyendo los residentes en los Proveedores de Servicios de Comunicaciones (CSP).

I. Descripción de la computación en el borde (EC)

A. Descripción de la estructura conceptual y física, y funciones de la EC

1. Consideraciones generales sobre la EC

La EC presenta un modelo desconcentrado de procesamiento de datos que está emergiendo principalmente debido a las exigencias del mercado en cuanto a los grandes y crecientes volúmenes de información que se genera en los usuarios y su procesamiento con muy bajo retardo, la cantidad de dispositivos que generan información a procesar, a la caída de costos y a la evolución de las tecnologías de acceso inalámbrico y particularmente la 5G. Estos cambios no pueden ser soportados por los Centros de Datos tradicionales y generan un estrés para su procesamiento y almacenamiento en la nube. Con la EC el almacenamiento y procesamiento de datos se realiza marginalmente en los equipos terminales como las computadoras personales y similares, y principalmente en centros de alta capacidad cerca de ellos y con estructuras generalmente complejas, logrando procesar eficientemente grandes flujos de datos, reducir los retardos en la devolución de la información favoreciendo las acciones en tiempo real, y mantener información sensible en un perímetro seguro. Es, en esencia, la descentralización del almacenamiento y procesamiento de datos.

La EC no se define en general en relación con el borde de la red del prestador del servicio de telecomunicaciones, que es uno o más límites de una red donde se transiciona a la red de comunicaciones operada por una empresa o institución. El borde se refiere principalmente a un área indefinida pero que asegura que los centros de cómputo instalados en esa área se encuentren conectados a los equipos del usuario de ellos a una “distancia” de entre 1ms. y 50 ms. de retardo.

Este fenómeno es observado por Gartner, la que predice que para 2025 el 75% de los datos producidos y usados por las empresas se procesarán fuera de los Centros de Datos centralizados.

La EC ubicada en Centros de Datos públicos constituye un avance sobre los Centros de Datos residentes ya que en ella se puede utilizar la potencia de procesamiento distribuido y ajustado a las

necesidades que surgen en sitios distintos o en equipos móviles o nómades. Se logra una mayor eficiencia en el uso de la infraestructura debido a las economías de escala, y a su vez se generan nuevos servicios y aplicaciones como la movilidad asistida o autónoma, la industria manufacturera o extractiva avanzada, entre otras.

Como se observa, este concepto de EC es más amplio que lo que es, estrictamente hablando, la EC en Centros de Datos de alta capacidad que se encuentran localizados en diferentes instituciones o empresas, formando una red fija segura y con muy bajo retardo, como sucede en hospitales, fábricas, centros logísticos, entre otros, en que además la información sensible se mantiene dentro de los límites físicos de ellas. También existe en menor grado ese procesamiento local de alcance limitado en dispositivos, como son los POS o Puntos de Venta, algunos dispositivos inteligentes de la Internet de las Cosas como los sensores y actuadores, los terminales de computación y los ruteadores, entre otros, que junto con realizar procesamiento de información tienen comunicaciones con el mundo exterior.

La situación cambia cuando se emplean redes en el borde, tanto dentro como en el exterior de los predios, en las cuales se encuentra alojada una mayor capacidad de procesamiento, constituyendo el primer eslabón de una cadena de múltiples centros que suele llegar hasta la computación en la nube. En general, y en forma creciente, son propiedad de empresas especializadas que ofrecen distintas facilidades como servicio, tal como se prestan en la nube, a veces con las mismas interfaces y procedimientos, pero cerca del usuario final. Es en esta etapa del desarrollo de las redes de EC, en que se observa que el despliegue de la 5G pasa a desempeñar un papel decisivo con su muy alta capacidad de transmisión de datos y muy bajos retardos, agregada con éxito a las conexiones fijas por cable de fibra óptica entre los terminales y los centros de procesamiento. Aplicaciones críticas como la robotización de una planta de producción, los automóviles asistidos o autónomos (que tienen también gran capacidad de cómputo a bordo), o algunas aplicaciones médicas como la cirugía asistida o remota, entre otras, y móviles por naturaleza, requieren una alta capacidad de procesamiento y retardos muy bajos, y al mismo tiempo recursos de procesamiento compartidos para reducir los costos que resultarían si se instalaran centros propios.

Esta nueva estructura de procesamiento de datos genera el entorno que estimula el desarrollo de nuevas funciones y procedimientos productivos dentro de las organizaciones, así como de nuevos modelos de negocio a través de los cambios internos de las empresas y el desarrollo de evolucionadas formas de comunicación con la comunidad.

Los centros de la EC propios pueden ser parte de arquitecturas centralizadas o distribuidas, donde pasan a trabajar coordinadamente con otros centros de EC, y más allá, con la Computación en la Nube, constituyendo entonces una Nube Híbrida¹. Este trabajo coordinado provee simultáneamente potencia total de cálculo, potencia distribuida en los bordes para atender cuestiones de respuestas en tiempo real y abundante cantidad de datos, confidencialidad controlada y reducción de costos de transporte de información que pueda ser procesada localmente. Al mismo tiempo permite integrar los grandes Centros de Datos en la nube para la realización de tareas de mayor envergadura, en un diseño económicamente óptimo.

La denominación de EC se usa frecuentemente más que la Computación en la Niebla. Es que cuando los recursos de procesamiento y almacenamiento de datos se acercan a la generación de los datos, existen distintos grados de acercamiento. La EC podría ser la ubicada en un predio y con capacidad limitada (procesamiento limitado de información en un local comercial), mientras que a veces se usa la Computación en la Niebla cuando está ubicada cerca de los datos, pero con más poder de almacenamiento y procesamiento (p.e. en la misma ciudad que donde ellos se generan en el caso de Ciudades Inteligentes). Los proveedores de los servicios de la EC se acercan, por el momento, más a

¹ Una nube híbrida es un modelo informático mixto en el que las aplicaciones se ejecutan en diferentes entornos: nubes públicas y nubes privadas, incluidos los Centros de Datos en el borde.

la llamada Computación en la Niebla. Una nube híbrida es un modelo informático mixto en el que las aplicaciones se ejecutan en diferentes entornos: nubes públicas y nubes privadas, incluidos los Centros de Datos en el borde. datos por lo que razonablemente lo que no es Computación en la Nube (CC) se podría incluir en la EC, y así se procede en este documento y en la generalidad de los trabajos sobre este tema. Más aún, en un documento de 2023 de la Agencia de la Unión Europea para la Ciberseguridad, "Fog and Edge Computing in 5G"², se hace la distinción entre ambas, pero en forma consistente con lo indicado: "La computación en la niebla es una arquitectura que está una capa debajo de la computación en la nube.

..., donde el objetivo principal es reducir la carga de trabajo de los dispositivos perimetrales y en la nube ofreciendo recursos de red y hardware a ambas partes..." y "la computación en el borde es la adición más reciente a los paradigmas informáticos cubiertos en este informe. Permite que los dispositivos y servidores perimetrales amplíen las capacidades de la nube en el borde para resolver procesos computacionales y almacenar datos muy cerca del usuario". En definitiva, se trata solamente del grado impreciso de proximidad al borde, por lo que la distinción más generalizada en la literatura, de computación en la nube y EC, es la más adecuada para este trabajo.

Junto a la Computación en el Borde (Edge Computing) aparece el concepto secundario de la Pasarela en el Borde (Edge Gateway). Mientras que la primera se ocupa principalmente del almacenamiento y procesamiento de los datos, la pasarela tiene como funciones gestionar la comunicación incluyendo el control de tráfico y la conversión de protocolos, la seguridad y el encaminamiento entre las redes local y externa o WAN.

También en el caso de EC constituida por centros distribuidos, la EC requiere a nivel físico la gestión dinámica de la Red de Área Extendida o WAN para optimizar los recursos de computación y de transporte de la información, mejorando la ejecución de las aplicaciones, minimizando la latencia y mejorando la experiencia del usuario. La definición por software de la WAN (SD-WAN) sustituye las rígidas redes tradicionales de WAN por redes de enlaces gestionados centralmente por software, o redes virtuales, de acuerdo con la política que se requiera. Esta virtualización desacopla la red física del plano de control, permitiendo la centralización de la configuración de los servicios y el control de tráfico. Las SD-WAN no sustituyen las WAN, sino que solamente habilitan la configuración más eficiente de los enlaces, adaptada a los requerimientos de la EC.

Esta transformación de la gestión de las redes físicas es acompañada por la necesidad de proveer también una seguridad más amplia en la nube, no ya desde entidades fijas como los servidores, los dispositivos a nivel de las organizaciones o similares, sino que se debe atender la seguridad considerando que el acceso puede tener lugar desde sitios variables, por lo que la seguridad debe estar relacionada a las personas, los equipos y demás terminaciones de las redes. Con este asunto como objetivo, Gartner publicó en 2019 un informe titulado "The Future of Network Security Is in the Cloud" (El futuro de la seguridad en la red está en la nube), en el que describe el concepto de Servicio de Acceso Seguro en el Borde (SASE, Secure Access Service Edge) como un método de acceso seguro definido por software y basado en la nube destinado a proteger los accesos desde cualquier lugar resultantes de la transformación digital de la empresa. SASE se basa en un modelo de seguridad Zero Trust, que no concede al usuario acceso a las aplicaciones y datos hasta que su identidad haya sido verificada, incluso si ya se encuentran dentro del perímetro de una red privada. Es una arquitectura de seguridad que reemplaza múltiples métodos de seguridad existentes en las redes tradicionales.

SASE y SD-WAN son virtualizaciones con focos distintos pero complementarios, considerando la conectividad dinámica de SD-WAN y el poder de seguridad ubicuo de SASE.

² <https://www.enisa.europa.eu/publications/fog-and-edge-computing-in-5g/@download/fullReport>.

Los Centros de Datos para la EC pueden resumirse desde otro punto de vista en:

- Centros de Datos propios. Son los ubicados en los predios de las organizaciones y son de su propiedad, sean comprados o arrendados. Cuando se consideran de borde son pequeños, ubicados lo más próximo posible a los usuarios de la organización. Los retardos son del orden de 2 ms.
- Centros de CSP para proveer la MEC, o EC multi acceso o móvil. Los CSP disponen de una malla extensa de sitios de radiobases que es el punto final más próximo a los clientes que demanden EC, pudiendo disponer de la conectividad en ambas direcciones y la energía. En la sección de análisis de la MEC se observa que los CSP están virtualizando sus redes, requiriendo disponer de una infraestructura similar a la de EC en sus radiobases. Otra alternativa es que el CSP ubique el Centro de Datos del borde en sus Puntos de Presencia. Los Puntos de Presencia son sitios de los CSP donde se alojan los equipamientos de frontera hacia la red de acceso, que pueden ser donde se alojan los controladores de radiobases, donde se alojaban los equipos de centrales de telecomunicaciones fijas, hoy generalmente transformados en centros digitales, y por tanto son sitios adecuados para instalar Centros de Datos en el borde para servir una ciudad pequeña, o una zona de una ciudad mayor. En este caso los retardos serán del orden de 5 ms. en las radiobases o del orden de 30 ms. o menos en los Puntos de Presencia.
- Centros de Datos regionales. Abarcan grandes zonas con retardos menores a 50 ms., aproximadamente, salvo compromisos que adoptan algunos hiperescaladores, u otros proveedores de EC, en cuanto a límites menores.

2. Edge Computing vs. Cloud Computing

Existe actualmente en el mundo algunas opiniones que consideran a las dos Computaciones EC y CC como competitivas, e inclusive a la EC como sustituto futuro de la CC. Sin embargo, ambas son complementarias, y cada una tiene prestaciones que se adaptan eficientemente a cada tipo de aplicación, e inclusive pueden compartir óptimamente su uso para una misma aplicación en sus distintas etapas de procesamiento. Por ejemplo, puede haber un procesamiento requerido en la EC por el bajo retardo, y al mismo tiempo almacenar datos en centros en la nube, que como son resultado de un preprocesamiento serán de menor volumen que los datos originales a procesar requiriendo menos capacidad de transporte, y tiene además un menor TCO el almacenamiento en los grandes Centros de Datos en la nube.

La CC constituye la puesta a disposición de los usuarios de infraestructura física gestionada por el proveedor, de servidores, almacenamiento, procesamiento, aplicaciones, y otras variantes de servicios a través de la Internet. La eficiencia en su empleo radica en las modalidades muy flexibles de contratación por arrendamiento o por uso, con la posibilidad de cambiar las capacidades a demanda, y alojando flexiblemente los recursos en distintas ubicaciones físicas por razones de seguridad o de calidad de servicio. Esta eficiencia es a su vez el resultado de la compartición de la infraestructura de equipamiento, servicios y aplicaciones, la cual es aprovechada por los proveedores para permitir a los usuarios reservar y liberar recursos en forma dinámica, si así lo desean, aparte de la reducción general de costos por la simple compartición de infraestructura. Una ventaja colateral, pero de impacto, es el de la reducción de la huella de carbono y el impacto ambiental en general, como el del menor consumo de agua o de abióticos, o por la búsqueda de su ubicación donde la energía sea limpia.

Las nubes son básicamente públicas y privadas. Cuando al usuario le resulta conveniente combina el uso de ambas en lo que suele llamarse Nube Híbrida. La Nube Pública ha sido la primera en desarrollarse y hoy en día, para algunos de sus proveedores, es la que tiene los volúmenes más importantes de recursos compartidos en cantidad y variedad, así como ubicaciones en sitios

estratégicos por conectividad, disponibilidad de energía a bajo costo, sostenibilidad, etc. Por contraposición, la Nube Privada se encuentra dentro del local del usuario o es tercerizada, aislada del exterior, y siempre para el uso exclusivo de todos los recursos por parte del usuario, empleando una red segura. Estos Centros de Datos permiten realizar trabajos sensibles con altos niveles de control y visibilidad, cumpliendo los requerimientos de seguridad del usuario. Obviamente estas redes privadas son más costosas y se justifican solo para lograr el cumplimiento de ciertos estándares del usuario. Las Nubes Híbridas emplean recursos de ambos tipos de nubes de acuerdo con sus objetivos de requerimientos y eficiencia en el uso de recursos.

La decisión final sobre el uso de la EC o de la CC, reside en los volúmenes de información, en el máximo retardo soportado por el usuario, y en otras condiciones como la seguridad y el cumplimiento de normas. Si los volúmenes son bajos y soportan retardos de transporte hasta el centro de procesamiento del orden de aproximadamente 50 ms. o más, el uso de las CC es más eficiente que la EC. Por tanto, el uso de un modelo u otro, o de la combinación de ambos, depende estrictamente de cada caso particular.

También la EC es mandatoria en sitios de baja calidad de comunicaciones de red, pero siempre considerando la posibilidad del uso de la CC para almacenamiento o procesamiento de datos con menores exigencias de retardo, entre otras condiciones, sea por redundancia o por reducciones de costos.

Y últimamente se han expandido las redes públicas de EC, como se puede ver en general en las secciones "II.C. Organización territorial de la EC en los hiperescaladores" y "IV. Tendencias en el despliegue previsto de la EC en LAC, y en otras regiones", para el caso de Latinoamérica y Caribe, en que se agrega a las prestaciones descritas para la EC en general, la posibilidad de extender servicios que requieren bajo retardo y gran volumen de datos a amplias Zonas dentro de una misma Región del proveedor de nube.

La CC tiene costos más bajos que las soluciones de EC propia, debido a la conversión de CAPEX de hardware de TI para esta EC, por el OPEX de suscripción para el almacenamiento en la CC. Los retardos en poner en servicio recursos adicionales en la CC también tienden a ser mucho menores que en las soluciones de EC propia, ya que los usuarios pueden escalar rápida y fácilmente con solo colocar la orden en el operador, sin tener que pasar por largos procesos de aprovisionamiento. La EC pública tiene las mismas ventajas que la CC frente a EC propia, pero en menor proporción debido a las menores economías de escala.

La configuración final de la estructura de almacenamiento y procesamiento de datos de un usuario sobre nubes públicas y privadas, así como en el Borde o en Centros de Datos principales, depende de múltiples factores.

3. Tecnologías emergentes relacionadas a la EC

Estas estructuras de datos están ligadas a múltiples tecnologías³ emergentes que emplean, o a las que impulsan en su desarrollo. Todas estas tecnologías tienen una relación sinérgica entre ellas, cada una depende de las demás para su aplicabilidad y desarrollo, y en conjunto constituyen un nuevo y gran escalón en el desarrollo tecnológico que debería ser considerado y favorecido en las políticas para asegurar un mayor bienestar para los ciudadanos.

En esta sección se hace principalmente referencia a la EC que no es propia de la organización, provista por centros instalados en sus locales, ya que ésta se ve impracticable para soportar las tecnologías que se analizan.

³ Cuando se habla de tecnologías en este documento, en realidad se está haciendo referencia a constelaciones de tecnologías individuales que surgen de su clasificación en tipos, los cuales tienen, cada uno, un objetivo general que requiere de múltiples tecnologías individuales en conjunto.

Como ejemplo, las PyMEs pueden tener Centros de Datos, en general pequeños, los que son inadecuados para soportar la digitalización ya que puede implicar aplicaciones que requieren mayor capacidad de procesamiento, velocidad de respuesta, transporte de información de banda ancha, costos optimizados por las economías de escala en la disponibilidad de los recursos y flexibilidad en la puesta en servicio y cierre de estos recursos, así como cualquiera de las tecnologías emergentes. En esa situación, estas empresas suelen recurrir a Centros de Datos en el borde y exteriores a la organización, como son los centros públicos. Por ejemplo, una PyME manufacturera que incorpora sensores y dispositivos de IoT en las líneas de producción para controlar su calidad, y detectar errores y ajustar en tiempo real los parámetros del proceso, o una PyME comercial que incorpora inteligencia artificial para procesar información de sus ventas, clientes y tendencias para hacer ofertas alternativas en tiempo real, optimizar los precios, programar futuros abastecimientos, o usar chatbots para dar respuestas de alta calidad y en forma automática a menores costos, o que realiza ventas en línea que requieren proveer rápidamente información de alta calidad a sus clientes para la toma de la decisión de compras, o realizar ofertas de acuerdo a su perfil o a reseñas.

La 5G⁴ puede ser la principal de las tecnologías emergentes, aunque ya en proceso avanzado de prestaciones, por ser el sustento de las comunicaciones que habilitan la movilidad (robots, automatizaciones, otros) con gran capacidad de transmisión y bajo retardo en el acceso, adecuados al tiempo real. A su vez la Internet de las Cosas de las aplicaciones críticas hace uso simultáneo de las comunicaciones 5G y de la EC centralizada o distribuida. Más allá de esto, la 5G permite la interacción con un número prácticamente ilimitado de dispositivos bajo una misma radiobase, permitiéndoles así un acceso casi local a los Centros de Datos de EC en forma muy flexible para la incorporación o retiro de dispositivos, y también la compartición del transporte por parte de múltiples aplicaciones que se emplean en diferentes sitios distribuidos geográficamente, como las cadenas de tiendas. De esta forma provee un alto y eficiente poder de aprovechamiento de las redes de Centros de Datos en el borde, independientemente de la ubicación o reubicación de los dispositivos. La red 5G es en definitiva lo que otorga transparencia al sustrato de las comunicaciones entre todos los dispositivos, equipamientos, Centros de Datos privados, públicos del operador o de terceros, a velocidades y retardos tales que el grado de proximidad se hace indistinguible con relación a la distancia física. Todo, desde el punto de vista práctico, está junto. Las comunicaciones con la nube, a más “distancia” temporal, se reservan, eficientemente, para procesamientos o almacenamientos necesarios y que no son tan críticos con la distancia.

La Internet de las Cosas (IoT) es también una constelación de tecnologías que convergen bajo el paraguas de cosas conectadas en general⁵ a través de la Internet, sin procesamiento humano, y que mediante la aplicación de múltiples tecnologías (Inteligencia Artificial, su subgrupo de Machine Learning, dispositivos ciber físicos como los robots, etc.) generan un potencial muy importante para el desarrollo económico y social. La llamada Industria 4.0, la logística avanzada también llamada a veces 4.0, la salud avanzada, y otros verticales se ven directamente impactados por la IoT. La EC es esencial para el desarrollo de la IoT en sistemas grandes o en aplicaciones críticas debido a la cantidad de información a procesar (en los grandes), o principalmente la cantidad de información y los bajos retardos requeridos en las críticas. Para la IoT existen entonces razones técnicas, pero también de eficiencia, seguridad y confidencialidad en el procesamiento y transporte de la información recogida por los dispositivos. La EC provee también, como se vio, una reducción drástica de costos de transporte (EC vs CC).

Los Gemelos Digitales son un importante avance de la digitalización en cuanto a representar virtualmente partes de activos, así como sistemas y procesos del mundo real. Se diferencian de las simulaciones, a pesar de que ambos son modelos de la realidad simulados digitalmente, en cuanto a que éstas se emplean principalmente para el diseño y la optimización del desarrollo en forma previa a

⁴ Hoyes la 5G. En 2024 será emitido el Release 18 del 3GPP con las especificaciones de 5GAdvanced, como camino intermedio hacia la 6G.

⁵ En algunos sistemas propietarios, que no operan sobre las redes normalizadas por la UIT y el 3GPP, puede no usarse la Internet.

construir el objeto real. Las simulaciones no requieren necesariamente de la EC debido a que no son sensibles al retardo, entre otras cuestiones. Por otro lado, los Gemelos Digitales y la EC van indefectiblemente juntos debido a que están ligados con la realidad, en tiempo real. Para ello, el Gemelo y la entidad real intercambian información mediante sensores inteligentes en la realidad, y órdenes o información del modelo digital hacia la realidad y viceversa. El Gemelo Digital puede actuar desde las pruebas iniciales y durante toda la vida útil del activo. Los hiperescaladores se encuentran avanzando en esta tecnología, como puede ser Amazon Web Service (AWS) con su producto AWS IoT TwinMaker para la creación de Gemelos Digitales a partir de sistemas en el mundo real⁶, Azure Digital Twins⁷ con su plataforma de Internet de las cosas (IoT) para crear la representación digital de cosas, lugares, procesos empresariales y personas reales, o Google Cloud con Supply Chain Twin⁸ para la representación digital de una cadena de suministros.

La Inteligencia Artificial ha comenzado a ser usada en el procesamiento en el borde reduciendo la carga de cómputo y ampliando la capacidad de las aplicaciones. Sus orígenes se remontan a apenas 10 años atrás cuando en lo que es un mojón del Aprendizaje Profundo, Google construyó una red neuronal de 16.000 procesadores y un billón de conexiones para reconocer imágenes. La mayor atención surge a fines de 2022 con el lanzamiento de ChatGPT de OpenAI⁹, un chatbot que puede mantener conversaciones muy parecidas a las humanas y que puede “generar” respuestas convincentes a preguntas técnicas complejas. A partir de ese momento han surgido a la luz miles de aplicaciones generativas de texto e imágenes para usos generales o específicos. Al mismo tiempo su uso para crear, sustituir o mejorar la eficiencia de ciertas tareas de las organizaciones se está acelerando, conduciendo a un aumento en los requerimientos de procesamiento a partir de Grandes Datos, a veces con retardos mínimos, para lo que la EC es la tecnología casi única para soportarlos. Basta considerar el ejemplo de un sistema de seguridad en el trabajo, en que se analizan imágenes de los trabajadores en actividades peligrosas, y que se implante un sistema de análisis de imágenes para tomar decisiones automáticas tempranas de detención de procesos ante un riesgo de accidente. Machine Learning, como un subgrupo de la Inteligencia Artificial, destinado a disponer de algoritmos y modelos estadísticos que hacen que la máquina pueda aprender y mejorar sus resultados a partir de datos, tiene los mismos requerimientos que pueden ser satisfechos por la EC. La analítica de Grandes Datos requiere una potencia de procesamiento similar a las anteriores familias tecnológicas, y por tanto también puede requerir la EC.

Recientemente hay un empuje de la Edge AI, o Inteligencia Artificial en el borde. Esta tendencia ha reactivado el mercado de chips con la elaboración de chips específicos¹⁰ para Inteligencia Artificial, llamados AI Chips, IPU o AIU, una carrera en la se encuentran Google, Amazon, Meta, IBM y principalmente Nvidia, que no solamente tiene aproximadamente el 70% de la venta de estos chips, sino también es líder en modelos de IA generativa entrenable. Estos chips surgen debido a que la rápida evolución de la Inteligencia Artificial y el Aprendizaje Profundo han dejado atrás los hardware existentes, incluyendo los de procesamiento de gráficos (GPU). Sobre esta base tecnológica se desarrolla actualmente la Edge AI, que es la incorporación de la Inteligencia Artificial en el ambiente de la EC, permitiendo una mayor eficiencia y eficacia en su procesamiento que usando centros en la nube. La Edge AI permite la toma de decisiones en procesos complejos casi sin retardo, en los sitios de EC próximos, o más directamente en los dispositivos como pueden ser los robots, las cámaras de control de calidad o los vehículos autónomos. De esta forma, agregando esta capa de fuerte inteligencia en el borde, los dispositivos están en condiciones, no solamente de levantar y enviar información, sino también de tomar acciones sobre ella debido a su modelo de Machine Learning integrado a su estructura

⁶ <https://aws.amazon.com/es/iot-twinmaker/>.

⁷ <https://azure.microsoft.com/es-es/products/digital-twins>.

⁸ <https://cloud.google.com/solutions/supply-chain-twin?hl=es-419>.

⁹ Generative Pretrained Transformer chat de www.openai.com.

¹⁰ Estos chips tienen una muy alta densidad de transistores. Por ejemplo, el recientemente presentado por IBM, llamado AIU (Artificial Intelligence Unit), tiene 23.000 millones de transistores de 5nm.

de procesamiento. En definitiva, con esta importante mejora de la capacidad de procesamiento para Inteligencia Artificial a través de chips específicos, los dispositivos que los contienen y las comunicaciones de alta velocidad (fibra o 5G) se abre el camino para una expansión a más sectores y regiones como resultado del desarrollo de la Edge AI.

El uso de contenedores¹¹ permite el despliegue estándar de las aplicaciones en diferentes plataformas de hardware independientemente de sus configuraciones. La contenedorización es un proceso de virtualización en que todos los componentes requeridos para correr una aplicación (código de la aplicación con sus bibliotecas, los archivos de configuración y el entorno de ejecución) son incluidos en un paquete de software estándar. De esta forma permite la ejecución en un sistema operativo compartido con otras aplicaciones y con independencia de las diferencias que puedan existir y de la infraestructura subyacente.

Si bien no es una tecnología reciente, es de extrema importancia cuando se multiplican los sitios de residencia de las aplicaciones.

Hay dos tecnologías de códigos abiertos que permiten el uso de los contenedores y que se destacan por la amplia difusión en los proveedores de servicios en la nube, constituyendo estándares de facto: Docker y Kubernetes. Docker es una tecnología que permite al desarrollador empaquetar software en los llamados contenedores con todo lo que el software necesita para correr en cualquier sistema operativo. El contenido del empaquetado lógico de Docker es suficiente para crear el entorno de ejecución y configurar el servidor donde va a ser instalado. A partir del contenedor se genera una imagen que puede ser instalada en segundos en un servidor. Por otra parte, Kubernetes es una herramienta del plano de control, de orquestación de contenedores que permite administrar, coordinar y programar contenedores en gran escala.

Otra ventaja importante, aparte de la simplificación de los procesos de instalación y actualización del software en múltiples servidores, es el importante ahorro de recursos de procesamiento. La alternativa al uso de contenedores es el empleo de máquinas virtuales, para las cuales se requieren crear emulaciones de sistemas operativos con gran consumo de recursos, ya que como se suelen correr varias aplicaciones en cada máquina virtual se dificulta el escalado horizontal de las aplicaciones.

Este es un campo de negocio para todos los hiperescaladores que proveen múltiples servicios para contenedores, como Amazon que tiene tanto su propio servicio para correr aplicaciones contenedorizadas y que automatiza la disponibilización de recursos, como de gestión de contenedores con Kubernetes.

Las mallas de servicios y los datos en el borde, a su vez permiten desplegar y recuperar datos y servicios distribuidos a través de contenedores y bases de datos en el borde. Estas mallas presentan una única interfaz que abstrae el enrutamiento y la gestión de servicios e interfaces de datos. Este habilitador crítico hace posible las consultas masivas para poblaciones enteras dentro del borde, en lugar de en cada dispositivo.

Aparte de éstas, que pueden ser las más destacadas, otras tecnologías pertenecen a este ecosistema de la EC.

La EC no solamente responde a los requerimientos físicos de ciertas aplicaciones y a la reducción de los costos asociados, sino que provee beneficios adicionales que también impulsan su desarrollo.

Existen situaciones de baja calidad y continuidad de las redes de comunicaciones en sitios especiales aislados o sometidos a condiciones ambientales complejas. En estos casos el procesamiento se realiza localmente, procurando el envío de la menor cantidad de información posible, y solamente aquella que no requiere integridad, cuando la red está disponible.

¹¹ <https://azure.microsoft.com/es-mx/resources/cloud-computing-dictionary/what-is-a-container>.

A un nivel territorial mayor, la EC evita el envío de los datos “crudos” a Centros de Datos (Cloud Computing) a veces situados fuera del dominio establecido legalmente. Cuando la información viaja a esos grandes centros, no es posible conocer su ubicación a priori por la existencia de sitios alternativos en caso de fallas. Un sitio doméstico puede tener su respaldo en cualquier parte del mundo. De esta manera los datos “crudos” se procesan localmente en la EC y solamente viajan datos procesados y anonimizados, y asegurando el cumplimiento de las regulaciones de protección de datos en cuanto a territorialidad, seguridad y confidencialidad. Las Zonas de los hiperescaladores presentan la independencia de otras Zonas como una de las ventajas para asegurar también que no sea violada dicha territorialidad de los datos, lo que se desarrolla más adelante.

4. Resumen de los beneficios que impulsan la EC

En síntesis, los principales beneficios que impulsan la EC junto a la 5G serían los siguientes:

- Por ser una estructura descentralizada y distribuida en múltiples sitios según la conveniencia para las organizaciones, proveyendo en cada uno alta capacidad de procesamiento de datos con muy bajo retardo, típicamente de menos de 5 ms., permite aplicaciones que no pueden ser provistas por los tradicionales grandes centros en la nube.
- Reduce notoriamente los costos del transporte de datos hasta los grandes centros en las aplicaciones que requieren procesar Grandes Datos, como ser la emergente Inteligencia Artificial, empleada en una forma rápidamente creciente. Con la EC solamente se transmiten datos preprocesados y en consecuencia en un volumen significativamente menor.
- Cuando se emplea la EC en nubes públicas, en lugar de en centros residentes, se mantienen los beneficios de los grandes Centros de Datos en cuanto a reducción de costos por compartición de la infraestructura de hardware y software, y las facilidades para el escalamiento o desescalamiento.
- Debido a la proximidad y los menores despliegues de infraestructura de transmisión, son redes con mayor resiliencia.
- Mejoras en la confidencialidad, ya que las organizaciones que manejan datos personales sensibles como en el sector financiero o en el de la salud, podrán garantizar más seguridad procesando datos en el origen o cerca del origen.
- La EC junto a la 5G son esenciales para el desarrollo de la Internet de las Cosas. Para el caso de la IoT masiva (mIoT) la 5G provee la posibilidad de la conectividad en el acceso simultáneo de cantidades casi ilimitadas de dispositivos, y la información generada puede ser procesada con mínimo retardo. Si se trata de las aplicaciones críticas de la IoT (cIoT) como la automatización industrial, las aplicaciones médicas, la movilidad asistida, entre otras, la 5G provee la ubicuidad del acceso con retardos menores de 5 ms permitiendo a la EC realizar el procesamiento de la información en tiempo real.
- Permite, junto a la 5G, la introducción de tecnologías avanzadas como las de reconocimiento de imágenes tanto para seguridad personal como para la seguridad física. Es posible prever predictivamente situaciones de peligro para el personal, con el uso combinado del reconocimiento de imágenes y la Inteligencia Artificial, lo que solo se logra con gran capacidad de procesamiento y almacenamiento y retardos muy bajos en el acceso.
- Asegura el cumplimiento de las regulaciones nacionales en cuanto a la territorialidad del procesamiento y almacenamiento de datos.

- Permite una mayor observabilidad total en tiempo real de lo que sucede en las aplicaciones y la conectividad, permitiendo a los profesionales de DevOps¹² tomar mejores decisiones para beneficio de los usuarios.
- El carácter descentralizado da lugar también a su inserción en un entorno multinube, y en general híbrido, que genera más dificultades en la gestión de los recursos que en los grandes Centros de Datos. Por ello surge el requerimiento de adoptar una disciplina y práctica de cultura de gestión financiera de la nube denominada FinOps¹³.

Así, el diseño de la arquitectura a emplear en EC debe tomar en consideración todos los beneficios y las dificultades de su implementación, como calidad de las redes, condiciones de seguridad y confidencialidad, disponibilidad de sitios de EC de terceros, capacidades propias de gestión, entre muchas otras.

B. Principales características de las acciones del ETSI con relación a la MEC¹⁴

El Instituto Europeo de Normas de Telecomunicaciones (ETSI¹⁵) es un foro sin fines de lucro “adecuado para la elaboración de las normas de telecomunicación que faciliten la estandarización del sector, y por lo tanto el avance hacia el Mercado Único Europeo. En el ETSI participan como miembros no solo las Administraciones, sino también los operadores de red, la industria, los centros de investigación y los usuarios de los servicios de telecomunicación. Los objetivos del ETSI se centran en la elaboración y el mantenimiento y actualización de normas técnicas a nivel europeo en los campos de las telecomunicaciones, tecnologías de la información y sistemas de radiodifusión y televisión”.

Como ya se ha visto la EC es principalmente un concepto que engloba una familia de tecnologías que permite llevar el almacenamiento y la capacidad de procesamiento al borde de las redes, cerca de los usuarios. Por otro lado, la Computación en el Borde Multi Acceso¹⁶, o MEC, es una arquitectura de EC normalizada que ayuda al mejoramiento de las funcionalidades de las redes de datos, como la 5G o las redes fijas de alta velocidad en general, habilitando aplicaciones y servicios en una forma eficiente y confiable.

En conjunto, la normalización resultaba imprescindible para asegurar el perfecto funcionamiento de las aplicaciones y servicios de las TI que necesitaban ser trasladadas al borde, así como la rapidez, eficiencia y seguridad en el despliegue, lo que a su vez fortalece toda la cadena de valor de la TI y los CSP. Estos últimos se benefician al asegurar que puedan usar sus propias redes para desarrollar la computación en el borde en el contexto de la red 5G, considerando su camino emprendido en la virtualización de las funciones de red (NFV) en la nube, lo que genera el ambiente para el alojamiento de servicios de TI en el borde. Por otra parte, el complejo ecosistema del borde está constituido por proveedores de servicios en la nube, de equipamiento de nodos de datos en gabinetes, o de plataformas en las cuales montar los servicios requeridos, entre múltiples modalidades de soportar la EC, lo que adicionalmente obliga a disponer de una normalización.

¹² Combinación de Desarrollo y Operaciones. Según Microsoft “DevOps permite que los roles que antes estaban aislados (desarrollo, operaciones de TI, ingeniería de la calidad y seguridad) se coordinen y colaboren para producir productos mejores y más confiables. Al adoptar una cultura de DevOps junto con prácticas y herramientas de DevOps, los equipos adquieren la capacidad de responder mejor a las necesidades de los clientes, aumentar la confianza en las aplicaciones que crean y alcanzar los objetivos empresariales en menos tiempo”.

¹³ FinOps es un término que se refiere a agregar la visión Financiera a las prácticas de DevOps a través de un equipo o una persona que se ocupa de asesorar a los profesionales de DevOps para reducir los costos optimizando los recursos de la nube, que permitan la velocidad, la capacidad y demás requerimientos de las funciones y aplicaciones manteniendo la calidad de los resultados. La FinOps Foundation (<https://www.finops.org/>) es una fundación que tiene como misión profundizar la disciplina de la gestión financiera de la nube.

¹⁴ MEC: Multi-Access Edge Computing y que también solía llamarse al principio Mobile Edge Computing. Es una evolución de la EC que usa la movilidad, así como las redes fijas o convergentes, la nube y la computación en el borde.

¹⁵ www.etsi.org.

¹⁶ <https://www.etsi.org/technologies/multi-access-edge-computing>.

Es, en definitiva, una evolución de la computación en la nube que usa la movilidad y/o las redes fijas, los servicios en la nube y la computación en el borde, como caso de uso de los proveedores de servicios.

