



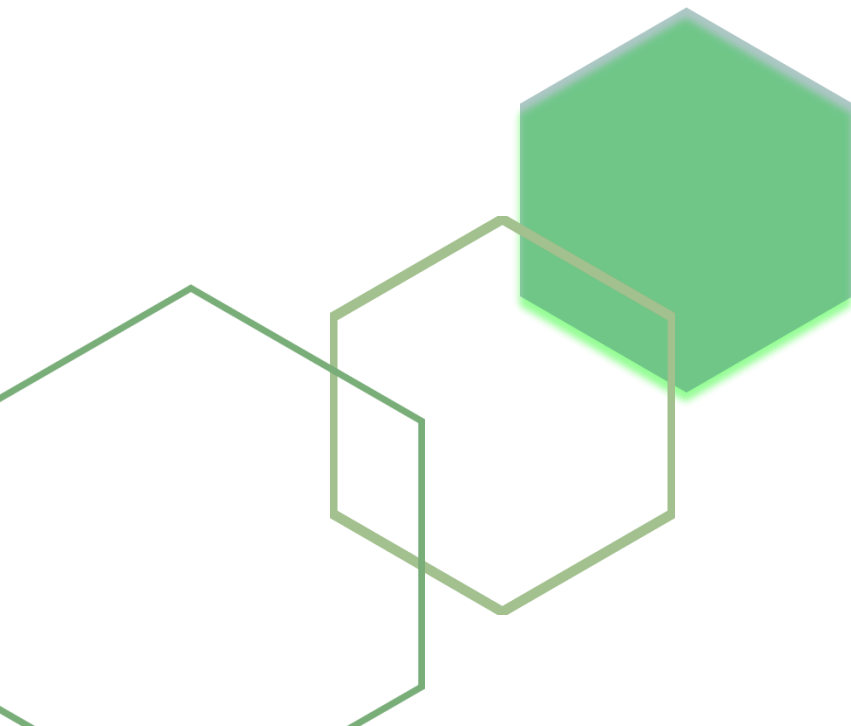
INSTITUTO FEDERAL DE
TELECOMUNICACIONES



Estudio de Cloud Computing en México

Instituto Federal de Telecomunicaciones

Julio 2020



INSTITUTO FEDERAL DE TELECOMUNICACIONES

Unidad de Política Regulatoria

Dirección General de Compartición de Infraestructura

Análisis de los servicios de Cloud Computing en México

Contenido

1. Términos y Acrónimos de Cloud Computing.....	5
2. Introducción	6
3. Transformación digital	10
3.1. Evolución de la Industria	10
3.2. Tipos de servicios en la nube	17
4. Cloud Computing.....	18
4.1. Modelos de despliegue de Cloud Computing	20
4.2. Características esenciales del Cloud Computing.....	22
4.3. Modelos de servicios.....	24
4.4. Cambiando el paradigma presupuestario	27
5. Infraestructura de Cloud Computing	28
5.1. Data Centers.....	29
5.1.1. Subsistemas y Nivel de Clasificación	29
5.2. Capacidad de procesamiento y almacenamiento de Centros de Datos	33
5.2.1. Localización	33
5.2.2. Cambios de arquitectura tecnológica	35
5.3. Dimensionamiento y consumo de energía	38
5.3.1. Tamaño y Rendimiento de Data Centers	38
5.3.2. Consumo de energía eléctrica.....	41
5.4. Conectividad y Gestión de Redes.....	46
5.4.1. Virtualización.....	47
5.4.2. Interconexión de Data Centers	50
5.4.3. Redes de Fibra Óptica	53
5.4.4. Parámetros de calidad.....	56

5.4.5. Conectividad de Internet.....	61
6. Mercados y Regulación	68
6.1. Regulación del Tráfico de datos fronterizos.....	70
6.2. Seguridad.....	72
6.3 Política de Cloud First.....	74
7. Conclusiones y recomendaciones	76
ANEXO 1: Empresas Cloud Computing en México	80
ANEXO 2: Clasificaciones para Data Centers.....	101
ANEXO 3: Indicadores de Eficiencia	105
ANEXO 4: Redes de Fibra Óptica en México.	109
ANEXO 5: Data Centers Internacionales.	112
ANEXO 6: Puntos de intercambio (IXP) en México.	120

Objeto del Estudio

En el Programa Anual de Trabajo 2019 el Instituto Federal de Telecomunicaciones incluyó el “Estudio de Cloud Computing en México”, con el objeto conocer la situación actual, así como presentar una evaluación prospectiva de los servicios asociados al Cloud Computing en México.

Para ello, entre otros aspectos, se identifican los tipos de configuración de los servicios en la nube, recursos de red e infraestructura necesarios para su desarrollo, se expone la evolución de las actividades que realiza esta nueva plataforma de servicios y su vinculación con las redes de telecomunicaciones. Además, se busca identificar diferentes regulaciones y mejores prácticas que promuevan la innovación, así como el desarrollo de la infraestructura y operaciones de las redes actuales y futuras, con el fin de implementar el uso de los servicios en la nube a gran escala en el país.

Para ello, se tuvo a bien recabar información relevante de fuentes públicas tanto de organizaciones públicas como privadas, de origen nacional e internacional, como la Agencia Internacional de Energía (*IEA*, por sus siglas en inglés), Unión Internacional de Telecomunicaciones (*ITU*, por sus siglas en inglés), Banco de México, Instituto Nacional de Estadística y Geografía (*INEGI*), Cisco, IBM, entre otras, así como a partir de entrevistas con los principales actores en la operación, implementación y desarrollo del Cloud Computing como lo son Amazon, Microsoft y Google, que a nivel mundial están habilitando aceleradamente infraestructura, redes y servicios de este nuevo ecosistema digital, así como la visita al Data Center de Alestra en la ciudad de Querétaro.

El documento se divide en siete secciones. La primera comprende los acrónimos para la comprensión de la lectura del documento cuando estos son usados, la segunda sección presenta una introducción de cómo se concibe el Cloud Computing y los principales elementos que lo integran. En la tercera sección se describe la evolución de la industria hacia la transformación digital, mostrando una serie de servicios de diversa índole que se implementan y acceden a través de los servicios en la nube. En la cuarta sección se explica el marco teórico, las definiciones y las características del Cloud Computing, incluyendo tipos de despliegue y modelos de servicio. En la quinta sección se expone la relevancia de la infraestructura para la correcta prestación de los servicios en la nube, de manera particular los Data Centers, los elementos que los conforman, su clasificación y se analiza el impacto que tienen factores como la demanda de energía eléctrica, las redes de fibra óptica y los parámetros de calidad.

En la sexta sección se realiza una síntesis del papel que juegan las nuevas tecnologías de la información y comunicación en el mercado mexicano. Así como

algunas políticas públicas implementadas en otros países para incentivar el acceso a nuevas tecnologías como la política de datos fronterizos y la política de “Cloud First”, adoptada por algunos gobiernos. Finalmente, en la séptima sección se presentan algunas recomendaciones derivadas de las conclusiones obtenidas del estudio que se presenta.

El estudio pretende servir como una guía dirigida a personas interesadas, con o sin conocimiento o experiencia en las Tecnología de Información y Comunicaciones (TIC), para acercarlos a los principales temas y conceptos que conforman esta nueva realidad del Cloud Computing, tan presente en la vida cotidiana.

1. Términos y Acrónimos de Cloud Computing

Para una mejor comprensión de los términos y acrónimos se presenta a continuación una relación de los utilizados con mayor frecuencia a lo largo del documento.

Tabla 1: Acrónimos y Términos.

Términos utilizados	
ANSI	American National Standards Institute. Instituto Nacional Americano de Estándares.
AS	Autonomous System. Sistemas Autónomos. Redes de Internet administrativamente separadas.
CC	Cloud Computing. Cómputo en la Nube, Servicios en la Nube o Nube, indistintamente.
CCTV	Circuito Cerrado de Televisión
DC	Data Center. Centro de Datos.
DCN	Data Center Net. Red de Centro de Datos
DFO	Distribuidor de Fibra Óptica
DSL	Digital Subscriber Line. Línea digital de suscriptor.
EB	Exabytes (10 ¹⁸ bytes)
GIE	Grupo de Interés Económico
Instituto o IFT	Instituto Federal de Telecomunicaciones.
IPv6	Protocolo de Internet versión 6 definida en el RFC 2460.
ISP	Internet Service Provider. Proveedor de Servicios de Internet.
IXP	Internet Exchange Point. Punto de Intercambio de Internet.
NOC	Network Operations Center. Centro de Operación de Red.

Términos utilizados	
OTT	Over The Top. Servicios que se proveen sobre el servicio de acceso Internet.
PSC	Proveedor de Servicios de Comunicaciones.
Rack	Estructura que permite sostener o albergar un dispositivo (router, switches, servidores, etc.). Su unidad de medida es la <i>unidad rack</i> o U, la cual es usada para describir la altura del equipamiento. Una <i>unidad rack</i> equivale a 4.445 (cm) de alto.
SDN	Software Defined Networking. Redes definidas por software.
SLA	Service Level Agreement. Acuerdo de Nivel de Servicio.
TIA	Telecommunications Industry Association. Asociación de la Industria de Telecomunicaciones.
TIC / IT	Tecnologías de Información y Comunicaciones, Tecnologías de la Información. Information Technologies.
ITU / UIT	International Telecommunications Union. Unión Internacional de Telecomunicaciones.
UPS	Uninterrupted Power Supply. Sistema de Alimentación Ininterrumpida.
ZB	Zettabyte (10^{21} bytes)

2. Introducción

El IFT reconoce la creciente importancia que está teniendo el sector de las tecnologías de información y comunicaciones (TIC) desde hace al menos un par de décadas, como herramienta para la mejora de la productividad, generando la posibilidad de la prestación de nuevos servicios puestos a disposición para la optimización de los recursos informáticos. Recientemente, esta evolución se ha materializado en las denominadas tecnologías disruptivas, como es el caso de Cloud Computing o Cómputo en la Nube.

La comunicación digital ha generado un progreso tecnológico a gran escala, creando nuevos dispositivos electrónicos, nuevas formas de interacción humana, propiciando el desarrollo de nuevas oportunidades laborales, generando distintas formas de educar, formando múltiples plataformas de entretenimiento, de información y contenido multimedia. Lo anterior ha transformado las comunicaciones entre dispositivos o máquinas a comunicaciones entre todo, rompiendo la concepción de que un **usuario final** se trata solamente de una persona, pues también puede ser una máquina, un sensor, una aplicación o dispositivos comunicándose entre ellos, lo que crea un ecosistema de **hiperconectividad**, es decir, que prácticamente todo está interconectado.

Cambio de hábitos de consumo de servicios digitales durante la pandemia COVID-19

Como consecuencia de la pandemia epidemiológica COVID-19, un gran número de países han impuesto medidas de restricción de movilidad de las personas, prácticamente trasladando un gran número de actividades a sus casas bajo esquemas de colaboración a distancia, llamadas telefónicas, video conferencias, capacitaciones y clases en línea, entre otras que se venían haciendo en los sitios de trabajo o centros educativos. Esta migración forzada de actividades, así como el incremento de algunas otras que en alguna medida ya se venían realizando, aunque en menor escala como comercio electrónico, servicios financieros y sobre todo entretenimiento, son posibles gracias a que actualmente se cuenta con la tecnología y las redes de telecomunicaciones, que por sus características de dimensionamiento principalmente de las redes fijas, han soportado los incrementos significativos de tráfico de todo tipo para sobrellevar la contingencia, pero que también muestran que el crecimiento esperado de la demanda por servicios digitales y el tráfico de datos para los próximos meses o inclusive años se aceleró de forma imprevista.

Para ello los Servicios en la Nube y su soporte a través de su infraestructura y gestión de los Data Centers es crucial para la vida en este nuevo orden digital: el teletrabajo, la escuela en casa, el entretenimiento, las redes sociales, la banca electrónica, compras en línea, etc. Al menos el 20% de toda la información que está en Internet la gestiona un Data Center. Las expectativas de crecimiento del mercado global del Cloud Computing se estimaban para el 2023 alcanzaría un valor superior a los 174,000 millones de dólares (mdd), según Market Watch¹, sin embargo con las necesidades de confinamiento es altamente probable que estas expectativas sean superadas mucho antes.

Diversos análisis recientes han capturado estos cambios de hábitos generados a partir de la contingencia sanitaria. La mayor práctica de *home office* aumentó el consumo de aplicaciones de video conferencia en el orden del 230%. La modalidad de tráfico remoto también intensificó el tráfico de las VPNs (red privada virtual) en 35%. El consumo en *gaming* creció 100% respecto a un día normal inicialmente (los que más crecieron fueron PlayStation y Blizzar en aproximadamente 120%) para luego estabilizarse en 40%².

En el caso de Skype al tercer día de la cuarentena creció en 650% y Webex y Zoom crecieron 430%, y en series, películas y otros consumos por streaming (Flow, Netflix o Youtube), el comportamiento fue similar al de un fin de semana de lluvia, con un crecimiento de 20%. El incremento importante en estos volúmenes de tráfico tratándose de video, también fue posible porque dichas plataformas decidieron bajar la velocidad de tasa de transferencia para disminuir problemas de conectividad³. Analizando otro ángulo del mismo fenómeno se observa un incremento del 28% en la productividad de las empresas⁴, lo cual se explica por los ahorros en tiempos dedicados al transporte de las personas y costos de operación de las oficinas. La adopción del comercio electrónico sumará aceleradamente usuarios y empresas de tal forma que el crecimiento esperado en tres años se tendrá en un año. Las categorías de ventas online que más actividad presentan son la comida a domicilio (40%), moda (30%), supermercado (28%) y en servicios los bancarios (55%), telefonía móvil (51%) y pagos de servicios (41%)⁵.

Lo anterior hace evidente la necesidad de contar con la información relevante y oportuna del comportamiento del mercado de las TIC, como la Encuesta Nacional sobre Disponibilidad y Uso de las Tecnologías de Información en los Hogares (ENDUTIH), realizada de manera conjunta entre el INEGI, la SCT y el IFT, que nos permite identificar los cambios y el grado en que ocurren para el diseño de las políticas públicas y de competencia que anticipen la capacidad de respuesta del sector a la nueva realidad.

Para realizar o soportar estos incrementos en los flujos de información, por lo general las organizaciones han requerido para su operación de infraestructura informática propia, equipos robustos que soporten las plataformas necesarias para el alojamiento de información, la instalación de aplicaciones propias, alojamiento de

¹ Medina, A. (2019). Un viaje al centro de un data center: el epicentro de la vida digital. Recuperado de <https://www.forbes.com.mx/un-viaje-al-centro-de-un-data-center-epicentro-vida-digital/>

² Gomez, D. (2020). Coronavirus e internet: cuánto aumentó la demanda de datos desde que comenzó el periodo de aislamiento en Argentina. Recuperado de <https://www.infobae.com/economia/2020/03/25/coronavirus-e-internet-cuanto-aumento-la-demanda-de-datos-desde-que-comenzo-el-periodo-de-aislacion/>

³ Vázquez, R. (2020). Durante la crisis del COVID-19 los medios tradicionales ganan credibilidad. Recuperado de <https://www.forbes.com.mx/durante-la-crisis-del-covid-19-los-medios-tradicionales-ganan-credibilidad/>

⁴ Forbes. (2016). Home office aumenta 28% la productividad de las empresas. Recuperado de <https://www.forbes.com.mx/home-office-aumenta-28-la-productividad-de-las-empresas/>

⁵ AMVO. Asociación Mexicana de Venta Online. Reporte 2-2020 sobre el impacto del COVID-19 en la venta online en México. Recuperado de <https://www.amvo.org.mx/>

sitios web, adquisición de licencias, así como las instalaciones físicas para alojar toda esta infraestructura.

Todo esto puede tener costos considerables debido a la adquisición, instalación, operación y el mantenimiento de dicha infraestructura, lo cual normalmente interfiere en la misión primordial de la organización.

Dado que las comunicaciones generan un volumen considerable de datos, es necesario apoyarse en **Data Centers**, los cuales son instalaciones con ambiente controlado que alojan sistemas informáticos, de procesamiento, almacenamiento, con fuentes de alimentación y bancos de energía de respaldo. Los **Data Centers** comúnmente formaban parte de la infraestructura de los mismos usuarios o clientes que los utilizaban para proveer sus soluciones tecnológicas. Por ejemplo, una tienda en línea, necesita un **Data Center** para alojar su sitio web, guardar la información de sus pedidos, almacenar los perfiles de sus compradores, etc.

Sin embargo, desarrollar un **Data Center** no es tarea fácil, además de que conlleva una inversión de alto riesgo. Implica tener un espacio designado para toda la infraestructura, mantener el ambiente controlado, e incurrir en todos los gastos de operación, como energía eléctrica y mantenimiento. Anteriormente, se recomendaba diseñar el **Data Center** sobredimensionado para poder atender un mayor número de clientes, lo que implicaba incurrir en gastos de infraestructura tecnológica que no se aprovechaba y resultaba poco eficiente. En el caso contrario, si el **Data Center** se dimensionaba de manera muy justa, y los clientes comenzaban a incrementar junto con la demanda de capacidad, el **Data Center** era insuficiente, por lo que se debía poner en marcha una planeación para el incremento de su capacidad, lo que podía generar un mayor costo y tomar un tiempo considerable de implementar.

Es aquí cuando el **Cloud Computing** surge como un modelo de negocio atractivo ofreciendo infraestructura tecnológica a través de la **Nube**, es decir, a través de Internet, cuyo fin primordial es ofrecer servicios de almacenamiento, aplicaciones, procesamiento y recursos informáticos a través de computadoras interconectadas y gestionadas por software, de manera inmediata, bajo demanda y cobrando solo por lo que se usa.

El **Cloud Computing** también permite agregar a distintos socios y colaboradores que conectan su infraestructura de gran escala para brindar los servicios a los clientes que requieran capacidad de procesamiento, almacenamiento y aplicaciones. Por lo que las áreas de tecnología de la información en las empresas están migrando a estos servicios tecnológicos, pues así se despreocupan por mantener un espacio con infraestructura informática en las condiciones necesarias y se concentran en su giro de negocio, ya que ahora pueden hacer uso de la infraestructura de su

proveedor de **Cloud Computing** en el momento que la requieran y con tanta capacidad como necesiten.

Con la solución del **Cloud Computing**, por ejemplo, un negocio puede planear sus ventas concentrándose en las promociones, sin prestar tanta atención a la planeación de la capacidad de su infraestructura informática, puesto que, si llega a requerir más almacenamiento en algún momento, basta con solicitarlo a su proveedor de **Servicios en la Nube**.

Con el surgimiento del **Cloud Computing** y su rápida adopción se presentan nuevas necesidades, como redes con nuevas capacidades para cursar el creciente tráfico de datos y que estas se mantengan estables para que los proveedores de **Servicios en la Nube** puedan brindar con calidad sus soluciones tecnológicas. Además, entre lo que destaca, se encuentra el requerimiento de bajas **latencias** de interconexión, es decir, la velocidad que tarda en llegar un dato de un lugar a otro, que depende, además de la calidad del medio de transmisión, de la distancia física que hay entre una instalación a otra, o entre la instalación y el cliente, o entre un usuario final y algún sensor.

La distancia entre puntos debe ser tal que se cuente con una baja **latencia** o retardo, con el fin de brindar un servicio de alta calidad, y no solo por parte del proveedor de **Servicios en la Nube** a sus clientes, como cuando estos últimos solicitan los archivos alojados en un **Data Center**, sino también entre proveedores de **Servicios en la Nube** que tengan algún tipo de sociedad comercial. Tal puede ser el caso de una película ofrecida por una plataforma de **Video Streaming**, tan populares hoy en día; si hay una gran **latencia** entre la instalación donde se encuentra este archivo de video y el usuario que lo solicita, el video tardará más tiempo en poderse reproducir o presentará intermitencias, lo que genera una experiencia de usuario deficiente.

Reducir la **latencia** de las redes es imperativo para el **Cloud Computing**, pero no es el único requerimiento para poder establecer esta integración tecnológica como un modelo de negocio viable. Otro elemento importante es la **energía** necesaria para operar, pues toda la infraestructura utilizada para ofrecer **Servicios en la Nube** consume cantidades considerables de energía eléctrica, y la generación de ésta tiende a variar en costos dependiendo del lugar donde se encuentre ubicada la infraestructura.

No menos importante, otra característica a tener en cuenta es la **Gestión del Tráfico** de datos, es decir, quién y cómo se encarga de recolectar y agregar todas las solicitudes o envíos de información hacia/desde Internet o la **Nube**. Todas estas solicitudes o envíos de información se conjuntan en algún punto y son transferidas en gran cantidad a través de redes de fibra óptica, principalmente, es este punto

el que interconecta todas las infraestructuras o redes que buscan acceder a la red principal o backbone de Internet para intercambiar tráfico entre ellas. Esta **Gestión del Tráfico** resulta importante para el estudio, dado que, con la tendencia a migrar a **Servicios en la Nube**, el tráfico de datos cada vez es mayor, y se vuelve crítico saber de qué forma se está gestionando el tráfico y la ruta que sigue, pues muchas veces tiende a salir del país y regresar por la falta de puntos de interconexión internos, lo que resulta en elevadas **latencias** debido a la distancia física y mayores costos por no contar con alternativas de intercambio de tráfico internos.

Las empresas más valiosas del mundo se han convertido en importantes proveedores de **Servicios en la Nube**, donde alojan sus aplicaciones y ofrecen a sus clientes una gama muy amplia de servicios adaptables a sus necesidades.

Si bien los **Servicios en la Nube** contribuyen a generar mejores condiciones de desarrollo para diversos sectores en la economía a través de incrementos en la productividad y eficiencias, en México no se ha intensificado su uso generalizado como en otros países con niveles de desarrollo similares, ni se ha implementado alguna política específica que tenga como fin el fomento de su desarrollo y adopción. Por lo anterior, en el documento se identifican diversos aspectos relacionados con la provisión de servicios digitales en el país que podrían resultar propicios para el establecimiento de un mayor número de proveedores de los servicios digitales y de tecnologías de información y comunicaciones necesarios para un mayor flujo de tráfico asociado a los **Servicios en la Nube**, y que a su vez contribuyen a establecer condiciones estructurales para fomentar el desarrollo de proveedores de acceso a Internet.

3. Transformación digital

3.1. Evolución de la Industria

La transformación digital, la cuarta revolución industrial, industria 4.0, economía digital o comercio digital, entre otras, es como se les ha denominado a las tendencias tecnológicas que surgen en el mercado global y se pueden considerar como la forma en que las organizaciones promueven cambios en sus modelos de negocios y ecosistemas aplicando competencias digitales.

El mundo ha experimentado y sigue experimentando en la actualidad una serie de revoluciones industriales. El concepto "Cuarta Revolución Industrial" fue acuñado

por Klaus Schwab fundador del Foro Económico Mundial en el contexto de la edición del Foro Económico Mundial 2016⁶.

Tabla 2: Evolución de la Industria del 1.0 a 4.0.⁷

La Primera Revolución Industrial (Fin S XVIII inicio S XIX)	Durante este período de tiempo la manufactura evolucionó, desde enfocarse en el trabajo manual realizado por personas y ayudado por animales de trabajo, hasta una forma de trabajo más eficiente realizada por personas a través del uso de motores impulsados a vapor y agua, así como otros tipos de maquinaria.
La Segunda Revolución Industrial (Inicio S XX)	La introducción de la electricidad permitió a los fabricantes aumentar su productividad y ayudó para que la maquinaria industrial fuera más móvil. Durante esta fase se introdujeron conceptos de producción en masa como la línea de ensamblaje.
La Tercera Revolución Industrial (Finales 50's)	Los fabricantes comenzaron a incorporar más tecnología electrónica, y finalmente informática, en sus fábricas. Durante este periodo, los fabricantes comenzaron a experimentar un cambio que disminuyó el énfasis en la tecnología analógica y mecánica y lo aumento en la tecnología digital y el software de automatización.
La Cuarta Revolución Industrial, o Industria 4.0. (a partir de 00's)	La industria 4.0 pone el énfasis en la tecnología digital de las últimas décadas a un nivel completamente nuevo con la ayuda de la interconectividad a través del Internet de las cosas (IoT), el acceso a datos en tiempo real y la introducción de sistemas ciber-físicos. La industria 4.0 ofrece un enfoque más integral, interconectado y holístico de la fabricación. Conecta lo físico con lo digital, y permite una mejor colaboración y acceso entre departamentos, socios, proveedores, productos y personas. La industria 4.0 permite a los propietarios de negocios controlar y comprender mejor cada aspecto de su operación y les permite aprovechar los datos instantáneos para aumentar la productividad, mejorar los procesos e impulsar el crecimiento.

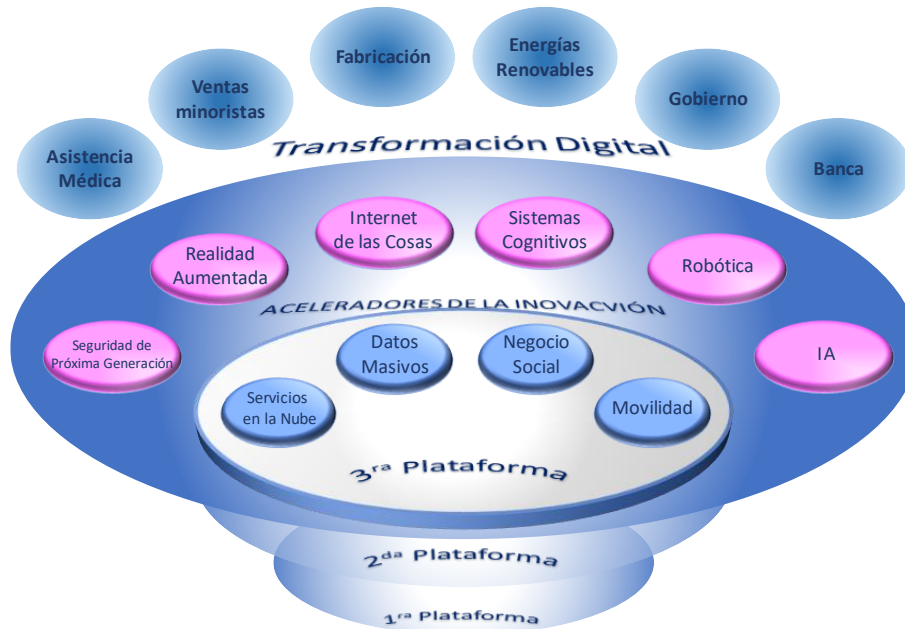
Conceptualmente la transformación digital se basa en un modelo con tres capas fundamentales: 1) la primera consta de las infraestructuras y redes que soportan las tecnologías de información y telecomunicaciones; 2) la segunda consta del hardware y software que conforman las aplicaciones y servicios de TIC, y 3) la tercera consta de los servicios, incluyendo plataformas digitales que soportan las redes sociales, análisis de datos masivos (*big-data*), cómputo en la nube (*cloud computing*), entre otros servicios habilitados por Internet.

Este último modelo de capas o plataformas contiene los servicios que soportan a los denominados aceleradores de innovación como la seguridad de información, realidad aumentada, internet de las cosas, sistemas cognitivos, robótica y la inteligencia artificial, que a su vez inciden en la digitalización de servicios de salud, negocios, energía, gobierno, banca, etc.