Cuando los proveedores de servicio comienzan la virtualización de su infraestructura, alejándola del equipamiento físico y llevándola a la nube, se produce un desacoplamiento generando un ecosistema más amplio que permite ejecutar simultáneamente las cargas de trabajo de movilidad con las cargas de trabajo que sean producidas por los usuarios.

De esta forma el ecosistema que genera la MEC, proveyendo latencia ultra baja y gran capacidad de transmisión de datos, así como interacciones sobre la red, puede ser usado para permitir un acceso nativo de aplicaciones en tiempo real.

Por lo anterior la MEC en entornos de virtualización del operador permite que la RAN sea de esa forma accesible para desarrolladores de aplicaciones autorizados y de proveedores de contenido, poniendo la EC tanto a disposición del nivel de aplicación, como del nivel inferior de funciones de red y procesamiento de información¹⁷.

Todo indica que próximamente la EC se mueva exactamente a la frontera de las redes a través múltiples modalidades, muchas de ellas interconectadas y trabajando simultáneamente en determinadas aplicaciones y servicios.

Frente a esta evolución el ETSI ha desarrollado la normalización técnica y de arquitectura para la MEC.

Su foco se encuentra en la MEC en cuanto a que genera un ecosistema residente en el borde de las redes inalámbricas de los operadores móviles, apoyada en un mecanismo de la apertura de sus Redes de Acceso (RAN) a terceras partes autorizadas para que puedan desplegar flexible y rápidamente servicios y aplicaciones a los clientes móviles, las empresas y a los segmentos verticales. Es un desarrollo natural sobre las estaciones radiobase en la convergencia de las telecomunicaciones y las Tecnologías de la Información, aliviando la carga de transmisión en el núcleo de la red y atendiendo directamente los requerimientos en el borde. El ETSI entiende también que la MEC es analizada incluyendo los accesos fijos físicos e inalámbricos.

La estandarización industrial de la MEC y el despliegue de sus plataformas habilitarán nuevos flujos de valor para los proveedores, los operadores y terceras partes que harán uso de ella.

El ETSI indica que el rol y actividades¹⁸ de su Grupo MEC de Especificación de la Industria (ISG) tiene como propósito estandarizar un entorno abierto que permita la integración eficiente y sin problemas de aplicaciones de proveedores de equipamiento y de servicios, y de terceros, a través de plataformas de computación en el borde para el acceso múltiple de múltiples proveedores.

Esta iniciativa Multi-access Edge Computing tiene el objetivo preciso de facilitar la integración y el fortalecimiento de la cadena de valor que incluye operadores móviles, desarrolladores de aplicaciones, iniciativas Over the Top (OTT), proveedores independientes de software (ISV), proveedores de equipos de telecomunicaciones, proveedores de plataformas de TI, integradores de sistemas y proveedores de tecnología.

En resumen, el trabajo de la iniciativa MEC tiene como objetivo unir los mundos de las telecomunicaciones y la nube de TI, proporcionando capacidades de TI y computación en la nube dentro de la RAN (Radio Access Network).

¹⁷ What is multi-access edge computing (MEC)?, RedHat. 22 de julio de 2022. [https://www.redhat.com/en/topics/edge-computing/what-is-multi-access-edgecomputing#:~:text=Multi%2Daccess%20edge%20computing%20\(MEC\)%20is%20a%20type%20of,an%20improve%20the%20customer%20experience.](https://www.redhat.com/en/topics/edge-computing/what-is-multi-access-edgecomputing#:~:text=Multi%2Daccess%20edge%20computing%20(MEC)%20is%20a%20type%20of,an%20improve%20the%20customer%20experience.)

¹⁸ <https://www.etsi.org/technologies/multi-access-edge-computing.>

El modelo de trabajo de este Grupo de Especificación incluye el desarrollo de la normativa y la realización de informes.

El ISG de MEC también creó en 2019 el Grupo de Trabajo para el Desarrollo del Ecosistema y el Despliegue de MEC, denominado WG DECODE¹⁹.

Las actividades del WG DECODE tienen su foco en acelerar el desarrollo del ecosistema en el borde y el inicio de las ofertas a los consumidores a través de las siguientes actividades:

- i) Facilitar la ruta de implementación para proveedores, operadores y desarrolladores de aplicaciones al proporcionar implementación de API de software;
- ii) Desarrollar un marco de pruebas y cumplimiento, y un entorno de sandbox para ser utilizado en el desarrollo de aplicaciones.

Todas estas facilidades están disponibles a través de ETSI FORGE²⁰ y, para el caso del sandbox MEC, empleando un portal dedicado²¹.

El TM Forum²² visualiza que la iniciativa MEC buscará garantizar la Calidad de Servicio sobre redes 5G según muchos operadores y proveedores, aprovechando los avances en virtualización de funciones de red (NFV) y segmentación de red. Múltiples organizaciones de estandarización y grupos de código abierto están colaborando con el ETSI para que ésto suceda.

Por ejemplo, ETSI está trabajando con el 3GPP, el NGMN²³, la Open Networking Foundation, la Open Network Automaton Platform y el propio TM Forum para resolver problemas relacionados con el control de la redirección del tráfico para aplicaciones MEC, la disponibilidad de recursos de red y la orquestación de segmentos de extremo a extremo. Adicionalmente los CSP y sus proveedores están trabajando juntos en varios proyectos, iniciativas, foros y grupos de trabajo para allanar el camino hacia el borde.

El 3GPP se encuentra trabajando en EC y coordinando con el ETSI desde la Release 15, congelada en 2019²⁴. El ETSI define las especificaciones relevantes a la EC, pero agnósticas a la tecnología de acceso, que puede ser fija o móvil. Mientras tanto el 3GPP trabaja en el soporte para las funcionalidades de la EC que deben ser soportadas por las redes 5G, por ejemplo, a través del mejoramiento del núcleo de red 5G en la Release 17 (2022).

En cuanto a casos de uso, se encuentra una gran cantidad de CSP en el mundo que están ofreciendo MEC, en general en combinación con 5G. Algunos casos de estos avances se observan brevemente en la sección "IV. Tendencias en el despliegue previsto de la EC en LAC, y en otras regiones." y en la sección "II. B. Los CSP como usuarios de la MEC y proveedores de servicios sobre ella".

A modo referencial se presenta un caso de uso provisto por Verizon en EE. UU.²⁵, entre múltiples casos a ver más adelante en la sección "IV. Tendencias en el despliegue previsto de la EC en LAC, y en otras regiones.", en que para el área manufacturera ofrece una solución para la eficiencia y la seguridad laboral que resume de la siguiente manera a partir de su servicio 5G Ultra Wideband²⁶ y MEC:

¹⁹ Deployment and Ecosystem Development working group.

²⁰ <https://forge.etsi.org/>.

²¹ <https://try-mec.etsi.org/>.

²² TM Forum. Noviembre de 2020. "How to Build and Operate at the Edge".

²³ NGMN Alliance (Next Generation Mobile Networks Alliance).

²⁴ <https://www.3gpp.org/technologies/edge-computing>.

²⁵ <https://www.verizon.com/business/en-gb/solutions/5g/edge-computing/multi-access-edge-computing/>.

²⁶ 5G UWB es el servicio de 5G usando bandas milimétricas combinadas con bandas medias (banda C) que permiten velocidades del orden del orden de 2 Gbps.

- El servicio ofrecido puede soportar Realidad Aumentada e Inteligencia Artificial para entrenamiento y monitoreo peligrosos a través de las siguientes prestaciones.
- Permite la supervisión continua y de baja latencia de los procesos y alerta si la actividad de la máquina representa un riesgo o si las condiciones ambientales circundantes son peligrosas. Esta prestación se ve fortalecida a través del empleo de Inteligencia Artificial concentrando el procesamiento en el borde o en los propios terminales.
- Aprovecha la Realidad Aumentada para capacitar a los operadores, lo que permite un rendimiento eficiente y optimiza la seguridad. La Realidad Aumentada es usada también para facilitar las operaciones de mantenimiento poniendo recursos a disposición del técnico superpuestos a la imagen de la realidad.
- Aprovecha los datos de los sensores y cámaras de Internet de las cosas (IoT) para ayudar a predecir y administrar el rendimiento y mejorar el uso de la máquina.
- Se combina con casos de uso de gemelos digitales (DT) y mantenimiento predictivo (PdM) para una solución de eficiencia total del sistema.
- Se integra con otras tecnologías, incluidas cámaras, sensores biométricos y sensores ambientales, para una solución de sistema completa.

C. Evolución de la EC y de los proveedores de las tecnologías directamente relacionadas

El mercado de la EC ha adquirido un momento importante y todo indica que el crecimiento en las inversiones y gastos en ella se profundizarán considerando el empuje de la digitalización y de las tecnologías que requieren sus prestaciones. Muchas industrias o sectores visualizan las oportunidades en este mercado extendido: hiperescaladores, proveedores de nube de menor escala, CSP, proveedores de infraestructura, equipamiento y software, proveedores de sistemas de gestión, de seguridad y de aplicaciones, fabricantes de chips, entre otros. La EC es actualmente indispensable para la digitalización.

Un informe de Deloitte²⁷ de fines de 2022 analiza el comportamiento de estos competidores. Deloitte Global predice²⁸ un crecimiento del mercado empresarial para 2023 del 22% para EC, 4% en equipamiento de redes y 6% en TI en general. Estiman que irá destinado inicialmente a gastos en hardware, migrando luego a software y servicios. De todas maneras, estos crecimientos son a partir de un stock muy bajo en este momento, como sucede en todas las tecnologías emergentes.

En la base de las industrias mencionadas, se encuentran las siguientes que se van a analizar: nubes públicas de los hiperescaladores, CSP, proveedores de infraestructura de computación en el borde y proveedores de plataformas de gestión.

Los hiperescaladores están trasladando al borde las ventajas de sus economías de escala y su oferta diversificada y flexible en cuanto a servicios y modalidades de contratación. Están extendiendo su red de Centros de Datos acercando los mismos a los usuarios, escalándolos y adaptándolos a los servicios requeridos para la EC. El criterio aplicado es que todos aquellos servicios que se pueden prestar desde la nube se sigan prestando, instalando en los Centros de Datos en el borde solamente los que requieren la reducción de retardo o cargas de trabajo masivas. Resulta así que el tamaño menor de los centros en el borde se adapta mejor a su ubicación, o en las propias organizaciones o en socios del negocio. En muchos

²⁷ Battle for the Enterprise Edge: Providers prepare to pounce on the emerging enterprise edge computing market <https://www.deloitte.com/global/en/our-thinking/insights/industry/technology/technology-media-and-telecom-predictions/2023/edge-computing-battle-for-the-enterprise-edge.html>.

²⁸ A partir de resultados de IDC, Gartner, Omdia, TBR, HPE, AvidThink, Precedence Research, Grandview y STL Forecasts.

casos este acercamiento se realiza a través de acuerdos con los CSP, los centros de los distribuidores de contenido, los propietarios de torres de comunicaciones, entre otros. También están asociándose, desde el punto de vista comercial, con empresas integradoras o similares que les permitan llegar a los diferentes verticales para la venta de equipos completos para su alojamiento dentro de las organizaciones.

Los CSP están preparados para la prestación de servicios empaquetados de EC y conectividad segura y resiliente fija y móvil. Adicionalmente prestan sus servicios tradicionales de conectividad entre los Centros de Datos de los hiperescaladores en el borde hasta los centros residentes en las organizaciones, sus terminales o los dispositivos, como pueden ser los de la IoT. La móvil es indispensable en muchos casos de aplicaciones críticas en dispositivos móviles o que sean más seguras usando accesos inalámbricos como es el caso de ambientes industriales o de logística. En su evolución hacia la 5G que es desplegada sobre una infraestructura virtual crean el ambiente para agregar la prestación de servicios de EC, sumado al bajo retardo y alta capacidad de acceso de la 5G. La tendencia es a desarrollar la infraestructura, las plataformas y los servicios, con cierta preponderancia de los hiperescaladores, para prestar la llamada Mobile Edge Computing o MEC, la que se analiza en la sección "I.B. Principales características de las acciones del ETSI con relación a la MEC." Adicionalmente pueden desarrollar sus propias aplicaciones para los segmentos B2B y B2C. En esta decisión de negocio es importante la relación sólida y de confianza, preestablecida de los CSP con sus clientes empresariales u organizaciones en general.

Por su parte, los proveedores de infraestructura de computación en el borde encuentran una nueva oportunidad de negocio con los CSP considerando cómo se ha ido desarrollando el mercado a través de las interfaces abiertas y virtualizadas. El CEO de uno de los principales proveedores indicaba que hay siete millones de radiobases en el mundo, y que cada una de ellas puede convertirse en un Centro de Datos en el borde. Se observa que múltiples proveedores tanto de equipamiento de red de CSP como de IT se encuentran trabajando para ser proveedores de infraestructura para la EC, tanto para los CSP como para las organizaciones. Para los casos de redes privadas de 5G junto con EC se dan asociaciones de proveedores con especialización en cada área.

Plataformas de gestión. La EC se sustenta en general en múltiples infraestructuras como ser centros privados, nubes públicas de borde o de gran escala, nubes privadas, dispositivos, entre otros, muchas veces pertenecientes a diferentes proveedores, por lo que se requiere un escalón de abstracción y disponer de una plataforma de gestión de la operación de la EC que sea agnóstica al sustrato, y que incluya herramientas de gestión administrativa de la EC. Estas plataformas programables usan APIs para el control de la red física subyacente. Existe además una tendencia a que las API se vuelvan estándar permitiendo mayor apertura en relación con los suministradores.

Las empresas que pertenecen a estas industrias o sectores identificados como los más importantes para soportar el vertiginoso desarrollo de la EC, constituyen un universo muy importante y en plena evolución en cantidad y en profundidad de su inserción en ese desarrollo. Un documento de STL Partners contiene los detalles de 100 empresas de EC a observar en 2023²⁹ ³⁰, incluyendo startups, en crecimiento y consolidadas. Se hace referencia a algunas de ellas para observar la variedad de Infraestructura de EC. Agregación y orquestación de EC, IaaS, sistema operativo, plataformas multi nube, virtualización y contenedores. Aplicaciones y Software. Analítica, APIs, software de aplicaciones, PaaS, productos y servicios ofrecidos y las importantes colaboraciones entre ellas para satisfacer la demanda de EC de las organizaciones.

²⁹ <https://stlpartners.com/articles/edge-computing/edge-computing-companies-2023/>.

³⁰ En las descripciones de los segmentos de actividad se consideran los alcances que puede tener cada empresa. Por más detalles, o para observar más empresas, se ruega ir al sitio de STL Partners indicado en la nota anterior.
Instalaciones. Bienes raíces, Centros de Datos, sistemas de racks, energía y acondicionamiento, mantenimiento y operación de sitios.
Hardware. Servidores, ruteadores y conmutadores, mantenimiento y operación de hardware, procesadores, dispositivos terminales.
Redes. Interconexión, plataformas IoT, pasarelas de borde, conexiones a la nube, distribución de aplicaciones y contenido.

- A5G Networks, Inc. es una *startup* que actúa en redes, aplicaciones y software e integración y servicios y que en 2022 se asoció con Vodafone Idea (proveedor líder de telecomunicaciones en la India) para habilitar la industria 4.0 y la computación de borde móvil (MEC). Han establecido una red privada piloto en Mumbai utilizando el espectro 4G existente.
- Accenture. Siendo una empresa de consultoría estratégica, servicios tecnológicos y externalización, está desarrollando intensa actividad con relación a la EC en los siguientes campos: Instalaciones, Redes, Infraestructura de EC, Aplicaciones y Software e Integración y Servicios. Por ejemplo, Accenture está colaborando con Mars, la compañía líder en confitería, alimentos y cuidado de mascotas, para modernizar sus operaciones globales de fabricación utilizando IA, nube, tecnología de borde y gemelos digitales. Las dos compañías han estado probando el uso de gemelos digitales para optimizar los procesos de producción y mejorar el rendimiento, la fiabilidad, la calidad y la eficiencia energética de los equipos.
- Adtran - ADVA. Opera en Hardware, Redes, Infraestructura de EC, Aplicaciones y Software e Integración y Servicios. Indica que en 2022 lanzó su nuevo producto Ensemble Cloudlet para la nube local. Esta solución abierta está optimizada para el alojamiento económico de cargas de trabajo virtualizadas y en contenedores y la implementación automatizada. La solución proporciona una ruta escalable para que los CSP, MSP o SI³¹ ofrezcan alojamiento de cargas de trabajo del cliente combinado con servicios de red administrados.
- Amazon Web Services (AWS). Es uno de los hiperescaladores con intereses en los segmentos de Instalaciones, Hardware, Redes, Infraestructura de EC, Aplicaciones y Software e Integración y Servicios. Según STL Partners Amazon Web Services ha lanzado dos productos en 2022. Esto incluye Amazon Elastic Kubernetes Service (EKS) Anywhere Curated Packages, que son paquetes de software que amplían las funcionalidades principales de Kubernetes. Esto también incluye IoT FleetWise de AWS, que ayuda a las empresas automotrices a recopilar, transformar y transferir datos de vehículos a la nube casi en tiempo real. AWS IoT RoboRunner es un servicio de AWS for Robotics que desbloquea nuevos casos de uso para la automatización robótica al ayudar a las flotas de robots a trabajar juntas sin problemas. Además existe la disponibilidad de AWS IoT TwinMaker, un servicio que facilita a los desarrolladores la creación de gemelos digitales de sistemas del mundo real, como edificios, fábricas, líneas de producción y equipos. AWS también ha anunciado su trabajo con varios operadores de telecomunicaciones, incluidos Telenor, Spark New Zealand y Vodacom. AWS se asoció con Verizon para habilitar MEC pública para los clientes 5G. Se observa el foco directo de este hiperescalador en los gemelos digitales, la robótica y los vehículos autónomos.
- American Tower. Opera en Instalaciones, Redes e Infraestructura de EC. Esta empresa originalmente dedicada a la construcción y compra de infraestructura para compartición, también ha entrado al mercado muy amplio de la EC. Indica que en 2022 agregó a su red un total de seis Centros de Datos en el borde con la adquisición de CoreSite y DataSite. En el último año, American Tower también ha firmado acuerdos para agregar nuevas instalaciones de ubicación de centros de EC en los sitios de torres existentes.

³¹ Cloud Service Provider, Managed Cloud Provider y System Integrator.

- Atos. Opera en Instalaciones, Redes, Infraestructura de EC, Aplicaciones y Software e Integración y Servicios. En 2022 lanzó varios acuerdos con CSPs como British Telecom y Verizon.
- Azion. Trabaja en los segmentos de Redes, Infraestructura de EC, Aplicaciones y Software e Integración y Servicios. Azion permite a los clientes crear aplicaciones una vez y ejecutarlas en cualquier lugar: el borde de la red, dispositivos remotos, múltiples locales o multinube.
- Cisco. Trabaja en los segmentos de Redes, Infraestructura de EC, Aplicaciones y Software e Integración y Servicios. En 2022, Cisco ha desarrollado su plataforma de aplicaciones de borde 'Great Bear'.
- Rakuten Symphony. Trabaja en los segmentos de Instalaciones, Redes, Infraestructura de EC, Aplicaciones y Software e Integración y Servicios. Según STL Partners Rakuten Symphony con Symworld Cloud están trabajando con socios ISV³² para proporcionar aplicaciones de borde para casos de uso como control de calidad, análisis de video habilitado para IA, mantenimiento y operaciones de planta y seguridad y control de acceso como parte de la automatización de fábricas de alimentos. En cuanto a Ciudades Inteligentes, sus aplicaciones de borde han permitido la gestión del tráfico, la medición en rampa (control de acceso a autopistas) y la gestión de incidentes de control de señales y la programación inteligente de autobuses, taxis y transporte público para un gobierno local. En el transporte, se están enfocando en puertos con seguimiento y liberación de contenedores, logística automatizada y seguridad y control de acceso.
- Verizon. Este CSP opera además en el ecosistema de EC en Redes, Infraestructura de EC e Integración y Servicios. En 2022 Verizon lanzó Private MEC con AWS Outpost y con Microsoft Azure Stack Edge. Verizon lanzó también varias soluciones MEC nuevas: Analítica de Multitudes, Pagos sin cajeros, Acceso Acelerado (por reconocimiento facial)", entre otras. Verizon también amplió 5G con MEC empleando AWS Wavelength en 19 mercados metropolitanos de EE. UU.

Se observa a través de estos ejemplos un vibrante mercado de múltiples industrias alrededor de la EC y a través del despliegue de tecnologías avanzadas para las cuales es imprescindible.

D. Soberanía de los datos y la EC

Los sitios donde los datos son almacenados o procesados adquieren una importancia creciente en los países tanto desde los puntos de vista personal como principalmente empresarial o gubernamental. El concepto principal es el de la soberanía de los datos en cuanto a las limitaciones territoriales impuestas en múltiples países, las que pueden ser incumplidas inintencionalmente a través del uso cada vez más intenso de la nube para el almacenamiento y procesamiento de datos en uno o más países fuera del propio donde se generan o de países con los que existan acuerdos de protección de datos.

Más adelante se analizará la estructura de las redes de Centros de Datos donde reside el procesamiento y el almacenamiento, y se observa que, en principio, si no se adoptan políticas especiales, los propietarios de los datos pueden desconocer la ubicación física real de ellos.

La soberanía pasa a ser una cuestión esencial para ser tomada en consideración en el uso de la nube. Es una preocupación a nivel nacional, cuando la violación de la seguridad o la privacidad ocurre fuera del territorio de origen de los datos.

³² Independent Software Vendor.

La soberanía de los datos impacta en la EC y en los servicios en la nube. El aseguramiento del cumplimiento de la normativa de mantener los datos dentro de fronteras para su almacenamiento y procesamiento se puede garantizar si estas operaciones se realizan cerca de donde se generan los datos, o sea empleando la EC. A estos efectos la computación en la nube, en centros lejanos, no permite asegurar en general el mantenimiento de los datos dentro de fronteras, o de las fronteras de países con los que existen tratados. En algunos casos directamente se incumple si quien contrata el almacenamiento no considera el país donde serán alojados. Los Centros de Datos tradicionales, empleando recursos virtuales, pueden tener respaldos en otros países, o estar ellos mismos ubicados fuera del país.

En LAC se observa un creciente interés por la inversión o la contratación de servicios en el borde para cumplir con los avances en las exigencias de los usuarios, y al mismo tiempo para cumplir con las disposiciones nacionales de protección de datos.

Las siguientes son algunas de las muchas regulaciones en el mundo en cuanto a la soberanía de los datos por interés nacional o regional y donde se muestran los condicionamientos clave que se han mencionado.

En Europa existe una extensa evolución en materia de protección de datos personales³³. Por su trascendencia a nivel internacional se destaca el Convenio 108 del Consejo de Europa de 1981, para la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal, que se constituyó en el primer instrumento internacional jurídicamente vinculante adoptado en el ámbito de la protección de datos. Uruguay fue el primer país no europeo en suscribirlo en 2013.

Este Convenio, junto a su Protocolo Adicional sobre las autoridades supervisoras y flujos transfronterizos, fue actualizado en 2018 bajo la denominación de Convenio 108+³⁴ y los países, incluyendo los no pertenecientes a la Unión Europea, han sido invitados a suscribir el Protocolo y sólo los Estados que hayan ratificado, aprobado o aceptado el Protocolo estarán condicionados por las obligaciones derivadas del Convenio enmendado. Incluye principios básicos de la protección de datos y facilita el intercambio de datos entre las Partes a través de mecanismos de colaboración entre autoridades de protección de datos. En su Artículo 14 establece las condiciones para permitir el flujo transfronterizo de datos personales entre las Partes que lo suscriban. Establece que un Estado Parte no podrá, con el solo propósito de proteger los datos personales, prohibir o someter a autorización especial la transferencia de dichos datos a un destinatario que se encuentra sujeto a la jurisdicción de otra de las Parte del Convenio 108+. No obstante, dicha parte podrá hacerlo si existe un riesgo real y grave de que la transferencia a otra parte, o de esa otra parte a un tercero, pueda llevar a incumplir las disposiciones del Convenio 108+.

A julio de 2023, en LAC, el Tratado 223 (Protocolo que modifica el Convenio) ha sido ratificado por Uruguay (05/08/21) y por Argentina (17/04/23), entre tres países no europeos³⁵.

³³ <https://www.europarl.europa.eu/factsheets/es/sheet/157/la-proteccion-de-los-datospersonales#:~:text=El%20Convenio%20n.&text=108%20del%20Consejo%20de%20Europa,de%20la%20protecci%C3%B3n%20de%20datos>.

³⁴ <https://rm.coe.int/convention-108-convention-for-the-protection-of-individuals-with-regar/16808b36f1>.

³⁵ <https://www.coe.int/en/web/conventions/full-list?module=signatures-by-treaty&treatynum=223>.

El Reglamento General de la Protección de Datos (RGPD³⁶) de la Unión Europea fue aprobado el 27 de abril de 2016 y puesto en efecto en mayo de 2018, siendo uno de los instrumentos vigentes de la Unión en materia de protección de datos. Las reglas para las empresas y las organizaciones, así como la guía de su aplicación, se encuentran en el sitio³⁷ de la Unión Europea para la protección de datos. En cuanto a lo que interesa en este caso, referente al almacenamiento o procesamiento fuera de las fronteras, establece que, en el mundo globalizado actual, existen grandes cantidades de transferencias transfronterizas de datos personales, que a veces se conservan en servidores en varios países distintos. La protección que ofrece el Reglamento general de protección de datos (RGPD) viaja con los datos, lo cual significa que las normas que protegen los datos personales seguirán aplicándose independientemente de dónde vayan a parar los datos. Esto también se aplica cuando los datos se transfieren a un Estado no perteneciente a la Unión Europea (UE). Más detalles sobre la aplicación del RGPD para las transferencias de datos de la UE a Estados no pertenecientes a la UE y no signatarios del Convenio 108+, se pueden observar en el enlace a este Reglamento.

En LAC, Brasil ha aprobado la Ley General de Datos Personales³⁸ en 2018, modificada en 2019, la que prevé el tratamiento de datos personales, incluso en medios digitales, por una persona física o por una persona jurídica de Derecho público o privado, con el fin de proteger los derechos fundamentales de libertad y privacidad y el libre desarrollo de la personalidad de la persona física.

Se observa el amplio alcance previsto en esta Ley.

En cuanto a la Transferencia Internacional de Datos, el Artículo 33 del Capítulo V establece los únicos casos en que está permitida la transferencia internacional de datos personales, aparte de los casos de cooperación jurídica internacional, de autorización expresa de la autoridad.

Los artículos 34, 35 y 36 reglamentan los numerales I y II del artículo 33.

En general se observa un tratamiento similar al de la Unión Europea y en particular el concepto de que la regulación local debe viajar con los datos, es decir que tengan el mismo grado de protección que domésticamente.

E. Sostenibilidad de la computación en la nube y en el borde

En esta sección se realiza un análisis breve de los impactos de las TIC, incluyendo los Centros de Datos, sobre la contención del calentamiento global.

Las acciones más importantes sobre el cambio climático surgen de la Convención Marco de Naciones Unidas sobre el Cambio Climático (CMNUCC), un tratado adoptado en 1992 y que entró en vigor en 1994, siendo una de las tres Convenciones surgidas de la Cumbre de la Tierra celebrada en Rio de Janeiro en 1992. Las Conferencias de las Partes (COP) son las cumbres anuales realizadas por la CMNUCC, integradas por todos los países que ratificaron la Convención.

De todas las COP realizadas, las más destacadas son las siguientes:

- La COP 3 celebrada en 1993 en Kyoto, donde se adopta el llamado Protocolo de Kyoto en que se incluye el compromiso de la reducción de 5% en las emisiones de gases de efecto invernadero (GEI) para el año 2012 y con respecto a 1990.

³⁶ Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679>.

³⁷ https://commission.europa.eu/law/law-topic/data-protection/reform/rules-business-and-organisations_es.

³⁸ http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm.

- La COP 21 de 2015 en París, de donde surge el Acuerdo de París que entró en vigor en 2020, donde se estableció limitar el calentamiento global para el año 2100 muy por debajo de 2° con respecto a los niveles de la época preindustrial, y mantener los esfuerzos para que no supere 1,5°.

La Iniciativa Exponential Roadmap³⁹ reúne a innovadores, disruptores y transformadores que toman medidas en línea con el objetivo de controlar el calentamiento global en menos de 1,5°C para el año 2050. El objetivo es acelerar la acción y las soluciones climáticas exponenciales a través de proyectos innovadores, con la misión de reducir a la mitad las emisiones antes de 2030. La Iniciativa Exponential Roadmap es un socio acreditado de Race To Zero⁴⁰ de las Naciones Unidas y socio fundador de los Líderes de la Cadena de Suministro de 1,5°C⁴¹ y el Hub Climático de PyMEs⁴². La iniciativa ha creado el 1,5°C Business Playbook, una guía para la acción climática exponencial. Su nombre proviene de que plantea que para lograr el objetivo del Acuerdo de París es necesario la reducción de las emisiones de GEI a la mitad cada década a partir de 2020.

En el cumplimiento del objetivo de la reducción del 50% para el 2030 el sector digital tiene el potencial de reducir directamente las emisiones de combustibles fósiles en un 15% e indirectamente apoyar una reducción adicional del 35% a través de la influencia de las decisiones de los consumidores y las empresas y la transformación de los sistemas. Entre esas decisiones se encuentra la reducción del consumo por transporte a través de la sustitución de viajes con reuniones virtuales o los sitios de coworking, la digitalización determinante para aumentar la eficiencia energética y de los materiales y permitir la economía circular, la inteligencia artificial y la IoT optimizando los procesos, entre otros.

Pero al mismo tiempo que las TIC ayudan a reducir las emisiones de los GEI, ellas son responsables de sus propias emisiones. Según un estudio, publicado por ScienceDirect en 2021, todos los sectores de las actividades a nivel mundial deben adoptar medidas para el control de las emisiones de gases de efecto invernadero en el marco del acuerdo de París, incluyendo a las TIC, ya que estas últimas son responsables de las emisiones mundiales de GEI entre el 1,8% y el 2,8%. Existen diferencias pronunciadas y mucho debate sobre los supuestos subyacentes detrás de los estudios revisados por pares, que podrían sugerir que, ajustando por el truncamiento de las vías de la cadena de suministro, las emisiones globales de las TIC son tan altas como 2.1% a 3.9%.

La ARCEP y la ADEME de Francia, a principios de 2022, emitieron un documento⁴³ sobre el impacto ambiental de la digitalización y su visión prospectiva. Entiende la ARCEP que los estudios actuales se basan en metodologías no armonizadas, no transparentes y que alcanzan solamente el efecto parcial sobre la huella de carbono. Utiliza un procedimiento más abarcativo a través de la evaluación del ciclo de vida (ACV) basado en Normas y repositorios públicos específicos y lo más completos posible para lograr un desglose de lo digital en tres ladrillos físicos: terminales, redes y Centros de Datos (éste es el aspecto multicomponente de ACV), una mayor precisión al emplear 11 indicadores ambientales aparte de la huella de carbono incluyendo por ejemplo el agotamiento de los abióticos, y un análisis que integra los impactos generados durante todas las etapas del ciclo de vida de cada uno de estos tres ladrillos, es decir, las fases de fabricación, distribución, uso y final de vida útil.

Así, según este documento, las TIC representan actualmente entre el 3 y el 4 % de las emisiones de gases de efecto invernadero (GEI) en todo el mundo y el 2 % de la huella de carbono a nivel de Francia, en todas sus fases. En cuanto a la participación de cada tipo de participantes de las TIC en la huella de

³⁹ <https://exponentialroadmap.org/>.

⁴⁰ <https://unfccc.int/climate-action/race-to-zero-campaign>.

⁴¹ <https://exponentialroadmap.org/supply-chain-leaders/>.

⁴² <https://smeclimatehub.org/>.

⁴³ EVALUATION DE L'IMPACT ENVIRONNEMENTAL DU NUMERIQUE EN FRANCE ET ANALYSE PROSPECTIVE. Note de synthèse réalisée par l'ADEME et l'Arcep. 19 janvier 2022.

carbono, los dispositivos representan el 79% de la huella, seguidos de los Centros de Datos (más del 16%) y las redes (alrededor del 5%). En cuanto a la fabricación de los equipos, esta fase representa el 78% de su total, mientras que la fase de uso representa el 21%.

Las medidas de corrección de la tendencia al crecimiento de las emisiones de GEI y consumo de los abióticos, se ha extendido al sector de las TIC.

Múltiples reguladores han comenzado a considerar la sostenibilidad, como el ComReg de Irlanda, la CNMC de España, la Ley de ARESEP de Costa Rica que impone obligaciones a la SUTEL, el BIPT de Bélgica, la OFCOM del Reino Unido, entre otros.

También se destaca lo indicado por la OECD en un reciente documento⁴⁴, como, por ejemplo: "...los objetivos medioambientales, como la reducción de las emisiones de gases de efecto invernadero y del consumo de recursos naturales, y la protección de la biodiversidad, deben traducirse, siempre que sea posible, en objetivos políticos de regulación de la comunicación. Si bien el área objetivo es muy reciente y solo emerge a través de los primeros reguladores de comunicación en el área de la OCDE, es probable que gane importancia para los reguladores del futuro."

Si bien, como se indica en el documento de la OECD y surge de las actividades de los reguladores, los trabajos en el sentido de proteger la sostenibilidad trabajando directamente sobre las TIC se encuentran en sus inicios, se observa que los Centros de Datos, que representarían más del 16% de la emisión de los GEI (según ADEME-ARCEP) de las TIC, son foco de atención en este momento. ARCEP y ADEME también indican⁴⁵ que, si no se toman medidas para reducir la huella ambiental de la tecnología digital, la huella de carbono (de las TIC) podría triplicarse entre 2020 y 2050 en el escenario tendencial.

En este marco la Unión Europea aprobó en 2020 el Pacto Verde Europeo, con el objetivo de impulsar el uso eficiente de los recursos a través de una economía limpia y circular y detener el cambio climático, revertir la pérdida de biodiversidad y reducir la contaminación. Tiene un alcance muy amplio llegando a todos los sectores de la economía: energía, agricultura, industrias, transporte, TIC y otras, convirtiendo a la Unión Europea en climáticamente neutra en 2050.

En este marco surge el Pacto para los Centros de Datos Climáticamente Neutros⁴⁶, lanzado a principios de 2021, en que más de 100 operadores de Centros de Datos y Asociaciones de operadores incluyendo a los proveedores de servicios de telecomunicaciones, acuerdan hacer climáticamente neutros sus Centros de Datos (DC) para el 2050. Entre otros operadores se encuentran los hiperescaladores AWS, IBM, Azure y Google.

Tienen cinco áreas principales de cumplimiento:

- Comprar energía eléctrica 100% libre de carbono.
- Priorizar la conservación del agua.
- Reusar y reparar servidores (conservación de abióticos-economía circular).
- Tener metas medibles de eficiencia energética.
- Buscar caminos para reciclar el calor. Sistemas de energía circular.

Tiene una política especial para PyMEs en que permite la autoevaluación, mientras que para los demás exige auditoría externa.

⁴⁴ <https://www.oecd.org/publications/communication-regulators-of-the-future-fo2209e6-en.htm>.

⁴⁵ https://www.arcep.fr/uploads/tx_gspublication/note-synthese-au-gouvernement-prospective-2030-2050_mars2023.pdf.

⁴⁶ <https://www.climateneutraldatacentre.net/>.

El 6 de julio de 2023 el 95% de la primera cohorte de signatarios, que representan casi el 75% de la capacidad de DC en Europa, han sido certificados como cumpliendo con sus obligaciones en el marco de la autorregulación con respecto a estas cinco áreas. La certificación confirma que estas empresas cuentan con los procesos y medidas establecidos para recopilar y analizar los datos necesarios para demostrar el cumplimiento de los objetivos y cumplir con los objetivos para 2030.

Coincidiendo con lo estimado por la ARCEP-ADEME, estima que el consumo de energía de los Centros de Datos es entre el 15% y el 25% del consumo de las TIC.

Indica que entre 2010 y 2022 el consumo total de energía para los DC se ha mantenido casi constante a pesar de que el tráfico de Internet se multiplicó por 12 y las cargas de trabajo se multiplicaron por 8, debido a las mejoras de la eficiencia y la virtualización a través de la migración de los DC propios a los ubicados en la nube y a los hiperescaladores.

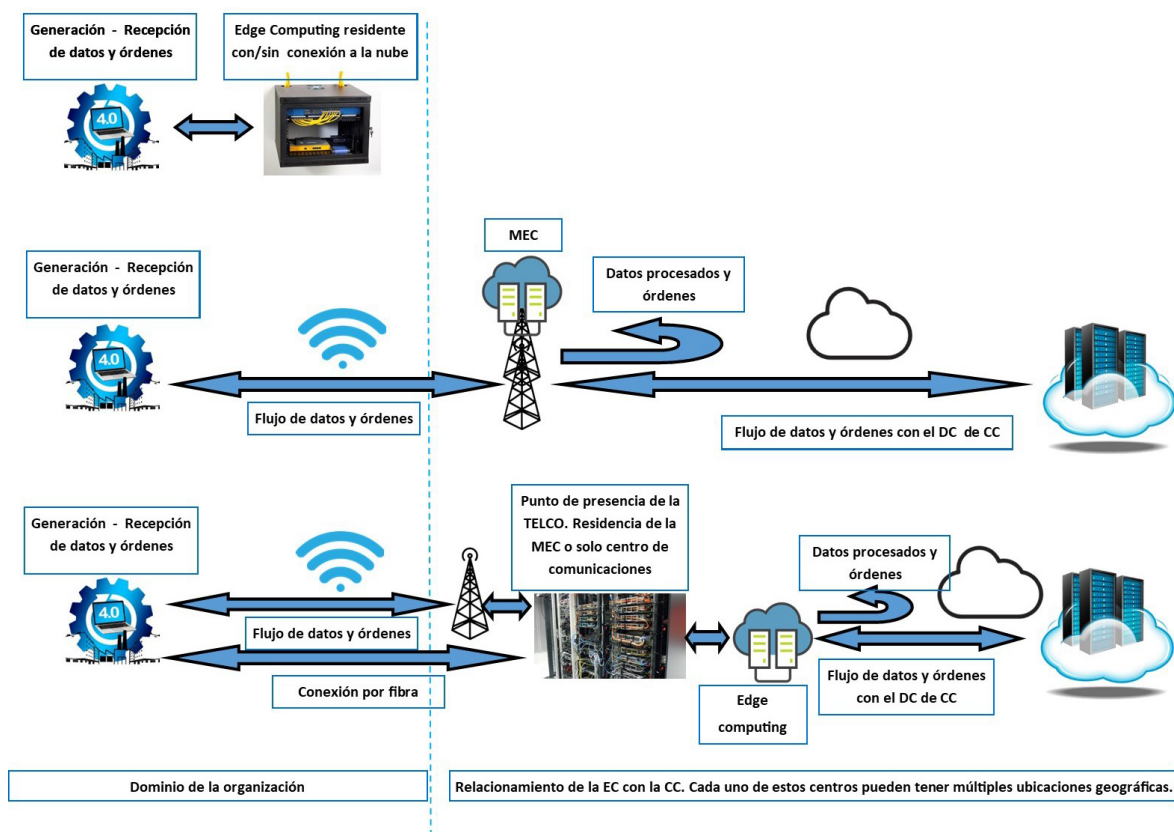
Se hace notar que la migración hacia la EC puede producir el efecto de crecimiento del consumo y el deterioro del impacto ambiental de los DC considerando sus menores economías de escala como se analiza en el capítulo III. Aspectos económicos de la EC. Este es un asunto que necesariamente debería ser incluido en las agendas para la neutralidad climática.

F. Resumen gráfico de la estructura de la EC

La gráfica siguiente es un resumen esquemático de los tipos básicos de computación en el Borde que han sido analizados en esta sección.

- El Tipo 1 corresponde al caso de las organizaciones que tienen su propio Centro de Datos donde realizan su procesamiento y almacenamiento. Con el avance del uso de la nube, por múltiples razones, estos centros emplean también servicios y aplicaciones de los Centros de Datos de la nube (CC). En este gráfico no se representa la conexión a la nube.
- El Tipo 2 es el Multi Access Edge Computing, en que la Computación en el Borde se desarrolla dentro del perímetro del CSP. Si bien ese esquema presenta solamente la conexión inalámbrica, progresivamente 5G, es posible que también sea cableada. Los datos son procesados en el Borde y solamente siguen a la nube aquellos que deben almacenarse o que requieren procesamiento no encontrado o no requerido en el borde, mientras que los datos procesados o las órdenes vuelven a los usuarios.
- El Tipo 3 es una modalidad en que los usuarios se conectan directamente al Punto de Presencia de la red del CSP, que es donde se produce la concentración de los accesos. De allí las conexiones se realizan al o los Centros de Computación en el Borde, casi siempre públicos y propiedad de terceros. En este caso los datos reciben el mismo tratamiento que en el Tipo 2.

Diagrama 1
Tipos básicos de estructura de la EC



Fuente: O. de León.

II. Evolución de los mercados relacionados a la EC

A. Descripción de los submercados y su papel en la EC

El mercado de la EC se está desarrollando íntimamente relacionado a otros mercados cuyos productos y servicios le dan soporte, y simultáneamente la EC es requerida para que ellos puedan generar valor y desarrollarse. Éstos son submercados de la EC, participando parcialmente en su desarrollo mutuo. Basta ver, por ejemplo, el mercado de la IoT para aplicaciones críticas, el que no se desarrollaría si no fuera por las prestaciones de bajos retardos y alta capacidad de comunicación de datos que soporta la EC. Al mismo tiempo la EC se desarrolla en algunas organizaciones principalmente porque es requerida por la IoT crítica, y no todo lo que es IoT requiere de la EC, por ejemplo, en las aplicaciones masivas, no críticas.

La EC se encuentra en los principios de una larga e impredecible evolución en que las organizaciones tienen necesidades de procesamiento y almacenamiento de información que surgen de casos de uso. Cada caso tiene sus particularidades que pueden incluir total o parcialmente las siguientes, o algunas adicionales que van surgiendo con la evolución de las organizaciones:

- Bajo retardo de transporte de la información a procesar entre los dispositivos que la proveen y el Centro de EC, y el retorno necesario, generalmente de menos de 5 a 30 ms. Refiere a todas las aplicaciones críticas de la Internet de las Cosas (IoT), como la automatización de la industria, la seguridad y la prevención de accidentes a partir del procesamiento de imágenes, entre otras.
- Bajo tiempo de respuesta entre los requerimientos de clientes de la organización y el suministro de la información (imágenes, audios, videos, transacciones, otras). Se requiere, por ejemplo, en los sistemas de ventas en línea a los efectos de prestar un servicio rápido y de alta definición en la presentación de los productos.

- Generación de grandes volúmenes de información a procesar como es el caso del uso cada vez más intenso de la Inteligencia Artificial o la Machine Learning. La EC reduce substancialmente los costos del transporte de datos a largas distancias y lo vuelve más resiliente. Por otra parte, esos procesamiento de la información requieren capacidades muy importantes, las que están haciendo evolucionar el hardware de la EC. Se ha pasado de los procesadores tradicionales o CPU, a los procesadores gráficos o GPU, con capacidad de múltiples procesamiento paralelos, y muy recientemente a los AI Chips especialmente desarrollados para Machine Learning, como se describió en la sección "I.A.3. Tecnologías emergentes relacionadas a la EC".
- Seguridad de la información. En muchas aplicaciones, y por diversas razones, es necesario dar alta seguridad a la información, lo que se logra manteniendo los datos en una red de EC, lo más restringida posible al sitio de su generación, al menos hasta que sean anonimizados.
- Protección de datos personales de acuerdo con la legislación de cada país, y como se ha analizado en la sección "I.D. Soberanía de los datos y la EC".
- Distintos tipos de relacionamiento con el proveedor de los servicios de comunicaciones que pueden ir desde la provisión de una solución empaquetada de comunicaciones y EC, con o sin acceso personalizado a la gestión u obtención de datos de la red de comunicaciones, hasta el suministro puro de la red de acceso y transporte.
- Requerimientos del suministro por parte del proveedor de servicios de nube en el borde, de aplicaciones o servicios propios para el caso de uso.
- Provisión de una solución completa por venta o arrendamiento para el caso específico de uso.

Esta apertura de múltiples casos posibles, casi uno por cada organización, indica que el mercado de los suministradores de la EC debería evolucionar, por razones de eficiencia, hacia soluciones más uniformes por el lado de los suministradores, y requerimientos también más uniformes por el lado de las organizaciones. Un estudio de Gartner⁴⁷ de fines de 2022 analiza los detalles de esta evolución esperada, observando la situación actual en que la EC está siendo soportada por una variedad de suministradores que están entrando a satisfacer la demanda en este nuevo mercado. En esta evolución no se toma en consideración los casos de Centros de Datos propios de las organizaciones y ubicados en sus predios, algo que, si bien se integra a la EC, tiene un origen distinto y extremadamente más limitado que la concepción actual de la EC.

Los principales resultados de este estudio se resumen en lo siguiente:

- Existe una amplia gama de proveedores de TI tradicionales y entrantes que están generando nuevas ofertas adaptada a la EC.
- Identifica del orden de ocho o más submercados de EC que intervienen en las soluciones extremo a extremo, con soluciones que a veces los superponen.
- Los primeros usuarios de EC a menudo se adaptan al principio a través de casos de uso frente a plataformas y arquitecturas, pero cuando la tendencia continúa madurando y la extensibilidad⁴⁸ se vuelve central, tal cual es requerida por la rápida evolución de los requerimientos, las ofertas y el mercado en general se consolidarán en un marco de capacidades / funcionalidades requeridas.

⁴⁷ "Market Guide for Edge Computing". <https://www.gartner.com/doc/reprints?id=1-2BTD6JNA&ct=221128&st=sb>.

⁴⁸ La extensibilidad del software de una solución es una cualidad que permite a los equipos de TI realizar las modificaciones y adaptaciones de acuerdo con las necesidades evolutivas impuestas por los objetivos de la organización sin afectar las funciones principales.

Este documento de Gartner también indica como principales recomendaciones para los líderes de Infraestructura y Operaciones (I&O) de los proyectos de EC tomar en consideración estos conceptos:

- Orientar inicialmente los desarrollos de TI a la extensibilidad, basados en un plan estratégico que se centre en el empleo de plataformas, marcos de trabajo e interoperabilidad.
- Tomar en consideración que los proveedores, y los submercados en los que operan, se irán consolidando durante el desarrollo progresivo de la EC. Para ello es importante trabajar con socios potencialmente fuertes en sus mercados, y al mismo tiempo mantener clara una estrategia alternativa de migración.
- Habilitar la flexibilidad eligiendo soluciones basadas en el ecosistema de los socios para la solución total, en lugar de solo en proveedores específicos u ofertas de proveedores específicos.

El mercado amplio de la EC es muy dinámico y fluido mientras se adapta eficientemente a la demanda a través de la reorientación de las ofertas de los prestadores tradicionales de servicio de TI, incluyendo los de la nube, o aparecen suministradores nuevos atraídos por la evolución esperada de la EC, y adicionalmente ir abandonando el suministro requerido de “soluciones únicas” para cada caso de uso.

Gartner considera que: “A medida que surjan estándares y patrones repetibles, los submercados superpuestos de computación en el borde se resolverán en una serie de mercados maduros y bien delineados en los próximos seis años más o menos. Hasta entonces, se discutirá la computación en el borde como un mercado de submercados superpuestos. Gartner cree que la superposición entre los submercados se reducirá a medida que los proveedores se centren en capacidades específicas y se asocien con otros proveedores para obtener soluciones totales.”

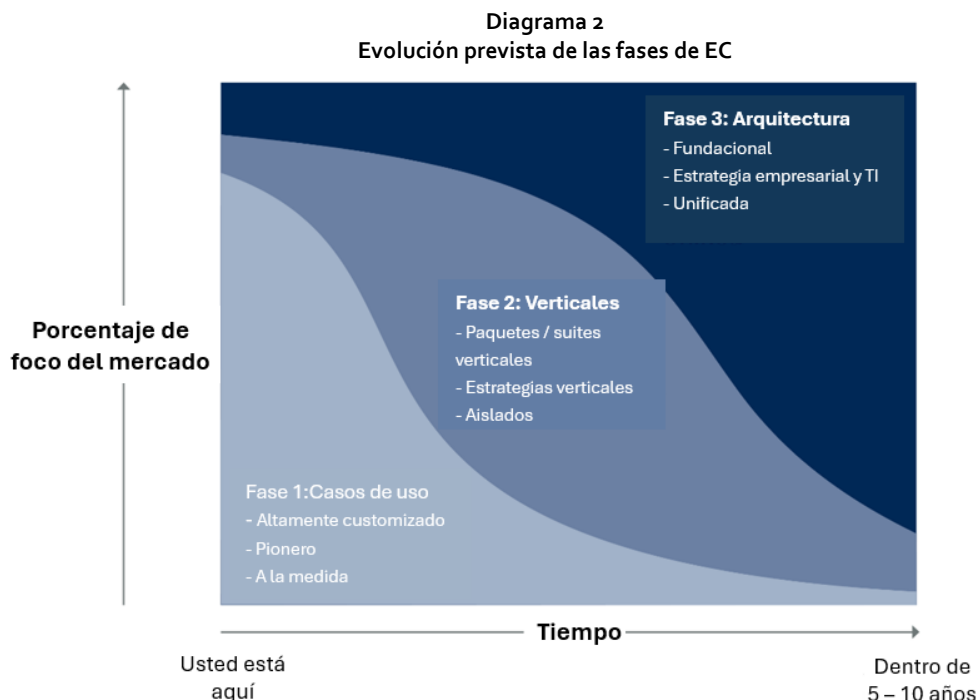
Distingue tres fases en la evolución prevista del mercado de EC, a partir de 2022 en que todavía no está orientado a las fuertes estrategias maduras como las que abarcan las recomendaciones indicadas, sino solamente a la solución de casos de uso.

- Fase 1 o de Casos de Uso. Los casos de uso en los que se basa el mercado son muy diversos considerando todas las variables que intervienen en el requerimiento de uso de la EC. Además, las decisiones para el empleo de la EC provienen de necesidades de las áreas de negocio (por ejemplo, mejorar el acceso de los clientes a las ofertas en línea) o de operación (como el caso de la introducción de IoT), por lo que son principalmente conducidas por los líderes de dichas áreas, y no tendrán la clara orientación dada por las mejores prácticas que podría darle el área de TI con una visión más estratégica. En esta fase la solución de EC está altamente personalizada, es primera en su tipo y hecha a medida.
- Fase 2 o de los Verticales. “Después de que las empresas comienzan a implementar soluciones para casos de uso específicos de EC, inevitablemente identifican nuevos casos de uso y cargas de trabajo para sus entornos que tienen otros requisitos de EC (baja latencia, datos locales de alto volumen, tolerancia a la desconexión, privacidad de datos). Dentro de las industrias verticales, esta expansión requerirá soluciones similares, impulsando suites y paquetes de computación de borde orientados verticalmente.” En esta fase se encontrarán los paquetes o conjuntos de software y aplicaciones destinados a un fin ÚTIL para verticales, o silos de organizaciones.
- Fase 3 o de la Arquitectura. “Muchos requisitos tecnológicos para la EC abarcarán verticales. En lugar de invertir en estas tecnologías y servicios, los proveedores centrados en industrias verticales específicas eventualmente aprovecharán las soluciones horizontales. Los requisitos dentro de las empresas también se extenderán desde el borde de la tienda orientada al cliente hasta el borde de la fábrica y el borde del lugar de trabajo, poniendo más requisitos en procesos y tecnologías holísticos y horizontales para mejorar la integración y

la seguridad, y para reducir los costos. Las estrategias y arquitecturas de EC impulsarán el mercado de EC en la Fase 3.” Las estrategias de TI serán el eje del desarrollo de la EC de las organizaciones, lo que se estima que dominará el mercado en 5 a 10 años.

“Las empresas con visión de futuro se centran hoy en día en las estrategias y arquitecturas de EC, pero aún no son los principales compradores en el mercado de EC. A medida que la estrategia, la extensibilidad y las arquitecturas para la computación de borde maduren, el mercado de la EC se sacudirá y se resolverá en submercados específicos y ecosistemas de socios que ofrecen soluciones totales.”

El siguiente diagrama ilustra la visión de Gartner sobre esta evolución en fases.



Fuente: Gartner.

En este trabajo se identifican ocho submercados en relación con la EC, que están impulsando el mercado de EC, el que evoluciona hacia un mercado más sólido y con soluciones estratégicas más que para casos de uso, sea por ser sus habilitadores o por requerir de la EC para su desarrollo:

- **Gestión y orquestación del borde.** “Las soluciones de administración y orquestación en el borde (EMO) proporcionan capas de control sobre la administración de servidores y dispositivos, la administración de redes y seguridad, la pila de software de infraestructura y las propias aplicaciones. Las soluciones EMO proporcionan para los servidores de borde, las pasarelas de borde, y otros nodos de borde, lo que las soluciones unificadas de gestión de puntos terminales⁴⁹ proporcionan para dispositivos móviles y clientes”.
- **Plataformas de Internet de las Cosas (IoT).** Con el constante aumento del uso de la Internet de las Cosas, principalmente en las aplicaciones críticas que pueden requerir muy bajos retardos y gran capacidad de conexión, las plataformas que gestionan el ecosistema de la IoT están íntimamente ligadas a la EC para lograr los objetivos de la solución.

⁴⁹ Se entiende como puntos terminales de la red en toda su extensión.

- Gestión de los datos en el borde. “Los datos de alto valor que se generan en el mundo físico, es decir fuera de los Centros de Datos en general, requieren progresivamente un rápido procesamiento que debe ubicarse cerca, en el espacio y el tiempo, del punto de generación. Este es el papel de las capacidades de gestión de datos en el borde.
- Analítica de datos y Machine Learning. “Los avances continuos de la analítica en el borde y Machine Learning se producirán a través de avances en hardware, software y estándares de desarrollo/implementación. Con el costo de la latencia de datos, la conectividad de red inestable y las demandas cada vez más rígidas de soberanía y gobierno de datos, los líderes de TI deben aumentar el conocimiento de las capacidades de la analítica de borde y ML. Esto asegurará que alineen la mejor solución técnica para respaldar las demandas vitales de resultados comerciales.” Uno de los sectores emergente importante es el de analítica de video empleando Inteligencia Artificial, el que se usa para proveer seguridad, incluyendo la laboral, mejorar la eficiencia de los procesos, entre otros, siendo además una aplicación que requiere la EC debido a los retardos mínimos exigidos y los volúmenes de información a transportar hasta el centro de procesamiento. Un caso típico es de los comercios minoristas que automatizan las compras realizadas por sus clientes, el pago a realizar y emplear los registros para optimizar los productos y su ubicación.
- Soluciones de servidores de la EC. Este submercado está constituido por una variedad de proveedores incluyendo los hiperescaladores, proveedores de EC a través de redes de Centros de Datos, proveedores de plataformas de EC que emplean nubes híbridas, y varios más.
- Infraestructura de comunicaciones para el borde. Es provista por los CSP para soluciones de acceso fijo o móvil, y puede estar empaquetada con EC en la llamada MEC. También existen sistemas propietarios de acceso como LORAWAN, SIGFOX y otros.
- Servicios en el borde de Centros de Datos y Redes de Distribución de Contenido (CDN).
- Soluciones en el borde para los verticales.

B. Los CSP como usuarios de la MEC y proveedores de servicios sobre ella

Antes de entrar en lo relativo a la EC es conveniente considerar que los CSP están entrando en una fase de evolución hacia la nube, en este momento principalmente en la provisión de servicios a las grandes empresas. Las redes de comunicaciones entre múltiples sitios, inclusive en diferentes países, se construyen sobre infraestructura de enlaces con cierto grado de flexibilidad y de rutas alternativas. Cuando las empresas comienzan a llevar sus datos y procesamiento a la multi nube, incluyendo centros ubicados en sus sitios o en nubes públicas o privadas, sea en el borde o en centros lejanos, todos interconectados entre ellos, los CSP pueden comenzar a escalar en valor al incorporar la nube en sus servicios, sean propias o de terceros. Por encima de ello pueden prestar servicios de transporte y de gestión multi nube, entre nubes con diferentes perfiles de servicios, de forma que los usuarios tengan servicios más resilientes, a menor costo (se estima que estas redes virtuales costarán 15% a 25% menos para el CSP), con menores tiempos de provisionamiento y con una visibilidad sobre esta red que no tenían antes los clientes. Esta evolución está en general en sus primeras etapas, y más avanzada en algunos CSP, incluye grandes inversiones y cambios culturales en los CSP que requieren la incorporación de recursos humanos que no solamente conozcan de la nube, sino que también puedan asesorar a los usuarios. Con una visión de futuro se puede decir que los CSP migrarán la prestación de sus servicios actuales, integrados con otros que surgirán, de las infraestructuras actuales hacia modelos de negocio de infraestructura en la nube, incluyendo las interconexiones internacionales. Es una transformación trascendental de los CSP, que dejan de estar atados al hardware de red a gestionar la conexiones en la nube, un asunto que requeriría un extenso

y profundo análisis signado por las incertidumbres para poder concluir sobre el futuro a partir de las condiciones actuales iniciales del cambio.

Lo que sigue muestra los pasos que se están dando en este momento y que van inicialmente en ese sentido.

Las telecomunicaciones, especialmente a través de la 5G, emplean la EC para la gestión de sus redes a través de la virtualización, y a su vez soportando soluciones que permiten generar nuevas fuentes de ingresos. Precisamente este doble uso de la EC surge, como lo dice el ETSI en la justificación de su trabajo del desarrollo de la MEC, de que la virtualización de la red crea una plataforma de EC que es eficiente para emplearla para otros fines.

Algunos indican que los CSP están viviendo el “momento AWS”, que fue cuando Amazon se vio obligada en la década del 90 a empezar a construir su propia red de procesamiento de la información, así como las capacidades humanas, para soportar flexiblemente su núcleo de negocio de venta en línea. Desarrollada esta infraestructura observó que podía venderla a terceros sobre la base de la flexibilidad total y a través de una tarjeta de crédito. En el caso de los CSP este camino se ve potenciado por estar en el pasaje de toda la información a ser procesada, y ya tener los clientes contratando sus servicios para poder procesar su información. Al igual que pasó con Amazon, su propia infraestructura requiere inversiones complementarias para incorporar el servicio de EC a sus propios servicios.

Por otra parte, los CSP observan que la necesidad de prestar servicios de alto consumo de comunicaciones cerca de los usuarios es una oportunidad considerando sus bases de clientes que en conjunto alcanzan a casi toda la población, su relación de confianza con ellos y sus flujos de caja asegurados. La EC es para los CSP la oportunidad de insertarse sólidamente en el crecimiento de los mercados ligados a la digitalización que muestra aplicaciones con grandes perspectivas de futuro como los juegos, la automatización, la inteligencia artificial, el metaverso y múltiples más. Las intensas acciones de los hiperescaladores, los proveedores de EC, los proveedores de equipamiento, entre otros que también ven las oportunidades de la EC, impulsan aún más a los CSP a participar con infraestructura propia o en sociedad con estos proveedores.

Así se genera el ecosistema que se ha analizado en la sección: “B. Principales características de las acciones del ETSI con relación a la MEC”, donde se observa su estandarización por el ETSI, y los mecanismos para favorecer el desarrollo de la EC en las propias redes de los CSP. La MEC es una tecnología complementaria de la 5G, colocando los recursos de procesamiento y almacenamiento en el borde de la red pública de acceso (RAN), y por tanto cerca de los usuarios en general (personas, empresas, dispositivos y otros).

En una primera revisión de lo que está sucediendo se detecta que este ecosistema se manifiesta de tres maneras: con los CSP y los proveedores de EC trabajando cada uno por su lado en algunos mercados, colaborando en la generación de productos y servicios convergentes en otros, o los CSP creando su propia EC para integrar a sus servicios de telecomunicaciones. Estos avances son muy flexibles como se puede ver a través de múltiples ejemplos en el mundo.

Por ejemplo, SK Telecom ha instalado, en sociedad con Dell Technologies de EE. UU., una red de MEC llamada Petasus, que incluye desde la virtualización de su red hasta las aplicaciones para sus clientes. En el año 2023 SK Telecom testeó su solución MEC en los EE. UU., en Arlington, usando espectro de ATSC 3.0⁵⁰ y su plataforma de borde para medios. Deutsche Telekom, así como Claro en Brasil, están desplegando su propia infraestructura de EC. Por otro lado, Telefónica está colaborando con AWS para su producto convergente.

⁵⁰ ATSC 3.0 es también conocido como NextGen TV y es el último estándar desarrollado para la emisión de televisión. Corea ha sido el primer país en desplegarlo en el 75% de la población. EE. UU. lleva más del 50% desplegado.

Las redes MEC privadas proveen los mismos recursos informáticos, pero en las instalaciones de los usuarios y empleando una red de acceso 5G privada, que en general es provista por el CSP. Éstas pueden ser empleadas para aplicaciones críticas, de muy bajo retardo y alta resiliencia, mejorando la seguridad y la privacidad.

En particular, Verizon puede ser tomada como referencia de la potencialidad de la MEC con la que está desempeñando una intensa actividad a nivel mundial, y a principios de 2022 era una de las dos empresas de EE. UU., junto a INTEL, que estaban entre las 10 con más patentes de EC en el mundo, y ubicada en el 8º lugar.

Mantiene asociaciones en cuanto a la EC con Amazon Web Services, Google Cloud Platform y Microsoft Azure, y considera generar ingresos en el borde a través de dos andariveles: en la nube pública donde comparte ingresos con los hiperescaladores, y a través de su modelo de MEC privada con relación directa con sus clientes y donde no comparte ingresos.

Ha fundado el 5G Future Forum⁵¹ junto a América Móvil, Vodafone, Telstra, KT y Rogers con el objetivo de desarrollar la EC que permita la interoperabilidad global de aplicaciones y servicios a través de diferentes operadores, y ya han comenzado a desarrollar especificaciones.

Es de interés observar su estrategia a través de las siguientes sugerencias que efectúa a sus potenciales clientes para seguir el camino hacia un negocio en tiempo real⁵²:

- “Al igual que con cualquier evolución tecnológica importante, cuanto mejor y antes se prepare para 5G y MEC, más suave debería ser la transición y más rápido se podrían aprovechar los beneficios que traerán. Los siguientes son pasos que se deberían seguir:
- Evaluar y establecer una línea base de su entorno, para que pueda desarrollar un plan de transformación ejecutable.
- Comenzar la transición de los entornos operativos tradicionales a los programables, bajo demanda y definidos por software.
- Implementar soluciones de visibilidad de aplicaciones para obtener información sobre su rendimiento.
- Identificar los procesos de negocio y las aplicaciones que más se beneficiarían de la capacidad de respuesta en tiempo real y la baja latencia.
- Crear un proyecto piloto para evaluar el valor de 5G y MEC para la organización.”

A los efectos de tener una visión global de lo que sucede en los CSP, con relación a la EC, también se tiene una encuesta realizada por el TMForum⁵³ en su versión más reciente de octubre de 2022⁵⁴. Incluye una muestra de 208 participantes de todo el mundo con el 81% perteneciendo a operadores. Los 33 operadores encuestados responden por más del 70% de los ingresos mundiales del negocio de los operadores de telecomunicaciones, y los encuestados corresponden a 3% de Latinoamérica, el 1% de Norte América, 11% de África, 37% de Europa, 12 % del Oriente Medio y el 36% de Asia y Asia Pacífico (incluyendo India y China). Se analizan a continuación los principales resultados referidos a los servicios de la nube (EC).

⁵¹ <https://www.5gff.org/about-us/>.

⁵² <https://www.verizon.com/business/solutions/5g/edge-computing/>.

⁵³ Es una alianza de más de 800 empresas globales trabajando en conjunto para romper las barreras tecnológicas y culturales, integrada por proveedores de servicios digitales, proveedores de tecnologías, consultoras e integradores de sistemas. https://www.tmforum.org/about-tforum/?_gl=1*okmzqe*_ga*MTg4NDQzOTAw.My4xNjkzMjUzNzZm*_ga_KQXXWgEY6Y*MTYY5MzMyNjIwNi4zLjAuMTY5MzMyNjIwNi4zLjAuMA..*_ga_W21R8NVK4E*MTY5MzMyNjIwNi4zLjAuMTY5MzMyNjIwNi4zLjAuM.

⁵⁴ <https://info.tmforum.org/rs/021-WLD815/images/Benchmark%20Report.pdf?alid=eyJpIjoiS28rQUtLUU1XK3JJdmgiZiIsInQiOiJnZVZlaml4Skxoc3hpPVokxbW42bE53PT0iFQ%253D%253D>.

A pesar de ser reciente, sus resultados pueden verse alterados en el último año, considerando los cambios permanentes de tecnologías y estrategias de los operadores. Por tanto, es una referencia importante, pero debe ser considerada como vigente al momento de su realización en 2022, al igual que sucede con otras encuestas que se encuentran en este documento.

A nivel macro se observa una CAGR del 19% de los ingresos de la EC por parte de los CSP, entre 2018 y 2021.

Se observa una fragmentación extrema en el mercado de la EC por parte de los operadores, adoptando múltiples estrategias de acuerdo con lo que van observando como oportunidades. Los extremos son, por un lado, los operadores que aprovechan su experticia en la provisión de servicios gestionados para elaborar su propia oferta de la nube, y por otro, aquellos que no desean tener sus propios activos y optan por asesorar a sus clientes para desarrollar sus propias infraestructuras de entornos híbridos o multi nube.

Pero también hay todo tipo de estrategias en el medio. Por ejemplo, muchas empresas de telecomunicaciones se están asociando con hiperescaladores e implementando la nube pública dentro de sus Centros de Datos, ya sea en una base de ingresos compartidos, o mayorista / minorista. Y luego hay un puñado de operadores, como Jio (India) y Rakuten (Japón y otros), que han desarrollado sus propias plataformas nativas de la nube y ahora apuntan a vender esa experiencia y tecnología a otras empresas de telecomunicaciones.

En cualquier caso, los CSP tienen claro la oportunidad de crecimiento de ingresos a través de esta línea de negocio, resultando de la encuesta de TM Forum que el 89% considera que este negocio representa una oportunidad significativa (52%) o moderada (37%) de crecimiento y solo el 11% considera que no es significativa.

Orange Business Services le otorga especial importancia a su estrategia de aumentar su presencia en cuanto a sus capacidades de nube, incluyendo tercerización de la infraestructura de datos, gestión de servicios sobre infraestructura híbrida, o asesoramiento tanto para que sus clientes inicien el camino hacia la nube, como para que puedan construir su propia infraestructura.

Telefonica Tech ha adoptado el camino de asociarse con AWS, como fue anunciado a principios de 2022, instalando AWS Outposts en su red saliendo juntos al mercado para ofrecer los servicios convergentes de EC y comunicaciones.

BT Global Services soporta la migración de sus clientes a la nube de varias maneras, incluyendo el servicio de conectividad multi nube ofrecido en sociedad con Equinix, dándole acceso a aplicaciones y servicios de varias terceras partes en multi nube a través de una conexión única que provee más seguridad, simplicidad y eficiencia.

En EE. UU., aparte de Verizon, AT&T tiene una sociedad con Microsoft Azure, IBM Cloud Satellite y Google Cloud Platform ofreciendo servicios que emplean sus redes 4G y 5G, mientras que T-Mobile ofrece sus servicios propios convergentes de 5G en la avanzada modalidad Stand Alone, y EC, principalmente para redes privadas.

Por otro lado, la evolución en el consumo de servicios en la nube hacia nuevas modalidades en el borde, incluyendo la multi nube, hace necesario proveer simultáneamente redes de comunicaciones más flexibles y eficientes en su uso. Como ya se vio, las redes WAN definidas por software (SD-WAN) son el soporte ideal para las nuevas redes de computación. Inclusive, los sistemas de gestión de las redes de EC pueden incluir la gestión de la SD-WAN.

En esta línea, Telenor espera crecimientos importantes de ingresos por SD-WAN y redes privadas, y en particular ha anunciado una red de conexión a la nube usando SD-WAN y alcanzando a más de 500 sitios.

La seguridad y la soberanía de los datos, ya analizadas en la sección “I. D. Soberanía de los datos y la EC”, es una cuestión crítica que es solucionada por la EC, y en particular le dan una ventaja a los CSP.

Por su parte, TM Forum, en su documento “Telco revenue growth: taking it to the next level”, de fines de 2022, incluye los siguientes casos.

La orientación nacional o regional, tradicional de los CSP, tanto en infraestructura como en gestión, puede ser una desventaja considerando el poder y competencia global de los hiperescaladores en la nube. Por otra parte, esta misma orientación puede ser una oportunidad en el campo de la gestión de datos, ya que los gobiernos requieren cada vez más que las empresas y las instituciones públicas respeten las leyes y regulaciones locales de soberanía de datos.

Un caso destacable en Europa es el de Telenor, Deutsche Telekom y Orange, los que integran un grupo de proveedores de servicios que se posicionan como proveedores nacionales o regionales de confianza de servicios en la nube, y que mantienen los datos dentro de las fronteras nacionales. Orange, por ejemplo, ha creado una empresa llamada Bleu con Capgemini para abordar los requisitos de soberanía de datos de los departamentos estatales y administrativos públicos franceses.

Las cuestiones relativas a la protección de datos a nivel nacional o internacional se extienden a varias partes del mundo, como el caso de la India en necesariamente Airtel e IBM se han asociado para proporcionar servicios de EC a las empresas a través de los 120 Centros de Datos de Bharti Airtel en 20 ciudades de la India.

Los CSP también están trabajando para la creación de las nubes específicas para determinados sectores económicos, y que soporten las regulaciones y especificaciones establecidas, como es el caso de la industria automovilística en Alemania. Deutsche Telekom es una de las fundadoras de Catena-X, una nube de la industria automotriz financiada por el gobierno, autodefinida como el primer ecosistema colaborativo de datos abiertos para la industria automotriz del futuro, que vincula a los actores globales en cadenas de valor de extremo a extremo. El objetivo compartido es un intercambio de datos global estandarizado basado en los valores europeos.

Los hiperescaladores, a su vez están trabajando en nubes específicas para determinados sectores lo que puede significar una competencia para los CSP o una oportunidad de asociarse para los servicios convergentes.

Estas actividades han motivado a los CSP, aprovechando los cambios importantes que trae la 5G, a hacer sus servicios de crear, provisionar y usar tan simples como lo hacen las empresas de la nube. Como ya se ha mencionado, el desacoplamiento de la red física de la capa de comando está permitiendo esta evolución hacia la creación de una plataforma de gestión en la nube de sus servicios de comunicaciones.

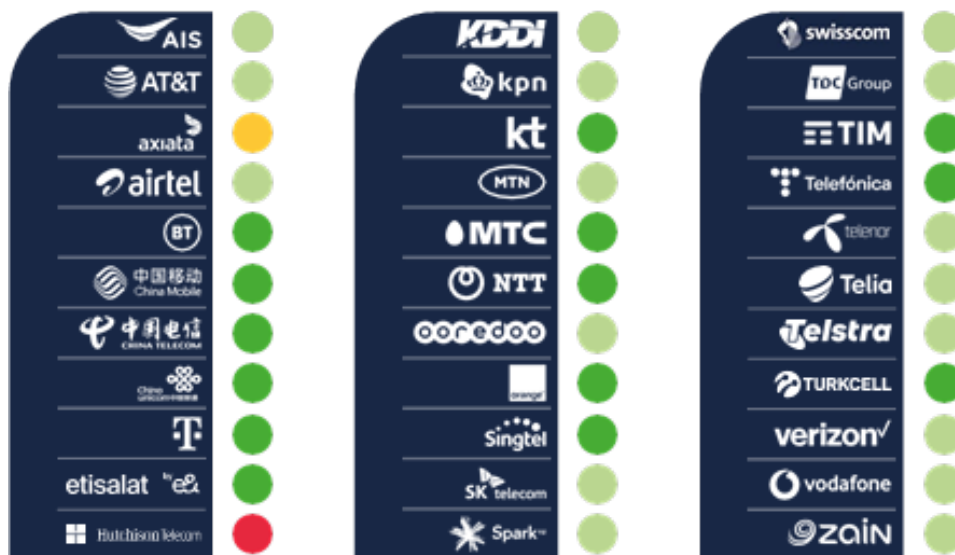
En este sentido el documento de TM Forum menciona que en Japón, Rakuten Symphony ha desarrollado una sola plataforma unificada, basada en datos, y habilitada para IA, que proporciona automatización sin intervención (Zero Touch) para su red nativa en la nube.