⁶ Schwab, K. (2020). The Fourth Industrial Revolution: what it means, how to respond. Recuperado de <https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/>

⁷ Epicor. (2020). What is Industry 4.0—the Industrial Internet of Things (IIoT)? Recuperado de <https://www.epicor.com/en-us/resource-center/articles/what-is-industry-4-0/>

Figura 1: Plataformas de transformación digital.



Fuente: Elaboración propia IFT.

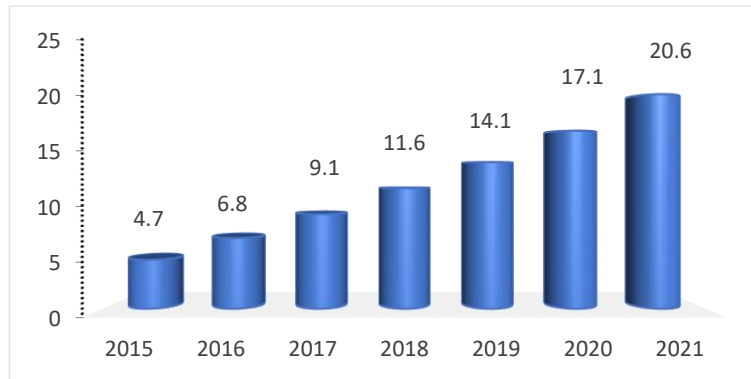
La innovación digital impulsada por los servicios en la nube, datos masivos, los nuevos modelos empresariales y la movilidad están cambiando el funcionamiento y alcance de la ciencia, los gobiernos, las ciudades y sectores como los de salud, comercio y la agricultura, lo que muestra una tendencia clara a un entorno de **hiperconectividad**.

La hipótesis es que las redes de proveedores de servicios establecidos tendrán restricciones en su capacidad de banda ancha debido al incremento de tráfico por la demanda de los servicios digitales, llámese OTT, datos móviles, *big-data*, la adopción del protocolo de Internet IPv6 y el *Cloud Computing*.

La inmersión en la Transformación Digital es inevitable, ante esta realidad todos los sectores que intervienen se enfrentan a importantes retos, entre ellos el alto incremento de tráfico, la consecuente expansión de los Data Centers y su exigencia de capacidad en términos de banda ancha, lo cual a su vez creará una significativa demanda de infraestructura crítica, principalmente en las redes de telecomunicaciones.

El supuesto del incremento de tráfico es soportado por diferentes fuentes, como por ejemplo las proyecciones de tráfico global emitidas por Cisco⁸, las cuales hacen ver que el tráfico anual global de Data Centers crecerá tres veces para 2021 y el tráfico IP de los Data Centers crecerá a una tasa de crecimiento anual (CAGR) del 25% entre 2015 y 2021. De los 4.7 ZB del tráfico total de 2015, el 92% del tráfico provino de la nube.

Gráfica 1: Pronóstico de Tráfico Global por año de Data Centers en Zettabytes.



Notas: 25% CAGR (2016 – 2021)

Fuente: Cisco (2018). Índice Global de Nube de Cisco, 2016-2021

Este crecimiento de tráfico puede atribuirse al aumento de la migración hacia las arquitecturas en la nube debido a que su capacidad para escalar es más rápida y eficiente que la de los Data Centers tradicionales.

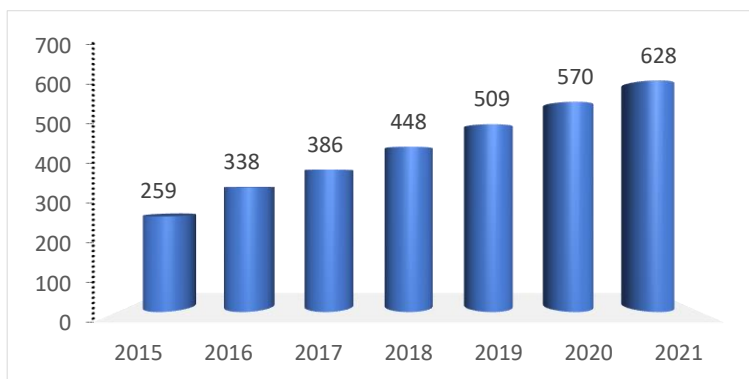
Por otro lado, se estima que la cantidad de Data Centers de Hiperescala⁹ aumentará de 259 a finales de 2015 a 628 para 2021¹⁰.

⁸ Carter, C., Hunt, E. (2016). Cisco Global Cloud Index Projects Cloud Traffic to Nearly Quadruple Representing 92 Percent of Total Data Center Traffic by 2020. Recuperado de <https://newsroom.cisco.com/press-release-content?articleId=1804748>

⁹ Las empresas de hiperescala, definidos como los que tienen un tamaño mayor a 400,000 pies cuadrados (37,161.21 m²).

¹⁰ Jain, S. (2016). Global Data Center Growth: Supporting the Next Data Deluge. Recuperado de <https://blogs.cisco.com/sp/global-data-center-growth-supporting-the-next-data-deluge>

Gráfica 2: Pronóstico de crecimiento de Data Centers de Hiperescala.



Fuente: Cisco (2018). Índice Global de Nube de Cisco, 2016-2021

Esto es un indicativo claro no solo del crecimiento de los Data Centers en general, sino de una creciente dependencia de los mismos en la nube con una entrega más rápida de servicios y datos, mayor rendimiento, y mejoras en las eficiencias operativas.

La importancia evolutiva del crecimiento de los Data Centers radica en la consolidación del poder de procesamiento de datos que representan. Para 2020, se estiman incrementos en todos los conceptos operativos, es decir, en proporción de servidores, capacidad de procesamiento, almacenamiento y tráfico, conforme se muestra en la siguiente tabla.

Tabla 3: Evolución de la capacidad de los Data Centers de Hiperescala.

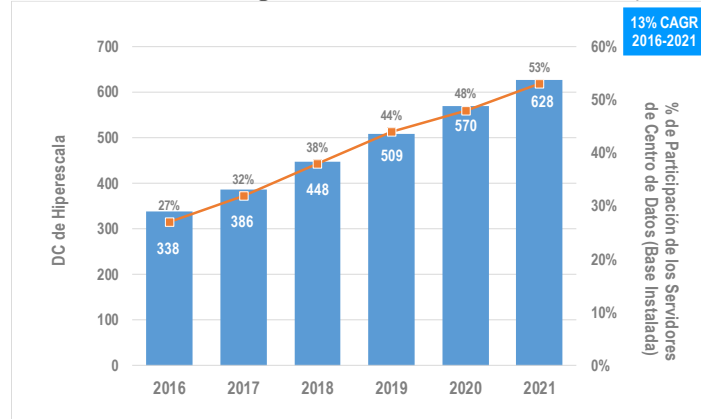
Concepto	2015	2020 (estimado)
Porcentaje de Data Centers de Hiperescala del total de servidores de datos.	41%	47%
Porcentaje de potencia de procesamiento de Data Centers proveniente de Data Centers de Hiperescala.	39%	68%
Porcentaje de datos almacenados en Data Centers de Hiperescala.	49%	57%
Porcentaje de tráfico de Data Centers llevado a cabo por Data Centers de Hiperescala.	34%	53%

Fuente: Cisco (2016). Índice Global de Nube de Cisco, 2016-2021

Para mayor referencia, una relación de los Data Centers más representativos a nivel mundial se encuentra en el ANEXO 5: Data Centers Internacionales, de este documento.

En resumen, la relación entre el crecimiento de los Data Centers y su participación en la proporción de servidores se puede observar en la siguiente gráfica: Crecimiento global de Data Centers de Hiperescala.

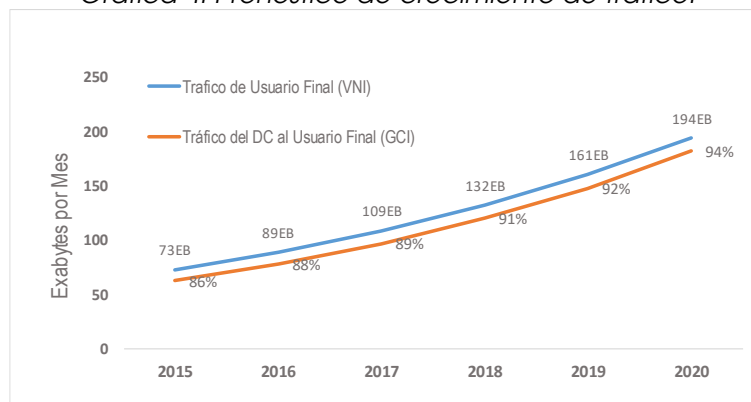
Gráfica 3: Crecimiento global de Data Centers de Hiperescala.



Fuente: Cisco (2016). Crecimiento Global de Data Centers

Aunado a lo anterior, en los próximos años se pronostica que cantidades cada vez mayores de tráfico de clientes estarán en contacto con un Data Center. En 2015, el 86% del tráfico total del usuario / dispositivo final,¹¹ transitó por un Data Center y se espera que este tráfico alcance el 94% para 2020, como se muestra en la siguiente gráfica.

Gráfica 4: Pronóstico de crecimiento de tráfico.



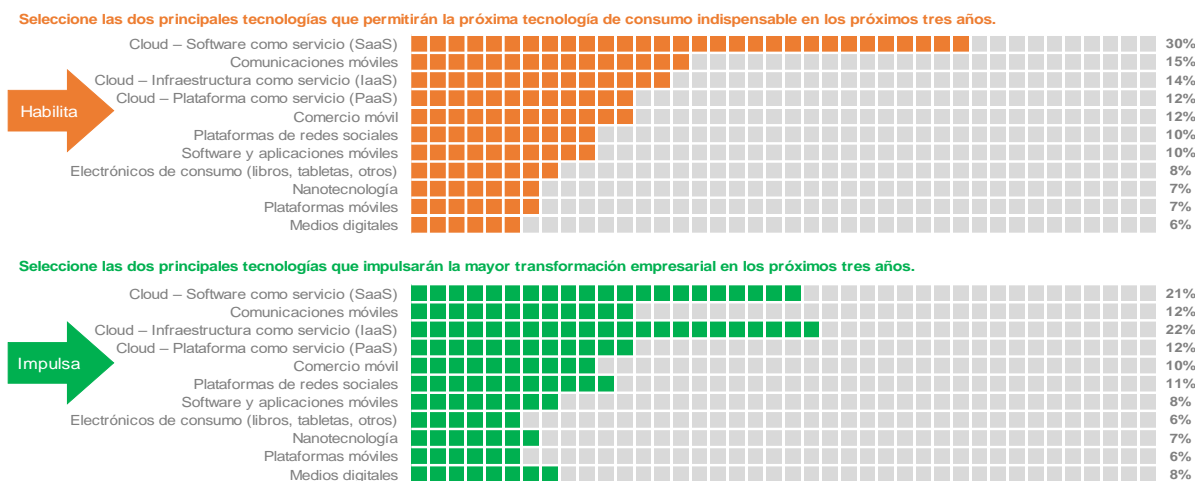
Notas: VNI: Índice de redes virtuales (VNI, por sus siglas en inglés)

Fuente: Cisco (2016). Crecimiento Global de Data Centers

¹¹ Jain, S. (2016). Global Data Center Growth: Supporting the Next Data Deluge. Recuperado de <https://blogs.cisco.com/sp/global-data-center-growth-supporting-the-next-data-deluge>

Por otro lado, el Cloud Computing encabeza la lista de opinión sobre las tecnologías más disruptivas, y propiciará la transformación digital en los negocios y ecosistemas industriales en el mundo¹².

Figura 2: Tecnologías disruptivas a futuro.



Fuente: ITU Academy, Strategic Aspects for Internet Governance and Innovations

Una vez explicado el fenómeno de la transformación digital se puede concluir que esta pretende cambiar todo a su paso, redefiniendo nuevos esquemas comerciales, educativos, médicos, económicos, sociales, políticos y de entretenimiento entre otros más, que impactarán a los gobiernos y a sus ciudadanos.

Los pronósticos de tráfico antes descritos auguran el crecimiento intensivo de nuevos servicios basados en la nube y las infraestructuras subyacentes que los soportan, como lo son los Data Centers de cualquier dimensión, por lo que se considera relevante realizar un análisis de los servicios de *Cloud Computing*, su evolución y dimensionar su impacto tanto en el contexto actual como prospectivo en las redes de telecomunicaciones en nuestro país. Al respecto se revisarán las definiciones, características técnicas y operativas de los principales elementos, así como aspectos económicos y regulatorios que inciden en el desarrollo de los servicios en la nube.

¹² ITU. Course Strategic Aspects for Internet Governance and Innovations.

3.2. Tipos de servicios en la nube

El mercado ya está proporcionando una variedad de servicios para implementar y consumir la nube, en función de las necesidades de una gran variedad de consumidores (Mr Nasser Kettani, Jules Essoh, Henri Numbi Ilunga, 2014-2017):

Tecnologías de implementación: ayudan a las organizaciones a implementar el Cloud Computing de diferentes formas, dependiendo de si despliegan nubes privadas (las organizaciones pueden transformar su propia infraestructura de TI y sus procesos de suministro de TI utilizando conceptos de computación en la nube, para fines propios), o de si su intención es convertirse en proveedores de servicios en la nube (suministrar servicios en la nube a terceros o en nombre de terceros, por ejemplo, algunos organismos gubernamentales han creado infraestructura compartida que se suministra como servicio en la nube a diversos organismos gubernamentales).

Servicios en la nube de usos múltiples: algunas organizaciones, locales y multinacionales, ofrecen un conjunto de servicios en la nube, en la mayoría de los casos, servicios en la nube públicos para todos. Los servicios pueden ser desde la disponibilidad de infraestructuras, plataformas de desarrollo, o programas y aplicaciones.