También Telenor se está alejando del modelo integrado de telecomunicaciones y evolucionando hacia plataformas nativas de la nube sin capas, incluida una nueva plataforma de conectividad como servicio.

Es importante observar el concepto de “conectividad como servicio” o CaaS.

El siguiente diagrama permite ver la evolución hacia la nube de los 33 CSP encuestadas en este informe de TMForum, mostrando el estado actual de este mercado altamente fluido y diverso, como ya se ha visto, en cuanto a verticales objetivo y tipo de soluciones.

Diagrama 3
Evolución de los CSP hacia la nube



Códigos:

- Poca o ninguna evidencia de una estrategia claramente articulada para abordar este mercado.
- Una estrategia declarada para construir en esta área, pero poca evidencia de clientes que pagan.
- Establecido como una línea de negocio, pero con ingresos modestos (menos del 1% de los ingresos totales).
- Un servicio principal a nivel de grupo o empresa operadora, ya sea por derecho propio o incluido con otros productos.

Fuente: TM Forum knowledge benchmark. Octubre de 2022. Telco Revenue Growth: taking it to the next level.

C. Organización territorial de la EC en los hiperescaladores

Los operadores de la nube están en constante evolución en cuanto a los servicios prestados, de acuerdo con los cambios de la demanda. Sus centros se encuentran bajo un control flexible en cuanto a la integración, por lo que su extensión a diferentes sitios es parte de su operación habitual. Por ello, tienen una facilidad especial para satisfacer las necesidades de la computación en el borde, sin limitarse a determinados verticales, simplemente estableciendo Centros de Datos en el borde y proveyendo las aplicaciones y servicios que se requieran, especiales para el borde o que sean parte de su cartera en la nube.

En definitiva, todos ellos pueden proveer servicios equivalentes, aunque por distintas modalidades.

Los proveedores de servicios en la nube de menor tamaño también ofrecen todos los servicios que son demandados por los usuarios, aunque no con toda la granularidad que ofrecen los hiperescaladores.

La demanda en el borde se origina en múltiples requerimientos entre los que se encuentran los siguientes.

- Provisión de capacidad de procesamiento, memoria y almacenamiento en el borde bajo modalidades de contratación similares a las de la CC.
- Combinación de capacidades entre la EC y la CC, y eventualmente con la EC local o propia, o con múltiples sitios. Esta situación es común cuando la organización tiene aplicaciones que requieren ser procesadas en el borde y otras que son más eficientes si se procesan en la nube.

- Adaptación a la provisión de MEC. Los CSP adoptan un par de caminos principales: desarrollar su propia EC o asociarse con proveedores de la nube para la prestación de servicios convergentes de 5G (o accesos fijos) y EC. Algunos CSP optan por desarrollar ambas modalidades para segmentos distintos de usuarios.
- Distribución de contenido. Éste es un requerimiento general que incluye el contenido de video, imágenes (ventas en línea), juegos, entre otros que requieren bajo retardo y alta capacidad de transmisión.
- Las mismas capacidades anteriores, pero para ambientes de poca o inestable conectividad. Son los casos de industrias extractivas, organizaciones en sitios de débil cobertura, entre otros.
- Empleo de Inteligencia Artificial (p.e. visión computada en seguridad), Machine Learning y similares. Exigen mucha capacidad de procesamiento y enormes transferencias de información.
- Más en general todos los sistemas de IoT crítica como la automatización industrial, la medicina de alta capacidad, entre otras, tienen los mismos requerimientos.

La organización territorial de los hiperescaladores es esencial para poder responder adecuadamente a estos requerimientos.

Desde el punto de vista de su organización territorial, las redes de los conocidos como grandes proveedores de servicios en la nube como Google Cloud, Microsoft (Azure), Amazon (AWS), IBM Cloud, Oracle Cloud, Alibaba Cloud⁵⁵ u otros, se estructuran principalmente en Regiones, Zonas y Borde. Las Regiones son las ubicaciones geográficas donde residen los Data Centers principales del proveedor de nube pública. Deberían ser el primer objetivo de los clientes en cuanto a elegir Regiones próximas a los clientes internos y externos, a los efectos de reducir retardos y costos. Las Zonas son ubicaciones precisas y aisladas dentro de una Región, donde se ubican Data Centers desde los que se originan y se operan los servicios provistos. Los usuarios suelen preferir replicar sus servicios en varias Zonas a los efectos de asegurar el cumplimiento de normas sobre datos, reducir la latencia por la proximidad a los clientes, y otras razones. Las ubicaciones de borde, ("Edge Locations"), son sitios donde los proveedores ubican el hardware y la capacidad de procesamiento y almacenamiento de acuerdo con los servicios específicos que ofrezcan de acuerdo con la demanda.

Cada proveedor les da distintos nombres a estos nodos, pero conceptualmente desempeñan funciones similares en cada uno de sus niveles de agrupamiento de los Data Centers hasta el Borde. Se presentan algunos ejemplos de ellos.

AWS⁵⁶ llama: 1. Región a un área geográfica separada; 2. Zona de Disponibilidad (AZ) a múltiples locaciones aisladas en una Región; 3. Zona Local⁵⁷, donde se provee la posibilidad de que el cliente disponga de sus recursos de computación y almacenamiento cerca de los usuarios finales; 4. AWS Outposts, que son bastidores que ofrecen una infraestructura de AWS completamente administrada, servicios de AWS nativos, API y herramientas para prácticamente cualquier cliente en las instalaciones locales (cualquier Centro de Datos, espacio de ubicación compartida o instalación local), habilitando las aplicaciones que necesitan ejecutarse en las instalaciones debido a la baja latencia, el procesamiento local de datos o las necesidades locales de almacenamiento de datos; y finalmente 5. Zonas Wavelength, que permite a los desarrolladores crear aplicaciones que ofrezcan latencias ultra bajas a dispositivos 5G y usuarios finales. Wavelength implementa servicios estándar de computación y almacenamiento de AWS en el borde de las redes 5G de los operadores de telecomunicaciones, de forma que el tráfico de las aplicaciones de los dispositivos 5G lleguen a los servidores ubicados en la

⁵⁵ <https://www.alibabacloud.com/es/global-locations>.

⁵⁶ <https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/using-regions-availability-zones.html>.

⁵⁷ <https://aws.amazon.com/es/about-aws/global-infrastructure/localzones/faqs/>.

Zona Wavelength sin salir de la red del operador de telecomunicaciones. Los clientes pueden ser externos o internos del operador. Una zona Wavelength es una extensión lógica de una Región y es administrada por el plano de control en la Región. Las primeras Regiones fueron lanzadas con dos AZ, pero las nuevas son lanzadas con tres o más. Actualmente⁵⁸ tiene 31 Regiones en todo el mundo y una en Sao Paulo, lanzada en 2011 y con tres AZ. AWS Snowcone y AWS Snowball son modalidades de EC de AWS que se emplean para los sitios remotos y de baja conectividad.

Google Cloud Platform está constituida básicamente por i) Regiones, que consisten en Zonas, generalmente cercanas entre ellas en un perímetro con un tiempo de respuesta de menos de 1 ms. para el 95%⁵⁹ de las peticiones de servicios, y casi siempre de menos de 5 ms.; ii) Zonas, que no siempre están constituidas por un único Data Center, pertenecen a una Región y son dominios independientes en cuanto a fallas; y iii) Ubicaciones de Borde, donde se ofrecen las conexiones a Google Cloud en el borde. Actualmente⁶⁰ tiene 37 Regiones, 112 Zonas y 187 ubicaciones de la Red de Borde. Esta plataforma provee un conjunto importante de servicios de computación en la nube que ofrece Google, que se ejecuta en la misma infraestructura que Google utiliza para sus productos de usuario final, como YouTube, Gmail y más. Uno de esos servicios es el de Google Kubernetes Engine (GKE), un servicio de Kubernetes administrado, destinado a implementar y operar aplicaciones alojadas en contenedores a gran escala con la infraestructura de Google.

Microsoft Azure⁶¹ se compone de 1. Región, que es un conjunto de Data Centers ubicados en un perímetro con determinada latencia y conectados a través de una red dedicada de baja latencia; 2. Zonas de Disponibilidad (AZ) constituidas por uno o más Data Centers construidos en infraestructuras de energía, enfriamiento y conectividad independientes, con un mínimo de tres AZ en las Regiones que las soportan; y 3. Zonas de Borde, que son extensiones de Azure, ubicadas en áreas densamente pobladas. Una Zona de Borde admite máquinas virtuales (VM), contenedores y un conjunto selecto de servicios de Azure que permiten ejecutar aplicaciones sensibles a la latencia y de rendimiento intensivo cerca de los usuarios finales. Azure Front Door (clásico) es un punto de entrada global y escalable que usa la red de borde global de Microsoft para crear aplicaciones web rápidas, seguras y ampliamente escalables. Con Front Door (clásico), se pueden transformar las aplicaciones empresariales y de consumo globales en aplicaciones personalizadas sólidas y de alto rendimiento con contenidos que llegan a una audiencia global a través de Azure. Azure Stack Edge es un rack que se monta localmente en la organización.

IBM Cloud Satellite⁶² permite que las operaciones de procesamiento y almacenamiento de la información corran sobre los servicios de IBM Cloud a través de cualquier entorno, incluyendo desde la Nube al Borde. Para esta operación IBM emplea Red Hat OpenShift, una plataforma de Kubernetes que habilita a los desarrolladores crear, implementar y administrar aplicaciones basadas en contenedores en cualquier infraestructura o nube, incluidos Centros de Datos privados y públicos o ubicaciones de borde. Como se observa, es un concepto de red que se apoya en la estructura de OpenShift independientemente de la estructura física subyacente que tiende a facilitar el camino en el viaje a la Nube Híbrida. Satellite provee herramientas basadas en API que se pueden usar para representar un Data Center local, un proveedor de Nube Pública o una red de Borde como ubicación de un nodo de Satellite. Se puede también generar un nodo de Satellite con equipos propios que cumplan determinadas condiciones. A continuación, estos nodos proporcionan la potencia de cómputo para ejecutar los servicios de IBM Cloud, como cargas de trabajo en clústeres gestionados de Red Hat

⁵⁸ <https://aws.amazon.com/es/about-aws/global-infrastructure/>.

⁵⁹ También expresado como P95 RT < 1 ms.

⁶⁰ <https://cloud.google.com/about/locations?hl=es-419>.

⁶¹ <https://azure.microsoft.com/en-us/explore/global-infrastructure/>.

⁶² <https://newsroom.ibm.com/Innovating-at-the-Edge-How-Hybrid-Cloud-Overcomes-the-Complexity>.

OpenShift u otras herramientas. Si bien se pueden usar los nodos que se desee, debido a la supervisión de actividades maliciosas o las actualizaciones de los nodos que son gestionadas por una Región multizona soportada por IBM Cloud Satellite, es conveniente que el cliente emplee la Región más cercana para reducir la latencia entre IBM Cloud y los nodos de Satellite. La latencia máxima recomendada es de 200 ms. RTT entre la IBM Cloud y el nodo empleado por Satellite. En Sudamérica⁶³ la única Región actual es Sao Paulo.

Oracle Cloud Infrastructure (OCI)⁶⁴. La OCI está alojada en regiones y en dominios de disponibilidad. Una región es un área geográfica localizada, mientras que un dominio de disponibilidad es uno o más Centros de Datos que se encuentran en una región. La mayoría de los recursos de la OCI son específicos de una región, como una red virtual en la nube, o específicos de un dominio de disponibilidad, como una instancia informática. Los dominios de disponibilidad están aislados entre sí, por lo que es poco probable que una falla en uno de ellos afecte a los demás de una misma región. Los dominios de disponibilidad dentro de la misma región se conectan entre sí mediante una red de gran capacidad y de baja latencia. La OCI abarca actualmente a 44 regiones geográficas de nube pública y comercial interconectadas. Cada Región ofrece más de 100 servicios de la Infraestructura de Nube de Oracle, incluyendo aplicaciones en la Nube e interconexión con Microsoft Azure. A su vez, las regiones se agrupan en dominios (comerciales, gubernamentales y dedicados), y quien arriende en un dominio puede acceder a todas las regiones de ese dominio, pero no a otros. En Latinoamérica existen las siguientes regiones del dominio comercial: Sao Paulo (2), Santiago (1) y Querétaro (1), y próximas regiones en Colombia, Santiago y México. Tiene interconexión con Azure solamente en Sao Paulo. En cuanto al borde, dispone de la Oracle Roving Infrastructure⁶⁵ (Infraestructura "itinerante"), que acelera el despliegue de cargas de trabajo fuera de los Centros de Datos. Los dispositivos reforzados de Oracle ofrecen servicios de computación en la nube y almacenamiento en el perímetro de las redes y en ubicaciones desconectadas. Estos dispositivos pueden estar sincronizados con los recursos informáticos y el almacenamiento de objetos en la OCI.

D. Distintos aspectos sobre las proyecciones para el futuro

Múltiples factores y organizaciones intervienen en todo el mundo en estos momentos iniciales de los despliegues y uso de la EC, consistentemente con la CC. Surgen así incertidumbres y tendencias iniciales que pautan esta evolución inicial, similares a las que suceden con todas las tecnologías emergentes, multiplicándose este efecto ya que la EC está estrechamente ligada a otras tecnologías emergentes de alto impacto, como soporte de ellas y como resultado por su uso o por sus requerimientos, como son la inteligencia artificial y el subgrupo de la Machine Learning, los desarrollos del uso de los cúbits y las tecnologías cuánticas, la 5G, la Internet de las Cosas y sus aplicaciones críticas, la movilidad asistida y autónoma, la medicina remota, la realidad aumentada e inmersiva, entre otras. Surge así una vertiente muy importante, como es la Edge AI, por el impacto que tendrá en el desarrollo de la Inteligencia Artificial y la Machine Learning desplegándose en el mismo borde, sea en los dispositivos finales o en sitios muy cercanos. Los nuevos chips AI, también llamados IPU, especialmente adaptados para la Machine Learning con densidades de transistores nunca alcanzadas, y producidos por hiperescaladores como Amazon, Google, IBM, Qualcomm y principalmente por Nvidia, superan largamente las capacidades de los chips GPU usados hasta ahora permitiendo la incorporación de la Inteligencia Artificial en el propio borde, en los dispositivos, en la llamada Edge AI.

Para tratar de echar luz sobre estos cambios profundos se presentan los resultados de algunas investigaciones recientes.

⁶³ <https://cloud.ibm.com/docs/satellite?topic=satellite-sat-regions>.

⁶⁴ <https://www.oracle.com/es/cloud/>.

⁶⁵ <https://www.oracle.com/es/cloud/roving-edge-infrastructure/>.

IDC, en un estudio de febrero de 2023⁶⁶, indica que los gastos de las empresas y los proveedores en hardware, software y servicios para soluciones en el borde mantendrán un crecimiento importante desde \$208 billones (+13,1% sobre 2022) hasta 317 billones en 2026, o sea con un 15% CAGR. Indica que la transformación digital se fortalece cuando los proveedores tienen la capacidad de trasladar aplicaciones y gestión de datos al borde y los clientes aceleran los planes de adopción de estas infraestructuras.

Estos avances que indica IDC, resultan de la combinación de los esfuerzos de desarrollo, y de generar ofertas atractivas, de los proveedores de los diferentes submercados ya analizados.

IDC ha identificado más de 400 casos de uso a través de múltiples industrias y dominios. Tres casos de uso, que significarían las mayores inversiones en 2023, serían fundacionales para la prestación de servicios en el borde por parte de los CSP: redes de distribución de contenido (CDN), virtualización de las funciones de red (NFV) y la computación en el borde multi acceso (MEC). En conjunto representan el 20% de las inversiones en el borde en 2023, o sea unos \$44 billones. Estos casos se encuentran entre lo que Gartner denomina la Fase 1, o de casos de uso, en los inicios de la evolución hacia la Fase 3, en que el foco es el de la Arquitectura, en que los proveedores se encontrarán en ecosistemas de socios que ofrecerán soluciones totales.

En cuanto a las empresas de adopción temprana, incluyendo el sector público, los casos de uso con mayores inversiones para 2023 serían: gestión de activos de producción, operaciones autónomas, operaciones omnicanal⁶⁷ y supervisión de fletes. Estas inversiones combinadas ascenderían a \$22 billones en 2023. Pueden observarse en dicho documento otras aperturas por sectores.

La EC, y principalmente en Europa, es una de las áreas de inversión más resilientes y atractivas, estimando IDC que estaría creciendo a tasas de dos dígitos en los próximos 5 años. En cuanto a la distribución geográfica de las inversiones en EC observa que Estados Unidos será el líder en gastos de borde durante todo el período de pronóstico de cinco años, con más del 40% del total mundial, seguido de Europa Occidental y China. Por otro lado, en cuanto a tasas de crecimiento, América Latina y China experimentarán el crecimiento más rápido del gasto durante el pronóstico de cinco años con CAGR de 18.1% y 18.0%, respectivamente. Respecto de China, más adelante se muestran indicadores claros de esta tendencia.

Se observa, a partir de este estudio de IDC que la EC tiene y mantendrá una intensa actividad de despliegue en los próximos años.

Un reciente estudio de McKinsey⁶⁸ de julio de 2023 analiza las perspectivas de desarrollo de las nuevas tecnologías, y entre ellas la EC como una de las 15 tecnologías de vanguardia. Según McKinsey, la EC, se destaca por el crecimiento firme de la innovación. No obstante, la escasez de talento podría estar limitando el crecimiento, algo común a todas las tecnologías emergentes. La cantidad de posteos de ofertas de trabajo para CC y EC ha crecido 12% en 2021-2022, mientras que el crecimiento para el desarrollo de software de nueva generación fue del 29%⁶⁹. Independientemente de este crecimiento, en ambos casos existe un déficit de talento, especialmente para la gestión de infraestructura de la nube.

La migración hacia la nube se ha ido enlenteciendo, aunque las organizaciones sigan haciéndolo. Las ventajas ya analizadas de la EC están provocando este movimiento hacia el borde, y en muchos casos provocando la llamada “repatriación” desde la nube, referida a algunas aplicaciones de las múltiples que tienen las organizaciones en la nube.

⁶⁶ <https://www.businesswire.com/news/home/20230215005133/en/New-IDC-Spending-Guide-Forecasts-Edge-Computing-Investments-Will-Reach-208-Billion-in-2023>.

⁶⁷ Una operación omnicanal significa que el cliente tiene la misma experiencia de compra independientemente de dónde compre (en línea o no) y a dónde la envía o pasa a recogerla.

⁶⁸ “Technology Trends Outlook 2023” <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-top-trends-in-tech>.

⁶⁹ Fuente: Plataforma de datos organizacionales propietaria de McKinsey, que se basa en datos de perfiles profesionales públicos no identificados. Sobre 150 millones de posteos relevados.

En cuanto a la “repatriación”, una encuesta de Uptime Intelligence⁷⁰ destaca que el hecho de que el 34% de los encuestados haya migrado algunas aplicaciones desde la nube pública no significa, necesariamente, que la nube pública deje de interesar a esas organizaciones. La nube pública continúa siendo un modelo complementario de la EC, viable para todas las aplicaciones.

Cuantificando esta migración, solo el 6% de los participantes de esta encuesta declararon que habían abandonado la nube pública por completo debido a la repatriación de la nube. Alrededor del 23% indicó que la repatriación no tuvo impacto en el uso de la nube pública, y el 59% indicó que el uso de la nube se había reducido un poco.

Estos resultados indican que la nube y la computación en el borde tienden a operar conjuntamente, cada una de ellas explotando sus características especiales, como se ha analizado en el capítulo I.

Con relación al futuro de la CC y la EC, el estudio de McKinsey de tendencias tecnológicas indica que existen algunas incertidumbres, debidas principalmente a que es una tecnología emergente, entre ellas algunas que son comunes a otras tecnologías emergentes.

- El impacto que podrá tener la regulación según como se posicione el negocio, o la disponibilidad de interoperabilidad y uniformidad de los estándares. En este sentido la MEC se ha desarrollado fuertemente para reducir estos posibles impactos negativos.
- También es importante en las decisiones a tomar cómo se posicionarán los proveedores en el borde, tanto los hiperescaladores como las empresas menores, y cómo interactuarán con los CSP, incluyendo la provisión de paquetes de computación y conectividad.
- La IA, como demandante de cercanía y capacidades de transmisión y procesamiento, influirá en el desarrollo de la EC según como se posicione en el negocio y sea afectada por la regulación.
- También tendrá impacto el desarrollo de los Chips AI, tanto en el procesamiento de los Centros de Datos como en los mismos dispositivos, como las cámaras inteligentes. La IA influirá también en su empleo para optimizar la EC y no solo en cuanto a su uso aprovechando la EC.
- La cantidad de centros en el borde y las extensas mallas de interconexión entre ellos y con la nube, provocarán la necesidad de una atención muy especial para evitar la vulnerabilidad.
- Toda la infraestructura requerida para la EC es a su vez productora de impacto ambiental, como es en las TIC en general, y como se ha analizado en la sección “E. Sostenibilidad de la computación en la nube y en el borde”. Este punto también genera cierta incertidumbre para la futura aplicación de la EC.
- Existe una fuerte tendencia a la reducción de costos de los sensores, lo que implica despliegues masivos, y a su vez con capacidades crecientes de generación de datos para transmisión y procesamiento, inclusive con IA. A esto se suma la digitalización de la gestión en general. Se desconoce con certeza el impacto de esta evolución en la EC, y también en la CC.

Este análisis de 2023 lo cierra un ejecutivo⁷¹ de McKinsey con estas palabras: “Se espera que la demanda de EC aumente significativamente en los próximos años, y los casos de uso que requieren capacidades de borde se multiplicarán. Esto puede crear oportunidades para una serie de proveedores de tecnología que actualmente juegan solo marginalmente en el espacio de la nube pública para capturar valor al aumentar ahora las capacidades de EC”.

⁷⁰ 2022 Data Center Capacity Trends Survey. <https://journal.uptimeinstitute.com/high-costs-drive-cloud-repatriation-but-impact-is-overstated/>.

⁷¹ Andrea Del Miglio, senior partner, Milan.

Respecto de la evolución futura de la EC, una observación de China muestra avances muy importantes en los últimos tiempos según un trabajo⁷² de TechMonitor de febrero de 2022, actualizado en agosto de 2023.

A partir de los datos provistos por la empresa de investigación de la innovación, GreyB, en el último año desde marzo de 2021 se ha producido un crecimiento muy grande de innovación en la industria de la EC. Siete de las diez organizaciones que han registrado más patentes en EC son chinas, al igual que las diez primeras universidades. Esta innovación parece estar conducida por el rápido despliegue de la 5G y su empleo en masivas redes de potencia, así como en el fuerte inicio de la vigilancia empleando la inteligencia artificial extendida en todo el país. La 5G es una componente básica de la EC.

Las dos primeras empresas en cantidad de patentes son dos empresas de redes de potencia de China, con 456 y 171 patentes respectivamente, la tercera es INTEL con 166 y la octava es Verizon con 98 patentes. Este último dato muestra como algunos CSP están trabajando internamente en la aplicación con desarrollo propio de la EC.

En cuanto a las universidades, la primera es la Universidad de Beijing de Correos y Telecomunicaciones con 189 patentes, y la décima es la Universidad del Centro Sur con 44 patentes. También acá, en los últimos años, se han producido cambios importantes involucrándose las universidades en la actividad de las empresas y organizaciones, tanto en el desarrollo de productos como asesorando o dirigiéndolas.

Si bien China empezó los servicios de 5G a fines de 2019, para enero de 2022 lideraba en cuanto a cantidad de radiobases, y estaba segunda en cantidad de habitantes por radiobase 5G, según el Observatorio Europeo de 5G⁷³. Las cantidades de habitantes por radiobase en países seleccionados son: Corea, 319; China, 1.531; Japón, 2.516; países de la Unión Europea, 3.988 y EE. UU., 6.590. Adicionalmente China se ha enfocado en redes 5G Stand Alone (5GSA) que libera su desarrollo dándole todo el potencial, sin depender de generaciones anteriores. Esta decisión estratégica hacia la innovación digital para la próxima década está abriendo puertas para que empresas como Tencent, Huawei y ZTE estén desplegando infraestructura para soportar la MEC.

A nivel más global, sobre el futuro de la EC, las proyecciones de la Linux Foundation⁷⁴ de 2021 para 2028 indican que el 38% de la infraestructura de EC va a estar en Asia-Pacífico, que el 29% estará en la Unión Europea y el 21% en los EE. UU.

Un área que también está adquiriendo una relevancia creciente es la de la gestión de la EC la que se vuelve compleja y desafiante cuando se multiplican los sitios tanto de EC en sí, como los de CC combinados. La emergencia reciente de la Edge AI, con la gran cantidad de sitios en los dispositivos hace aún más compleja la gestión. Los sistemas de gestión intervienen en varias capas en las instalaciones de hardware y de software y su optimización, la supervisión y el mantenimiento y actualización de todos los nodos. A los efectos de disponer de estos sistemas de gestión existen varias alternativas que dependen de varios factores como ser el tamaño de la organización, la extensión y complejidad de la red de EC, si dispone de EdgeAI, entre otros. Pueden ser desarrollados internamente, se puede empezar con un sistema de terceros, pero con posibilidades de expansión con recursos propios, o directamente tercerizado.

⁷² "What China's lead in edge computing means for the world" <https://techmonitor.ai/focus/what-chinas-lead-in-edge-computing-means-for-the-world>.

⁷³ <https://5gobservatory.eu/observatory-overview/5g-scoreboards/>.

⁷⁴ <https://stateoftheedge.com/reports/state-of-the-edge-report-2021/>.

III. Aspectos económicos de la EC

A. CAPEX y OPEX en las diferentes modalidades de nube vs. servicios prestados y calidad

Como ya fue analizado, la EC es uno de los mayores avances en la infraestructura de la computación definiendo un nuevo modelo de almacenamiento y procesamiento de la información próximos al usuario, lo que optimiza el uso de recursos, reduciendo el retardo de adquisición y entrega de la información, aumentando la seguridad de la información sensible, facilitando el cumplimiento de condiciones regulatorias y de gestión sobre los datos, reduciendo los costos de transporte, entre otros, mejorando la experiencia del usuario, lo que a su vez está permitiendo un importante impacto en las organizaciones. Esta forma de tratar la información ha sido impulsada por el gran aumento de información entregada por los dispositivos y procesos de las organizaciones, y la necesidad de su procesamiento en muy bajos tiempos acordes al tiempo real requerido para las respuestas. Finalmente se están generando nuevos y más exigentes modelos de negocio de las organizaciones.

Es cierto que para la EC, en los casos de residencia en el predio, es necesario un CAPEX mayor que en la CC, en la cual la inversión en hardware y software se convierte en una parte en general importante de pagos recurrentes u OPEX. Cuando se emplean infraestructuras compartidas para la EC el comportamiento de ellas es similar a la CC en cuanto a economías de escala, pero a menor nivel. Pero otros costos también se deben considerar, como son los de transporte de la información desde el borde a los Centros de Datos, el almacenamiento de los datos y el mantenimiento del sistema, entre otros, los que también forman parte del OPEX a pagar al operador del Centro de Datos. El análisis de todos ellos, y sus valores en ambas alternativas es el que permite en definitiva ver cuál permite el menor costo total

de propiedad (TCO⁷⁵) durante el ciclo de vida de la solución a adoptar, debido a que el TCO toma en consideración el valor presente del flujo de todos los gastos de CAPEX y OPEX.

Con relación a los costos de transporte de la información, cuando los volúmenes son importantes, la EC presenta ventajas sobre la CC, considerando las contrataciones requeridas de grandes capacidades, en general muy variables, para el transporte desde el borde a los Centros de Datos, y viceversa.

En cuanto al almacenamiento de datos, los costos de ambas opciones, la EC y la CC, deben ser evaluados considerando aquellos datos a ser almacenados en múltiples sitios próximos a los usuarios, o su almacenamiento centralizado y a mayor distancia. Habrá razones operativas para el acercamiento y la multiplicidad de sitios, y razones de costos a considerar. Por razones operativas hay datos que deben necesariamente estar en el borde para su procesamiento en tiempo real, o casi tiempo real, como los de la robotización, servicios médicos, etc., otros que son Grandes Datos y que sería costoso enviarlos a procesar lejos de la fuente, y otros que siendo indiferentes a su sitio de almacenamiento pueden ubicarse en un lado u otro por razones de escala. Esta evaluación es la más delicada porque incluye además hasta la valoración de la experiencia del cliente. Como se ve, la ubicación del almacenamiento de los datos está principalmente ligada a su procesamiento, o a su volumen.

Los costos de gestión y mantenimiento, que podrían verse aumentados en la EC considerando su despliegue distribuido, están siendo reducidos debido a las automatizaciones que se están aplicando, y que incluyen hasta la incorporación de los sistemas de mantenimiento que permiten la predominancia de la autorreparación. La progresiva reducción de la intervención de técnicos es acompañada por la reducción de costos.

Por otra parte, es importante considerar el costo de la seguridad y del cumplimiento de normas vigentes en el país. La principal razón del aumento de costos de la seguridad en la EC es que su carácter distribuido aumenta la superficie de ataque agregando dificultades para garantizar la seguridad. Debido precisamente a esta cuestión, las tecnologías de defensa frente a los ataques están evolucionando y reduciendo sus costos, debiendo preverse caídas importantes en los próximos años.

Como resumen inicial, si bien en la EC es necesario incurrir en un CAPEX mayor, el ahorro se genera con el tiempo al tener que incurrir en menores OPEX, mejorando por otra parte las prestaciones, o simplemente permitiendo cumplir con los requerimientos que no se alcanzan con la CC. Una comparación final de costos debe tomar en consideración que la solución óptima final es en general una combinación de computación en el borde y en la nube, bajo el formato de Nube Híbrida.

Considerando que la EC es una familia de tecnologías emergentes, impulsada por otras familias también emergentes como la IoT, la IA, la Machine Learning, entre otras muchas, siguiendo la evolución conocida de ellas, se esperan importantes reducciones de costos a medida que los avances y las economías de escala profundicen su impacto. Se espera que la EC sea progresivamente adoptada por organizaciones de menor tamaño, que a su vez vayan viendo en la EC la solución a los requerimientos de los clientes y usuarios.

Estamos en los inicios de esta evolución, lo que a su vez se ve en los avances de las empresas proveedoras de los servicios en América Latina y el Caribe, incluyendo desde los hiperescaladores hasta las pequeñas empresas, así como los del equipamiento necesario para redes de EC privadas propias. Los CSP están incursionando fuertemente en el nuevo modelo de negocio, derivado de la MEC, que es la EaaS (Edge as a Service), en que la computación en el borde se desarrolla en sitios de uso público y compartido pero que cumple con todos los requerimientos de la EC.

⁷⁵ Costo total de la propiedad. Total Cost of Ownership. Es la estimación financiera del costo total de un activo incluyendo los costos de capital y los gastos durante todo su ciclo de vida, incluyendo los períodos de inactividad.

B. Estimaciones de costos de Edge vs. Cloud

Comenzamos con un análisis general de la generación de los costos en las diferentes opciones informáticas, sin entrar en profundidad en las cuestiones técnicas relacionadas a los detalles de las variadas ofertas de servicios en la nube que cada proveedor realiza.

Se incluye en este análisis desde la versión del centro de procesamiento propio hasta las alternativas que se centran en los servicios prestados.

Los costos de construir y operar una infraestructura informática propia incluyen un conjunto de rubros entre los que se encuentran:

- CAPEX. Compra o arrendamiento con opción a compra del hardware de procesamiento y de comunicaciones, compra de licencias de software y los costos relativos al sitio de ubicación del equipamiento central.
- OPEX. Costos recurrentes de arrendamientos varios, de energía, de comunicaciones, y otros.
- OPEX. Operación y mantenimiento de toda la plataforma, incluyendo la sustitución o actualización de partes defectuosas o actualizaciones de software de base.
- Costos generales ocultos de sobre aprovisionamiento frente a eventuales necesidades de crecimiento, considerando los retardos propios de esta modalidad, o de la rigidez frente al desescalamiento.

Esta estructura de costos afecta negativamente la digitalización de las PyMEs en primer lugar, y significa una carga para empresas mayores, todas las que además deben enfrentar cambios permanentes y actualizaciones para adaptarse a ellos y al mercado.

Por esa razón, desde hace años se ha desarrollado una fuerte tendencia hacia la contratación de los recursos informáticos fuera de la empresa, y principalmente en los grandes Centros de Datos en la nube, la que últimamente se está enfocando en traer nuevamente los recursos a las organizaciones, o principalmente a los Centros de Datos en el borde, a veces en una coordinación de centros privados y públicos en la nube o en el borde, lo que se llama Nubes Híbridas.

Un paso que se ha dado inicialmente fue el de alojar equipamiento propio en Centros de Datos, bajo la modalidad de colocation, compartiendo todos los recursos necesarios para la operación: energía ininterrumpida, enfriamiento, seguridad, conectividad redundante, entre otros, y asumiendo la responsabilidad de su funcionamiento, mantenimiento y actualizaciones. En esta alternativa la organización debe incurrir en el CAPEX necesario para la compra del sistema de procesamiento y conexión, así como los mismos costos de software que en la instalación propia, su mantenimiento y actualización.

En esta alternativa los costos compartidos en el Centro de Datos, y donde se producen los ahorros, se suelen clasificar en varios rubros: espacio físico de piso o de unidades de rack para alojar el equipamiento, consumo de energía/enfriamiento, capacidad de los enlaces redundantes de conectividad a Internet, mano de obra de tareas varias (instalación, conexión, etc., según los casos), almacenamiento de respaldo de datos, salas de operación remota del equipamiento instalado, otros.

La Computación en la Nube (CC), en que ya no es necesario disponer de equipamiento propio, ha permitido a las organizaciones sacar su equipamiento fuera de su recinto, y reducir su CAPEX al compartir los recursos informáticos en base a tres modelos principales de servicios:

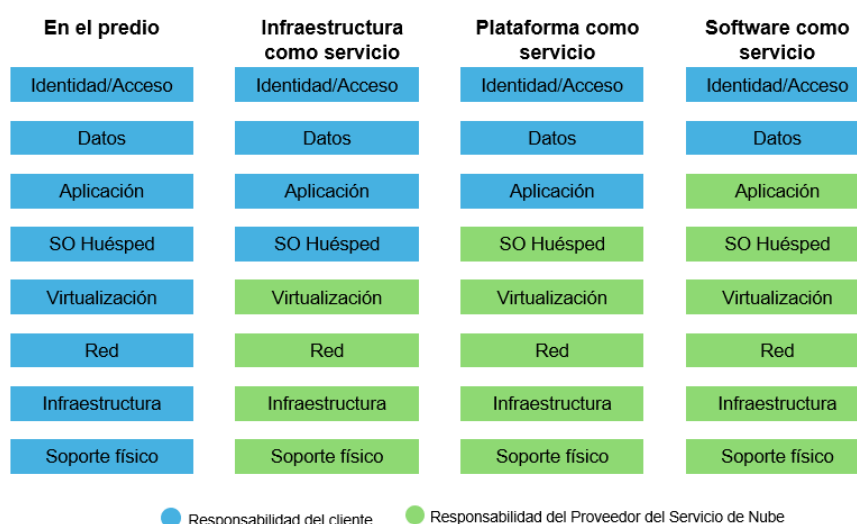
- Infraestructura como servicio (IaaS): Este modelo proporciona acceso a pedido a la infraestructura de nube, como la red de acceso, el almacenamiento, los servidores de procesamiento, la virtualización que da acceso a las máquinas virtuales y la capa física

(energía, piso, etc.). Esto sustituye la adquisición, operación y mantenimiento de la infraestructura propia. Este servicio convierte en OPEX al CAPEX en que se incurre tanto en la infraestructura propia como en su instalación en un Centro de Datos.

- Plataforma como servicio (PaaS): Este modelo ofrece una plataforma de procesamiento con todas las herramientas de infraestructura y software subyacentes (Run Time, Middleware y Sistema Operativo) necesarias para desarrollar, ejecutar y administrar aplicaciones.
- Software como servicio (SaaS): Éste es el modelo más profundo de la Computación en la Nube al ofrecer directamente las aplicaciones que proporciona y mantiene el proveedor de servicios, eliminando la necesidad de que los usuarios finales implementen software localmente.

La definición de esta estructura de capas, aunque no su concepto, la que permite visualizar los modelos, puede diferir de un proveedor a otro. Ilustrativamente se presenta la propuesta de Oracle⁷⁶.