Algunos de los servicios en la nube están diseñados para el consumo y otros para las necesidades empresariales (y gubernamentales). En la mayoría de los casos, los servicios de grado empresarial y los servicios de grado de consumidor se rigen por diferentes principios, cláusulas contractuales, acuerdos de nivel de servicio (SLA por sus siglas en inglés), modelos de negocio y reglas de seguridad y privacidad.

Nubes integradas verticalmente: son servicios a gran escala (por ejemplo, redes sociales o motores de búsqueda), que se proporcionan a los usuarios y que se ejecutan en la infraestructura de la nube desarrollada e integrada del proveedor.

Servicios innovadores en la nube: se ejecutan en la infraestructura de la nube multiuso. En esta categoría, las compañías de todos los tamaños, proveedores de software independientes (ISV por sus siglas en inglés), desarrollan servicios en la nube para que otros los usen, sin embargo, necesitan infraestructura en la nube para construir y operar sus propios servicios en lugar de crear los suyos. Esta categoría es probablemente la más dinámica en el mercado, ya que las organizaciones pequeñas, medianas y grandes pueden usar la infraestructura de la nube lista para usar para brindar sus propios servicios a sus clientes, ya sea para consumidores o para empresas.

4. Cloud Computing

Tratar de definir el *Cloud Computing* no es una tarea sencilla, por lo que se realiza un análisis de diferentes referencias para dar una idea más precisa de su significado y alcance.

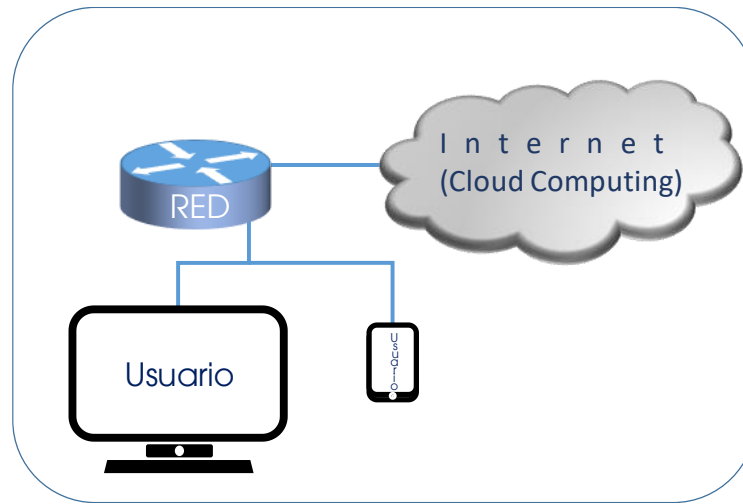
Referencias como "*utility computing*", "*servicios en red*", "*servicios de computación a la carta*", "*súper computación*" o "*computación elástica o escalable*", son ejemplos del constante cambio de las variantes de uso de las tecnologías de información y del tratamiento o procesamiento de información como un fundamento del ecosistema digital.

Quizá el rasgo más característico y fundamental es el uso irrestricto de Internet, razón principal para insertar en diagramas de flujo de red el dibujo de una "nube" para simbolizar la "Internet". Esta representa una incógnita y se ubica como punto medio entre computadoras interconectadas a la red de redes, lo que permite el flujo de información y la comunicación entre usuarios.

Toda esa zona es a lo que se le conoce como "*ciberespacio*" o "*Internet*", un entorno virtual e intangible, un gran canal de telecomunicación. Si bien existe una gran confusión entre el acceso a Internet, a los servicios, o incluso las aplicaciones (en dispositivos móviles), y el Cloud Computing, no se deben confundir estos términos. El hecho de que uno acceda a un servicio a través del Internet o a través de una aplicación no significa que el *back-end*, o servidores fuente de la información, que proporciona el servicio, cumple con las características principales y se ejecuta como Cloud Computing.

El *Cloud Computing* trasciende de la metáfora para describir al Internet, y puede representarse gráficamente de la siguiente manera:

Figura 3: Nube de Internet.



Fuente: Elaboración propia

El término de Cloud Computing fue definido por el NIST¹³, como:

"Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e. g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."

This cloud model promotes availability and is composed of five essential characteristics, three service models, and four deployment models".

"El Cloud Computing es un modelo para permitir el acceso ubicuo, conveniente y bajo demanda de la red a un conjunto compartido de recursos informáticos configurables (por ejemplo, redes, servidores, almacenamiento, aplicaciones y servicios) que se pueden aprovisionar y lanzar rápidamente con un mínimo esfuerzo de gestión o Interacción del proveedor de servicios."

Este modelo de nube promueve la disponibilidad y está compuesto por cinco características esenciales, tres modelos de servicio y cuatro modelos de despliegue."

Traducción Libre

De acuerdo al estándar ISO/IEC DIS 17788:2015 el Computo en la nube es:

"(A) Paradigm for enabling network access to a scalable and elastic pool of shareable physical or virtual resources with self-service provisioning and administration on-demand"

¹³ Mell, P., Grance, T. (2011). The NIST Definition of Cloud Computing. Recuperado de: <https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-145.pdf>.

"(Un) Paradigma para establecer el acceso de red a una fuente de recursos físicos o virtuales de una forma escalable y elástica con auto aprovisionamiento y administración bajo demanda"

Traducción Libre

Así, el Cloud Computing es más que una tecnología absoluta o total por sí misma; es decir, se trata más de un modelo de servicios de tecnología de información para el acceso, la asignación, el control y la optimización de los recursos, entendida como una puesta a disposición de recursos tecnológicos para el disfrute del cliente en diversas modalidades de servicio e implementación.

En torno al mismo término, *Cloud Security Alliance*¹⁴ describe al cómputo en la nube como:

"Modelo a la carta para la asignación y el consumo de computación. La nube describe el uso de una serie de servicios, aplicaciones, información e infraestructura compuesta por reservas de recursos de computación, redes, información y almacenamiento. Estos componentes pueden orquestarse, abastecerse, implementarse y desmantelarse rápidamente, y escalarse en función de las dimensiones para ofrecer unos servicios de tipo utilidad."

De las definiciones anteriores, podemos decir que el Cloud Computing es el ecosistema de recursos tecnológicos de la información y la comunicación, que ofrece servicios escalables, compartidos y bajo demanda en diferentes modalidades y a diversos usuarios a través de Internet.

Siguiendo la definición del NIST, que es una de las que ha conseguido mayor aceptación por empresas y gobiernos, el cómputo en la nube se compone de cuatro modelos de despliegue, cinco características esenciales y tres modelos de servicio, que se explican a continuación.

4.1. Modelos de despliegue de Cloud Computing

Según su definición, el Instituto Nacional de Estándares y Tecnología (NIST) identifica cuatro **modelos** de implementación o despliegue (Mell P, Grance T. , September 2011) :

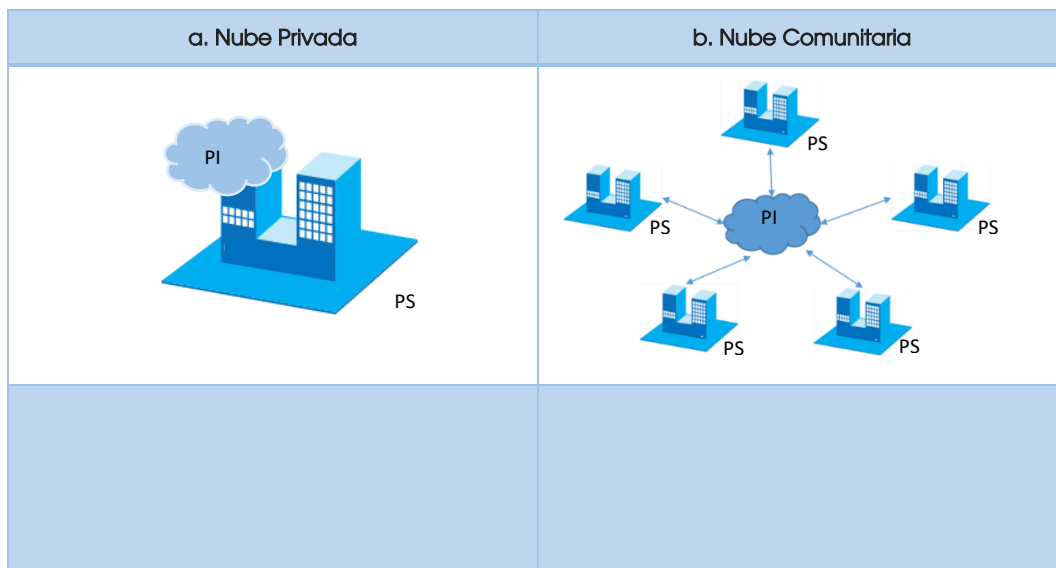
- a. **Nube privada:** la infraestructura de la nube es de uso exclusivo de una sola organización conformada por diversos consumidores (por ejemplo, unidades de negocios). La nube puede pertenecer, ser operada y administrada por la

¹⁴ Guía de Seguridad de Áreas Críticas para la computación en la Nube V.4, de la Cloud Security Alliance (es una organización no lucrativa, cuya misión es promover las buenas prácticas para afianzar la seguridad en Cloud Computing, y formar en el uso seguro de las tecnologías de la información y la comunicación), (2018). Recuperado de <https://www.ismsforum.es/ficheros/descargas/csaguide-v4-es-20181525872082.pdf>.

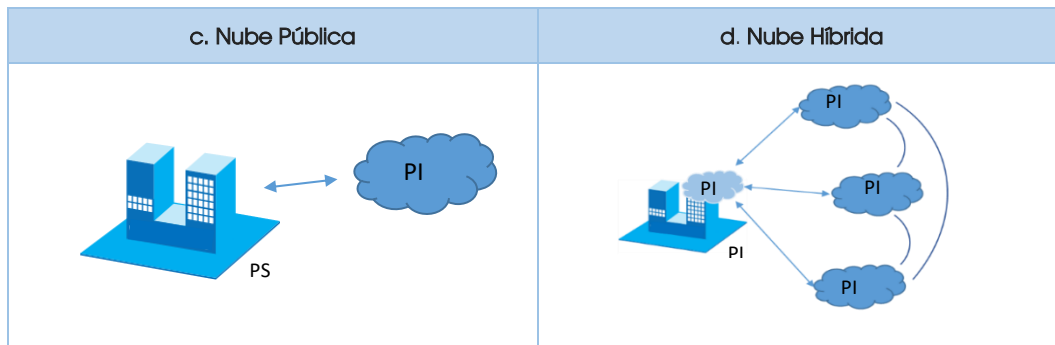
organización, un tercero o una combinación de ambos, y puede existir dentro o fuera de las instalaciones.

- b. **Nube comunitaria:** la infraestructura en la nube se proporciona para uso exclusivo de una comunidad específica de consumidores o de organizaciones que tienen requerimientos o propósitos comunes (por ejemplo, misión, requisitos de seguridad, políticas y consideraciones de cumplimiento). La nube puede pertenecer, ser operada y administrada por una o más de las organizaciones de la comunidad, un tercero o alguna combinación de ellas, y puede existir dentro o fuera de las instalaciones.
- c. **Nube pública:** la infraestructura de la nube es de uso abierto al público en general. La nube puede pertenecer, ser operada y administrada por una organización comercial, académica o gubernamental, o una combinación de ellas. Existe en las instalaciones del proveedor de servicios de la nube.
- d. **Nube híbrida:** la infraestructura de la nube es una combinación de dos o más infraestructuras de la nube (privada, comunitaria o pública) que se mantienen como entidades únicas, pero que están unidas por tecnología estandarizada o patentada que permite la portabilidad de datos y aplicaciones (por ejemplo, *cloud bursting*¹⁵ para equilibrio de carga entre nubes).

Figura 4: Modelos de despliegue de Cloud Computing.



¹⁵ Cloud bursting. Aplicación que funciona en una nube o Data Center privado y utiliza recursos de una nube o Data Center público cuando la demanda de capacidad aumenta.



Notas: PI = Proveedor de Infraestructura; PS = Proveedor de Servicios

4.2. Características esenciales del Cloud Computing

Esta sección describe las **características principales** del servicio siguiendo las definiciones del NIST.

- a. **Amplio acceso a la red**: los recursos físicos y virtuales están disponibles a través de una red y se accede a través de mecanismos estandarizados que permiten el uso por parte de plataformas heterogéneas. El enfoque de esta característica clave es que los usuarios pueden acceder a los recursos físicos y virtuales desde cualquier lugar, siempre que sea accesible a través de la red y mediante una gran variedad de dispositivos, como por ejemplo teléfonos móviles, tabletas, computadoras portátiles y estaciones de trabajo.
- b. **Servicio medido**: la entrega medida de los servicios en la nube es tal, que el uso se puede monitorear, controlar, reportar y facturar. Esta es una característica importante que se necesita para optimizar y validar el servicio de computo en la nube entregado. El enfoque de esta característica clave es que el cliente paga efectivamente por los recursos que utiliza. Desde la perspectiva de los clientes, *Cloud Computing* ofrece valor a los usuarios al permitir el cambio de un modelo de negocio de baja eficiencia y uso de activos a uno de alta eficiencia.
- c. **Autoservicio bajo demanda**: función en la que un cliente del servicio en la nube puede configurar capacidades informáticas, según sea necesario, automáticamente o con una interacción mínima con el proveedor de servicios en la nube. El enfoque de esta característica es que la computación en la nube ofrece a los usuarios una reducción relativa en los costos, el tiempo y el esfuerzo necesarios para realizar una acción, ya que le otorga al usuario la capacidad de hacer lo que necesita, en el momento en el que lo necesita,

sin requerir recursos humanos adicionales, interacciones del usuario, o gastos generales.

d. Elasticidad, escalabilidad y rapidez: una función en la que los recursos físicos o virtuales se pueden ajustar de forma rápida y elástica, en algunos casos de forma automática, para aumentar o disminuir rápidamente los recursos. Para el cliente del servicio en la nube, los recursos físicos o virtuales disponibles para el aprovisionamiento a menudo pueden comprarse en cualquier cantidad en cualquier momento de forma automática, sujeto a las restricciones de los acuerdos de servicio. Por lo tanto, el enfoque de esta característica clave es que la computación en la nube significa que los clientes ya no necesitan preocuparse por los recursos limitados y es posible que no deban preocuparse por la planificación de la capacidad.

e. Agrupación de recursos: una función en la que se pueden agregar los recursos físicos o virtuales de un proveedor de servicios en la nube para servir a uno o más clientes de servicios en la nube. El enfoque de esta característica es que los proveedores de servicios en la nube pueden admitir la tenencia múltiple al mismo tiempo que utilizan la abstracción para ocultar la complejidad del proceso del cliente. Desde la perspectiva del cliente, todo lo que saben es que el servicio funciona, mientras que generalmente no tienen control ni conocimiento sobre cómo se proporcionan los recursos o dónde se ubican los recursos. Esto descarga al proveedor parte de la carga de trabajo original del cliente, como los requisitos de mantenimiento. Incluso con este nivel de abstracción, debe señalarse que los usuarios aún podrían especificar la ubicación en un nivel más alto de abstracción (por ejemplo, país, estado o Data Center).

Multitenencia¹⁶: los recursos físicos o virtuales están asignados de manera que los elementos de computación y datos de los clientes están aislados e inaccesibles entre sí, es decir, es la convivencia de datos de múltiples clientes en un servidor que pueden manipularse simultáneamente sin comprometer la confidencialidad de su contenido.

¹⁶ La multitenencia no es considerada por el NIST como una característica esencial del modelo de Cloud Computing, pero generalmente se discute y considera como tal debido a que es un concepto implícito en las características de agrupación de recursos, servicio medido y elasticidad, escalabilidad y rapidez. Asimismo, la Cloud Security Alliance (CSA) identifica a la multitenencia como un elemento clave del modelo Cloud. (Bhowmik, Sandeep. Cloud Computing. Cambridge University Press, 2017.)

4.3. Modelos de servicios

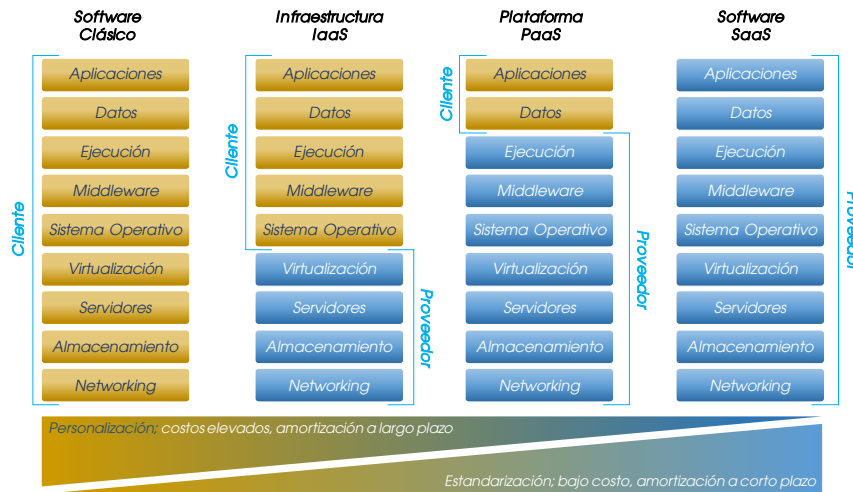
Una práctica común en la industria es permitir a una empresa o institución alojar recursos de computación y almacenamiento, o servicios tecnológicos fuera de sus instalaciones. Esta práctica existía antes de la era del *Cloud Computing*, y ayudó a las empresas a centrarse en las prioridades clave de operación mientras trasladaba la gestión de parte de la infraestructura de TI a terceros.

Por lo general un proveedor de servicios en la nube puede brindar sus servicios en tres combinaciones diferentes de negocio, por medio de distribución de software, alquiler de entornos, ejecución y disponibilidad de aplicaciones o infraestructura. Estos modelos se conocen de la siguiente manera:

- a. **Software como servicio (Software as a service – SaaS)**
- b. **Plataforma como servicio (Platform as a service – Paas)**
- c. **Infraestructura como servicio (Infrastructure as a service – IaaS)**

Una forma de visualizar el uso de los modelos de servicios antes mencionados es representando la gradualidad de los elementos o niveles de cómputo que se van migrando o trasladando desde el cliente al proveedor del *Cloud Computing* mediante la responsabilidad compartida y de esa misma forma cómo el cliente tiene la flexibilidad de adoptar el esquema que más le convenga, como se puede ver en la siguiente tabla "Personalización vs estandarización". Las ramificaciones y los beneficios del **SaaS** frente a otros tipos de servicios se muestran gráficamente, como se puede apreciar, con el software local tradicional, el cliente tiene que administrar la mayoría de las actividades relacionadas con la configuración y ejecución del software. En el extremo opuesto del espectro, **SaaS** se encarga de todo esto por parte del proveedor de software. Hay esquemas mixtos en el medio; por ejemplo, Infraestructura como Servicio (**IaaS**) y Plataforma como Servicio (**Paas**).

Tabla 4: Personalización VS estandarización.



Fuente: Elaboración propia con información de IBM¹⁷

a) Software como servicio (SaaS)

El modelo de servicio SaaS, consiste en un despliegue de software en el cual las aplicaciones y los recursos computacionales se han diseñado para ser ofrecidos como servicios de funcionamiento bajo demanda. De esta forma se reducen los costos tanto de software como hardware, así como los gastos de mantenimiento y operación. El cliente de estos servicios es por lo general personas particulares, empresas o instituciones.

La seguridad es controlada por el proveedor del servicio. El cliente del servicio únicamente tiene acceso a la edición de las preferencias y a unos privilegios administrativos limitados.

En el segmento de SaaS, el cómputo en la nube ha demostrado ser útil como un modelo de negocio. Ejecutando el software mediante servidores externos especializados en procesamiento de datos en lugar de los servidores locales del cliente, los costos se reducen. Por otra parte, al eliminar los gastos de mantenimiento, de licencias y de hardware necesario para mantener servidores locales, las empresas son capaces de ejecutar aplicaciones de forma mucho más eficiente desde el punto de vista informático.

De esta manera se le evita al cliente tener que instalar la aplicación en sus propios ordenadores, tener un CPU para ello, evitando así asumir los costos de soporte y el

¹⁷ Gupta M. (2014). Cloud brokerage: Requirements to support service contestability—application stacks. Recuperado de <https://www.ibm.com/blogs/cloud-computing/2014/09/19/cloud-brokerage-requirements-support-service-contestability-application-stacks/>

mantenimiento de hardware y software. Este es el caso de aplicaciones web como *Gmail, Office 365, Dropbox o Google Docs*.

Cisco estima que para el 2021, el 75% de la carga de trabajo en la nube y las instancias de cómputo serán SaaS, con un aumento de 71% respecto del 2016¹⁸.

b) Plataforma como servicio (PaaS)

La plataforma de cómputo en nube PaaS permite a los usuarios acceder a aplicaciones alojadas en servidores externos por medio de la infraestructura de la nube. De esta manera, un cliente puede usar aplicaciones alojadas en la nube, sin el costo y la complejidad de mantener múltiples capas de hardware y software. Un ejemplo de cliente de estos servicios son los programas desarrolladores de aplicaciones informáticas.

Bajo este modelo de plataforma como servicio, el servicio se entrega bajo demanda, desplegándose el entorno (hardware y software) necesario para ello. De esta forma, se reducen los costos y la complejidad de la compra, el mantenimiento, el almacenamiento y el control del hardware y el software que componen la plataforma.

El cliente del servicio tiene control parcial sobre las aplicaciones y la configuración del entorno, ya que la instalación de los entornos dependerá de la infraestructura que el proveedor del servicio haya desplegado. La seguridad se comparte entre el proveedor del servicio y el cliente. Este es el caso de proveedores como Google App Engine, Heroku, Openshift, Cloud Foundry y muchos más.

De acuerdo con CISCO, para el 2021, el 9% del total de las cargas de trabajo en la nube y las instancias de cómputo serán Plataforma como Servicio (PaaS), un aumento del 8% con respecto al año 2016.

c) Infraestructura como servicio (IaaS)

El último segmento del cómputo en la nube, la infraestructura IaaS, representa en gran medida la columna vertebral de todo el concepto. La infraestructura es la que permite a los usuarios desarrollar y usar el software y las aplicaciones. El servicio va dirigido hacia el cliente final y a los desarrolladores.

En lugar de comprar y mantener centros de datos o servidores propios, los clientes compran los recursos como un servicio completamente externo. Los proveedores

¹⁸ Cisco. Cisco Global Cloud Index: Forecast and Methodology, 2016–2021. 2018 Cisco and/or its affiliates. All rights reserved. This document is Cisco Public Information.

cobran los servicios según la base establecida y por la cantidad de recursos consumidos.

Es un modelo en el cual la infraestructura básica de cómputo (servidores, software y equipamiento de red) es gestionada por el proveedor como un servicio bajo demanda, en el cual se pueden crear entornos para desarrollar, ejecutar o probar aplicaciones.

Por lo general, el cliente mantiene la capacidad de decisión del sistema operativo y del entorno que instala. Por lo tanto, la gestión de la seguridad corre principalmente a cargo del cliente.

Existen muchos ejemplos de proveedores de Infraestructura como servicio, como es el caso de *Amazon Web Services (AWS)*, *Microsoft Azure*, *Google Compute Engine*, *GoGrid*, *Rackspace* o *ElasticHosts*.

4.4. Cambiando el paradigma presupuestario

Uno de los principales factores a considerar para la adopción del *Cloud Computing* tiene que ver con la estructura de costos, basado en los gastos operacionales (OpEx).

Lo que aporta el modelo OpEx es la capacidad de pagar simplemente por la infraestructura que necesitan para ejecutar las aplicaciones que requieren, basadas únicamente en el consumo de recursos.

Con el modelo tradicional de adquisiciones a través de gastos de capital (CapEx), la compañía compra infraestructura basada en expectativas futuras que en ocasiones se sobredimensiona por lo que se termina comprando capacidad de infraestructura que estará inactiva durante el periodo planeado. Además, por la propia evolución tecnológica, la infraestructura sufre un rápido proceso de depreciación que generalmente ocurre más rápido de lo que las empresas pueden pagar con un préstamo comercial tradicional. Desde una perspectiva financiera, el modelo de adquisiciones a través del CapEx tradicional no es la manera óptima de proporcionar los recursos tecnológicos que la empresa necesita para ejecutar aplicaciones y el almacenamiento de datos.

Es por ello que cada vez con mayor frecuencia ejecutivos y personal financiero de entidades de todo el mundo, solicitan a sus profesionales de TI la adopción de un modelo OpEx para proporcionar la infraestructura necesaria para ejecutar las aplicaciones de la empresa y almacenar sus datos. Estos modelos ciertamente se aplican a la infraestructura de nube pública, pero también pueden aplicarse a la

infraestructura local, ya sea a través del proveedor que suministra la tecnología o a través de fuentes de financiamiento creativas.

Asimismo, se requiere que la selección de una infraestructura esté diseñada de tal manera que pueda ampliarse o reducirse fácilmente con la mayor eficiencia y el menor costo posible. Es por ello que la infraestructura y la virtualización modernas definidas por software permiten a las empresas hacer más eficiente sus compras de infraestructura y escalar más fácilmente su infraestructura cuando es necesario.

No es difícil entender por qué para las empresas es más fácil obtener la aprobación presupuestaria para gastos operativos que para grandes gastos de capital con expectativas inciertas. De hecho, el modelo operativo de muchos departamentos de TI se está alejando de los gastos de capital (CapEx) y se acerca más hacia un modelo financiado principalmente por gastos operativos (OpEx).

El éxito reciente en las áreas de TI se debe al traslado de los recursos corporativos a la infraestructura de nube pública, por ejemplo, en un modelo de servicio de tipo SaaS, donde todos los elementos son operados por un proveedor externo, los recursos de la nube se pueden facturar mensualmente o de manera periódica, como un servicio de telefonía, por lo que el cliente solo le asigna presupuesto sobre demanda al servicio de acuerdo al uso específico que le dé.

Debido a sus ventajas desde el punto de vista financiero, serán preferibles los productos y servicios que requieran poco o ningún gasto de capital, donde el control y la oferta de un modelo esté basado en OpEx, como el alquiler o el arrendamiento, la facturación basada en el uso mensual y la expansión en pequeños segmentos predecibles según la necesidad.

En esta sección se explicaron las principales definiciones de Cloud Computing, así como los modelos de despliegue, características y tipos de servicios bajo los que se ofrece con el fin de comprender los fundamentos de su desarrollo y operación, así como el cambio de paradigma respecto a la consideración de los costos de capital y operación que representa un avance sin precedentes en la manera de prestar servicios digitales en la actualidad y en un futuro.

5. Infraestructura de Cloud Computing

Una vez revisada la información que nos permite comprender los conceptos del Cloud Computing, sus características, modelos de despliegue y de servicios, es oportuno analizar todos aquellos elementos que permiten su funcionamiento, tanto en el plano de infraestructura como en su operación. En esta sección se pretende

relacionar todos aquellos elementos, con sus funciones y características que se consideran relevantes para la implementación de los servicios en la nube.

5.1. Data Centers

Algo vital para el nuevo ecosistema digital es la **infraestructura subyacente** basada principalmente en **Data Centers**. El alcance del Cloud Computing se extiende de forma global con la finalidad de acercar los servicios a los clientes y otorgar mayor disponibilidad y confiabilidad en los servicios digitales y de TI.