Diagrama 4
Capas de informática en la nube



Fuente: Oracle Cloud Infrastructure.

El comienzo de la oferta de la Computación en la Nube, en sus diversos modelos basados principalmente en la estructura de capas indicada, ha permitido la reducción de costos de la digitalización. De todas formas, al igual que con la coubicación, en este caso también el precio final a pagar depende de múltiples factores que a su vez dependen fuertemente de las aplicaciones a correr. Pueden clasificarse en ciertas categorías, para las cuales los precios, a su vez, pueden diversificarse considerando los diferentes requerimientos:

- Constitución de las instancias de cómputo⁷⁷, cuyos precios dependerán de los requerimientos de memoria y de capacidad de procesamiento.
- Almacenamiento de datos, en que el vector de costos es el volumen en GB.

⁷⁶ <https://docs.oracle.com/es-ww/iaas/Content/cloud-adoption-framework/cloud-computing-layers.htm>.

⁷⁷ Una instancia de computación en la nube es un recurso de servidor que a veces se usa indistintamente como máquina virtual y que está definida por determinados parámetros de cómputo y memoria. Se usa la instancia para ejecutar cargas de trabajo.

- Costos de red, dependiente de la capacidad, de la redundancia y otros factores.
- Diferentes plataformas como servicio.
- Múltiples ofertas de *software* como servicio en que se incluye hasta la aplicación.

La estructura final de costos a incurrir es muy compleja y variable de acuerdo con los proveedores, por lo cual es casi imposible efectuar una comparación sin emplear un caso concreto. Por ejemplo, el costo puede abrirse en costo de reserva de capacidad en determinados horarios y costos por uso encima de la capacidad, o una combinación de ambos. La capacidad comprometida genera menores costos unitarios que la capacidad por uso, por lo que la combinación de ambas surge de estudios estadísticos de uso de las aplicaciones. Por ello es necesario que exista una gestión, propia de la organización, de servicios y costos de la nube a los efectos de optimizar los resultados, lo que es analizado en la sección “C. Gestión de costos de los servicios en la nube”.

En este complejo escenario, un comparativo interesante de costos es entre usar una red Edge propia de la organización (incluyendo el CAPEX y el OPEX, alcanzando toda la operación y mantenimiento), una Edge basada en los servicios de un proveedor de la Nube, u operar directamente en la Nube.

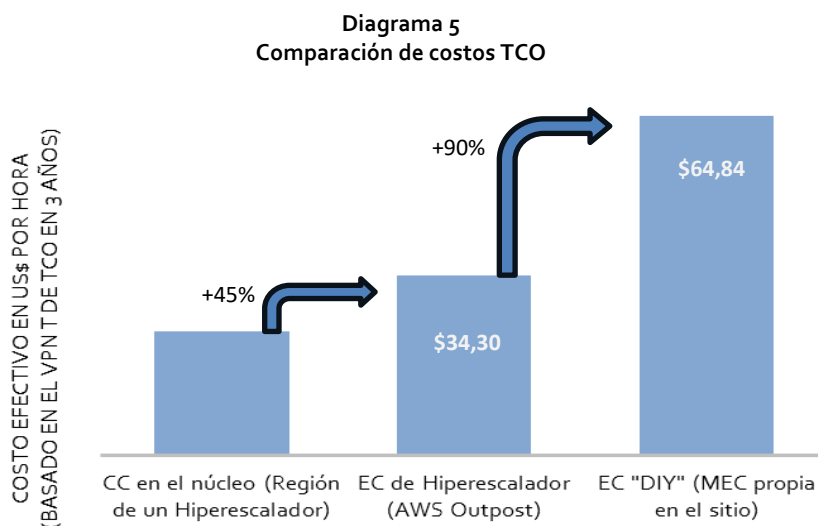
Un resumen de estos costos se observa en un documento de Edge Computing News⁷⁸.

El primer gran movimiento para la reducción de costos para el tratamiento de la información ha sido a través de los hiperescaladores que ofrecen la reducción del CAPEX y del OPEX de las organizaciones a través de la compartición de infraestructura, mecanismo exitoso en múltiples actividades. El movimiento hacia el borde implica necesariamente un paso hacia un entorno de operación más costoso pero que puede ser controlado satisfaciendo las necesidades operativas o comerciales que motivan la EC.

El camino hacia el borde puede ser hacia una infraestructura propia de la organización, o hacia las varias redes de borde de las empresas que están surgiendo para prestar este servicio, incluyendo a los hiperescaladores. Mobile Experts realizó un estudio con importantes proveedores de la nube, publicado en julio de 2020 por el citado Edge Computing News, comparando el TCO de la MEC⁷⁹ de una Industria con una gran carga de almacenamiento y computación, empleando tecnologías de producción avanzadas y alto consumo de los empleados, junto con el uso de 5G, para la alternativa de emplear una solución de nube en el borde de un hiperescalador, o la de desarrollar la llamada DIY (Do It Yourself), la que implica construir toda su infraestructura propia, en este caso combinada con una red 5G privada. La gráfica que sigue es ilustrativa de la comparación de TCO.

⁷⁸ <https://www.edgecomputing-news.com/2020/10/29/analysis-economics-of-edge-computing/>.

⁷⁹ MEC: Multi-Access Edge Computing o Mobile Edge Computing. Es una evolución de la EC que usa la movilidad, la nube y la computación en el borde.



Fuente: Edge Computing News.

Se observa que construir una red propia tiene un TCO a tres años de aproximadamente un 90% más que emplear servicios en el borde, en este caso provisto por AWS. A su vez esta última opción es aproximadamente del 45% más costosa que emplear los servicios de la nube, que de todas formas no es una solución viable para el cliente que requiere la computación en el borde por razones operativas o comerciales.

Independientemente del proveedor de Edge, considerando que la competencia tiende a uniformizar los precios, las diferencias de costos indicadas son similares.

Si el período de evaluación se extiende a más de tres años, puede empezar a mejorar el TCO de la opción de infraestructura propia, pudiendo llegar a ser inferior en algunos casos. Ésto está muy condicionado a que en la propia es necesario sobre invertir por razones imprevistas, existen dificultades y costos adicionales por las necesidades de actualización, limitaciones en el talento, entre otros múltiples aspectos.

El análisis de este comparativo simplemente ilustrativo, junto a los comentarios de la sección “1. CAPEX y OPEX en las diferentes modalidades de Nube vs. servicios prestados y calidad”, sirven para percibir la esencia de los aspectos económicos y técnicos de la EC. Además, muestra el papel importante de la EC desarrollada en las redes de los proveedores especializados, entre los que se encuentran los hiperescaladores y los CSP que desarrollan sus propios Centros de Datos, y así aprovechar las economías de escala y la compartición.

Vale destacar que algunas industrias como la petrolera, la financiera y otras similares por el requerimiento de seguridad de la información y porque sus requerimientos informáticos son muy especiales, con baja posibilidad de compartir, optan por llevar las principales funciones informáticas al borde propio (DIY) y hacen un uso complementario de la nube pública.

Más avanzada aún en el mejoramiento de los servicios y en la reducción de costos es la llamada arquitectura Serverless⁸⁰, sin servidor. Con esta nueva generación de la Nube no es necesario registrarse en el servidor virtual, ni disponer de una conexión directa a un servidor, sino simplemente correr la API funcional sin preocuparse de la infraestructura.

⁸⁰ El nombre “Serverless” no es totalmente correcto ya que se necesitan servidores, aunque se los llama así haciendo referencia a que el usuario no ve ni gestiona los recursos del servidor.

Permite reducir los costos creando funciones para propósitos simples y comandadas por eventos, las que generan costos solamente mientras se usan. Estas funciones son Sin Estado⁸¹ (Stateless) y efímeras, que no necesitan de servidores permanentes para correr las aplicaciones como son los casos de IaaS, PaaS o SaaS. Esto permite costos por uso en una forma más granular de acuerdo con las necesidades, evoluciones más simples desde aplicaciones heredadas y menos horas de desarrollo.

Un estudio de McKinsey Digital de 2020⁸², que la llama Nube 2.0, define esta nueva evolución tecnológica. Entiende que esta arquitectura es adecuada para las organizaciones que desean emprender el viaje digital, empezando con la CC clásica expandiéndose luego a la arquitectura "Serverless".

Esta arquitectura se desarrolla en tres etapas a partir de la oferta inicial de la nube clásica, a la que llama Cloud 1.0, y en la que se ofrecen soportes de infraestructura como servicio.

- i) Estas ofertas ya se han analizado: la infraestructura como servicio (IaaS), la plataforma como servicio (PaaS) y el software como servicio (SaaS). Estas son las bases para las próximas siguientes dos etapas en la evolución de la arquitectura sin servidor: backend como servicio y función como servicio.
- ii) El Backend como servicio (BaaS) es un conjunto de servicios basados en la nube, como almacenamiento en la nube, notificaciones push, autenticación y autorización, y servicios de ubicación que se proporcionan a los desarrolladores a través de API. Estos servicios están destinados a los desarrolladores para que los integren y utilicen.
- iii) Finalmente, la Función como servicio (FaaS) se refiere a funciones basadas en eventos (lógica de negocio) que se ejecutan en contenedores efímeros (muy flexibles en los cuales su capacidad facturable se configura bajo demanda, y también según sea necesario). Es una profundización del concepto Serverless. Los proveedores de nube administran FaaS, y se pueden acceder a través de API. FaaS es un nivel de aumento en la abstracción por encima de PaaS / contenedorización y un nivel de disminución en la abstracción por debajo de SaaS. (Nota: También se le llamaba fPaaS).

Este modelo de computación en la nube es provisto por los mismos proveedores de servicios en la nube como son los casos de AWS Lambda, Google Cloud Functions, Azure Functions o IBM OpenWhisk.

Gartner estimaba que la mitad de las empresas a nivel global estarían usándola en 2025.

Un ejemplo de este modelo es por ejemplo el caso de una importante fábrica de refrescos que usa la Serverless Cloud en sus máquinas dispensadoras. Cada vez que se compra una botella, la pasarela de pago realiza una llamada a un AWS API Gateway y activa una función de AWS Lambda para completar la transacción. De esta forma el pago por transacción le ha permitido reducir los costos.

Este modelo de procesamiento Serverless puede aplicarse en la computación en el borde agregando las ventajas vistas para este modelo a las propias de la computación en el borde.

En resumen, las particularidades de la computación Serverless son principalmente las siguientes:

- La arquitectura Serverless permite que el sistema del software se constituya con múltiples funciones que pueden correr procesos efímeros e independientes entre ellas. La comunicación entre las funciones es mediante API. De esta manera la introducción de modificaciones se puede realizar en forma más simple sin tener que actualizar la aplicación completa.

⁸¹ Por lo general, cuando se comienza a trabajar en un mundo de microservicios y contenedores, se suele empezar con aplicaciones sin estado, las que no exigen la persistencia del almacenaje de los datos, y el servidor actúa sin considerar peticiones anteriores. Esto facilita que diferentes peticiones sean tratadas por diferentes servidores.

⁸² <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/tech-forward/cloud-20-serverless-architecture-and-the-next-wave-of-enterprise-offerings>.

- Provee capacidades adicionales de ambientes de desarrollo y testeo sin afectar la provisión del servicio o la aplicación.
- Las instancias que habilitan las funciones son efímeras y desaparecen luego de corridas o de que se produzca una falla. Es un sistema tolerante a fallas debido a su diseño simple.
- El uso de los recursos es muy elástico, lo que permite optimizar costos al pagar solamente durante el uso de las funciones.
- Es la solución adecuada cuando el software se corre en forma aleatoria en el tiempo, sujeto a acontecimientos externos, no es predecible la intensidad o frecuencia de su uso, puede requerir escalamiento y desescalamiento frecuentes, puede trabajar sin estado en sucesivas invocaciones, entre otros.
- Es esencial que se adopte una cultura DevOps.
- Es muy adecuado para los bajos retardos y su soporte en el borde.
- La IoT y la Machine Learning pueden soportarse muy bien en la Stateless Computing debido a la reducción de los tiempos de respuesta.

El costo final para la computación Serverless se constituye básicamente en tres ítems:

- Número de invocaciones a las funciones.
- Complejidad y duración del trabajo de procesamiento.
- Memoria requerida para el proceso.

C. Gestión de costos de los servicios en la nube

Tanto la CC como la EC han reducido los costos del procesamiento y almacenamiento de la información, que son la base de la digitalización en general, y que la favorecen para las PyMEs que no tienen condiciones para disponer de una infraestructura propia. En un primer momento la CC fue la principal herramienta de compartición de infraestructura (aparte de la ubicación en los Centros de Datos), con la consecuente reducción de costos, y la eficiente evolución de los costos hacia una flexibilidad mayor que permite a los clientes elegir entre el pago inicial y el pago recurrente. Así aparecen los términos XaaS, o sea varios tipos de activos ofrecidos como servicios: capacidad de almacenamiento, software, otros.

En 2020, surgía un incremento del interés por la CC a partir de la necesidad de reducir costos fijos por la pandemia, tal que en una encuesta de Deloitte⁸³ el 80% de los consultados respondía que la Computación en la Nube sería una de las dos tecnologías más relevantes, junto a la ciberseguridad. La migración a la nube estaba motivada por la búsqueda de un modelo elástico basado en el consumo.

Luego, un informe de Deloitte⁸⁴ de diciembre de 2020, desarrollado por dos de sus expertos, indicaba lo siguiente.

“Sin embargo, a medida que aumenta el gasto en la nube sin poner los controles y la gobernanza necesarios en torno al consumo, el gasto en desperdicio se está convirtiendo en un problema importante. Como informa Flexera, tales desperdicios pueden representar hasta el 30% del gasto total

⁸³ La encuesta incluyó a 1.089 ejecutivos directamente involucrados en la transformación de sus organizaciones, y se llevó a cabo en cuatro regiones del mundo en junio-julio de 2020. <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/process-and-operations/us-save-to-thrive.pdf>.

⁸⁴ <https://www2.deloitte.com/us/en/blog/deloitte-on-cloud-blog/2020/maximizing-value-from-cloud-investments.html>.

en TI, y los ejecutivos citan evitar el desperdicio como un desafío importante⁸⁵. Para lograr ahorros de costos de manera efectiva y ayudar a hacer realidad la promesa transformadora de la nube, los CIO deben planificar estratégicamente nuevas migraciones a la nube, optimizar el gasto en la nube y monitorear y gobernar de cerca sus implementaciones en curso en la nube”.

Visto desde otro punto de vista, no ya como porcentaje de gasto en TI sino como porcentaje de sobre gasto específico en la nube, este mismo informe indica que según un estudio de Gartner⁸⁶, las organizaciones que no optimizan el gasto en la nube pueden sobre gastar alrededor del 70% o más. Para evitar el gasto excesivo e innecesario, es esencial comprender el gasto en la nube y optimizarlo tanto como sea posible. Esa optimización comienza con el control del desperdicio de recursos.

Ese desperdicio tiene varios orígenes que deben ser controlados permanentemente por personal de la organización.

- i) Puede haber recursos contratados que dejan de usarse en el medio de sistemas grandes y que por diversas razones no se dan de baja generando gastos inútiles. Estos recursos se suelen llamar “huérfanos”.
- ii) Debido a la falta de conocimiento, o de atención, de los mecanismos flexibles de contratación con los proveedores de la nube, se aplican los mismos procedimientos que para equipamientos propios, sobrecontratando por si son necesarias expansiones, incurriendo así a costos excesivos.
- iii) Los costos de licencia de software y las suscripciones redundantes a veces no se administran de manera óptima, lo que también pueden generar desperdicio.
- iv) La carencia de talento actualizado y adecuado a estas nuevas modalidades de TI operando en la nube, o en forma mixta local y la nube, puede producir costos excesivos por deficiencias de contratación o de diseño ineficiente de los sistemas.
- v) Otros factores derivados de la carencia de talento pueden contribuir a un mayor gasto incluyen la falta de visibilidad del uso de los recursos y el conocimiento de los costos de los componentes de la nube; arquitectura, marcos y mecanismos de gobernanza aislados y no orientados al trabajo con la nube; precios complejos para suscripciones a la nube; e inclusive herramientas inadecuadas de los proveedores de servicios en la nube.

Por esta razón las organizaciones deben establecer una política de control de costos en la nube y disponer de una persona o sector, dependiendo del tamaño, responsable de la optimización del uso de estos recursos, (ingeniería de la Nube, o similares). FinOps es la disciplina que se está desarrollando, y que ya se ha mencionado, en cuanto a la gestión financiera de la nube a través del seguimiento y optimización del costo de la nube. Se incorpora a los equipos de DevOps a través de profesionales especializados orientados principalmente a la optimización de la ingeniería de la nube.

Según este estudio de Gartner, para optimizar los costos es necesario definir procedimientos de control de gastos, incluyendo el uso de tecnologías de supervisión y control. Entre estas medidas que apuntan precisamente a esos desperdicios de los recursos, se encuentran las siguientes:

- Cerrar instancias no utilizadas identificando las instancias en ejecución que no están en buen estado o que no se utilizan.
- Reemplazar el almacenamiento ineficientemente costoso evaluando los requisitos para disponer de los parámetros adecuados como el tipo de almacenamiento requerido, las opciones de redundancia o la frecuencia de acceso para seleccionar el almacenamiento correcto.

⁸⁵ <https://info.flexera.com/SLO-CM-REPORT-State-of-the-Cloud-2020>.

⁸⁶ <https://www.gartner.com/document/3901768?ref=solrAll&refval=226009619&qid=e4ed4e83cc3b8edb80d6d1c3d6>.

- Ajustar de tamaño de instancias y los recursos adecuados de las máquinas virtuales para determinar el tamaño correcto de la instancia para optimizar el costo y el rendimiento y reducir el sobre aprovisionamiento.
- Automatizar la optimización utilizando la automatización para crear alertas, informar sobre el desperdicio y activar acciones automatizadas para resolver ineficiencias.
- Migrar a regiones o tipos de instancias de menor costo: cambiar las series de instancias de mayor costo a series de instancias anteriores y de menor costo.

Las herramientas para supervisar y controlar el desperdicio incluyen el escalado automático y las prestaciones de autoservicio. Así las empresas pueden implementar aplicaciones que generen información sobre el uso, para administrar y optimizar el gasto en la nube. Estas aplicaciones a menudo tienen un precio como porcentaje del gasto en la nube (a menudo del 1% al 3%). En estos procedimientos se debe cuidar que no se altere la productividad ni que se inserten procesos de aprobación manual que solo ralentizarán las operaciones.

Se hace notar que los proveedores de servicios en la nube (incluyendo la EC pública) presentan estructuras de precios a partir de las cuales elaborar las diferentes instancias ajustables a las necesidades en cuanto a CPU, memoria, almacenamiento y redes de transporte. Las instancias son a su vez escalables considerando los casos de variaciones en cuanto a requerimientos. A su vez, los diferentes proveedores presentan ofertas con diferentes parametrizaciones. Por ejemplo, el almacenamiento puede contratarse con un precio fijo mensual por una cantidad fija de unidades, o, para prever picos de uso, agregando un costo variable adicional si se excede la capacidad.

Un estudio de 2021 de Flexera⁸⁷ presenta también diversas consideraciones sobre la importancia de la gestión de los activos de TI (ITAM. IT Assets Management), tomando en cuenta que existe un porcentaje alto de empresas que, habiendo migrado a la nube, no alcanzan total satisfacción al no tener una gestión centralizada.

Se destacan los posibles incumplimientos de las reglas de licenciamiento del software que se usa en el sitio cuando además se trabaja en la nube, reglas que varían según el proveedor. Tampoco son las mismas reglas con los hiperescaladores en que lo que es válido para uno puede no ser válido para otro. Por ello es necesario tener un control estricto sobre estos cumplimientos. Tan importante es el hecho de que, al no tener una visión clara de los activos involucrados, el sistema se encuentra más expuesto a riesgos. Estos son algunos de los temas de los que es importante que se ocupen los responsables de ITAM.

D. Resumen comparativo del impacto de la EC y su relación con la CC

A continuación, se presenta un cuadro cualitativo comparativo de la EC y la CC:

Cuadro 1
Comparación cualitativa de EC vs. CC

Concepto	<i>Edge Computing (EC)</i>	<i>Cloud Computing (CC)</i>
Soporte de tecnologías emergentes apoyadas en ambas modalidades por razones operativas o económicas.	Ambas son complementarias y pueden ser suministradas por los mismos Operadores de Servicios de Nube, incluyendo los hiperescaladores, operadores menores, nubes propias o de terceros provistas por los CSP, entre otras alternativas. EC puede ser privada con tendencia clara a moverse total o parcialmente a la nube en la configuración Híbrida.	

⁸⁷ <https://info.flexera.com/CM-WP-Cloudy-IT-Visibility>.

Concepto	Edge Computing (EC)	Cloud Computing (CC)
Arquitectura	Es un modelo de computación en general distribuida y próxima al origen o destino de los datos, a donde se trasladan las aplicaciones y servicios que requieren de las particulares prestaciones de la EC: bajo retardo de respuesta, menores costos de transporte de datos, cumplimiento de normas, mejoramiento de la seguridad, entre otros. Eficiencia menor que en la CC y mayor que en los centros residentes.	Prestación de aplicaciones y servicios de TI en grandes Centros ubicados en sitios estratégicos (energía limpia, disponibilidad de agua de enfriamiento, otros). Eficiencia.
Casos de uso. Ejemplos.	Servicios solamente soportados por la EC. Realidad Aumentada y Realidad Virtual (AR/VR), vehículos autónomos, Industria 4.0, medicina avanzada, juegos en la nube, Metaverso, otros.	
Costos totales. TCO (Costo Total de Propiedad).	Debido a su menor escala es más costoso el procesamiento y el almacenamiento, así como las aplicaciones y servicios. Permite casos de uso que no permite la CC, y/o se obtiene una reducción importante de costos de transmisión. La 5G juega un importante papel en el borde. Es flexible en la contratación, como la CC.	La CC en los grandes Centros de Datos tiene TCO menor debido al alto grado de compartición de infraestructura, aplicaciones y servicios, así como la flexibilización de la contratación.
5G y Redes definidas por software (SDN).	Habilitan aplicaciones y servicios más resilientes, automatizados y flexibles (SDN), o que serían imposibles por las exigencias de velocidad de datos y bajo retardo (5G) lo que no provee la CC.	No se ve impactada.
Retardos de respuesta y volumen de la información a transferir al Centro.	Su proximidad al usuario reduce drásticamente la distancia hasta el Centro, lo cual tiene dos ventajas ineludibles para un creciente número de casos de uso. La reducción del retardo por distancia en la respuesta es la habilitante de las aplicaciones en tiempo real, y permite una mayor productividad en los usuarios. La reducción de la distancia reduce los costos de los enlaces de banda ancha hasta el Centro para los casos de Grandes Datos.	No aplica.
Protección de datos personales.	Las redes de EC siempre se desarrollan dentro del territorio del país, y la organización puede y debe garantizar el cumplimiento de las normas nacionales que exigen medidas de seguridad contra diversos riesgos.	Para esta modalidad en los casos de transporte transfronterizo, a Centros procesadores ubicados en el extranjero o cuando el respaldo del Centro de destino se encuentra en el extranjero, se requiere que la transferencia se realice solo si existe un nivel de protección apropiado de acuerdo con la regulación del país del controlador de los datos.
Operación sin conexión permanente o de calidad a Internet.	La red de EC dispuesta en un sitio lejano, o en general con baja disponibilidad de Internet permite el procesamiento de la información y las respuestas sin necesidad de estar conectada a Internet, o tener alta calidad de comunicación. Cuando dispone de Internet se realiza la transferencia de datos preprocesados, si fuera necesario.	No es adecuada para estos casos.
Ciberseguridad. Consideraciones generales.	El trabajo de la EC en contacto con los dispositivos, generalmente de baja capacidad de procesamiento, y por tanto con protocolos de seguridad simplificados, así como la transmisión directa entre nodos, reduce las condiciones de seguridad. No obstante, como al mismo tiempo los datos son procesados localmente, y solo viajan datos preprocesados, existe un balanceo en cuanto a seguridad, al reducirse los riesgos en los transportes de larga distancia. El empleo de SASE (Secure Access Service Edge) provee una seguridad fuerte y flexible en el borde.	En este caso la seguridad se asegura con protocolos de encriptación de extremo a extremo y las políticas y controles más fuertes que en el borde, aunque los datos deben viajar largas distancias.
PyMEs	Las PyMEs requieren exclusivamente la EC, por las mismas razones que las grandes, cuando deben procesar Grandes Datos o cuando necesitan muy bajos retardos de respuesta. Adicionalmente debido a su menor escala de compra obtienen beneficios por las economías de alcance ^a que provee la MEC, al ofrecer el empaquetamiento de los servicios de comunicaciones con los de EC. Es el modelo principal para PyMEs en cuanto a la optimización de los costos ya que además generalmente no se encuentran en condiciones de emplear un Centro propio.	

Fuente: Elaboración propia con base en múltiples fuentes.

^a La economía de escala surge de los ahorros unitarios cuando se aumenta la producción de un determinado producto o servicio. La economía de alcance (economy of scope) se obtiene cuando una empresa reduce los costos comunes y conjuntos atribuibles a un servicio o producto si aumenta la variedad de servicios o productos. Las economías de escala se observan a través del decrecimiento de la curva de los costos medios. Las economías de alcance se observan comparando el costo de la producción conjunta en una sola empresa vs. la producción individual de los mismos productos y servicios en empresas separadas.

IV. Tendencias en el despliegue previsto de la EC en Latinoamérica, el Caribe y en otras regiones

A. América Latina

1. Aspectos generales de la EC en la región

Considerando que el despliegue de la Computación en el Borde en el mundo es una evolución en curso de la Computación en la Nube, impulsada por múltiples aplicaciones, es de interés considerar los avances de ambas modalidades de computación.

A medida que aumenta la demanda por una mayor capacidad de almacenamiento y procesamiento de datos, así como la necesidad de reducir fuertemente el retardo para poder tomar acciones en tiempo real, se vuelve imprescindible acercar la capacidad de procesamiento al usuario, sea ésto a través de un solo sitio o multi sitio. Se ha visto que los principales avances se encuentran en la provisión de nodos de procesamiento y almacenamiento de la información constituidos en redes especializadas, como forma de aumentar el poder computacional en paralelo y a demanda, para alimentar múltiples sitios que cumplan las condiciones típicas de la EC. Se suma a estos requerimientos la reducción de costos de transporte de información buscando un óptimo de TCO⁸⁸.

Sin cumplir estas condiciones las aplicaciones críticas de la IoT, así como los procesos más avanzados de la digitalización, serían inviables.

América Latina muestra un interés en el empleo de la EC para mejorar sus operaciones, así como los procesos en verticales como las industrias en general, la logística, el comercio, servicios públicos, etc.

⁸⁸ Costo total de la propiedad. Total Cost of Ownership. Es la estimación financiera del costo total de un activo incluyendo los costos de capital y los gastos durante todo su ciclo de vida, incluyendo los períodos de inactividad.

La atención prestada y los servicios brindados por la digitalización durante la pandemia han producido un gran interés en la inversión en las tecnologías emergentes (soportadas principalmente por la EC y la 5G), generando cambios en la gestión y los modelos de negocio impulsados por la inversión que asegure mejoras en la productividad y la resiliencia.

A los efectos de observar la situación en cuanto a la preparación general para el desarrollo de la EC en América Latina, se analiza la preparación de los países para las tecnologías de vanguardia. Se observa que a pesar de las condiciones socio económicas desafiantes algunos países muestran aptitudes destacables en cuanto a estas tecnologías. El cuadro 2 muestra los índices de los países de América Latina y el Caribe, de acuerdo con los criterios de la UNCTAC para el año 2022.

Cuadro 2
Índice de preparación para las tecnologías de vanguardia en América Latina y el Caribe en 2022,
por país, publicado en 2023

América Latina: preparación para las tecnologías de vanguardia por país 2022				
	Alto	Medio-alto	Medio-bajo	Bajo
Brasil	0,71			
Chile		0,65		
Barbados		0,62		
Costa Rica		0,61		
México		0,58		
Argentina		0,57		
Uruguay		0,57		
Panamá		0,54		
Colombia		0,54		
Bahamas			0,50	
Perú			0,49	
Trinidad y Tobago			0,47	
Ecuador			0,43	
República Dominicana			0,43	
Belize			0,43	
Jamaica			0,42	
Santa Lucía			0,41	
Surinam			0,40	
Paraguay			0,40	
San Vicente y las Granadinas			0,39	
Bolivia (Estado Plurinacional de)			0,38	
El Salvador			0,37	
Guyana			0,35	
Guatemala			0,34	
Venezuela (República Bolivariana de)				0,31
Honduras				0,30
Nicaragua				0,29
Haití				0,15

Fuente: UNCTAD (<https://unctad.org/publication/technology-and-innovation-report-2023>).

Este análisis de la situación general frente a las tecnologías de vanguardia permite ubicar en sus justos términos los desarrollos que se están produciendo en la EC en América Latina, y que se analizan en esta sección, los que pueden estar produciéndose principalmente en los países con mayores índices y en los sectores económicos más dominantes. No obstante, la EC es la más extendida de las tecnologías de vanguardia debido a que es aplicable tanto a las empresas grandes como a las PyMEs. Basta observar que las PyMEs pueden necesitar procesamiento y almacenamiento de la información con alta velocidad de respuesta, escalable y otras condiciones típicas de EC, por lo que son potenciales usuarios de ella. No es lo mismo para otras tecnologías de vanguardia, como la IA, la robotización, la IoT, la Machine Learning, entre otras que son solamente aplicables a empresas o instituciones de cierto gran porte.

En un estudio reciente⁸⁹ de Harvard Business Review Analytic Services (HBRAS), de noviembre de 2022, y enfocado exclusivamente en las tendencias que se visualizan en la región, se observan conclusiones en cuanto a la experimentación e innovación de la infraestructura digital. Es la encuesta, dirigida a 316 ejecutivos de Argentina, Brasil, Chile, Colombia, México y Perú, todos ellos países grandes y que se ubican en los primeros lugares del Índice de preparación para las tecnologías de vanguardia, HBRAS ha llegado a estos resultados.

- i) El 90% considera que las tecnologías emergentes, en general, colaboran con el éxito en el futuro, un 87% que las aplica hace tiempo entiende que han sido útiles en los últimos dos años, mientras que el 88% considera que ha sido rentable la inversión, principalmente con mejoras en la reputación de marca, el incremento de ingresos y de rentabilidad, o la retención de clientes.
- ii) También ese 90% considera que aumentará la inversión en las tecnologías emergentes en los próximos 12 meses.
- iii) En cuanto a la observación hacia el futuro, la percepción de que en el corto plazo se observe la rentabilidad de la inversión surge de que el 73% espera obtener un retorno de su inversión en tecnología emergente en un plazo de dos años, y el 36 % dice que tomará solo un año o menos, lo que indica altos niveles de confianza.
- iv) La computación de borde (edge computing), incluyendo la infraestructura propia, es la tecnología emergente más utilizada por organizaciones de la región, con una adopción del 65%. Le siguen la biometría (63%), la robótica avanzada 60%, la realidad virtual y aumentada (56%), 5G (55%) y la inteligencia artificial (54%). De todos estos datos resulta que la Edge Computing, junto a la 5G, tendrán motores importantes para su desarrollo considerando que son el soporte imprescindible de todas las demás innovaciones.
- v) El incremento de la productividad es el principal objetivo apuntado por la mitad de los encuestados.
- vi) Se destaca que crece el involucramiento de los directivos de más alto rango, desde el Presidente, y ya alcanza casi la mitad de los casos.

Este posicionamiento de los directivos de las empresas grandes viene acompañado de una tecnología imprescindible como es la 5G que está desplegándose en América Latina. De esta manera los CIO se encuentran, por un lado, en un ambiente de aumento importante de dispositivos y sistemas requeridos para la producción o la gestión, y por otro con las exigencias cada vez mayores del área comercial y de los consumidores.

Se observa así que la EC se verá principalmente fortalecida en las empresas grandes en el futuro próximo en América Latina. Este impulso necesariamente creará el ecosistema para que los avances se observen también en las empresas de menor tamaño, las que de por sí ya están requiriendo servicios

⁸⁹ <https://cl.nttdata.com/newsfolder/latinoamerica-el-go-de-las-empresas-aumentara-su-inversion-en-tecnologias-emergentes>.

que deberían prestarse en la EC. Y junto a este impulso las empresas deberán preparar sus cambios de infraestructura y de modelos de gestión, dedicando esfuerzos importantes tanto financieros como humanos. En medio de esta evolución hacia la EC es necesario prestar gran atención a los temas de seguridad, integridad y resiliencia de las redes de datos y de los centros de procesamiento.

Se entiende que esta aceleración de la EC en LAC se debe en buena medida a que la pandemia obligó a adoptar mecanismos para acelerar la digitalización y así mantener e impulsar los negocios.

Según un análisis de IDC Research's Edge View 2022 se espera para los gastos en Edge una CAGR del 18,1% en América Latina en el período 2021-2026.

Se distinguen en LAC algunos verticales que concentran el gasto en EC. En el año 2020, según IDC, el 57% del valor total del mercado se concentró en Manufactura, Banca, Minorista, Telecomunicaciones y Seguros. Se observan los pagos inteligentes en Banca y Minorista, así como la gestión de flotas y supervisión de cargas.

Se destaca que el avance en estas aplicaciones tecnológicas implica importantes cambios culturales en toda la trama organizativa de las empresas, un tema importante a ser considerado en LAC tanto desde el punto de vista de la gestión como de la formación del capital humano. Entre ellos:

- i) Apertura al aprendizaje a través de la asunción de riesgos en el sentido de la digitalización acelerada.
- ii) Transformar la cultura en el sentido de agregar dinamismo en las respuestas a clientes internos y externos, a través del empleo de tecnologías digitales, entre ellas la Inteligencia Artificial que se avizora como la de mayor impacto. En este caso su uso debería estar restringido a ser considerada solamente como una herramienta para el uso del plantel de la empresa, que acelere los procesos pero que los equipos de cómputo que la usan no tomen directamente decisiones finales. La Inteligencia Artificial es uno de los principales motores del uso de la EC por el intenso transporte de datos, en general los bajos retardos y la requerida potencia de procesamiento, lo cual podría hacerla inviable si es alojada en la nube.
- iii) Incorporación de tecnología a las reuniones, permitiendo un mayor contacto de calidad y una reducción de costos.
- iv) Empleo de los Grandes Datos para actuar prospectivamente en la planificación, sustituyendo los procedimientos actuales más vinculados y basados en el estado actual. El Machine Learning es esencial en estos procesos, y es un fuerte cliente de la EC.

Estos cambios profundos implican inversiones y recursos humanos que permitan y orienten los cambios, no solamente técnicos sino también organizacionales. El 73% de los encuestados por HBRAS dice que existe escasez de talento, y que los déficits de habilidades están impidiendo la adopción efectiva de las tecnologías emergentes. Este punto es de destacar a los efectos de la formulación de las políticas educativas formales y de formación permanente.

El posicionamiento de las organizaciones en LAC, que entienden y justifican el esfuerzo por las tecnologías emergentes, las que a su vez se apoyan o colaboran con la EC, surge como una gran tendencia según la encuesta analizada de la HBR, y es analizado por el Profesor de Estrategia y Director Académico del Índice de Competitividad por el Talento Global del INSEAD⁹⁰, Felipe Monteiro; "Impulsadas por una sólida infraestructura digital y el comportamiento del consumidor, las compañías latinoamericanas son conocidas por su "disposición a experimentar, una predisposición a probar cosas nuevas y deseos de ser los primeros en adoptarlas".

⁹⁰ <https://www.insead.edu/>.

Por ejemplo, el banco Itaú Unibanco, según Rafael Heringer⁹¹, está incursionando en la computación avanzada y sobre ella operar con algoritmos de Aprendizaje Automático para mejorar la precisión de los modelos utilizados actualmente para predecir la rotación de clientes, y capacitar un modelo para su retención, así como para la optimización de sus inversiones y el cumplimiento de condiciones regulatorias.

En el caso de LATAM Airlines, se emplea la Inteligencia Artificial, que se soporta principalmente en EC, para estimar con precisión el peso de un vuelo a los efectos de calcular también con precisión la cantidad de combustible y ahorrar en el consumo, según Juliana Ríos, directora de información de LATAM Airlines en Chile. El resultado, no solo es una mayor eficiencia operativa y un ahorro de costos para la aerolínea, sino también son “beneficios ambientales” que se acumulan a partir de un menor consumo de combustible. También se emplea esta plataforma para informar la ruta óptima a los pilotos y despachadores, con más reducción de costos, de acuerdo con la acumulación de inteligencia aplicada a los Grandes Datos acumulados.