Es por este motivo que la clasificación de los Data Centers se ha convertido en una exigencia para los clientes que desean emprender sus actividades en Cloud Computing.

5.1.1. Subsistemas y Nivel de Clasificación

Un Data Center es una construcción de gran tamaño donde se albergan los equipos electrónicos necesarios para mantener una red de computadoras, servidores y aplicaciones que brindan suficiente infraestructura dedicada para respaldar los servicios que ofrece.

Los Data Centers van más allá de la configuración que tendrían en el sitio de un cliente, incluyendo el espacio dedicado para los sistemas de telecomunicaciones. Requieren de una fuente de alimentación ininterrumpida, equipos de enfriamiento dedicado y uno o varios generadores de energía eléctrica para proteger el funcionamiento a pesar de cortes prolongados, lo que le permite incluir, entre otras cosas, sistemas de seguridad sofisticados.

La primera clasificación de los Data Centers se presenta bajo el estándar ANSI/TIA-942-A (Instituto Nacional Estadounidense de Estándares – Asociación de la Industria de Telecomunicaciones, por sus siglas en inglés), la cual proporciona una serie de recomendaciones y directrices para la instalación de su infraestructura.

De acuerdo con dicho estándar, la infraestructura de soporte de un Data Center está compuesto por cuatro subsistemas¹⁹:

- **Telecomunicaciones:** Cableado de armarios (racks), accesos redundantes, cuarto de entrada, área de distribución, backbone (acceso a internet),

¹⁹ Grupo cofitel, (2014). Data Center: El Estándar TIA 942. Recuperado de <https://www.c3comunicaciones.es/data-center-el-estandar-tia-942/>

cableado horizontal, elementos activos redundantes, alimentación redundante, panel y cables de conexión y documentación relacionada.

- **Arquitectura:** Selección de ubicación o layout, tipo de construcción, protección ignífuga (resistencia contra el fuego), barreras de vapor, techos y pisos, áreas de oficina, salas de UPS y baterías, sala de generador, control de acceso, Circuito Cerrado de Televisión (CCTV), Centro de Operaciones de Red (NOC).
- **Sistema eléctrico:** Número de accesos, puntos únicos de falla, cargas críticas, redundancia y topología de UPS, puesta a tierra, EPO (Emergency Power Off - sistemas de corte de emergencia), baterías, monitoreo, generadores, sistemas de transferencia (transfer switch).
- **Sistema mecánico:** Sistemas de climatización, presión positiva, tuberías y drenajes, enfriadores de agua (chillers), aire acondicionado de la sala de computadoras (CRACs) y condensadores, control de aire acondicionado de alta ventilación (HVAC), detección de incendios, aspersores, sistemas de extinción de incendios por agente limpio (NFPA 2001)²⁰, sistemas de detección de humo por aspiración (ASD) y detección de líquidos.

A su vez, el "Uptime Institute" clasifica a los Data Centers en cuatro grupos llamados **TIER**, indicando así su nivel de fiabilidad en función del nivel de disponibilidad, por lo cual, a mayor número de nivel de fiabilidad, mayor disponibilidad. La clasificación es aplicable de forma independiente a cada subsistema de la infraestructura señalado anteriormente, por lo que la clasificación global de un Data Center será igual a la de aquel subsistema que tenga el menor nivel. Esto significa que, si un Data Center tiene todos los subsistemas en nivel IV excepto el eléctrico que es nivel III, la clasificación global del mismo será de nivel III.

TIER I: Data Center básico. Un Data Center nivel I puede ser susceptible a interrupciones tanto planeadas como no planeadas. Cuenta con sistemas de aire acondicionado y distribución de energía; pero puede o no tener piso técnico, UPS o generador eléctrico; si los posee pueden no tener redundancia y existir varios puntos únicos de falla. La carga máxima²¹ de los sistemas en situaciones críticas es del 100%.

La tasa de disponibilidad máxima del Data Center es 99.671% del tiempo.

²⁰ NFPA (2018). NFPA 2001 Standard on Clean Agent Fire Extinguishing Systems. Recuperado de <https://www.nfpa.org/codes-and-standards/all-codes-and-standards/list-of-codes-and-standards/detail?code=2001>

²¹ Turner, W.P., Seader, J., Brill, K (2006). Tier Classifications Define Site Infrastructure Performance, 3-5.

TIER II: componentes redundantes. Los Centros de Datos con componentes redundantes son ligeramente menos susceptibles a interrupciones, tanto planeadas como las no planeadas. Estos centros de datos cuentan con piso falso, UPS y generadores eléctricos, pero están conectados a una sola línea de distribución eléctrica. Su diseño es en redundancia de N+1 (lo necesario más uno), lo que significa que existe al menos un duplicado de cada componente de la infraestructura. La carga máxima de los sistemas en situaciones críticas es del 100%. El mantenimiento en la línea de distribución eléctrica o en otros componentes de la infraestructura pueden causar una interrupción del procesamiento.

La tasa de disponibilidad máxima del Data Center es 99.749% del tiempo.

TIER III: mantenimiento concurrente. Las capacidades de un Data Center de este tipo le permiten realizar cualquier actividad planeada sobre cualquier componente de la infraestructura sin interrupciones en la operación. Actividades planeadas incluyen mantenimiento preventivo y programado, reparaciones o reemplazo de componentes, agregar o eliminar elementos y realizar pruebas de componentes o sistemas, entre otros. Para infraestructuras que utilizan sistemas de enfriamiento por agua significa doble conjunto de tuberías.

Debe existir suficiente capacidad y doble línea de distribución de los componentes, de forma tal que sea posible realizar mantenimiento o pruebas en una línea, mientras que la otra atiende la totalidad de la carga. En este nivel, las actividades no planeadas como errores de operación o fallas espontáneas en la infraestructura pueden todavía causar una interrupción del Data Center. La carga máxima en los sistemas en situaciones críticas es de 90%.

Muchos centros de datos TIER III son diseñados para poder actualizarse a TIER IV, cuando los requerimientos del negocio justifiquen el costo.

La tasa de disponibilidad máxima del Data Center es 99.982% del tiempo.

TIER IV: tolerante a fallas. Este Data Center provee capacidad para realizar cualquier actividad planeada sin interrupciones en las cargas críticas, pero además la funcionalidad tolerante a fallas le permite a la infraestructura continuar operando aun ante un evento crítico no planeado. Esto requiere dos líneas de distribución simultáneamente activas, típicamente en una configuración sistema + sistema; eléctricamente esto significa dos sistemas de UPS independientes, cada sistema con un nivel de redundancia N+1. La carga máxima de los sistemas en situaciones críticas es de 90% y persiste un nivel de exposición a fallas, ya sea por una alarma de incendio o porque una persona inicie un procedimiento de apagado de emergencia (EPO), los cuales deben existir para cumplir con los códigos de seguridad contra incendios o eléctricos.

La tasa de disponibilidad máxima del Data Center es 99.995% del tiempo.

Para poner en perspectiva la tasa de disponibilidad que se pretende para los distintos niveles, en la tabla de clasificación estándar se establece su significado expresado en el tiempo de parada anual del Data Center. Estos porcentajes deben considerarse como el promedio de cinco años.

Tabla 5: Clasificación estándar ANSI/TIA-942.

TIER	% Disponibilidad	% Parada	Tiempo anual de parada
TIER I	99.671%	0.329%	28.8 horas
TIER II	99.749%	0.251%	22 horas
TIER III	99.982 %	0.018%	1.6 horas
TIER IV	99.995%	0.005%	26.3 minutos

Fuente: Data Centers: Jobs and Opportunities in Communities Nationwide, 2017.

Según el Uptime Institute, hay que tener en cuenta que para un TIER IV se contempla que la única parada que se produce es por la activación de un procedimiento de apagado de emergencia, lo cual no sucede más de una vez cada cinco años.

Además de la clasificación anterior del Uptime Institute, existen otros organismos e instituciones que clasifican y certifican los Data Centers²², los cuales se encuentran relacionados en el ANEXO 2 del presente documento.

De acuerdo a estimaciones de CBRE, en México los Data Centers ocupan una superficie de más de 250,000 metros cuadrados distribuidos en 49 ubicaciones de seis estados del país: Ciudad de México, Estado de México, Querétaro, Nuevo León, Jalisco y Aguascalientes, siendo los primeros tres estados los que concentran el mayor número. Lo anterior convierte al país en el segundo mercado de Data Centers más grande de Latinoamérica, solo por detrás de Brasil, siendo los sectores de las telecomunicaciones, finanzas, comercio electrónico y gobierno los que han representado el mayor impulso a dicho mercado. Asimismo, expertos prevén el crecimiento de los Data Center en el país durante los próximos cinco a diez años para posteriormente comenzar una etapa de maduración, esto al haber un cada

²² DCD (2017). Los nuevos estándares de Centros de Datos: Tier 5. Ramos, Isidro. Recuperado de <https://www.dcd.media/opinion/los-nuevos-est%C3%A1ndares-de-centros-de-datos-tier-5/>

vez mayor reconocimiento sobre la relevancia del análisis de la información y su almacenamiento.^{23 24 25}

Cabe destacar que una de las principales ventajas para la construcción de Data Centers en México es que posee una ubicación geográfica estratégica, ya que el país es el punto intermedio para ofrecer conectividad entre Estados Unidos y el resto de América Latina, por lo que existe un gran potencial de desarrollo en este mercado para generar una mayor inversión en los próximos años.

5.2. Capacidad de procesamiento y almacenamiento de Centros de Datos

Como se ha señalado, los Data Centers son instalaciones equipadas principalmente con servidores y computadoras en red para las comunicaciones y el procesamiento y almacenamiento de datos. Los grandes Data Centers suelen consistir en carcassas apiladas con bastidores de servidores y equipos de TI en un piso elevado con sistemas de respaldo de energía y control de temperatura, y la gran mayoría de Data Centers de gran capacidad también tienen sus propios generadores de energía para equipos de calefacción y refrigeración.

Existen métricas establecidas para el desempeño de los Data Centers y sus redes asociadas con las cuales se pueden evaluar propiedades tales como su capacidad de almacenamiento, transporte, latencias, capacidad de balanceo de cargas, tolerancia a fallas y costos de construcción. Estas propiedades dependen en gran medida de la localización y equipamiento de las instalaciones, cuestiones que se tratan a continuación.

5.2.1. Localización

La motivación para ubicarse en un lugar determinado, hoy en día, difiere mucho de los criterios para tener acceso a materias primas o las consideraciones de mano de obra barata como sucede en diversas industrias. Los Data Centers no necesitan una gran cantidad de mano de obra no calificada como manufactura y centros de llamadas (call centers), y no están vinculados a la ubicación por las materias primas, por lo que físicamente pueden ubicarse en cualquier lugar donde tengan acceso a

²³ Medina, A. (2019). Un viaje al centro de un Data Center: el epicentro de la vida digital. Recuperado de <https://www.forbes.com.mx/un-viaje-al-centro-de-un-data-center-epicentro-vida-digital/>

²⁴ Cervantes, L. (2019). México el segundo mercado más importante de Data Centers en Latinoamérica. Recuperado de <https://elceo.com/bienes-raices/mexico-el-segundo-mercado-mas-importante-de-data-centers-en-latinoamerica/>

²⁵ Ventura, P. (2019). Data Center desarrolla 'conexiones' para ganar terreno en el mercado inmobiliario industrial. Recuperado de <https://www.elfinanciero.com.mx/empresas/data-center-desarrolla-conexiones-para-ganar-terreno-en-el-mercado-inmobiliario-industrial>

una cantidad significativa de energía y conectividad a redes de comunicaciones nacionales e internacionales de alta velocidad.

Además, influyen para la determinación de una ubicación entre países aspectos regulatorios como diferencias en la estructura tributaria, las leyes, la propiedad de bienes inmuebles, la regulación de seguridad de los datos o información por vulnerabilidades naturales, disturbios políticos, etc.

Sin embargo, hay una diferencia que no se elimina fácilmente con la tecnología o los factores financieros disponibles, la latencia: las señales de comunicación no pueden propagarse más rápido que la velocidad de la luz y esto se relaciona directamente con el retardo o latencia de las comunicaciones en función de la distancia.

Como ejemplo de restricciones de selección de sitios, consideremos una ruta desde Bombay, India, a la ciudad de Nueva York, Estados Unidos. La velocidad de la luz en el espacio libre es de 299,792 kilómetros por segundo (km/s). En fibra óptica la transmisión de la luz es menor, pero vamos a considerar la velocidad en el espacio libre por simplicidad en los cálculos. La distancia directa entre ambas ciudades es de aproximadamente 12,552.88 Km. y, suponiendo que una ruta de fibra sería 50% más larga que una ruta directa, la señal tarda aproximadamente unos 60 ms (1 milisegundo es igual a 0,001 segundos), para que se propague y llegue a su destino, es decir, 60 ms es en teoría la mejor latencia de señal posible entre estos dos puntos.

Hoy en día, la latencia esperada entre los Estados Unidos y la India es entre 250 y 500 ms. Las mejoras de la red, tanto en arquitectura como en topología, probablemente bajarán esta cifra con el tiempo a unos 200 ms. La forma en que esto limita la selección del sitio depende en gran medida de las aplicaciones que se utilicen²⁶.

Una de las aplicaciones más comunes sensibles a la latencia, es el video streaming, pues si se tiene una alta latencia debido a que el material audiovisual se encuentra alojado a una distancia física considerable, el contenido tardará más en encontrarse disponible para el solicitante, esto causa que el video se detenga por la falta de los paquetes necesarios para su reproducción, lo que genera una mala experiencia al usuario.