Otros casos agregan ejemplos de cómo la EC, en el mundo real de LAC, es el motor de aplicaciones disruptivas en múltiples verticales.

2. Despliegues previstos o en fase de desarrollo en LAC

En esta sección se analizan específicamente casos de despliegue de EC en LAC, en que se considera el 2023 como el inicio de los proyectos de despliegue de redes privadas de 5G y la EC. Este análisis se puede complementar con las informaciones relativas a Latinoamérica contenidas en la sección “A. Descripción de la estructura conceptual y física, y funciones de la EC.”, donde se analiza más en detalle la estructura, y también el despliegue territorial de los grandes prestadores de servicios en el Borde, como Google Cloud, Microsoft (Azure), Amazon (AWS) e IBM Cloud.

Si bien los operadores de telecomunicaciones están en la base estructural de los despliegues de la EC, lo que les permitiría aumentar su cartera de productos y aumentar su participación en el mercado de las TIC, como lo están haciendo con la IoT, se encuentran con que es una familia tecnológica en que los operadores de la CC se encuentran trabajando y ganando experiencia y base de clientes, desde hace muchos años. Por ello existe una tendencia importante, salvo excepciones, a la asociación de ambos para disponer de ofertas conjuntas a sus clientes. El ejemplo más potente ha sido el de Telefonía Tech con Amazon – AWS que se ha descrito.

Un informe⁹² de Frost & Sullivan de mayo de 2022 se orienta a los hiperescaladores y su estrategia en la región.

En una encuesta citada, de la misma F&S de 2021, alrededor del 52% de las empresas latinoamericanas, desde diversos sectores económicos, habían adoptado la política de emplear Infraestructura como Servicio (IaaS) para algunas de las aplicaciones, cargas de trabajo e infraestructura. Estimaba F&S que otro 36% adoptaría IaaS hasta el 2023, llegando así al 88%.

Destaca que, en el 2020, como consecuencia de la pandemia, las empresas de Latinoamérica invirtieron un 50,1% más en nuevas tecnologías que en el mantenimiento y actualización de las de TI legadas. En ese momento fue necesario mantener el flujo de relaciones con empleados y clientes empleando más intensamente la digitalización que ya estaba evolucionando, resultando en una aceleración percibida por todos los analistas en el momento actual.

⁹¹ Rafael Heringer, Director de Tecnologías de Itaú Unibanco.

⁹² Amazon Web Services en América Latina y su Estrategia de Crecimiento: todavía es Día 1. <https://www.frost.com/frost-perspectives/aws-en-america-latina-crecimiento/>.

F&S estima una CAGR del 40,1% en el período 2021 a 2028, un valor más alto que en otras regiones, incluyendo nubes públicas y privadas alojadas. Las nubes privadas alojadas se encuentran en las instalaciones o Centros de Datos de un proveedor en la nube, pero todo el equipamiento está asignado solamente al cliente. En este crecimiento de un sistema productivo aceleradamente muy conectado, por ejemplo, por la difusión de la 5G, entra la EC que es esencial para la mayoría de las actividades por sus menores costos, baja latencia, altas velocidades de transmisión, resiliencia, seguridad, entre otras especificaciones.

También, las predicciones de F&S aparte de lo indicado para la IaaS, indican que a partir de que en 2021 el 36% de las empresas de LAC ya habían comenzado actividades de EC, mientras que otro 45% las planificaba para 2022-2023, llega a que el 81% de las empresas utilicen EC para fines del 2023.

Uno de los ejemplos de uso específico de acercamiento al Borde de los hiperescaladores, aunque no precisamente en cuanto a la estructura principal de la EC, es el establecimiento de CloudFront Edge⁹³ de AWS en diversas ciudades de América Latina, como son: Rio de Janeiro, San Pablo, Fortaleza, Bogotá, Buenos Aires, Santiago, Lima y Querétaro, algunas ya operativas, y otras anunciadas, en febrero de 2022. De esta forma se logra la latencia baja y menores costos de transportes en esos centros estratégicos, facilitando el trabajo de desarrolladores, entregando contenidos dinámicos de alta definición para las áreas comerciales o de entretenimientos, entre otros. Los nodos de CloudFront hacen parte de una zona dentro de las regiones AWS, aunque la disponibilidad de los servicios depende de la región. El caché regional para estas Zonas de Sudamérica se encuentra en Sao Paulo.

Aquí surge un aspecto importante de la EC, y es que una de las razones para establecer estos CloudFront es permitir la residencia nacional de los datos de acuerdo con legislaciones vigentes en diversos países.

Esto es un ejemplo del fenómeno que se está dando en América Latina en cuanto a que varios proveedores han desplegado redes de varios Centros de Datos para satisfacer la demanda creciente.

Según BN Americas⁹⁴, en enero de 2023 Edge UOL, integrante del Grupo UOL, como proveedores de infraestructura y de orquestación de las partes del negocio del cliente, ha considerado los primeros Outposts⁹⁵ ("puestos de avanzada") de AWS en América Latina para proveer EC en la Región, en la que ya tiene clientes para servicios gestionados. Sin embargo, por el momento su foco de inversión en Outposts es Brasil donde anunció en mayo de 2023 la implementación en Sao Paulo de un Centro de Soluciones para el Cliente, donde testear y homologar servicios de EC con AWS Outposts.

Embratel, subsidiaria de Claro, está haciendo foco en el desarrollo conjunto de la EC y la 5G, considerando que las demandas de latencia, velocidad de datos, cantidad de dispositivos, nuevas aplicaciones como la automatización y costos, provocarán la migración generalizada del almacenamiento y el procesamiento a una red distribuida y próxima a los usuarios, o sea a la EC. Sus autoridades entienden precisamente que la 5G con la EC permitirán la habilitación de las tecnologías emergentes que se describieron en la sección 1. "América Latina en general" en el mercado corporativo. Su objetivo es desarrollar los Data Centers en el borde para el sector corporativo, junto a la 5G.

⁹³ Amazon CloudFront es un servicio de red de entrega de contenido (CDN) creado para ofrecer un alto rendimiento, seguridad y comodidad para los desarrolladores. Está constituida por 450 puntos de presencia y 13 caches de regiones de borde. <https://aws.amazon.com/es/cloudfront/features/?nc=sn&loc=z&whats-new-cloudfront.sort-by=item.additionalFields.postDateTime&whats-new-cloudfront.sort-order=desc>.

⁹⁴ 19 de abril de 2023. <https://www.bnamericas.com/en/news/brazils-edge-uol-advances-with-private-network-projects>.

⁹⁵ Los Outposts, como ya se vio en la primera sección, incluyen recursos informáticos y de almacenamiento que pueden ejecutar servicios nativos de AWS y se administran mediante las mismas API que impulsan la nube pública de la empresa, lo que ofrece una experiencia coherente en un entorno de nube híbrida.

Estas acciones están relacionadas con el anuncio de Claro Enterprise Solutions y VMware, para incorporar VMware SASE (Secure Access Service Edge)⁹⁶ a su portafolio de servicios gestionados. Este anuncio de junio de 2022 indica que permitirá a Claro aprovechar VMware SASE en América Latina para implementar un enfoque centrado en la nube para los servicios de seguridad, redes y EC. Claro entiende que sus clientes empresariales han acelerado el cambio a una fuerza de trabajo remota y distribuida, aumentando su necesidad de soluciones SASE para satisfacer sus demandas de seguridad. En esta situación VMware SASE será utilizada junto con los servicios administrados de Claro Enterprise Solutions en su cobertura de red, abordando los requerimientos de seguridad de red del ecosistema de múltiples nubes. Claro Enterprise Solutions alojará POPs SASE en toda América Latina, incluyendo Brasil, Colombia y México.

En mayo de 2022 Telefónica anunció la elección de AWS Outposts para que integre la infraestructura de núcleo de su red 5G SA nativa en la nube en Brasil para implementar servicios 5G autónomos, que optimicen los servicios 5G de Vivo. Los AWS Outposts permitirán a los clientes ejecutar servicios informáticos, de almacenamiento, de bases de datos y de otro tipo en entornos locales, bajo la infraestructura y modelos operativos nativos de AWS, mientras se conectan transparentemente a los servicios en la nube de AWS. Con los AWS Outposts, los clientes pueden utilizar las mismas API, panel de control, herramientas y hardware que AWS. El uso de estos servicios gestionados nativos de AWS va a apoyar al operador con la oferta de servicios de valor agregado en 5G como el Network Slicing y la Computación en el Borde.

Esta acción en Brasil está alineada con la decisión de Telefonica Tech a nivel corporativo, anunciada a principios de 2022⁹⁷, respecto de un acuerdo con Amazon en cuanto a la salida conjunta al mercado con AWS para mejorar y agilizar los servicios prestados a sus clientes basados en la nube.

Telefónica Tech ofrece a sus clientes empresariales, de PyMEs y del sector público, servicios profesionales en las diversas etapas de evaluación de la plataforma, diseño e implementación, y migración. Un paso importante es que lanza al mercado soluciones de valor agregado sobre la EC y redes privadas 5G nativas en la nube que operan en AWS Outposts. Esta acción tiene como objetivo el desarrollo sinérgico y conjunto de casos de uso basados en la EC y que alcanzan la Internet de las Cosas, el Machine Learning, el soporte a la industria 4.0 y al streaming de video y los juegos para los usuarios finales. Además, Telefónica Tech creará un centro de competencia de AWS para ayudar a los clientes en su transición a la nube de forma rápida y segura.

Es importante destacar que estos acuerdos orientados a la infraestructura de 5G de Telefonica + Outposts de AWS son partes de una convergencia de servicios, en la que también Amazon ofrecerá los suyos juntamente con los de Telefonica, escalando ambas, y principalmente esta última, en la cadena de valor por encima de la pura red de transporte. Es parte de una recuperación de la participación de los CSP en los ingresos globales de la Industria.

Estas acciones están íntimamente ligadas a la evolución intensa que existe en el mundo en cuanto a que, como resultado de la competencia en esta área, por ejemplo⁹⁸, Vodafone anuncia una colaboración piloto con Amazon Web Services (AWS) para sus clientes empresariales en España que estará disponible en varias ciudades, y brindará acceso a servicios y aplicaciones de latencia ultra baja gracias a la potencia de Multi-access Edge Computing (MEC) de Vodafone impulsado por AWS Wavelength.

⁹⁶ Los Outposts, como ya se vio en la primera sección, incluyen recursos informáticos y de almacenamiento que pueden ejecutar servicios nativos de AWS y se administran mediante las mismas API que impulsan la nube pública de la empresa, lo que ofrece una experiencia coherente en un entorno de nube híbrida.

⁹⁷ <https://www.telefonica.com/es/sala-comunicacion/telefonica-amplia-su-colaboracion-estrategica-con-amazon-para-el-desarrollo-cloud-y-el-hogar-digital/>.

⁹⁸ <https://www.telesemana.com/blog/2023/07/28/vodafone-espana-se-asocia-con-aws-para-lanzamiento-de-mec-en-sus-redes-4g-y-5g/>.

Además, en este caso es una combinación de 4G, 5G y MEC, respaldada por AWS Wavelength, que permite implementar recursos bajo demanda para escalar la capacidad, optimizar costos con opciones de pago por uso y acelerar el desarrollo de aplicaciones 5G.

Según lo publicado por BN Americas⁹⁹ en setiembre de 2022 sobre la EC, en LAC se observa lo siguiente.

IBM, a través de su filial Red Hat, comprada en 2019, es un jugador en el desarrollo de la EC en América Latina, anunciando varios compromisos con actores de verticales como banca, seguros, industria y telecomunicaciones, aunque no dio más detalles sobre contratos ni proyectos en curso. Observa este proveedor que Latinoamérica, normalmente más atrás en la incorporación tecnológica, luego de la pandemia está acelerando la adopción de la EC, tal como lo anunciaba F&S. Estos conceptos coinciden con las consideraciones de Harvard Business Review, y otros analistas.

La oferta de IBM permite implementar soluciones habilitadas en el borde a través de una gestión autónoma que permite orquestrar la escala, la variabilidad y la velocidad de cambio en los entornos de borde, en las soluciones y servicios industriales, y en las soluciones que ayuden a los CSP a modernizar sus redes en la nube y ofrecer nuevos servicios en el borde. Las soluciones de red de los CSP en la nube y el borde de IBM se ejecutan en la tecnología OpenShift de Red Hat, una plataforma multinube híbrida abierta que se ejecuta en cualquier lugar, desde cualquier Centro de Datos hasta múltiples nubes, y en el borde. La plataforma de IBM permite también a terceros desplegar sus productos y servicios generando un ecosistema. Por su parte Red Hat utilizará IBM Cloud for Telecommunications para implementar servicios en la nube, en el borde, o en instalaciones propias.

Otras empresas también están accediendo a este mercado de la EC en Latinoamérica, entre muchas de ellas se encuentran Azion¹⁰⁰, Cirion¹⁰¹, etc.

Azion está desarrollando una red de nodos de EC en América Latina, la que ya está operando en el borde en varios países, como Argentina (Buenos Aires), Chile (Santiago), Colombia (Bogotá), Perú (Lima), México (Querétaro) y principalmente en Brasil en unas 30 ciudades. Esos nodos están en las redes de última milla de los ISP y están conectados a los IXP, buscando explotar las principales ventajas de la EC, incluyendo la reducción de retardos de transporte a través de un uso óptimo de las redes de telecomunicaciones. Uno de sus casos de uso es el del Grupo Pernambucanas en Brasil en que informa que, mediante la instalación de las aplicaciones del Borde, lograron acelerar la plataforma de comercio electrónico, y mejorar la experiencia del cliente a través de menores tiempos de espera, por ejemplo, en la descarga de imágenes de alta calidad o la respuesta del sitio web en general, lo que no es realizable hoy en día si se usa CC con el Centro de Datos concentrado y lejano.

Otro informe de BN Americas¹⁰² muestra que Google tiende a complementar su estrategia de 2019 en cuanto a la verticalización en los segmentos minorista local, salud, telecomunicaciones, medios y servicios financieros. En Latinoamérica han adicionalmente observado un potencial importante en la agroindustria y en servicios públicos de energía y gas. Un ejemplo es su contrato con Casa dos Ventos Energias Renováveis S.A., que necesitaba mejorar sus simulaciones de demanda, incluyendo machine learning e inteligencia artificial.

En Colombia desarrolló un chatbot soportado en su plataforma para una empresa de servicios financieros para el agro, AgrodatoAI.

⁹⁹ <https://www.bnamericas.com/es/reportajes/america-latina-al-borde-de-la-computacion-en-el-borde>.

¹⁰⁰ <https://www.azion.com/es/>.

¹⁰¹ <https://blog.ciriontechnologies.com/es/beneficios-edge-computing-plataforma-lumen/>.

¹⁰² "How Google is planning to hang its Cloud over LatAm". 25 de febrero de 2022. <https://www.bnamericas.com/en/features/how-google-is-planning-to-hang-its-cloud-over-latam>.

Opera con soluciones sobre múltiples nubes, de acuerdo con las aplicaciones y las cargas de trabajo, incluyendo nubes híbridas públicas y privadas, usando Kubernetes¹⁰³ para correr su plataforma analítica en ellas.

Ya ha desplegado varias de las llamadas Regiones en Latinoamérica, entre ellas las primeras en Sao Paulo, Santiago y Querétaro (anunciada en julio de 2022) cada una comprendiendo tres Zonas. Las Zonas, a su vez, contienen las Ubicaciones de Borde para el acceso de borde a la Google Cloud. En Santiago, uno de sus Data Center está dedicado a servicios en la nube y almacenamiento de datos, proveyendo los siguientes IaaS: Compute Engine que es un servicio de procesamiento que permite crear y ejecutar máquinas virtuales, Kubernetes Engine y BigQuery¹⁰⁴. Desde Santiago Google presta servicios regionales.

Uruguay es otro país en el que está avanzando con un proyecto de Data Center que se encuentra en una etapa de redefinición por razones ambientales, principalmente por el consumo de agua.

Google ha anunciado que está trabajando en pruebas en el borde con operadores de Latinoamérica y casos piloto para redes privadas 5G como los autos conectados.

Azure¹⁰⁵ también está extendiendo su plataforma en Latinoamérica, con los siguientes despliegues de sus Azure Front Door POP, o sea los Puntos de Presencia, y que en este momento comprende las ciudades de Buenos Aires, Campinas (2), Rio de Janeiro (3), Sao Paulo (2) y Bogotá, en que en cada una puede haber más de un POP como se indican entre paréntesis. Más detalles de su estructura se incluyen en la sección "A. Descripción de la estructura conceptual y física, y funciones de la EC".

A fines de 2022 el operador brasileño TIM tenía la intención de instalar más de 60 Centros de Datos de EC en todo el país para fines de 2023, teniendo a esa fecha 40 ya desplegados. También, está cerrando sus dos Data Centers y migrando sus operaciones a la nube usando las plataformas de Google, Oracle y Microsoft: las tecnologías de analítica en Google Cloud Platform, sus bases de datos y procesos operativos en Oracle, y sus aplicaciones de front end y canales digitales en Azure.

Respecto de la región del Caribe, existe un estudio de junio de 2022¹⁰⁶, de IDC, comisionado por Schneider Electric, que presenta algunos datos de la evolución reciente y esperada en la región.

Indica que el volumen de datos en Latinoamérica se triplicará entre 2020 y 2025, o sea un CAGR del 24,7% sobre el promedio global estimado de 22,9%, lo que también se verá reflejado en el Caribe. Estas estimaciones son consistentes con otras que muestran un alto crecimiento en LAC. Como consecuencia sería necesario disponer en el Caribe de infraestructuras de procesamiento y almacenamiento de datos escalables, que por otra parte se encuentren cerca de los usuarios, y según esta encuesta, preferentemente como Data Centers modulares y prefabricados a los efectos de la reducción de plazos hasta la puesta en operación, y de los costos. IDC ha observado en este estudio que el crecimiento en TI sería del 9% de 2021 a 2022, estimando un CAGR del 12,5% en 2021-2025, debido a los procesos de digitalización, y, también lo expresa abierto por Software (15,1%), Servicios de TI (5,1%) y Hardware (12,1%). Estima que, en cuanto a Hardware, en el período 2021-2025 el crecimiento de IaaS, como parte importante del crecimiento de las inversiones en TI, crecerá a una CAGR del 45%, llegando en 2025 al 41% de la inversión en Hardware. También estimaba IDC que para fines de 2022 el 75% de las empresas implementarán sistemas unificados de gestión de sus nubes, redes y Data Centers para reducir costos y complejidad. La tendencia es a migrar la infraestructura de procesamiento hacia una estructura centrada en la carga de trabajo sobre tecnologías en la nube.

¹⁰³ Kubernetes automatiza las tareas operativas de la gestión de contenedores e incluye herramientas para el despliegue de aplicaciones, efectuar cambios en las aplicaciones en todos los sitios en que se encuentre, escalar las aplicaciones y monitorearlas.

¹⁰⁴ Es una Plataforma de Google, que se vende como servicio, que incluye el almacenamiento de datos y su procesamiento, incluyendo Machine Learning.

¹⁰⁵ <https://learn.microsoft.com/en-us/azure/frontdoor/edge-locations-by-region>.

¹⁰⁶ <https://geezam.com/idc-schneider-iaas-edge-2025/>.

B. Otras regiones

Se analiza en esta sección el estado actual y evolución prevista de la EC en el mundo en general.

En primer lugar, se destaca que en este mismo trabajo se han analizado múltiples aspectos sobre las perspectivas para el futuro de esta importante tecnología. se pasan a detallar las secciones donde se encuentran:

- “I.A.3. Tecnologías emergentes relacionadas a la EC”. Se puede observar el conjunto de tecnologías con gran poder de crecimiento y de extensión de sus aplicaciones, donde todas ellas requieren de la EC para su aplicación y expansión. Serán un fuerte motor de la EC. Basta mencionar una de ellas, la Inteligencia Artificial, en que su aplicación intensiva en el procesamiento de los grandes datos, la visión automatizada, entre otras, requiere reducir el transporte de la información y el retardo en la respuesta. Ambas facilidades las provee la EC.
- “I.B. Principales características de las acciones del ETSI con relación a la MEC.” A medida que los proveedores de servicios se alejan de los dispositivos físicos y se acercan a una arquitectura basada en servicios, el resultado es un desacoplamiento que permite un ecosistema más amplio donde se pueden ejecutar las cargas de trabajo de movilidad. Y en este entorno, el ETSI observa la necesidad de estandarizar la MEC, lo que ya está impulsando el fuerte involucramiento de los CSP en el mercado de la EC.
- “I.C. Evolución de la EC y de los proveedores de las tecnologías directamente relacionadas”. Se analiza la evolución desde el punto de vista de la cantidad de proveedores en competencia para este mercado naciente.
- “II. Evolución de los mercados relacionados a la EC.” Se describe la evolución prevista de estos submercados que en conjunto impulsan la EC, y del comportamiento de los CSP y los hiperescaladores.

Uno de los más recientes estudios sobre la nube en general, incluyendo la EC, es el resultado de una encuesta realizada por Flexera, publicada en 2023, que contiene los parámetros principales del estado actual, y de la opinión de los integrantes de la muestra respecto de su percepción de la evolución¹⁰⁷. Esta encuesta alcanza a 750 profesionales de TI y líderes ejecutivos de todo el mundo realizada a fines de 2022.

La muestra está integrada por la siguiente proporción de integrantes según el tamaño de la empresa:

Cuadro 3
Proporción de las empresas de los integrantes de la muestra por cantidad de empleados

Empleados	1-100	101-1 000	1 001-2 000	2 001-5 000	5 001-10 000	10 001-100 000	> 100 000
Porcentaje	6%	11%	16%	20%	15%	21%	11%

Fuente: Flexera.

Si bien abarca a empresas mucho mayores que las más comunes en LAC, es una referencia importante de las tendencias en el mundo que, considerando la escala de cada una, permitirá observar la aplicabilidad total o parcial de los resultados. El análisis de la sección anterior es una ayuda importante del comportamiento en LAC en comparación con lo analizado en esta sección.

¹⁰⁷ Flexera 2023 State of the Cloud Report. <https://info.flexera.com/CM-REPORT-State-of-the-Cloud>.

Entre los principales resultados se encuentran los siguientes:

- El 87% emplean múltiples nubes, de las cuales el 72% son nubes híbridas. Solo el 2% usa una nube única privada.
- Las cinco nubes más empleadas, son por su orden de mayor uso: Azure, AWS, Google Cloud Platform, Oracle Cloud Infrastructure e IBM Cloud.
- En cuanto a los tres mayores desafíos que les presenta el uso de la nube se observan los siguientes:
 - La gestión de los gastos en la nube, que como ya se analizó en la sección “III. C. Gestión de costos de los servicios en la nube.”, es el mayor desafío en el 82% de los casos.
 - La seguridad está en segundo lugar en general con el 79%, y en particular para las PyMEs con el 73%.
 - La falta de recursos y de experticia ocupa el tercer lugar de preocupación en el 78% de los casos.
 - En general la gobernanza, el cumplimiento de las normas de datos y la gestión de licencias de software son las siguientes preocupaciones, éstas con el orden del 70%.

Si bien los gastos en la nube son la principal preocupación, el hecho de que la innovación o el mejoramiento de los servicios y procesos en un mundo crecientemente competitivo son impensables sin usar la nube, lleva a las organizaciones a seguir esta evolución en el mundo a pesar de las incertidumbres económicas, pero buscando crear centros de control de gastos adoptando prácticas FinOps maduras. Cerca del 50% de los encuestados pertenecen a organizaciones que planifican mover su software en sus facilidades a los servicios de SaaS.

STL Partners realizó un estudio¹⁰⁸ publicado en enero de 2023, analizando el despliegue de la EC en el mundo. Respecto al importante papel de los CSP, luego de sus pruebas de concepto y asociaciones, se presentan los siguientes casos que muestran la variedad de situaciones, en forma similar a lo que se ha estudiado en LAC. Nota: Los casos de uso de cada CSP son solo algunos de sus servicios prestados¹⁰⁹.

- EE. UU., Lumen. Se encuentra planificando más de 100 sitios iniciales de EC en localidades de todo EE. UU. y se ha asociado con Azure. Casos de uso: minorista, servicios de comida, logística y transporte.
- EE. UU., Verizon. En sociedad con AWS desde 2020 ahora está ofreciendo 17 Wavelength Zones, y en 2021 también se asoció con Google. Casos de uso: reconocimiento facial con IA, realidad aumentada y virtual, distribución de contenido.
- Canadá, Bell. Se asoció con AWS para proveer Wavelength, inicialmente en Toronto y planificando su expansión. Con Google se asoció en 2021 para operar su núcleo de red (core) en una red de EC distribuida. Casos de uso: compras en línea y entregas con robots y drones.
- EE. UU., AT&T. En 2021 se asoció con Azure y lanzó las Zonas de Borde, y en el mismo año con Google para lanzar sus soluciones de EC en la red y en los locales de los clientes en 15 localidades. En 2022 se asoció con IBM para lanzar aplicaciones de IA en nubes híbridas. Casos de uso: IA y Machine Learning, analítica de video, realidad aumentada empresarial y robótica.

¹⁰⁸ Edge computing deployments: around the globe. <https://stlpartners.com/articles/edge-computing/edge-computing-deployments/>.

¹⁰⁹ El autor entiende que son algunos pues en este proceso de evolución se observan casos en que se prestan otros servicios a medida que pasa el tiempo.

- EE. UU., Cox Communications, una empresa de cable. No se asoció con un hiperescalador, creó una unidad de negocio especializada en EC, y en 2021 lanza el servicio de un suministro completo de EC desplegado usando Centros de Datos en la última milla. Casos de uso: CDN en el borde, analítica de video y vehículos autónomos.
- Europa, Vodafone. Han firmado con 180 empresas contratos para su Programa de Innovación con EC. Tienen también una asociación con AWS Wavelength. Casos de uso: Analítica de video, inspección de activos en tiempo real, realidad aumentada, drones, edición en medios empleando IA.
- España, Telefonica. Comenzó la actualización de 1.200 centrales telefónicas para transformarlas en centros de EC. Anunció su asociación para EC federativa con Singtel (Singapur) como parte del Foro de GSMA para las nubes de borde de las TELCO¹¹⁰. Casos de uso: movilidad asistida, entretenimiento y medios, servicios financieros.
- Alemania, Deutsche Telekom. En 2022 lanzó un laboratorio de investigación de 5G en Berlín. Tiene un MoU con SKT para desarrollar MEC en 5G en Corea. Casos de uso: automatización industrial e industria de juego.
- Japón, KDDI. Sociedad con AWS para desplegar Wavelength en sitios de 5G. Casos de uso: realidad aumentada y virtual en deportes, entretenimientos y turismo.
- Corea, SK Telecom. Lanzó su propia plataforma de EC en MEC 5G. Se asoció con Dell en 2022 para lanzar la solución MEC Petasus. En 2023 inició sus pruebas en EE. UU. Tiene una sociedad con AWS. Casos de uso: realidad aumentada y virtual y streaming de video juegos, así como interfases virtuales móviles.
- China, China Mobile. Desplegó su propia plataforma Open Sigma de EC en varias provincias. Se asoció con CAS – Vision y Huawei para anunciar el primer proyecto comercial de turismo cultural cobre 5G MEC. Casos de uso: gemelos digitales, Machine Learning.
- Singapur, Singtel. Asociación con Azure para lanzar Zonas de Borde con su 5G. Lanzó la plataforma Paragon para desarrolladores enfocados en 5G, EC y la nube. Casos de uso: realidad aumentada y virtual, vehículos autónomos y robótica.

En un estudio¹¹¹ anterior de esta misma empresa se observan otros CSP:

- Emiratos Árabes Unidos, Etisalat. Se ha asociado con Azure. Casos de uso: prestar servicios de ciudades inteligentes, IoT, seguridad pública y RAN virtual.
- Bélgica, Proximus (ex Grupo Belgacom). Asociación con Azure. Casos de uso: servicios para la industria manufacturera, realidad aumentada y virtual, salud, juegos y logística.
- Canadá, Rogers. Sociedad con Azure. Casos de uso: campus inteligentes, juegos y realidad aumentada y virtual.
- Italia, Telecom. Sociedad con Google.
- Indonesia, Telkomsel. Sociedad con Azure para servicios de la industria manufacturera, IoT, inteligencia artificial y realidad aumentada u virtual.
- Australia, Telstra. Sociedad con Microsoft Azure, para desplegar 500 sitios con servicios financieros y juegos.

¹¹⁰ <https://www.gsma.com/futurenetworks/wp-content/uploads/2022/03/GSMA-TEC-Value-Whitepaper-v13.pdf>.

¹¹¹ <https://stlpartners.com/articles/edge-computing/hyperscalers-telcos-edge-partnership-opportunities/>.

Estos ejemplos muestran la importancia, ya analizada en el estudio de la MEC, del despliegue de EC en el mundo a través de asociaciones con proveedores de servicios en la nube, o con infraestructura propia. En el caso de los hiperescaladores y los CSP, ambos necesitan del otro para poder mantenerse en la cúspide del mercado de la EC. Los movimientos hacia la prestación de servicios de computación en el borde se han visto altamente impulsados por la incursión de los hiperescaladores en el lanzamiento de soluciones para la infraestructura de los CSP, integrándola a sus propios núcleos, y habilitando la prestación de servicios de EC a los clientes.

Otros proveedores de servicios en la nube, que no son los hiperescaladores, que ya fueron analizados, así como los proveedores de equipamiento, también juegan su importante papel en la provisión del servicio o de la infraestructura, y el conocimiento, para que los clientes finales usen una infraestructura que no es del CSP. El estudio de STL Partners indica que el 67% de los CSP relevados se han asociado en algún grado con hiperescaladores.

Un caso particular es el de Cox Communications¹¹², una empresa de cable que presta servicios desde televisión hasta comunicaciones empresariales gestionadas. En lugar de asociarse con un hiperescalador, creó su propia unidad Cox Edge para brindar servicios de EC a sus clientes empresariales a través de más de 30 ubicaciones de borde en los EE. UU., Sao Paulo y varias ciudades en el mundo, además de prestar sus servicios de distribución de contenidos. Provee todos sus servicios independientemente de la nube que se emplee.

Complementando este análisis de la EC provista por múltiples CSP, se destacan las principales conclusiones de Frost & Sullivan a partir de un trabajo presentado en setiembre de 2023 en un evento de 5GAmericas¹¹³, de alcance global, para un período de proyección 2022-2025.

Se observan los resultados referentes al crecimiento, exclusivamente de la MEC, y sus impulsores.

- Los ingresos a nivel global se estiman en 16.282 millones de US\$ (MUS) para una CAGR del 115,9% entre 2021 y 2025.
- Si estos valores se abren para América Latina y el resto del mundo, resultan respectivamente valores de 293,1 y 15.989 MUS, y 239,7% y 115,1% para la CAGR. Este crecimiento mayor en América Latina ya se ha observado en otras tecnologías emergentes.
- En cuanto al crecimiento esperado abierto por verticales se tienen estos valores:

Cuadro 4
Ingresos de la MEC y su crecimiento, por vertical

Vertical	Ingresos en 2025	CAGR 2021-2025
Automóviles, transporte y logística	1 905	140,1%
Telecomunicaciones, medios y entretenimiento	5 822	160,7%
Industrias manufacturera y extractiva	7 057	95,3%
Otros (ciudades, salud, minorista, construcción, educación, energía, etc.)	370	46,4%

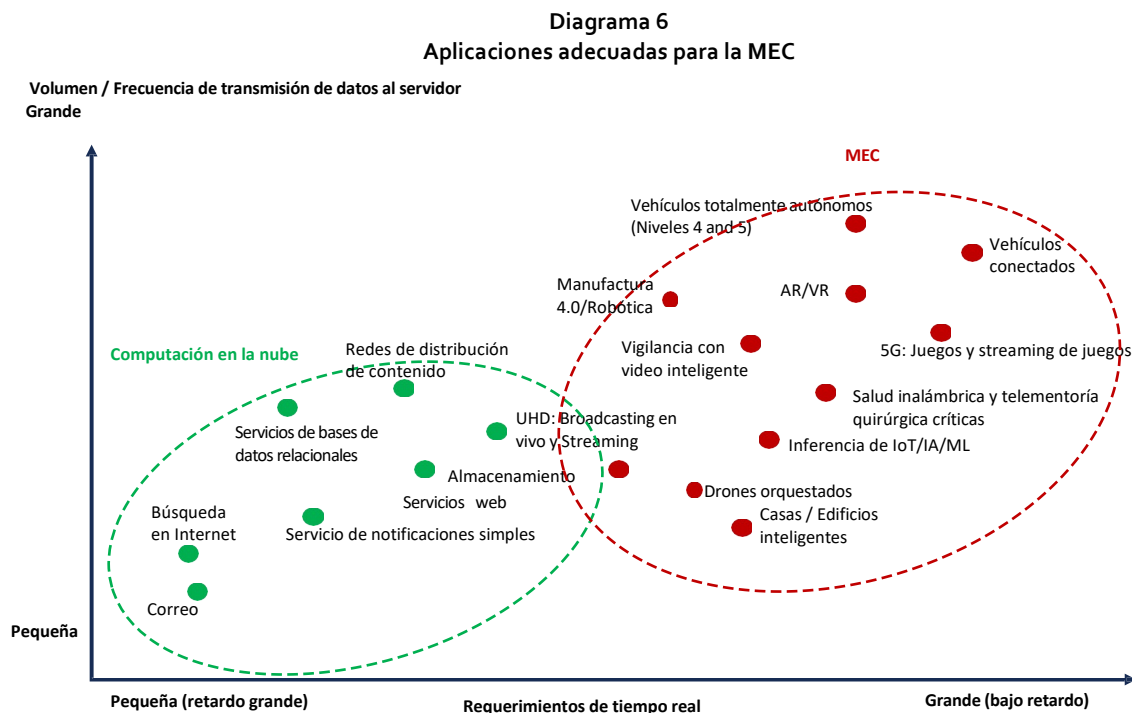
Fuente: Frost & Sullivan. 5G Americas.

¹¹² <https://www.coxedge.com/about-cox-edge>.

¹¹³ Serie de webinarios de 5G Americas, 26 de setiembre de 2023. 5G and Multi-Access Edge Computing: Market Perspectives and Growth Opportunities. Renato Pasquini, Vicepresidente de Investigación, IoT & Edge, de Frost & Sullivan. Estudio que abarca Latinoamérica, Norte América, Europa, Cercano Oriente y África, y Asia Pacífico. Alcanza los siguientes tipos de participantes: operadores de telecomunicaciones, proveedores de nube, proveedores de infraestructura de telecomunicaciones y TI, proveedores de infraestructura de soporte de centros de datos de borde, proveedores de servicios de ubicación de centros de datos de borde, desarrolladores de aplicaciones/soluciones de borde.

Estos valores indican el empuje de la EC y principalmente en América Latina.

F&S presenta además la siguiente gráfica que ilustra la adecuación de las diferentes aplicaciones para emplear la MEC. En el eje de las abscisas, donde dice “Grande (Bajo Retardo)” significa que existen grandes oportunidades de crecimiento cuando se exige un pequeño retardo, y contrariamente donde dice “Pequeña”. Se observa que en el primer rango se engloba lo que requiere MEC, y en el segundo lo que emplea la nube. En el eje de las ordenadas se observa el impacto de los volúmenes de los datos a transmitir, que cuando son grandes se requiere de la MEC.



Fuente: Frost & Sullivan, con traducción del consultor.

V. Impacto en países y en verticales seleccionados

A. Impactos en el sector productivo en general

La EC es una familia de tecnologías que se desarrolla sinérgicamente con otras familias (5G, IoT, IA y Machine Learning, robótica, otras) y que en conjunto están cambiando los modelos de negocio y la actividad económica y social en general. En la sección “D. Conjunción de la EC, la Internet de las Cosas y la 5G en el desarrollo productivo.” se observa más en detalle la interacción entre las tres primeras. Como ya se ha analizado, su poder proviene de permitir acercar el procesamiento y el almacenamiento a los usuarios, habilitando menores retardos y mayor capacidad para ellos, lo cual impulsa las tecnologías que están siendo adoptadas por las organizaciones para mejorar los servicios y la eficiencia productiva. Está siendo usada en despliegues de gran porte en telecomunicaciones, las industrias manufacturera y extractiva, las ciudades inteligentes, la salud, en habilitar una mejor y más avanzada digitalización de las PyMEs, entre otros.