Otras aplicaciones sensibles a la latencia serían las orientadas a la transacción, como la banca y las ventas minoristas, donde, debido a una tasa de transacción muy alta, la latencia debe ser muy baja y se requieren tasas de transmisión considerables para satisfacer la alta demanda.

²⁶ La calidad de la transferencia de información de un punto de origen a uno final (end to end) está dado por partes o secciones de red que pueden ser heterogéneas y de diferentes operadores. Esto se refleja en el modelo de referencia ITU-T Y.1540 que define los parámetros de desempeño de una red IP.

Es importante señalar qué bajo un enfoque tradicional, es decir, donde la carga de trabajo se encuentra centralizada en Data Centers ubicados en el núcleo de la red y se gestiona desde su origen sin que intervengan procesos que agilicen su manejo, se provoca que la capacidad de los dispositivos finales ubicados en el extremo opuesto de la red se vea limitada en su rendimiento a causa de la latencia, la cual puede exceder los 60 ms.

Asimismo, dentro de un enfoque regionalizado, los Data Centers se encuentran distribuidos en núcleos regionales, de tal forma que las cargas de trabajo se gestionan en un nodo más cercano a los clientes, y provoca que la capacidad de los dispositivos finales en el extremo opuesto de la red sea eficiente en su rendimiento y gestión. Bajo este enfoque se busca que la latencia no exceda los 60 ms.

Finalmente, considerando un enfoque localizado, las cargas de trabajo se gestionan de manera local, es decir, los componentes de procesamiento son colocados en el **borde de la red**, complementándose mediante la interacción con el núcleo regional, lo que provoca el óptimo rendimiento y gestión de los dispositivos finales en el extremo opuesto de la red. En este caso, la latencia disminuye considerablemente, pudiendo ser menor o igual a los 20 ms²⁷.

Si bien, algunas organizaciones podrían obtener beneficios económicos al seleccionar una ubicación fuera del país, por los altos niveles de latencia no siempre será una opción adecuada para todos los servicios, puesto que los nuevos servicios digitales requieren una mayor calidad de conectividad para operar correctamente o brindar una mejor experiencia de usuario.

5.2.2. Cambios de arquitectura tecnológica

Con el paso del tiempo, los Data Centers han ido evolucionando debido a cambios arquitectónicos claves en los bloques de construcción, red/conmutación, cómputo/servidor-almacenamiento, entre otros aspectos. A continuación, se presenta una breve descripción de los cambios que se están generando en su equipamiento.

La tecnología de los servidores está cambiando rápidamente. Las empresas de hiperescala, definidos como los que tienen un tamaño mayor a 400,000 pies cuadrados (37,161.21 m²), han aumentado en importancia e influencia. La

²⁷ EQUINIX. (2019). Global Interconnection Index. Recuperado de https://www.equinix.com/gxi-report/?ls=Advertising%20-%20Web&lscd=19q3_cross-vertical_digital-edge+indexvol3_programs-global_Equinix-run_equinix-owned-site_us-en_AMER_GX13_awareness&utm_campaign=us-en_equinix-owned-site_GX13_programs-global_awareness&utm_source=&utm_medium=equinix-owned-site&utm_content=digital-edge+index-vol3_Report_Further_Insights_pg11

velocidad de transmisión es un factor crítico a medida que aumentan los requisitos de rendimiento. Si bien en general la industria está operando a velocidades de 8 Gbps, las compañías ahora están buscando 16, 25 o 32 Gbps. Las restricciones térmicas que representan estas velocidades están impulsando cambios en la arquitectura de los bastidores.

Respecto a los protocolos de almacenamiento, el que utiliza Flash²⁸ es cada vez más frecuente para impulsar el empaquetado y el rendimiento de alta densidad. Vemos esta creciente demanda de mayor flexibilidad y rendimiento, que proporciona soluciones de interconexión innovadoras. Las capacidades de rendimiento del almacenamiento Flash moderno, son la única razón por la que es posible lograr un rendimiento aceptable desde una plataforma a Hiperescala.

El almacenamiento Flash es más rápido que el disco magnético debido a su naturaleza de estado sólido (no mecánico). Una sola unidad de estado sólido (SSD²⁹) podría entregar fácilmente el rendimiento IOPS³⁰ de 24 discos giratorios. Debido a este beneficio en el rendimiento, el almacenamiento flash es crítico para la hiperconvergencia. Por su parte los SSD también pueden proporcionar un caché de alto rendimiento que puede poner en funcionamiento una gran cantidad de capacidad. El uso de SSD como caché permite a las plataformas hiperconvergentes³¹ obtener un alto rendimiento y una gran capacidad al mismo tiempo.

Como lo señalado anteriormente, hay muchas otras manifestaciones que derivan en mejoras a la arquitectura de los Data Centers, las tendencias de conmutación indispensables para la convergencia en redes de transporte 5G, las arquitecturas en los circuitos como las unidades de procesamiento de gráficos, etc., las cuales obligan a reconfigurar las redes de telecomunicaciones actuales, diseñadas inicialmente para el tráfico de voz y de datos con características de conectividad no totalmente distribuidas (jerarquías de centrales), y no necesariamente masivas o de anchos de banda tan exigentes, por lo que dichas redes resultan no necesariamente óptimas para conducir los nuevos servicios digitales y de TI.

²⁸ Flash: Dispositivo de Almacenamiento de Datos

²⁹ Unidad de estado sólido, SSD por sus siglas en inglés (solid-state drive)

³⁰ IOPS (Input/Output Operations Per Second, pronunciado i-ops) es una medida del rendimiento de referencia común para los dispositivos informáticos de almacenamiento como unidades de disco duro (HDD), unidades de estado sólido (SSD) y redes de área de almacenamiento (SAN).

³¹ La hiperconvergencia elimina las incidencias de la gestión de la TI tradicional agrupando servicios de centro de datos como el servidor, el almacenamiento y la red, y permite que se gestionen en una única aplicación. A diferencia de la convergencia, la hiperconvergencia es una infraestructura definida por software que separa las operaciones de la infraestructura del hardware del sistema y las converge a nivel de hipervisor en un bloque único (y, por tanto, hiperconvergente). Los sistemas hiperconvergentes aprovechan la inteligencia definida por software para desglosar los silos de almacenamiento y procesamiento, y permiten que se ejecuten y se gestionen en la misma plataforma de servidor, lo que elimina las ineficiencias y acelera el procesamiento.

Tabla 6: Los Data Centers en México.

Poco a poco, las empresas de telecomunicaciones se están dando cuenta de la enorme responsabilidad que representa ofrecer servicios de Cloud Computing, además de la enorme inversión que requieren en la reconversión de su infraestructura. Sea cual fuere el motivo, algunas empresas ya han vendido algunos de sus Data Centers, a pesar de haber obtenido las diversas certificaciones de calidad, como se muestra en la siguiente tabla, donde ICREA³² reconoce desde 2013 las mejores prácticas de los Data Centers Mexicanos y de América Latina.

Tabla 7: Data Centers reconocidos con la NORMA ICREA 2019.

Nivel	Disponibilidad %	Data Centers reconocidos en 2019
Nivel I	95%	Insys, Universidad Autónoma de Nuevo León, Nexoit y el Hospital Regional de Tlalnepantla - Infraestructura Hospitalaria del Estado de México
Nivel II	99%	BConnect Services, GSI/Cometra, Universidad Autónoma de Nuevo León, NETMIND Solutions y Teleinformática en Servicios Avanzados.
Nivel III	99.9%	ICA Flour, Disaster Recovery, Nacional Financiera, EYC, Rapiscan y Auditoría Superior de la Federación - Centro de Datos Ajusco.
Nivel IV	99.99%	ENTEL (Chile), Cloud Data, Micronet, IBM de México, Data Center GlobeNet Barranquilla y ATALA IT Centro de Procesamiento de Datos "Gobernadores" y EPM-Millicom Cellular Data Center Tigo Medellín.
Nivel V	99.999%	Grupo Santander, Triara, KIO Networks, Axtel y Telmex Colombia.
Nivel VI	99.9999%	No aplica

Nota: Los niveles y porcentajes que se presenta en esta tabla son diferentes a los niveles que maneja el "Uptime Institute"

Fuente: Elaboración propia con información del ICREA

(<https://itcomunicacion.com.mx/mejores-data-centers-mexicanos/>)

Así, Equinix cerró la adquisición de dos de los cinco Data Centers de Axtel y Telefónica Movistar cedió 11 de sus centros de Datos incluyendo el de México a Asterion Industrial Partners, ambas operaciones se cerraron a finales de 2018.

La importancia que tienen los Data Centers en México es evidente y para entender el grado de presencia de éstos, basta con recordar dos eventos ocurridos en 2019. En mayo, una vez anunciado el final de Game of Thrones con una audiencia en streaming de cerca de 19 millones de personas a nivel mundial, la plataforma encargada de su difusión en México no soportó el alto volumen de usuarios que se conectó a la hora indicada, ocasionando intermitencias en el servicio y una baja calidad en la imagen.

En otro evento frustrante tres meses después, una falla eléctrica de uno de los Data Centers de PROSA³³ dejó por 7 horas a millones de usuarios sin sus tarjetas de crédito y débito en todo el país.

A pesar de los eventos señalados, no se puede negar que los Data Centers son hoy el corazón del ecosistema digital, y ofrecen sus servicios como insumo a empresas de diversos sectores.

³² **ICREA** "International Computer Room Experts Association" es una asociación Internacional sin fines de lucro formada por ingenieros especializados en el diseño, construcción, operación, administración, mantenimiento, adquisición, instalación y auditoría de centros de cómputo. (<https://icrea-international.org/>)

³³ **PROSA** es una empresa enfocada a los medios de pago que integra soluciones tecnológicas, además de ser el procesador de transacciones electrónicas más grande de América Latina.

5.3. Dimensionamiento y consumo de energía

5.3.1. Tamaño y Rendimiento de Data Centers

Como se ha mencionado a lo largo de este documento, los Data Centers cobran una importancia neurálgica en la economía moderna y la adopción del Cloud Computing. Sin embargo, la explosión de contenido digital, los datos masivos, el comercio electrónico y el tráfico de Internet también está haciendo de los Data Centers uno de los consumidores de energía de mayor crecimiento.

Esto incluye la concentración de cargas de trabajo en Data Centers de Hiperescala. El uso de energía por parte de los Data Centers también puede disminuir si se traslada más carga de trabajo a centros de Hiperescala, debido a la capacidad de estos últimos para escalar su arquitectura de manera apropiada de acuerdo al incremento de la demanda.³⁴

El Data Center Institute³⁵ clasifica los Data Centers en seis tipos de acuerdo a su tamaño, que se miden por espacio de cómputo o rendimiento de rack. El rendimiento de rack es el número de racks que pueden caber dentro de un espacio total de cómputo. El espacio de cómputo es el área, medida en pies cuadrados (sf) o metros cuadrados (m²), dentro de la instalación del Data Center que contiene bastidores de servidores y equipos de TI relacionados. Normalmente, un rack o bastidor se establece en 2.32 m² para permitir el espacio de pasillo y perímetro alrededor de la sala de servidores como se puede ver en la siguiente tabla.

Tabla 8: El tamaño de los Data Centers por rendimiento.

Tamaño	Rendimiento en Rack (Racks)	Espacio de computo (m ²)
Mega	> 9,000	> 20,903.18
Masivo	3,001 - 9,000	6967.73 - 20,903.18
Grande	801 - 3,000	1858.15 - 6967.73
Mediano	201 - 800	464.60 - 1858.15
Pequeño	11 - 200	23.31 - 464.60
Mini	1 - 10	0.09 - 23.31

Fuente: Data Centers: Jobs and Opportunities in Communities Nationwide, 2017.

Con cada vez mayor frecuencia, los Data Centers requieren sistemas de suministro de energía ininterrumpida para minimizar el tiempo de inactividad de los servidores y los sistemas de seguridad para sus usuarios. Los costos de infraestructura del Data

³⁴ Cisco Global Cloud Index Supplement: Cloud Readiness Regional Details

³⁵ C_TEC (U.S. Chamber Technology Engagement Center), Journal Article: Data Centers: Jobs and Opportunities in Communities Nationwide, 2017.

Center y las complejidades operativas aumentan con el nivel de confiabilidad o TIER (ver sección Subsistemas y Nivel de Clasificación).

Para satisfacer la creciente demanda de servicios en la nube, los Data Centers están creciendo en tamaño rápidamente. Generalmente, el tamaño de un Data Center se define por el área de cobertura y la relación volumétrica (la relación entre el área de cobertura y el número de racks). De acuerdo con esta área de cobertura, los Data Centers se pueden clasificar en tamaños enorme, grande, mediano y pequeño como se muestra en la siguiente tabla. (Chen, Gao, & Chen, 2016)

Tabla 9: El tamaño de los Data Centers por m².

Tamaño	Área de cobertura m2	Ejemplos
Enorme	> 30480	Microsoft Quincy
Grande	6096 - 30480	Oracle Austria
Mediano	1524 - 6096	Sinopec group
Pequeño	609 - 1524	SJTU

Fuente: A survey. Journal of Parallel and Distributed Computing, (2016).

De acuerdo con la cantidad de racks, también podemos dividir los Data Centers en escala mega (> 10,000 racks), masiva (3000–10,000 racks) y grande (800 – 3000 racks), mediana, pequeña y mini (<3000 racks). La relación volumétrica es un indicador más efectivo que puede reflejar la utilización de un Data Center. Cuanto menor es la relación volumétrica, mayor es la utilización de racks.

Desde que los Data Centers son la infraestructura fundamental que soporta muchas aplicaciones, incluyendo el Cloud Computing, esta relación volumétrica es un indicador fundamental clave para medir el rendimiento de los mismos.

Tabla 10: Relación volumétrica de Data Centers.