Como ejemplo de estos motores del desarrollo de la EC, la IoT estaba llegando a 10.000 millones de dispositivos en el mundo a fines de 2022, estimándose una incorporación del orden de 150.000 dispositivos por minuto, generando 73 Zettabytes¹¹⁴ de información a transportar y procesar para el año 2025¹¹⁵ Por otra parte, IBM estima que habrá 38.500 millones de dispositivos que generarán 90 Zettabytes para 2025, un resultado similar. Estos volúmenes de información, creciente, a ser procesada y generar respuestas, en general con retardo de menos de 5 ms., no son soportables por la CC por problemas de retardo y de costos de la transmisión, y será imprescindible disponer de una EC ubicua considerando su dispersión geográfica.

¹¹⁴ 10²¹ Bytes, o 10¹² Gigabytes.

¹¹⁵ <https://dataprot.net/statistics/iot-statistics/#:~:text=In%202021%2C%20there%20were%20more,in%20economic%20value%20by%202025.>

Antes de analizar las profundas transformaciones del sector productivo impulsadas por la EC, resulta de interés resumir en el cuadro 4 las opciones que están disponibles para las empresas y sus características principales¹¹⁶.

Cuadro 5
Prestaciones y usos de los diferentes tipos de EC

Tipos de EC	EC propia de la empresa	Coubicación	Nube privada	Nube pública
Definición	La empresa alberga en su local su propia infraestructura de procesamiento y almacenamiento.	La infraestructura propia de procesamiento y almacenamiento es albergada en un centro de datos de terceros.	El centro de datos de un tercero provee a la empresa procesamiento y almacenamiento dedicado.	Un tercero (hiperescalador) provee a la empresa la infraestructura de procesamiento y almacenamiento.
Beneficios	1. La empresa retiene el control total sobre las políticas y la seguridad. 2. Latencia ultra baja. 3. Protección de datos y soberanía.	1. Reducción de CAPEX y OPEX respecto de la instalación local. 2. Protección de datos y soberanía.	1. Menores costos de CAPEX de infraestructura que en coubicación. 2. Protección de datos y soberanía. 3. Escalabilidad.	1. Menor CAPEX de infraestructura con relación a la nube privada. 2. Acceso a la tecnología de los hiperescaladores. 3. Escalabilidad.
Inconvenientes	1. Altos CAPEX y OPEX. 2. Dificultad para prepararse para la fácil desactualización en el futuro. Difícil de lograr las habilidades internas.	1. Los costos de la infraestructura de procesamiento y almacenamiento se mantienen altos. 2. Persisten dificultades para el futuro y las habilidades.	Los costos de infraestructura de computación y almacenamiento, la que es dedicada, se mantienen altos.	1. Confianza en el modelo de negocio y precios del hiperescalador. 2. Riesgo de soberanía de datos.
Casos de uso	Operaciones de misión crítica en que la necesidad de escalar es limitada y que la necesidad de control cercano en tiempo real justifica mayores costos.	Aplicaciones importantes donde la inversión en el local no es viable pero la demanda puede escalar y la empresa quiere retener el control sobre el procesamiento.	Aplicaciones donde la escalabilidad y la gestión de costos es importante pero la sensibilidad (inc. latencia) y la customización son importantes.	Soporte al negocio donde la necesidad de reducir costos y gestionar la escalabilidad es razón de peso y es bajo el riesgo (percibido) de la soberanía de datos.
Dirección general del camino de las empresas en la EC. 				

Fuente: STL Partners.

Para analizar en profundidad el impacto que está produciendo la EC en el sector productivo, desde el punto de vista de la orientación, objetivos y cambios en las operaciones, se emplea como referencia un reciente estudio de Accenture¹¹⁷, de setiembre de 2023.

Es una investigación multi método. Incluye una encuesta cuantitativa con un análisis de clústeres, el que busca identificar distintos tipos de usuarios. También incluye entrevistas con adoptantes de la EC en una combinación de industrias y geografías. Se condujo entre noviembre de 2022 y enero de 2023, entrevistando a 2.100 ejecutivos de nivel C y de Dirección, a nivel global en 16 países y 18 industrias, y otras acciones de relevamiento de información.

A nivel macro las principales conclusiones son, consistentes con otros estudios, que el 83% cree que la EC será esencial para la competitividad en el futuro. Además, el 81% piensa que se debe actuar rápidamente para aprovechar todos los beneficios de la EC. Aun así, el 65% utiliza la EC en distintos grados en la actualidad, y solo la mitad la tiene integrada profundamente en su núcleo digital que abarca diferentes áreas.

También observa que en las empresas en que el desarrollo de la EC fue liderada exclusivamente por los responsables de la TI, su inserción suele no estar siendo eficiente para trabajar profundamente en la actualización a estos nuevos modelos en todas las áreas de la empresa en forma integrada.

¹¹⁶ <https://stlpartners.com/articles/edge-computing/when-will-edge-private-cloud-supplant-colocation/>.

¹¹⁷ <https://www.accenture.com/content/dam/accenture/final/accenture-com/document-2/Accenture-Leading-With-Edge-Computing.pdf#zoom=40>. Nota: en la transcripción se usa la abreviatura EC por razones de uniformidad cuando en este documento se empleen palabras como borde, perimetral o similares que hagan referencia a la computación en el borde.

La razón que se suele observar, en primera instancia, es que se desarrollan proyectos de EC ad-hoc para lograr resultados rápidos en áreas que se manifiestan con la necesidad de introducir más computación.

La mitad que aplica la EC integrándola profundamente en todas las partes de su negocio, adopta un enfoque general de transformación de su núcleo digital, orientado a los datos y la IA, y su procesamiento basado en la nube. Ven la EC como un diferenciador clave para llevar la inteligencia artificial a su negocio principal, integrada en sus productos, servicios y fuerza laboral. Y están viendo mejores resultados: innovación acelerada que conduce a nuevas oportunidades de ingresos, costos reducidos que resultan en una mayor eficiencia y mejores experiencias para los clientes.

Independientemente de la industria, lo que diferencia a estas empresas es su enfoque estratégico de la EC y su integración con su estrategia de nube más amplia.

A los efectos de identificar más precisamente los tipos de caminos adoptados con la incorporación de la EC a los procesos de las empresas, este trabajo de Accenture efectúa una clasificación en cuatro tipos, los que a veces son resultados de procesos de incorporación de la EC resolviendo principalmente cuestiones coyunturales y no estratégicas:

- Tipo 1. Ad Hoc, en el que entran 413 de los consultados, y corresponde a los despliegues centralizados en la conducción del sector de TI. En su evolución, estas empresas obtendrían mejores resultados procurando integrar en forma continua el procesamiento centralizado con el procesamiento en la nube.
- Tipo 2. Táctico, en que se incluyen los adoptantes de soluciones empaquetadas, destinadas a requerimientos específicos. Lo integran 275 de los consultados. En este caso las soluciones adoptadas en general son de difícil replicación en otras zonas de operación de la empresa, o no son adecuadas para una evolución posterior de la EC a otros sectores de la empresa. Si así lo fueran, por ejemplo, si se incorporan para dar solución a una necesidad puntual estando preparadas para su modernización y evolución, se simplifica el camino para su integración a una estructura mayor y de integración de la nube con el procesamiento centralizado.
- Tipo 3. Integrado, donde el despliegue se integra con la nube, la IA y escala ampliamente. Corresponde al tipo adoptado por 684 respondientes. Específicamente es el resultado de una planificación estratégica tendiente a disponer de una estructura de procesamiento escalable, en una única plataforma extensible para prestar servicios de diversas características, local y regionalmente, que incorpora transparentemente la nube, la EC y los centros de datos propios. La escalabilidad, simplificada al máximo, permite adaptar las capacidades, pero también incorporar progresivamente los servicios y aplicaciones que sea necesario, y sin afectación o con baja afectación, de lo que ya está corriendo en el sistema. Se ha determinado que aproximadamente el 50% de los respondientes que adoptaron la EC, lo hicieron con este Tipo. Y el 79% en general, planea hacerlo en un plazo de tres años.
- Tipo 4. Super integrado. Es un subconjunto del Tipo 3, y son las empresas que han atado la EC al negocio en el proceso de transformación. Son los que han profundizado más la transformación de su negocio obteniendo los mejores resultados. Utilizan el conjunto de tecnologías de su núcleo digital, empezando con la nube y agregando la analítica de datos, la IA, las aplicaciones y las plataformas. Adicionalmente se apoyan en un ecosistema de proveedores externos y en la formación de las habilidades internas. En este estudio son solamente el 6% de los que adoptaron la EC. "Reconocen que la EC no es una tecnología independiente, sino una topología que integra arquitecturas centralizadas y distribuidas necesarias para diferenciar su negocio principal ya sea que se trate de integrar inteligencia dentro de un producto o un lugar.

Los restantes 728 aún deben ejecutar sus planes con la EC. En el documento se denominan Pre-adoptantes.

Indica que aquellos de los Tipos 1 y 2 son los que han sido menos exitosos en cuanto a los resultados obtenidos, y analizando los cuatro tipos llega a estas conclusiones en cuanto a los resultados de la adopción de la EC.

- Tipo 1. "Logran resultados más bajos en todos los ámbitos y son más escépticos sobre el éxito futuro de la EC".
- Tipo 2. "Logran reducciones de costos utilizando la EC, pero se quedan atrás en relación con sus pares en otras áreas, como la innovación y la productividad".
- Tipo 3. "Obtienen los mejores resultados para mantener los datos en la EC y acelerar la innovación".
- Tipo 4. "Obtienen los mejores resultados en todos los ámbitos, con los más altos niveles de satisfacción con sus implementaciones de EC".

En cuanto a cómo perciben el valor agregado por la EC aquellas empresas que adoptaron la super integración, frente a todas las demás empresas de los otros tres Tipos, se observa que es sustancialmente mayor en todas estas ópticas: capacidad de conservar y gestionar eficientemente los datos en el borde, crecimiento de la eficiencia, transformar la experiencia de los empleados, capturar datos manteniendo al mismo tiempo la seguridad y la privacidad, acelerar la innovación, capturar datos para obtener mayores conocimientos, colaborar con otros jugadores de la industria, transformar la experiencia del cliente, cumplir con los objetivos de ambiente, sociales y de gobernanza, crecer el negocio y reducir los costos (muy relacionado con el crecimiento de la eficiencia).

B. Impactos en verticales seleccionados

Múltiples verticales están siendo impactados por la llegada de la EC, entre ellos las telecomunicaciones, una industria que está en el corazón del procesamiento de la información por ser quien la transporta entre los usuarios y los centros de datos.

A su vez las telecomunicaciones están siendo usuarias de la EC, sea sobre infraestructura propia (DIY) o de proveedores de la computación en el borde. Esto le permite a esta industria mejorar la eficiencia del ciclo productivo a través de la automatización del BSS y OSS, logrando mejorar su ROA permitiéndole monetizar sus activos. La EC entra a jugar también en la posibilidad de aumentar su cartera de productos a través de, desde el juego en la nube (mínimo retardo con EC y 5G) para consumidores, hasta la automatización de los procesos productivos, así como flexibilizar su oferta más allá del tradicional arrendamiento por ancho de banda, permitiendo modelos más complejos de acuerdo con la demanda. Una encuesta¹¹⁸ de Nokia, orientada al crecimiento de los ingresos por la 5G, plantea las siguientes conclusiones:

- Entre los 10 servicios empresariales, los que a su vez son el área de mayor potencial de crecimiento, que harán crecer los ingresos con la 5G se encuentran: aplicaciones y soluciones nuevas basadas en la EC (lugar 5), y la oferta de MEC¹¹⁹ como servicios (MEC as a Service en el lugar 9).

¹¹⁸ https://pf.content.nokia.com/too7fy-cns-monetization-5g-charging/The-Path-to-Successful-5G-Monetization?_ga=2.175189734.596288939.1687986617-1955968695.1687986617&_gac=1.12047814.1687987080.CjoKCOjwO-kBhDIARIsAL6Lordy8Dwnb2lvebbfSe3W-GjuqHdLKcbY8RMoVV6hMxRgmQ6eSiPYoNUaAsBuEALw_wcB

¹¹⁹ Multi Access Edge Computing, o Mobile Edge Computing.

- Una arquitectura nativa de la nube es clave para el despliegue de una red 5G capaz de monetizar las inversiones. Solamente así los CSP podrán alcanzar la flexibilidad y velocidad requeridas por los clientes. El 70% de los CSP encuestados están considerando implementar sistemas de monetización en la nube pública. Esto tiene muchos beneficios, incluida la escalabilidad ilimitada, una plataforma ideal para IA y análisis, y capacidades de EC para casos de uso de latencia ultra baja. Este 70% se abre en 87% en América del Norte, 57% en Europa y 65% en Asia-Pacífico.

Analizando más en general el impacto de la EC en los diferentes verticales, el documento analizado de Accenture¹²⁰ presenta el siguiente cuadro en que se observa un resumen de la situación para múltiples verticales a través de los porcentajes de avance en que se encuentran para cada tipo de implementación.

El Tipo 4 de los despliegues Super Integrados presenta valores muy bajos, lo que es consecuencia de que es la última etapa de la evolución hacia la EC, se está en la etapa de adopción, e implica una fuerte decisión estratégica que involucra una transformación muy importante y alta inversión.

Por otra parte, y en general, existe una disparidad grande en el desarrollo de la EC de acuerdo con los diferentes verticales, como se observa en el cuadro 6.

Cuadro 6
Estado de avance de la EC en diferentes verticales (2023)
(En porcentajes)

Áreas de actividad económica	Pre-adoptantes	Ad-Hoc	Táctico	Integrado	Super integrado
Aero espacio y Defensa	26	30	13	31	2
Aerolíneas, viajes y transporte	48	21	6	25	3
Automotriz	34	14	15	37	7
Bancaria	26	26	11	27	8
Mercados de capital, incluyendo banca de inversión	35	23	16	26	5
Química	31	10	6	52	6
Comunicaciones, Medios y Entretenimiento	29	19	18	34	5
Bienes de Consumo	34	25	7	34	6
Energía	31	27	13	29	4
Salud (Excluyendo la salud pública)	42	14	20	24	2
Alta tecnología	23	26	23	29	2
Equipamiento y Bienes Industriales	27	27	7	40	7
Seguros	34	13	21	32	3
Recursos Naturales, Metales y Minería	60	13	7	20	1
Farmacéutica, Biotecnología y Ciencias de la vida	31	20	10	39	3
Minorista	41	18	11	30	4
"Utilities"	39	13	12	36	4
Servicios Públicos	37	16	18	29	1
Total	35	20	13	33	4

Fuente: Accenture.

Nota: Los porcentajes de la columna "Super integrado" son parte de los de "Integrado".

¹²⁰ <https://www.accenture.com/content/dam/accenture/final/accenture-com/document-2/Accenture-Leading-With-Edge-Computing.pdf#zoom=40>.

A continuación, se efectúa una descripción más detallada del impacto de la EC en los diferentes verticales. Se toman en consideración múltiples fuentes, entre otras, la información provista por Accenture y por IBM.

IBM reporta los avances de la EC y las expectativas de crecimiento y de evolución en un estudio publicado en 2021¹²¹.

Considerando los disruptores de la EC, entendidos como aquellas empresas que invierten en EC del orden del doble que los demás (6,3% del presupuesto de tecnología frente al 2,7%), se observa que se encuentran más avanzados en el camino de la EC. Mientras los menos avanzados se encuentran solamente planificando o implementando casos, los disruptores ya están planificando implementar en 3 años soluciones de EC que se alineen con sus objetivos de negocios. Y en 5 años estiman que van a estar apalancando la EC en una forma totalmente alineada con sus objetivos de negocio e integrada efectivamente con los sistemas y procesos actuales de las empresas.

También consideran que en 3 años van a estar incorporando máquinas inteligentes que tomen decisiones sencillas y en el camino de acercarse a máquinas que tomen decisiones complejas en 5 años, aunque no todavía de misión crítica.

Si bien la nube se va a seguir empleando, estas empresas se encuentran firmemente desarrollando la EC, incorporando, además, las tecnologías complementarias que se han analizado.

Es importante hacer notar que efectivamente la incorporación de la EC es un camino que empieza con los Pre Adoptantes, y avanza hasta la EC Super Integrada.

Se analizan los impactos en algunos verticales destacados que han acelerado la adopción de la EC como consecuencia del empuje de las tecnologías que solo pueden funcionar con ella, como la IoT crítica, la automatización industrial, el empleo intenso de la IA, entre otras.

- **Vehículos autónomos.** Si bien no existe un uso global, ya existen alrededor de 30 países que los han autorizado, en algunos casos con un conductor supervisor, y en otros sin conductor. Estos vehículos se clasifican en 6 niveles de acuerdo con la SAE¹²². Esto es: Nivel 0: No automatizado, Nivel 1: Automatización mínima como el conocido control de velocidad de cruce, Nivel 2: Automatización escasa y que requiere atención humana constante, Nivel 3: Puede circular controlado por sí mismo con atención humana en condiciones severas o de fallas, Nivel 4: Altamente autónomo, en que no es necesario usar volante o pedales pero su circulación es reducida a zonas limitantes, y Nivel 5: Completamente autónomo sin limitaciones de zonas ni conductor asistente, por ejemplo los de Waymo¹²³ en San Francisco. Estos vehículos requieren que la información de sensores, cámaras, ubicación, información de tráfico y otros datos básicos sean capturada y procesada instantáneamente por lo que deben disponer de un poderoso procesamiento y almacenamiento a bordo, con comunicación intensa y de bajo retardo con Centros de EC en el exterior del vehículo.

¹²¹ <https://www.ibm.com/thought-leadership/institute-business-value/en-us/report/edge-computing-advantage> Información de una encuesta el IBM Institute for Business Value, publicada en mayo de 2021 y realizada en colaboración con Oxford Economics, encuestando a 1.500 ejecutivos con conocimiento directo de las estrategias, inversiones y operaciones de sus organizaciones en relación con el edge computing. Este grupo abarcaba 22 industrias con sedes corporativas que abarcaban 21 países. Este informe se centra en los 140 encuestados de productos químicos, petróleo y productos industriales.

¹²² Society of Automotive Engineers.

¹²³ <https://waymo.com/intl/es/waymo-one-san-francisco/#:~:text=Waymo%20One%20currently%20operates%2024,Waymo%20One%20take%20the%20wheel.>

- **Minorista.** Los usos más comunes son en asistentes robóticos, asistentes para optimizar el local de ventas, y software de edificios inteligentes. En el mercado minorista el objetivo es proveer un servicio excelente a los clientes manteniendo la eficiencia de la operación, lo cual puede ser logrado por la EC. Ya se ha analizado el caso de Casas Pernambucanas de Brasil, en que el principal motivo es el de dar rápida respuesta con imágenes de alta calidad a los clientes que compran en línea. Pero también ya existen, por ejemplo, cadenas de comidas en que integran la EC en sus locales para procesar información sobre el stock de mercadería, las horas de mayor afluencia para acelerar las entregas, las condiciones ambientales y otros aspectos de los edificios inteligentes, y la gestión robótica de insumos en sus depósitos. También, a través de cámaras con IA, algunos comercios realizan un seguimiento de los recorridos de los clientes, sus elecciones y puntos de interés, la ubicación de los vendedores y otros aspectos que les permiten optimizar la distribución de la mercadería y la de los asistentes. Se agregan puntos de pago inteligentes que detectan lo que se compra, e inclusive agrega sugerencias previstas por el sistema, y efectúa la cobranza en tiempo mínimo.
- **Redes inteligentes de potencia.** La EC no solamente permite mejorar la eficiencia de las redes de generación y suministro, sino también el uso por parte de clientes. Las instalaciones productivas disponen de sensores y actuadores sobre el consumo de energía que les permiten optimizarlo a través de planes con las compañías suministradoras, o de acciones directas sobre el mejor momento para correr máquinas altamente consumidoras de energía. La aplicación de inteligencia en las redes permite una gestión óptima y predictiva de la generación, transmisión y distribución. También ha demostrado una gran eficiencia en la gestión de fallas de distribución, especialmente en China, donde hay un despliegue muy fuerte y denso, que permite aislar totalmente las áreas afectadas. Vale la pena mencionar que las dos primeras empresas a nivel mundial en cuanto a cantidad de patentes de EC, son chinas y de redes de energía.
- **Manufactura.** Usos típicos: IoT, asistentes robóticos y seguridad y privacidad de los datos. Hoy ya existe una evolución hacia la convergencia de las TI y las tecnologías de la operación (OT) a través de la EC, en que trabajen juntos los datos y la analítica con los sistemas que comandan las máquinas, los dispositivos y los procesos. De esta forma se agregaría a la IoT y la robótica el empleo de la IA, directamente en la producción, y en la seguridad personal. Estos aspectos son la base, a su vez, del empleo de los gemelos digitales.
- **Seguridad general en las empresas.** La EC permite mejorar mucho la seguridad a través de la visión asistida por IA y del uso extensivo de datos biométricos. Con la visión asistida se pueden determinar las condiciones seguras de trabajo, y predecir y evitar accidentes con el análisis de los movimientos del personal y de las cosas en situaciones que son identificadas como peligrosas.
- **Transporte.** Las compañías aéreas ya emplean la EC para la optimización de los costos de sus viajes, mejoramiento de los servicios de atención a los pasajeros, entre otros. Se emplea la inteligencia artificial para realizar estimaciones precisas sobre carga de los pasajeros por su peso y equipaje, capacidad de carga, condiciones de la ruta y demás datos, para optimizar el combustible a cargar, implicando ahorros acumulados importantes.
- **Petróleo y Gas.** Casi el 75% de los respondientes de esta industria dicen que están trabajando en aplicaciones sobre EC en estaciones remotas en el mar. El poder de procesamiento es local y cuando hay conexión con tierra intercambian información con el centro de datos central. De esta manera mantienen una gestión de la operación y la seguridad empleando IoT, o automatizando respuestas ante incidentes cuando es necesario.

- **Mantenimiento predictivo.** Este tipo de mantenimiento implica el empleo de dispositivos IoT en maquinaria, móviles, motores, turbinas, así como cámaras, entre otros. La EC permite el procesamiento de los grandes datos empleando IA y Machine Learning, en tiempo real y sin costoso transporte en largas distancias. Permite importantes ahorros frente a los mantenimientos programados independientemente de la situación real de los equipos, o a los reactivos ante fallas. En Alemania se emplea este modelo en el mantenimiento de las vías de tren, en que un vehículo circula por las vías capturando las imágenes que, siendo procesadas en tiempo real junto con información del vehículo y la señalización, generan alarmas tempranas de potenciales fallas. Entre otras ventajas se encuentra la inexistencia casi total de interrupciones del servicio. La IA en el Borde permite una mejora en el mantenimiento predictivo del conjunto de equipos de una planta industrial. Puede determinar a través de la IA cuándo determinada pieza puede fallar, analizando la información de sensores sobre intensidad y tipo de vibraciones, las temperaturas en puntos sensibles (rodamientos, procesadores) y sonidos que no tienen el patrón aprendido con Machine Learning.

Para finalizar, puede resultar de interés un conjunto de recomendaciones de IBM, en este mismo documento en la sección Guía de Acción, para que las empresas puedan tener en consideración para lograr las promesas del camino hacia la EC Super Integrada.

C. Impactos en las Pequeñas y Medianas Empresas. Mejoramiento de la eficiencia en la PyMEs y los cambios tecnológicos y de gestión

El impacto de la CC y de la EC en las PyMEs, ha sido analizado brevemente y en forma general a lo largo del documento, como en las secciones “3. Tecnologías emergentes relacionadas a la EC.”, cuando las PyMEs manufactureras o comerciales necesitan crecer y sofisticar la digitalización, “D. Soberanía de los datos y la EC.”, en el tratamiento de los datos personales; en su tratamiento en cuanto a la sostenibilidad y su impacto positivo sobre los costos; “D. Resumen comparativo del impacto de la EC y su relación con la CC.” donde se hace una breve referencia a las ventajas de la EC; entre otras.

En este capítulo se analizará más en detalle estos impactos.

En primer lugar, es importante considerar la digitalización de la PyMEs que en su etapa de expansión y desarrollo requiere el empleo de los recursos de la CC, y más adelante, cuando sus requerimientos sean de aplicaciones y servicios avanzados, análisis de Grandes Datos o bajos retardos, naturalmente han de evolucionar hacia la EC. Según el World Economic Forum¹²⁴ las PyMEs representan globalmente el 90% de las empresas, proveen aproximadamente el 70% de todos los empleos, y de acuerdo con estimaciones, contribuyen hasta el 70% del PBI mundial. “A nivel agregado, estas empresas desempeñan un papel importante en la habilitación, limitación y configuración de la naturaleza del crecimiento, la innovación y la sostenibilidad en nuestras economías globales, regionales y locales”.

A pesar de la enorme oportunidad que representa el mercado de las PyMEs, de acuerdo con un reciente estudio de Accenture¹²⁵, se observa que las empresas de alta tecnología no están acompañándolas totalmente en su camino hacia la transformación digital, a pesar de que aquellas están fortaleciendo su inversión en productos y servicios de TI. Según Gartner¹²⁶, en un estudio publicado en octubre de 2021 estimaba que el gasto en TI de las PyMEs llegaría a 7,5% en 2021, con un CAGR de 6,1%

¹²⁴ <https://www.weforum.org/publications/future-readiness-of-smes-mobilizing-the-sme-sector-to-drive-widespread-sustainability-and-prosperity/>.

¹²⁵ Ben Quirt. Managing Director – Strategy & Consulting, North America. 13 de setiembre de 2022. <https://www.accenture.com/us-en/blogs/high-tech/smb-the-next-growth-opportunity-for-high-tech>.

¹²⁶ <https://www.gartner.com/en/documents/4007233>.

en 2020-2025, e indicaban una rápida recuperación en el mercado global. En particular, el gasto en software fue estimado con un crecimiento CAGR del 11,4% en 2020-2025.

Accenture observa que las empresas de alta tecnología en el tratamiento de los datos (como los hiperescaladores) tienen la posibilidad de entrar a este mercado y establecer flujos de ingresos recurrentes para el futuro. Sin embargo, muchas de estas empresas no pueden operar eficientemente en este mercado porque no comprenden completamente las necesidades y expectativas de las PyMEs. A medida que estas empresas continúan cambiando hacia modelos de negocio como servicio (aaS), se están dando cuenta de la importancia del mercado de las PyMEs y buscan mejorar su oferta principal para ellas.

El cuadro 7 destaca que existe una gran diferencia de percepción de la comprensión entre las PyMEs y las empresas proveedoras de la nube en sus diferentes sectores. Este aspecto es relevante pues es parte de las dificultades principales de las PyMEs en su viaje hacia la digitalización y el uso de la CC y de la EC. No obstante, según este mismo documento, el 47% de las PyMEs emplean servicios en la nube, comparado con el 87% de las empresas más grandes, como se ha visto en el Informe de Flexera de 2023.

Cuadro 7
Las PyMEs saben que existe una desconexión entre ellas y las empresas, brecha de comprensión de la relación entre ellas

	Confianza	Afectuosidad	Entendimiento	Relacionamiento
PyMEs	78% de las PyMEs confían en las empresas con las que trabajan ^a .	60% de las PyMEs creen que las empresas se preocupan por su éxito.	69% de las PyMEs piensan que las empresas están tratando de entender sus desafíos.	59% de las PyMEs tienen un fuerte relacionamiento con las empresas con las que trabajan.
Brecha entre PyMEs y empresas.	8%, vs. 32% en un estudio basado en EE. UU.	26% vs. 33% en EE. UU.	16% vs. 19% en EE. UU.	27% vs. 25% en EE. UU.
Empresas	86% de las empresas creen que las PyMEs con las que trabajan confían en ellas.	86% de las empresas creen que su compañía se preocupa por el éxito de la PyMEs.	85% de las empresas creen que las PyMEs son el mayor foco de su compañía.	86% de las empresas creen que tienen un fuerte relacionamiento con las PyMEs con las que trabajan.

Fuente: Cuadro traducido y elaborado por el autor sobre la información de Accenture.

^a Proporción de PyMEs que no compran productos a socios empresariales. Fuente: Accenture Proprietary Research. Encuesta Global PyMEs- Empresa. <https://www.accenture.com/us-en/insights/software-platforms/winning-small-medium-sized-world>.

Las PyMEs, al igual que las grandes empresas, requieren de un cada vez mayor procesamiento y almacenamiento de la información, y el empleo de soluciones informáticas. En este camino, cuando toman en consideración los costos de la construcción, mantenimiento y actualización de un centro de datos propio, la simplificación de los accesos a sus plataformas desde cualquier dispositivo autorizado, así como evitar la exposición a la no disponibilidad de un área específica para TI que pueda estar actualizada permanentemente, las conduce a trasladar sus operaciones informáticas a la nube.

En este mismo camino, el empleo de la EC reduce los tiempos de respuesta, así como el tratamiento de los Grandes Datos, colocando a estas empresas en pie de igualdad con las más grandes en cuanto a la calidad de sus procesos informáticos. Por otra parte, evitan la dependencia de un único centro de datos bajo su propia responsabilidad, pudiendo, en caso de falla, provocar impactos muy negativos en la producción y en la actividad comercial. La EC, por su calidad de recurso compartido, aun cuando la PyME contrate una red privada, puede disponer de una estructura eficiente, óptima y resiliente frente a fallas de los sistemas. De esta manera, al existir respaldos en la EC, los múltiples accesos de la PyME pueden seguir operando sin dificultad obteniendo datos, procesando, respondiendo requerimientos de los clientes, entre otros, no siendo la falla de un sistema informático una causa de parada de la producción ni de la actividad comercial. Estas actividades de las empresas se han vuelto

muy dependientes del procesamiento de la información en tiempo real, y sin interrupciones, por lo que una falla en un único sistema propio puede provocar grandes daños.

En cuanto a la seguridad de la información, de la cual son cada vez más dependientes los sistemas informáticos debido a ataques de diversos tipos, las PyMEs pueden así lograr un equilibrio reduciendo costos y riesgos a través de la EC.

Los centros propios, para una operación no extendida, por ejemplo, para almacenamiento de datos personales, pueden proveer una seguridad y privacidad importantes siempre que se disponga de sistemas de alto nivel de protección y que se mantengan actualizados. La situación cambia cuando el centro propio se usa para múltiples actividades y con accesos desde el exterior de la empresa a través de la Internet. En este caso es necesario un gasto importante de software y recursos humanos para el manejo sofisticado y actualizado de la seguridad, por lo que resulta razonable empezar a usar la EC provista por operadores especializados. Estos servidores de EC disponen de los sistemas de seguridad y privacidad adecuados para acceder a los servidores de la EC por parte de múltiples usuarios humanos o máquinas que suelen acceder desde las más vulnerables redes WAN, como son los empleados en la calle o en las sucursales, y los dispositivos de IoT ubicados en cualquier sitio.

Ésto se logra pues la EC tiene protecciones para esta operativa, y siempre se puede seguir almacenando información en el servidor propio, con respaldo en la nube, donde la operativa a través de canales exteriores a la empresa es mucho más restringida y escasa. Así, operando con una red privada en la EC, su propio centro de datos simplificado, y en la nube, se produce una integración de los recursos centrales que físicamente se encuentran extendidos, a través de una interfase de usuario unificada, y permitiendo el empleo de nuevas tecnologías como la IA, la Machine Learning, la analítica de datos, la IoT, entre otras, cada una en el entorno más adecuado.

A través de este camino las PyMEs pueden ir adaptando su operativa, agregando además un equilibrio entre seguridad y funcionalidades.

La proximidad de los servidores de EC a los usuarios hace que las PyMEs sean menos sensibles a la calidad de las redes en cuanto a capacidad, retardo e inestabilidades y fallas, un aspecto a destacar todavía en nuestra región. La gestión de los negocios, la operación y otros aspectos han ido evolucionando hacia requerimientos de rápida respuesta e independientes de la calidad de las redes.

Tal como ya se ha analizado en general, la mejor solución para las PyMEs es la combinación de la EC propia o provista por terceros y la CC, luego de analizada la conveniencia de su uso para cada una de las actividades de la empresa teniendo en cuenta los requerimientos de procesamiento y almacenamiento, las aplicaciones, medios de comunicación con el exterior, su resiliencia, capacidad y retardo, la seguridad y privacidad, las capacidades de los recursos humanos y materiales del área de TI, entre otras.

Por otra parte, considerando los escasos recursos de IT de las PyMEs es importante el análisis cuidadoso de las tareas a trasladar a la EC, y de cuáles servicios o aplicaciones nuevas es necesario

incorporar para mejorar la eficiencia del negocio. En estos periodos transitorios de cambios bruscos, la incorporación de una cantidad excesiva de tareas en la nube, y principalmente cuando se usa el modelo multinube, puede requerir una gestión importante, para evitar aplicaciones y servicios “huérfanos”, y otros, como se ha analizado en la sección de costos, gestión que también es costosa en recursos humanos y software.

Por ello es necesario, principalmente en las PyMEs cuando se emplean nubes híbridas, disponer de plataformas de gestión de nueva generación, que optimicen los costos y el funcionamiento independientemente del ambiente en que corren las aplicaciones y servicios, trabajando sobre la visibilidad y el control del uso.

Observando más ampliamente la EC, el intenso despliegue actual de la conectividad de acceso 5G, con algunas demoras coyunturales en nuestra región, puede impulsar más la puesta a disposición de cualquier empresa el empleo de avanzados servicios por parte de los operadores de EC como aplicaciones SaaS, IA, Machine Learning y similares. Si a ésto se agrega que los propios CSP están entrando fuertemente en el mercado de EC, como se analiza en la sección “B. Los CSP como usuarios de la MEC y proveedores de servicios sobre ella”, se crea el entorno para que los CSP puedan proveer la MEC a todas las empresas, pero particularmente a las PyMEs, a quienes pueden ofrecer paquetes muy atractivos de EC y conectividad 5G. Esta simplificación de las condiciones para la contratación de EC directamente a los CSP, sean servicios de terceros o propios, es un catalítico para el uso de la EC por parte de las PyMEs. Este efecto se observará en cuanto se extiendan estas ofertas en el mundo y particularmente en nuestra región.

Quizás, la oferta actual de servicios de Verizon¹²⁷ de EE. UU. a las PyMEs, sea la más ilustrativa en cuanto a la visión de cómo los CSP pueden impulsar el uso de la nube en general, y especialmente de la EC, simplificando su incorporación y gestión. Se analiza la oferta de esta empresa en particular, considerando que es la octava empresa en el mundo (incluyendo las de China), y el único CSP en las 10 primeras, en cantidad de patentes anuales de EC.

Ofrece un servicio denominado “Network as a Service” o NaaS, Red como Servicio, cuyas ventajas para las PyMEs se indican para sus potenciales clientes:

“La mayoría de los líderes de las PyMEs están centrados en sus negocios principales: sus clientes, sus empleados, sus mercados locales y nacionales. No hay suficientes horas en el día para pensar en su red u otros servicios de TI centrales, aunque sepa que es importante. Además, por lo general, no hay un gran equipo de TI para ayudar, si es que hay alguno.

Aquí es donde la Red como Servicio puede ayudar. Continúa permitiendo el acceso a software e infraestructura modernos de una manera rentable, sin requerir un personal de TI de tamaño empresarial para respaldarlo.

El modelo de Red como Servicio ofrece infraestructura y servicios de red mediante un modelo de suscripción, similar al software basado en la nube”.

Y en cuanto a las razones para que las PyMEs maximicen el impacto de la EC sobre sus negocios, adoptando como una alternativa de su evolución la llamada Red como Servicio, se encuentran las siguientes provistas por Verizon, las que son totalmente válidas en el mundo para las PyMEs y los CSP en general, y se encuentran en consonancia con los esfuerzos del ETSI, el TM Forum y el 3GPP, entre otras destacadas organizaciones que propician la EC multiacceso, o MEC.

Eficiencia operativa. Este modelo de servicio permite implementar una red de vanguardia en un modelo de suscripción o gastos operativos (OPEX) que tenga el tamaño adecuado para el negocio sin incurrir en las inversiones y gastos de una red propia que puede resultar prohibitivo para una PyME.

Ahorro de tiempo. La disponibilidad de la red requerida pasa solamente por contratar el servicio, pues esas redes compartidas ya se encuentran desplegadas y son escalables.

Máxima experticia. La dirección de la PyME puede concentrarse en su negocio y aprovechar el conocimiento de expertos confiables para guiar la transformación de su red.

Escalabilidad y flexibilidad. La Red como Servicio es muy flexible y puede ampliarse o reducirse de acuerdo con la evolución de los requerimientos, muchas veces impredecible.

¹²⁷ <https://www.verizon.com/business/resources/articles/s/what-are-the-benefits-of-cloud-infrastructure-for-smb/>.

Catalizador de la innovación. Indica, como ya se ha visto, que el empleo de la infraestructura en la nube es que es un medio para evolucionar innovando en la producción, en la gestión en general y en los servicios, y no es un acto único.

Esta visión del impacto de la EC sobre las PyMEs, a través del eficiente uso de la MEC, revela un importante camino de las PyMEs para su incorporación a la digitalización de impacto al mismo nivel que las grandes empresas.

D. Conjunción de la EC, la Internet de las Cosas y la 5G en el desarrollo productivo

El desarrollo de la IoT en las aplicaciones industriales, comerciales y de servicios, y principalmente en las primeras, se ve apoyado en los accesos finales de las telecomunicaciones, principalmente móviles debido a las características de los terminales, y en la EC. Los avances de la IoT gozan de tasas de crecimiento muy importantes, y en la medida que se han ido sofisticando requieren un mayor poder de procesamiento de bajo retardo y alta capacidad, lo que requiere necesariamente de la EC. La movilidad de los dispositivos como los robots o las máquinas en zonas que no permiten la conectividad fija, lo que es una situación muy generalizada en la automatización, hacen necesario tener movilidad con alta capacidad de datos, por lo que la 5G es la adecuada.

De esta forma la EC y la 5G tienen un alto impacto en el desarrollo de la digitalización en la producción ya que sin ellas en su pleno desarrollo no sería viable. Solo ambas operando simultáneamente pueden mantener un vínculo permanente, y de capacidad suficiente de transmisión y procesamiento, para mantener optimizada la operación de una planta industrial o de logística, asegurando además comunicaciones y respuestas de seguridad en tiempo real. En cuanto a ambas tecnologías, éstas podrán ser provistas bajo la modalidad de redes privadas, privadas dentro de operadores públicos o directamente públicas. La 5G brindará toda su potencia cuando se emplee el espectro de las bandas milimétricas, algo que por el momento está mayormente limitado a redes privadas, cobertura de zonas especiales como parques industriales o la cobertura de empresas que dependen esencialmente de disponer de velocidades de varios Gbps.

En conclusión, estas tres familias tecnológicas son el sustento básico de este desarrollo, pero además son el sustento de las otras familias que se apoyan en ellas, como la IA que es esencial en las operaciones complejas y que ahora tienen sus propios AI chips desarrollados para su instalación en los dispositivos, la Machine Learning como sub sector de la IA, los gemelos digitales que permiten mejorar el funcionamiento de las máquinas y bajar el riesgo de daño y reducir los defectos, potenciados con la IA, entre otras ya analizadas en este documento.

El requerimiento conjunto de EC y 5G es un factor desencadenante del interés de los CSP para la venta de estos paquetes a las grandes empresas, adicionalmente al interés que despierta para la PyMEs como ya se ha analizado. Adicionalmente la EC es muy importante en los despliegues de la 5G que en forma nativa se asienta en la nube. Inclusive la propia necesidad de soportar sus operaciones con EC le está dando a los CSP la infraestructura de base para prestar servicios de EC a terceros.

La combinación de estas tres familias se sustenta también en la CC para los tipos seleccionados de procesamiento y almacenamiento, superadas las instancias de procesamiento expeditivo en el borde, y la reducción de la información que resulta económicamente más eficiente almacenar en la nube.

Esta estructura básica es el resultado de las evoluciones tecnológicas de estos últimos pocos años y que están acelerando la digitalización y su uso muy eficiente.

El entramado de las comunicaciones entre los dispositivos de IoT y los centros de procesamiento, o entre ellos, sobre la infraestructura de telecomunicaciones, y principalmente 5G, requiere de un protocolo de mensajería estándar que asegure el entendimiento entre extremos y que provea una seguridad mínima independientemente del fabricante o modelo. Para ello se emplea el protocolo estándar de hecho, denominado MQTT¹²⁸, que fue diseñado “como un transporte de mensajería de publicación/suscripción extremadamente ligero que es ideal para conectar dispositivos remotos con una huella de código pequeña y un ancho de banda de red mínimo. Hoy en día, MQTT se utiliza en una amplia variedad de industrias, como la automotriz, la manufacturera, las telecomunicaciones, el petróleo y el gas, y otras”.

Según AWS: “MQTT utiliza un patrón de publicación o suscripción para desacoplar el remitente del mensaje (editor) del receptor del mensaje (suscriptor). En lugar de ello, un tercer componente, denominado agente de mensajes, controla la comunicación entre editores y suscriptores. El trabajo del agente consiste en filtrar todos los mensajes entrantes de los editores y distribuirlos correctamente a los suscriptores. El agente desacopla los editores y suscriptores. ...El editor y el suscriptor no conocen la ubicación de la red del otro y no intercambian información como direcciones IP o números de puerto”.

¹²⁸ www.mqtt.org.

VI. Requerimientos de política pública para el despliegue y adopción de la EC

Vista la importancia de la Nube en general, y más reciente y particularmente de la EC en el mundo, es esencial conocer el contenido de las políticas públicas relevantes para considerar su posible aplicabilidad en el desarrollo en cada país de LAC. En esta sección se analiza la planificación y estructuración de la política pública de la Unión Europea, considerando que representa un esfuerzo temprano de alto nivel y de amplio alcance, extensamente documentado, y que además abarca simultáneamente varios países, lo que permite ver su importancia en las consideraciones transfronterizas de la Computación en la Nube y el papel de la EC, lo cual ya ha sido desarrollado en algunos aspectos en este documento.

A. Alianza Europea para los datos industriales, Edge y Cloud

La “Alianza Europea para los Datos Industriales, Edge y Cloud”¹²⁹ es un proceso que empieza formalmente el 19 de febrero de 2020 con la publicación de la Estrategia Europea de Datos¹³⁰ que se podría considerar la columna vertebral del desarrollo del tratamiento amplio, eficiente y seguro de los datos.

El objetivo de esta Estrategia es convertir a la Unión Europea en líder de una sociedad orientada a los datos. La creación de un mercado único de datos permitirá que estos fluyan libremente por la Unión y entre sectores, en beneficio de las empresas, los investigadores y las administraciones públicas.

Un pilar clave de esta estrategia es la Ley Europea de Gobernanza de Datos^{131 132}, la que busca aumentar la confianza en el intercambio de datos, reforzar los mecanismos para aumentar la

¹²⁹ <https://digital-strategy.ec.europa.eu/es/policies/cloud-alliance>.

¹³⁰ https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_es.

¹³¹ <https://digital-strategy.ec.europa.eu/es/policies/data-governance-act>.

¹³² Reglamento (UE) 2022/868 del Parlamento Europeo y del Consejo de 30 de mayo de 2022 relativo a la gobernanza europea de datos y por el que se modifica el Reglamento (UE) 2018/1724 (Reglamento de Gobernanza de Datos) (Texto pertinente a efectos del EEE). <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R0868>.

disponibilidad de datos y superar los obstáculos técnicos a la reutilización de los datos, a los efectos de crear el mayor valor posible para la sociedad a través del empleo de las cuantiosas cantidades de datos valiosos generados en la Unión.

La entrada en vigor es muy reciente, en 2022, y luego de un período de gracia de 15 meses, es aplicable desde setiembre de 2023.

El valor que crea esta Ley surge del apoyo a la creación y el desarrollo de espacios de datos europeos comunes en ámbitos estratégicos, en los que participen tanto agentes privados como públicos, en sectores como la salud, el medio ambiente, la energía, la agricultura, la movilidad, las finanzas, la fabricación, la administración pública y las competencias.

“La iniciativa tiene como objetivo poner a disposición más datos y facilitar el intercambio de datos entre sectores y países de la UE con el fin de aprovechar el potencial de los datos en beneficio de los ciudadanos y las empresas europeos”.

Responde a una necesidad para el desarrollo, en cuanto a disponer de datos suficientemente anonimizados, según los casos, para mejorar múltiples actividades que se benefician de la buena gestión y puesta en común de los datos (datos sanitarios, ubicación de personas, entre otros).

Esta Ley (Reglamento), según lo establece en el Considerando (4), “no debe interpretarse en el sentido de que establezca una nueva base jurídica para el tratamiento de los datos personales para cualquiera de las actividades reguladas o que modifique los requisitos de información dispuestos en el Reglamento (UE) 2016/679.” Este último Reglamento es analizado en este documento en la sección I.D. Soberanía de los datos y la EC”.

Desde el lanzamiento de la Estrategia Europea de Datos de 2020 se han producido múltiples eventos de esta “Alianza Europea para los Datos Industriales, Edge y Cloud”, pasando por el anuncio de esta Alianza por parte de la Comisión, el 5 de mayo de 2021, por la primera reunión de trabajo del 16 de diciembre de 2021, y llegando al evento de la Asamblea General de la Alianza y el Foro de la Alianza del 4 de julio de 2023, cuando se publican los dos Primeros Resultados de los Grupos de Trabajo¹³³. Éstos son “... la hoja de ruta industrial estratégica actualizada de la computación en el borde, y los requisitos para una plataforma basada en la nube para el sector de la aeronáutica, la seguridad y la defensa.” La Hoja de Ruta actualizada es la que interesa a nuestros efectos.

Los documentos más relevantes, y que se van a analizar son el anuncio y la Declaración de la Alianza, y la Hoja de Ruta, por ser los que contienen los contenidos más pertinentes para este trabajo.

1. Anuncio de la Alianza Europea para los datos industriales, Edge y Cloud

Este documento es importante en cuanto a presentar los objetivos y su estructura, la que es abierta a las partes interesadas incluyendo empresas, representantes de los Estados miembro y expertos pertinentes.

Indica que “Las tecnologías de nube y de borde son facilitadores estratégicos de innovación para la adopción de tecnologías emergentes, como la inteligencia artificial, el Internet de las Cosas y 5G. Proporcionan la infraestructura para casos de uso altamente innovadores. Europa debe reforzar su posición de la industria de la UE en relación con las tecnologías en la nube y las tecnologías de vanguardia.

Las tecnologías de nube y de vanguardia son factores clave para la transformación digital de Europa. La Alianza tiene como objetivo reunir a las partes interesadas pertinentes del sector privado y público para definir conjuntamente hojas estratégicas de ruta de inversión que permitan la próxima generación de tecnologías informáticas altamente seguras, distribuidas, interoperables y eficientes en el uso de los recursos. Además, la Alianza servirá como plataforma de intercambio sobre temas de gobernanza de la nube, por ejemplo, en relación con la contratación pública de servicios en la nube.”

¹³³ <https://digital-strategy.ec.europa.eu/es/news/european-alliance-industrial-data-edge-and-cloud-presents-its-first-deliverables>.

“Para alcanzar los objetivos mencionados, la Alianza trabaja en la siguiente lista no exclusiva de tareas:

- reunir a los agentes pertinentes para preparar y actualizar hojas de ruta de inversión horizontales y tecnológicas específicas, en la nube y en los bordes;
- formular recomendaciones para garantizar la integración coherente de las inversiones con las previstas para el despliegue de espacios comunes europeos de datos en los ámbitos pertinentes;
- asesorar a la Comisión sobre los requisitos y normas aplicables a los servicios en la nube, incluida la contratación pública.”

Se observa que esta Alianza trabaja con la plena participación de las partes interesadas pertinentes, con cuidados especiales en cuanto a la aceptación de entidades sujetas al control de un tercer país, en cuanto a la presentación de la justificación de que no violan el marco de la protección de datos de la Unión Europea.

“La Alianza contribuirá a configurar la próxima generación de servicios e infraestructuras en la nube y de vanguardia seguros, hipo carbónicos e interoperables para Europa, tal como se prevé en la Estrategia Europea de Datos.”

2. Declaración de la Alianza Europea para los Datos Industriales, Edge y Cloud

Cualquier organización con actividades de relevancia significativa para la provisión de la nube y el procesamiento de datos altamente seguros puede unirse a la Alianza en cualquier momento firmando la Declaración, siempre que cumpla con los criterios de elegibilidad determinados en los Términos de Referencia¹³⁴.

Esta Declaración incluye sus objetivos principales, los documentos a emitir para sustentar el desarrollo de la nube, el borde y el procesamiento de los datos, así como los principios que la rigen. Estos asuntos, que han surgido del trabajo conjunto de los países y las partes interesadas, pueden ser considerados como referencias importantes para todos aquellos países que deseen emprender el camino de este desarrollo.

Se pasa a transcribir los aspectos clave de la Declaración.

En este párrafo se observa la importancia que se le otorga a la distribución y descentralización de las innovadoras capacidades de procesamiento, procurando alcanzar los 10.000 sitios en el borde para el 2030, seguros y sostenibles, que garanticen el acceso con baja latencia en cualquier lugar en que la organización se encuentre.

“La Alianza tiene por objeto reforzar la posición de la industria de la UE en materia de tecnologías y capacidades en la nube y en el borde. Debe hacerlo respondiendo a las necesidades de la industria y el sector público europeos en términos de tratamiento de datos personales sensibles o muy sensibles de empresas y del sector público, abordando al mismo tiempo la tendencia hacia una mayor distribución y descentralización de capacidades innovadoras de tratamiento de datos. La Alianza apoyará la consecución de los objetivos de la década digital de la UE para 2030: el establecimiento de centros de datos sostenibles y climáticamente neutros, altamente eficientes en el uso de los recursos y la energía, el despliegue de 10 000 nodos de borde de alta seguridad y climáticamente neutros en toda la UE y el aumento del porcentaje de empresas europeas que utilizan servicios avanzados de computación en nube en Europa hasta el 75 %”.

Aparte de los objetivos indicados anteriormente, se agregan los siguientes:

¹³⁴ <https://digital-strategy.ec.europa.eu/en/policies/cloud-alliance>.

- “El desarrollo y despliegue de ecosistemas de datos, informáticos e industriales que apoyen la transformación digital de la industria y el sector público europeos, y que cumplan plenamente el marco legislativo de la UE y las normas más estrictas de la UE en materia de tratamiento de datos.
- El desarrollo y despliegue de una oferta europea de infraestructuras y servicios en la nube y en el borde, competitivos de próxima generación, incluidas plataformas y servicios en la nube completamente interoperables, abiertos y de múltiples proveedores.
- El desarrollo y despliegue de una próxima generación de centros de datos europeos en la nube y en el borde, resilientes y sostenibles, así como de nodos relacionados a ellos, interconectados, neutros en carbono, altamente seguros y distribuidos, tanto para el sector público como para el privado.
- Ofrecer apoyo y recomendaciones para la definición de normas y requisitos comunes para los servicios en la nube que operan en el mercado de la UE, satisfaciendo en particular las necesidades de la administración pública y las empresas que tratan con datos personales sensibles o muy sensibles de las empresas y del sector público, incluidos los sectores de la seguridad y la defensa”.

En cuanto a las acciones y documentos concretos a producir para conseguir los objetivos encomendados, los definen de la siguiente manera:

- “Crear una plataforma para aprovechar las sinergias de inversión de la industria en la investigación, el desarrollo y el despliegue de la próxima generación de tecnologías en la nube y en el borde de la UE, eficientes en el uso de los recursos, interoperables, altamente seguras y fiables que satisfagan las necesidades de las empresas y el sector público europeos que se ocupan de datos personales sensibles y de empresas y del sector público altamente sensibles (como la nube de grado militar).
- Recomendaciones detalladas (por ejemplo, casos de uso; hojas de ruta estratégicas de tecnología e inversión; análisis de mercado y cortes en la cadena de valor) y servir de plataforma de emparejamiento para las empresas y las autoridades públicas sobre las inversiones de la UE¹³⁵ en capacidades de tratamiento de datos de próxima generación.
- Una plataforma permanente para la coordinación entre la Comisión y los Estados miembros de la UE sobre el uso de la nube en el sector público y la contratación pública en la nube, en la que participen representantes de todos los Estados miembros de la UE.
- Crear una plataforma de consulta de las partes interesadas a la Comisión Europea sobre reglas y normas técnicas comunes para los servicios en la nube que operan en el mercado de la UE, en particular con vistas a la elaboración del código normativo de la UE en la nube, y la aportación de conocimientos especializados sobre normas y requisitos comunes para la contratación pública de servicios en la nube, tal como se prevé en la Estrategia Europea de Datos, como parte de un proceso de consulta de la Comisión abierto a todas las partes interesadas.
- Crear una plataforma para desarrollar sinergias con los espacios comunes europeos de datos, en particular los que albergan datos personales sensibles o datos muy sensibles del sector público y de las empresas que requieren requisitos de seguridad elevados. Esta plataforma de coordinación debe reunir los requisitos técnicos para las infraestructuras y los servicios en la

¹³⁵ La inversión en tecnologías de procesamiento de datos en la nube, en el borde y otras tecnologías de procesamiento de datos está especialmente prevista en los instrumentos de financiación, como Horizonte Europa, el Programa Europa Digital, el Mecanismo «Conectar Europa» 2, el Fondo Europeo de Defensa y el Mecanismo de Recuperación y Resiliencia (iniciativa emblemática «Scale up»), junto con los programas nacionales y la inversión del sector privado.

nube o en el borde para esos espacios de datos, y garantizar las sinergias y la interoperabilidad entre los diferentes espacios de datos en estrecha coordinación con el Centro de Apoyo para el intercambio de datos y el Consejo Europeo de Innovación en Datos”.

El análisis profundo de los documentos relacionados a esta Alianza, más allá de lo expresado en este trabajo, podría permitir obtener antecedentes adicionales para la formulación de políticas para el desarrollo de la nube en el borde y de las tecnologías de vanguardia. El documento que se analiza a continuación contiene una estructura muy detallada de los principales aspectos a tener en consideración.

3. Hoja de ruta de la tecnología industrial europea para la próxima generación Cloud-Edge

Este documento fue preparado por el Grupo de Trabajo Cloud-Edge, y entregado a la Comisión el 4 de julio de 2023¹³⁶, el que juntó a los principales jugadores de la CC de Europa para la preparación de esta hoja de ruta, y trabajó para identificar inversiones para el desarrollo y el despliegue conjunto de la próxima generación de tecnologías en la nube y en el borde de la UE que satisfagan las necesidades de las empresas europeas.

Se presentan en él lo que se puede entender como una muy amplia guía maestra que contiene los conceptos y acciones necesarios para orientar en cortos plazos, compatibles con las exigencias y evolución actual de la computación en la nube y en el borde, en un continuo de soporte a los desarrollos, las tecnologías y los despliegues, que cualquier país puede considerar en cuanto a su adaptabilidad o no, en acciones y plazos, a sus necesidades de acompañar los avances globales.

La extensión y profundidad del documento lo hacen muy apto para observar soluciones con gran granularidad, lo que facilita la consideración de acuerdo con cada país.

Se presenta un muy breve resumen de su contenido y el enlace al documento original.

Esta hoja de ruta “proporciona una visión colectiva de los dominios tecnológicos y las dimensiones relacionadas que requieren una inversión estratégica para permitir el desarrollo y el despliegue conjuntos de servicios en la nube y en el borde competitivos, seguros, fiables y climáticamente neutros en toda Europa para construir el continuo nube-borde de la próxima generación. Además, este documento proporciona una descripción actualizada de las principales iniciativas, asociaciones, estándares y proyectos de código abierto europeos para el borde y la nube. También incluye un análisis de las oportunidades y desafíos de la soberanía digital.

La hoja de ruta ofrece las principales prioridades en dominios transversales, fundamentos tecnológicos y servicios específicos del sector impulsados por casos de uso y facilitadores relevantes, abordando dos casos de prioridades:

- Prioridades tecnológicas con actividades para la especificación, investigación, desarrollo e innovación de las capacidades continuas de la nube de próxima generación.
- Prioridades de despliegue con actividades para la coordinación y el despliegue de infraestructuras y servicios de la nube al borde para permitir el despliegue inicial de casos de uso de próxima generación a escala europea”.

Estas prioridades se agrupan, a su vez, en tres dominios:

- Convertirse en el líder en los dominios tecnológicos transversales que darán forma a las ofertas europeas de nube y borde en el mercado global, y que aseguren la neutralidad de carbono, ciberseguridad, interoperabilidad y otros.

¹³⁶ <https://joinup.ec.europa.eu/collection/egovernment/news/eu-industrial-technology-roadmap-cloud-edge>.

- Renovación y ampliación de los cimientos de infraestructuras en toda Europa, que incluye en el mismo orden de importancia el despliegue denso de los centros de procesamiento ubicuos y las redes 5G.
- Habilitación de servicios soberanos y sectoriales a los usuarios finales a través de la oferta de soluciones de software soberano y de acceso abierto, y un ecosistema de servicios de datos y aplicaciones avanzadas.

“Aunque la hoja de ruta contiene un gran número de prioridades tecnológicas y de despliegue importantes, algunas de ellas son especialmente críticas para ofrecer una solución para el continuo del borde a la nube en Europa que garantice:

- Competitividad: es decir, un bajo Costo Total de Propiedad (TCO), una solución que es conveniente y fácil de usar. Prioridades como la orquestación y la automatización serán esenciales para lograrlo.
- Soberanía digital: cierto control sobre la solución tecnológica, garantizando su mantenimiento y evolución. Prioridades como el desarrollo de código abierto o los mecanismos de federación permitirán la entrega de una solución basada en componentes tecnológicos europeos clave por parte de proveedores de servicios europeos, federados para proporcionar un servicio paneuropeo y, en el futuro, global.
- Mecanismo de puesta en marcha: para garantizar que la escala se alcance rápidamente para poder desempeñar un papel en el panorama mundial, será esencial la participación de sectores críticos en las primeras etapas, como los servicios públicos o los servicios de telecomunicaciones. Otros sectores más fragmentados, como el sector industrial, se beneficiarán de este impulso y aportarán la escala necesaria para consolidar la solución”.

B. Documento de trabajo de los operadores de red de telecomunicaciones de Europa

Este documento¹³⁷ hace foco en aspectos muy destacados de políticas que son complementarios de los analizados anteriormente, provenientes de la Alianza y su trabajo sobre las prioridades tecnológicas y de despliegue para conformar el continuo Nube – Borde que requiere para el desarrollo, para orientar las inversiones estratégicas requeridas, y también son la base sobre la cual se asientan las regulaciones del mercado de la nube en Europa.

Este documento, desde la óptica de los CSP, y también las referencias que se incluyen en otras partes de este estudio, son una fuente abundante de reglamentaciones e iniciativas que son imprescindibles para el desarrollo del modelo tan complejo del continuo Borde – Nube en condiciones de seguridad, eficiencia y eficacia.

A su vez, alguna de las referencias abre un abanico de nuevas regulaciones e iniciativas a ser desarrolladas en el futuro.

Se considera que las personas interesadas en las políticas, reglamentaciones e iniciativas, y especialmente los CSP, podrían disponer de valiosa información de referencia en este documento, al igual que en los ya analizados de la Alianza.

¹³⁷ Edge Cloud and Strategic Capacity Building in the EU Policy Context on Cloud. ETNO, enero de 2023. etno.eu.

La Unión Europea está considerando el desarrollo de su continuo de Nube – Borde con las orientaciones indicadas anteriormente, por lo que el acceso a él resulta ser un objetivo de política esencial que se encuentra en proceso de definición debido a la etapa de desarrollo en que se encuentra en la Unión y en el mundo.

Dice la ETNO: “La creciente atención de los responsables políticos ha llevado a la aparición de lo que puede describirse como un primer conjunto completo de regulaciones e iniciativas dirigidas al mercado de la nube en Europa. Varias iniciativas con impacto en la nube ya están adoptadas (DMA¹³⁸, NIS-Directive¹³⁹), actualmente están en preparación (IPCEI CIS¹⁴⁰, EUCS¹⁴¹, Data Act¹⁴²) o aún no se han propuesto (Cloud Rulebook¹⁴³, Cloud Marketplace¹⁴⁴, Data Act sector-specific regulations). En medio de este panorama de políticas en rápida evolución, este documento tiene como objetivo explicar qué es la nube de borde, por qué es relevante y cómo encaja en el contexto de políticas más amplio de la UE. Además, en este documento se explica el papel único de los proveedores de telecomunicaciones europeos en la consecución de un continuo Borde-Nube”.

La Asociación ETNO provee una visión desde los proveedores de infraestructura y servicios de conectividad que hacen viable el cambio de paradigma de los datos en que se estima que el 80% van a ser procesados en el borde, en lo que se suele denominar el continuo Borde-Nube.

Por ello el documento de la ETNO es también de interés en cuanto a la aplicación de la MEC por parte de los CSP, en el marco de la situación económica y social de LAC.

“El concepto de “Multi-Access Edge Computing” (MEC) permite a los proveedores de servicios acceder y utilizar el continuo Borde - Nube de una manera fácil para implementar y activar/ejecutar servicios. Una capa común de “federación de nube de borde”, que actualmente es el objetivo de IPCEI- CIS, es una parte esencial cuando se trata de abarcar servicios que utilizan la infraestructura de múltiples proveedores de infraestructura, incluidos los proveedores de telecomunicaciones”.

La Comisión Europea ha destacado la posición estratégica de los CSP: “Utilizando la experiencia de los proveedores de telecomunicaciones europeos combinada con la especialización de los proveedores europeos de servicios en la nube, Europa puede construir soluciones de telecomunicaciones en la nube, especialmente en el borde, aprovechando el despliegue conjunto de infraestructuras 5G y

¹³⁸ Digital Market Act. https://digital-markets-act.ec.europa.eu/index_en.

¹³⁹ La Directiva NIS2 es la legislación de la UE en materia de ciberseguridad. Proporciona medidas legales para impulsar el nivel general de ciberseguridad en la UE. <https://digital-strategy.ec.europa.eu/en/policies/nis2-directive>.

¹⁴⁰ Este proyecto que es indicado como en proceso de preparación a enero de 2023, ya ha sido aprobado por la Comisión el 5 de diciembre de 2023 por un monto de hasta 1.200 millones de euros en concepto de ayudas estatales de siete Estados miembros a un proyecto importante de interés común europeo sobre tecnologías de computación en la nube y en el borde. El proyecto llamado PIICE, de infraestructura y servicios en la nube de nueva generación (IPCEI CIS, por sus siglas en inglés), ha sido notificado conjuntamente por siete Estados miembros: Alemania, España, Francia, Hungría, Italia, Países Bajos y Polonia. https://ec.europa.eu/commission/presscorner/detail/en/ip_23_6246.

¹⁴¹ EU Cloud Certification Scheme. ENISA lanza un esquema de certificación de ciberseguridad para armonizar la seguridad de los servicios en la nube con la normativa de la UE, las normas internacionales y las mejores prácticas del sector. <https://www.enisa.europa.eu/> - <https://ec.europa.eu/newsroom/cipr/items/713799/en>.

¹⁴² Esta Ley de Datos, indicada como en proceso de preparación en enero de 2023, ha sido políticamente acordada el 28 de junio de 2023 entre el Parlamento y el Consejo. Ahora se encuentra en proceso de ser formalmente aprobada por ambos colegisladores. La Ley de Datos entrará en vigor el vigésimo día siguiente al de su publicación en el Diario Oficial y será aplicable a los veinte meses de su entrada en vigor. https://ec.europa.eu/commission/presscorner/detail/es/ip_23_3491.

¹⁴³ El Cloud Rulebook, o Código Normativo de la Nube se debe a que La UE tiene previsto elaborar un conjunto de normas, en forma de un código normativo de la UE en la nube y una guía sobre la contratación pública de servicios de tratamiento de datos. El código normativo proporcionará un marco Único europeo aplicable a las normas vinculantes y no vinculantes para los usuarios y proveedores de servicios en la nube en Europa. Para aumentar la eficiencia y la calidad de la contratación pública de servicios de tratamiento de datos en Europa, las orientaciones propondrán recomendaciones para aplicar políticas nacionales coherentes, complementadas por un conjunto completo de criterios esenciales para los servicios de tratamiento de datos que deben considerar los organismos del sector público durante el proceso de licitación. <https://digital-strategy.ec.europa.eu/es/policies/cloud-computing>.

¹⁴⁴ <https://digital-strategy.ec.europa.eu/en/library/building-european-cloud-marketplace-conceptualisation-study>.

Computación en el Borde, de una manera que cumpla con los requisitos ambientales y de seguridad del mercado europeo¹⁴⁵.

Manifiesta este documento de la ETNO la preocupación por la reducción progresiva de la competencia en el negocio de la nube: "Europa es competitiva en lo que respecta a la generación de datos industriales, por ejemplo, en el ámbito del IoT. Sin embargo, el almacenamiento y el procesamiento de datos siguen siendo llevados a cabo en gran medida por empresas no europeas, sobre la base de una infraestructura de la que los proveedores europeos solo tienen una parte marginal. En su revisión en profundidad del mercado de la nube que acompaña a la Estrategia Industrial de la UE (2021), la Comisión Europea señala que "el mercado de la infraestructura de la nube pública está convergiendo a nivel mundial en torno a cuatro grandes empresas no europeas", representando más del 80% de los ingresos globales en 2021¹⁴⁶". Expresa también la posición de la ETNO respecto de las dificultades de los CSP para competir con los hiperescaladores.

Europa presta una especial atención al mercado de los operadores de la nube, tanto desde el punto de vista del posicionamiento estratégico del complejo económico europeo en el mundo, como desde el punto de vista de la seguridad de los datos. En este marco se han abierto investigaciones del Supervisor de la Protección de Datos de la Unión Europea sobre un par de hiperescaladores. También en otros países como el Reino Unido y los EE. UU. se está investigando la competencia en los mercados digitales.

C. Aplicabilidad en la región

Los documentos y referencias incluidos en las secciones A y B incluyen los más recientes avances de la Unión Europea a partir de 2020 y 2021, y predominantemente en este último año. Son, además, un conjunto denso de principios, marcos normativos, procedimientos, recomendaciones, y demás asuntos a ser incluidos en la formulación de políticas abarcativas con relación a la computación en la nube y muy especialmente en el borde, para desarrollar una infraestructura ubicua, segura y eficiente de almacenamiento y procesamiento de datos. El marco para el desarrollo de esta infraestructura cumple gran cantidad de condiciones que se entiende que permitirán a la Unión Europea, no solamente soportar el rápido despliegue de tecnologías emergentes, sino también constituir un bloque de países con un mercado único, eficiente y seguro de datos, procurando un avance en la independización de infraestructuras extranjeras, y lograr el posicionamiento para un desarrollo global.

Si bien LAC no se encuentra en el mismo nivel de desarrollo económico y social, ni en la preparación para las tecnologías de vanguardia, como ya se ha presentado en el cuadro 2, dicha documentación de la Unión Europea puede ser una referencia importante para orientar los pasos requeridos en cuanto a las políticas públicas para desarrollar la EC, y más allá, para crear un mercado regional de datos altamente eficiente y con la participación de los CSP. Si bien el objetivo de este estudio es la EC, se analizan a continuación otros aspectos que son de posible impacto negativo en el desarrollo, más allá de la formulación de políticas específicas.

Como ya fue estudiado, en paralelo con las decisiones que los Gobiernos puedan tomar, la EC en LAC se encuentra en desarrollo dispar en los diferentes países impulsada principalmente por los requerimientos de los mercados y las acciones propias de los proveedores de la nube y los CSP, todo ésto estrechamente correlacionado con el nivel de la preparación para las tecnologías de vanguardia. En este estudio se ha visto que, si bien existe un interés en la EC en las grandes organizaciones y en determinados países, no es la misma situación en toda la región.

¹⁴⁵ Commission Staff Working Document, "Strategic Dependencies and Capacities", SWD (2021) 352, 2021.

¹⁴⁶ Commission Staff Working Document, "Strategic Dependencies and Capacities", SWD (2021) 352, 2021.

Los siguientes aspectos necesariamente deberían ser tenidos en consideración, con diferente grado de profundidad, según la situación de cada país y la demanda que provenga de las organizaciones de cada uno, ya que son la base que permiten la implantación de las nuevas tecnologías.

El principal reto tecnológico en LAC en relación con la EC, por su impacto observado en el capital humano requerido para el desarrollo, es la escasez de talento en TI que dificulta la adopción de las tecnologías emergentes por parte de las organizaciones (demanda que permite el avance de la productividad), así como el desarrollo de aplicaciones y servicios de procesamiento en la nube que permitan su empleo (oferta). En muchos países de LAC esta carencia incide inclusive muy profundamente en la innovación, la investigación y el desarrollo, así como en el desarrollo económico en general, donde se notan hasta importantes dificultades para la radicación de empresas en el proceso denominado “nearshoring”. Es una situación retroalimentada que retrasa el desarrollo del país. Se destaca que, si bien esta escasez se aprecia en todo el mundo a pesar de las medidas que se están adoptando para el cambio educativo formal y la formación permanente de los trabajadores, ella es más impactante en LAC. Esta situación es por demás compleja pues la formación de capital humano puede no estar equilibrada con la demanda en su evolución, produciendo la fuga al extranjero, en este momento facilitada por el trabajo remoto.

Otro aspecto de impacto en este desarrollo de la EC es el de la ciberseguridad y el de la protección de datos, cuyas políticas y herramientas se encuentran desarrolladas en forma dispar en LAC.

Finalmente, considerando el papel esencial de la conectividad en la EC y en su infraestructura para el despliegue de aplicaciones, servicios y otras tecnologías que se basan en ella, el disponer de una red de conectividad de la velocidad y resiliencia adecuadas a estos avances que se registran en el mundo, también es un factor para analizar y solucionar para alcanzar los desarrollos requeridos de la digitalización.

En conclusión, se entiende que para un mejor desarrollo de la EC en LAC, y con ello la productividad y la expansión del desarrollo económico, sería necesario atender al menos los tres asuntos mencionados, como soporte básico para los despliegues extendidos de la EC y las tecnologías relacionadas, y adoptar políticas específicas, para las cuales la extensa documentación presentada de la Unión Europea puede constituir una referencia amplia de una estructura sólida, completa y evolucionada para fomentar el desarrollo específico de la EC en nuestra región.

Anexo A1

Unión Europea, EC e IOT

Se incluye como anexo el enlace a un importante estudio realizado para la Comisión Europea, denominado “Study on the Economic Potential of Far Edge Computing in the Future Smart Internet of Things”, publicado a fines de 2023¹⁴⁷.

Según se indica en esta publicación: “Este estudio detalla tanto las oportunidades económicas como los impactos ambientales de un cambio de paradigma en el dominio del Internet de las Cosas (IoT) hacia la computación en el borde. Los objetivos del estudio están alineados con los de las metas de la Década Digital, y contribuyen a los objetivos del plan estratégico de Horizonte Europa para el desarrollo de tecnologías, estándares y experimentación para usuarios que apoyarán el EC para el futuro IoT inteligente”.

Su objetivo es entonces la aplicación de la EC para una IoT inteligente, por lo que los países interesados en la expansión de la IoT, los que ya se encuentran en un importante desarrollo, podrían encontrar información de valor en este documento.

¹⁴⁷ <https://op.europa.eu/en/publication-detail/-/publication/ff35c457-8f3b-11ee-8aa6-01aa75ed71a1/language-en/format-PDF/source-323701960>.

Este documento presenta los aspectos más relevantes de la computación en el borde, desde las múltiples ópticas que surgen del análisis de los mercados y de la investigación basada en información reciente.

Se abordan, en grandes líneas, las estructuras conceptual y física, y las funciones de la computación en el borde; su ecosistema con otras tecnologías emergentes y con los proveedores de servicios de comunicaciones; las condicionantes técnico-económicas en las diferentes modalidades, las combinaciones eficientes con la computación en la nube y los requerimientos del retardo y los grandes volúmenes de datos; las estimaciones y la gestión de los costos en las empresas y sus estrategias de desarrollo hacia un modelo superintegrado con la computación en el borde; las acciones de instituciones internacionales y autoridades nacionales en áreas como la normalización y la regulación vinculada, en aspectos como la soberanía de los datos, y la evolución de los mercados, del ecosistema tecnológico y de la estructura de los proveedores, incluidos, por ejemplo, los hiperescaladores.

Se analiza particularmente la situación en América Latina y el Caribe, su evolución territorial y las estimaciones del crecimiento y su relación con la preparación de los países para las tecnologías de vanguardia.