Nombre	Área de cobertura (m ²)	Número de Racks	Relación Volumétrica
Google Dalles	18,580.60	9090	2.0
Oracle Austria	7,618.05	2280	3.3
Cisco Richardson	3,158.70	1151	2.7

Fuente: A survey. Journal of Parallel and Distributed Computing, (2016).

Uno de los aspectos más importante de un Data Center es la Red de Centros de Datos (DCN, por sus siglas en inglés) ya que juega un papel fundamental en la informática y la comunicación. Conecta también los componentes físicos de los

Data Centers, como los servidores, conmutadores, etc., en una topología específica con cables y fibras ópticas, y la eficiencia y el rendimiento de un Data Center depende en gran medida de la DCN. Dado que SIGCOMM (la conferencia insignia del Grupo de Interés Especial de ACM sobre Comunicación de Datos) estableció por primera vez una sesión sobre redes de centros de datos en 2008, el diseño de la arquitectura se ha convertido en un campo de investigación muy activo para mejorar la eficiencia y el rendimiento de DCN. Se han diseñado y presentado muchas arquitecturas novedosas, y se han conectado muchos dispositivos novedosos a DCN, como antenas inalámbricas e interruptores ópticos. (Chen, Gao, & Chen, 2016)

Así, el rendimiento se convirtió en uno de los objetivos de los Data Centers, hasta llegar a los Data Centers ecológicos que desarrollan y emplean tecnologías de ahorro de energía como el diseño modular de los Data Centers, el uso de unidades de energía avanzadas y la renovación de recursos.

El uso efectivo de energía o PUE (Power Usage Effectiveness) es un indicador clave para determinar el rendimiento de los Data Centers y relaciona la energía utilizada por los equipos informáticos y la energía total utilizada por el Data Center (incluidos los equipos informáticos, la refrigeración y otros gastos generales). El valor esperado es 1, lo que significa que la energía consumida es aprovechada exclusivamente por los equipos informáticos, sin embargo, es difícil alcanzar este valor debido a los sistemas de enfriamiento que consumen una parte considerable de la energía eléctrica.

Tabla 11: Data Centers con indicadores PUE bajos.

Nombre	PUE
HP EcoPOD Data Center	1.05
Facebook Prineville Data Center	1.07
Yahoo New York Data Center	1.08
Capgemini Merlin modular Data Center	1.08
Google Saint-Ghislain Data Center	1.16
Microsoft Dublin Data Center	1.25

Fuente: A survey. Journal of Parallel and Distributed Computing, (2016).

De acuerdo con el área de cobertura, los 10 Data Centers más grandes del mundo para 2015 se enumeran en la siguiente Tabla, que van desde 228 mil metros cuadrados a 1.9 millones de metros cuadrados. Otros mega Data Centers, contienen campus de instalaciones múltiples o edificios de uso mixto donde el espacio de Data Center coexiste con un gran espacio de oficinas de terceros (es decir, hoteles de transportistas de grandes ciudades).

Tabla 12: Los diez centros de datos más grandes.

Nombre	Área de cobertura (m ²)
Range International Information Group	1,920,240
Switch SuperNAP	1,066,800
DuPont Fabros Technology	487,680
Utah Data Centre	457,200
Microsoft Data Centre	365,760
Lakeside Technology Center	335,280
Tulip Data Centre	304,800
QTS metro Data Center	301,752
NAP of the Americas	228,600
Next Generation Data Europe	228,600

Fuente: A survey. Journal of Parallel and Distributed Computing, (2016).

Una de las características de los grandes Data Centers es que son intensivos en capital y requieren importantes inversiones en tiempo y dinero para construir. Es muy común que, a lo largo del tiempo, se agreguen centros de datos nuevos y más grandes al mismo sitio o campus del primer Data Center. Por ejemplo, *Switch*, que diseña, construye y opera algunos de los Data Centers más avanzados, ha estado construyendo y expandiendo continuamente su núcleo central de Las Vegas. Una vez finalizados los 12 Data Centers, Switch Campus "Las Vegas" cubre casi 222,967.3 metros cuadrados, con un consumo de 315 MW de potencia.

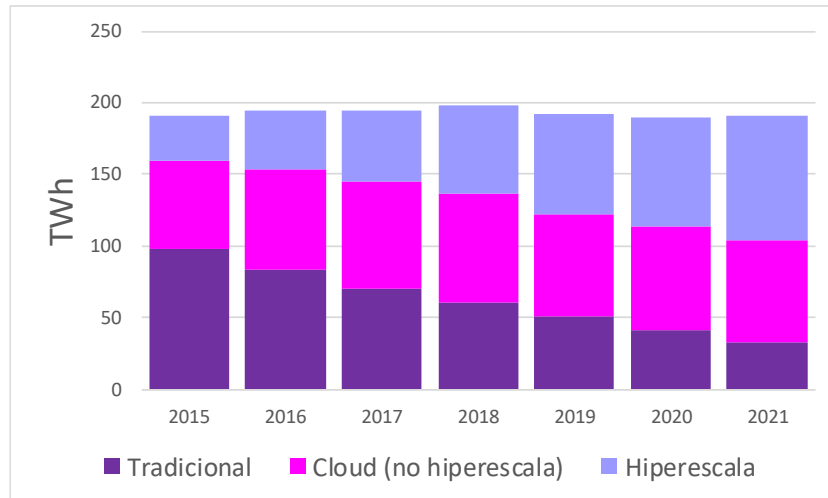
5.3.2. Consumo de energía eléctrica

El tamaño de los Data Centers se relaciona con el consumo de energía eléctrica que requieren. El fuerte crecimiento en la demanda de servicios de Data Centers continúa siendo compensado por las continuas mejoras en la eficiencia de los servidores, dispositivos de almacenamiento, conmutadores de red e infraestructura, así como por un cambio a una participación mucho mayor de Data Centers de Hiperescala, como se puede apreciar en la siguiente gráfica³⁶. Los Data Center de Hiperescala son centros de datos en la nube muy eficientes y de gran escala que funcionan a gran capacidad, debido en parte al software de virtualización que

³⁶ Elaboración propia con información consultada en el sitio de Internet (<https://www.iea.org/reports/tracking-buildings/data-centres-and-data-transmission-networks>)

permite a los operadores de Data Centers ofrecer una mayor producción de trabajo con menos servidores.³⁷

Gráfica 5: Demanda global de energía por tipo de Data Center.



Nota: Los Data Centers tradicionales son aquellos que no brindan servicios de Cloud Computing.

Fuente: Agencia Internacional de Energía (IEA, por sus siglas en inglés)

Derivado del rápido crecimiento y el potencial de una abrumadora demanda de electricidad de los Data Centers, el Congreso de los Estados Unidos solicitó un informe que estimó que los Data Centers de los Estados Unidos habían consumido alrededor de 61 mil millones de kilowatts-hora (KWh) en 2006 (1.6% de las ventas totales de electricidad de los Estados Unidos).³⁸

Con ayuda de este informe se generó un modelo que es usado para estimar el consumo de energía de los Data Centers (E), como la suma del consumo de electricidad de las siguientes cuatro categorías de equipos de Data Centers basados en una taxonomía ampliamente utilizada por la International Data Corporation (IDC)³⁹:

- Servidores (E^S),
- Almacenamiento (E^{ST}),
- Red (E^P), e
- Infraestructura (E^I)

³⁷ Data Centers and data transmission networks, International Energy Agency.

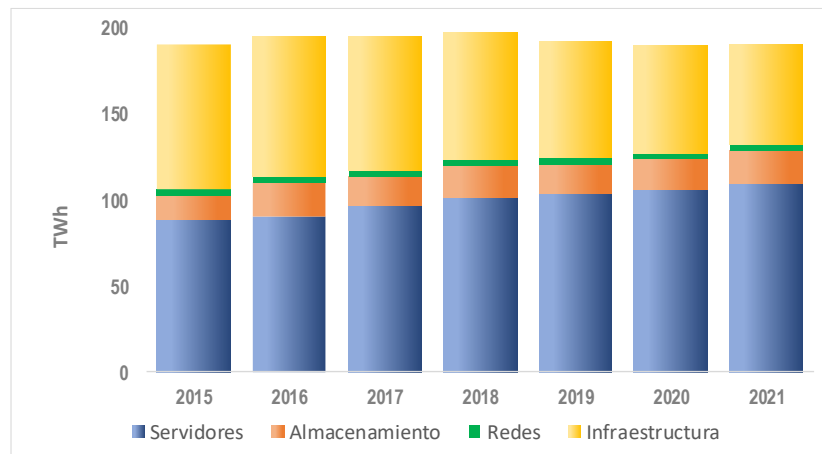
³⁸ IOPscience, (s.f). Data center growth in the United States. Recuperado de <https://iopscience.iop.org/article/10.1088/1748-9326/aaec9c/pdf>

³⁹ US Datacenter Census and Construction 2014–2018 Forecast: Realigning Workloads, Villars

$$E = E^S + E^{ST} + E^P + E^I$$

En relación con lo anterior se observa el siguiente gráfico,⁴⁰ que muestra el total de consumo eléctrico por cada una de las categorías o funciones de operación, así como el consumo de energía global de los Data Center por año, desde el 2015 hasta una estimación al 2021:

Gráfica 6: Consumo global de energía por categoría.



Nota: La infraestructura de Data Center se refiere a la energía consumida por equipo que no es de Tecnologías de la Información como equipo de refrigeración e iluminación. Información actualizada al 7 de enero de 2020.

Fuente: Agencia Internacional de Energía (IEA, por sus siglas en inglés)

A efecto de tener una idea más práctica sobre la dimensión y escala del consumo de energía eléctrica de una instalación de Data Center sigamos un ejemplo simplificado⁴¹.

Considerando que una computadora o un servidor estándar requiere en promedio de 460 watts/hora. Se puede comparar el dato del consumo de un servidor que esta encendido todo el día con el consumo promedio de un hogar también en un día, es decir, si multiplicamos los 460 Watts/hora por 24 horas obtenemos 10,800 Watts/hora o 10.8 KWh.

Un rack o bastidor se configura con 10 servidores en promedio, y entonces tenemos un requerimiento de energía de 4600 watts/hora.

⁴⁰ Elaboración propia con información de reportes de la Agencia Internacional de Energía (iea, por sus siglas en inglés). (<https://www.iea.org/reports/tracking-buildings/data-centres-and-data-transmission-networks>)

⁴¹ Allen, M. (2015). Data Center Power Costs and Requirements. Recuperado de <https://www.datacenters.com/news/data-center-power-costs-and-requirements>

Aplicando un factor de potencia del 80%, porque no toda la energía es aprovechada, nos quedan $(4600 \text{ watt/hora} \times 0.8) = 3680 \text{ watts/hora}$. Este número lo dividimos entre 1000 para convertirlo a 3.68 Kilowatts/hora (KWh).

Para obtener las horas útiles de un mes multiplicamos las 24 horas por 30 días, que resultan en 720 horas.

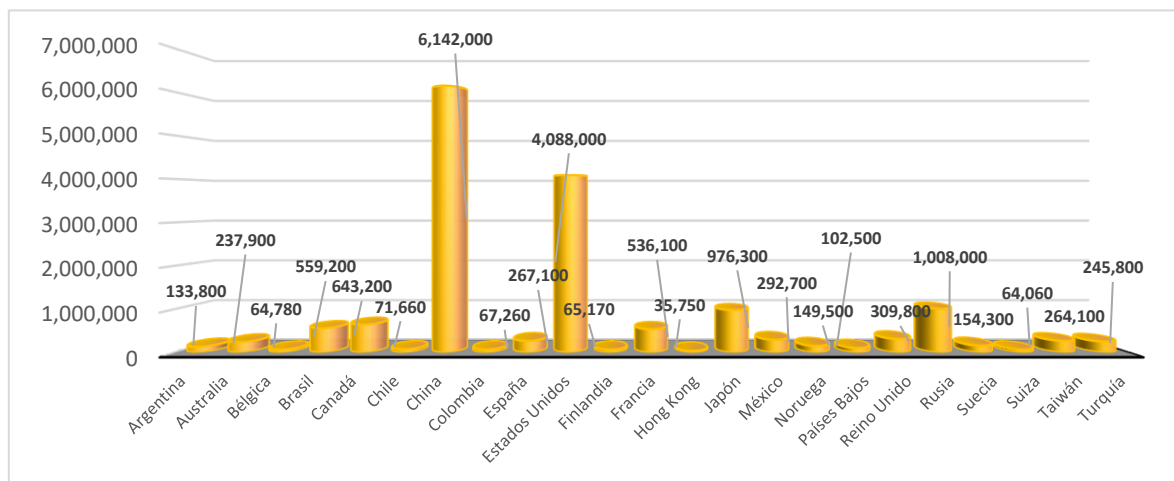
Para obtener el consumo de nuestro rack o bastidor, multiplicamos los 3.68 kilowatts/hora por las 720 horas totales mensuales que resultan en 2,649.6 KWh, al mes, o el equivalente al consumo de energía de 10 hogares en el mismo periodo como lo señalamos antes.

Con esta referencia podemos extrapolar o estimar lo que un Data Center de cualquier escala de las antes mencionadas podría demandar en términos de energía, lo cual nos lleva a magnitudes de los miles (KW), o hasta millones (MW) de watts de potencia al mes en función del número de racks o bastidores que ocupe.

El factor de energía es de primordial importancia para evaluar en qué lugar sería viable la inversión en un Data Center como soporte a los servicios en la nube. Se debe considerar la capacidad de generación y costo de la energía de cada región o país.

Por ejemplo, para México en 2018 se tiene el dato de generación de 292,700 KWh anuales, que se compara con países como Australia, España, Taiwán y Turquía, pero lejos de Brasil, Canadá, Francia, Japón y Rusia con más del doble de generación y aún más lejos de China y USA con más de diez veces su rango de generación de energía.

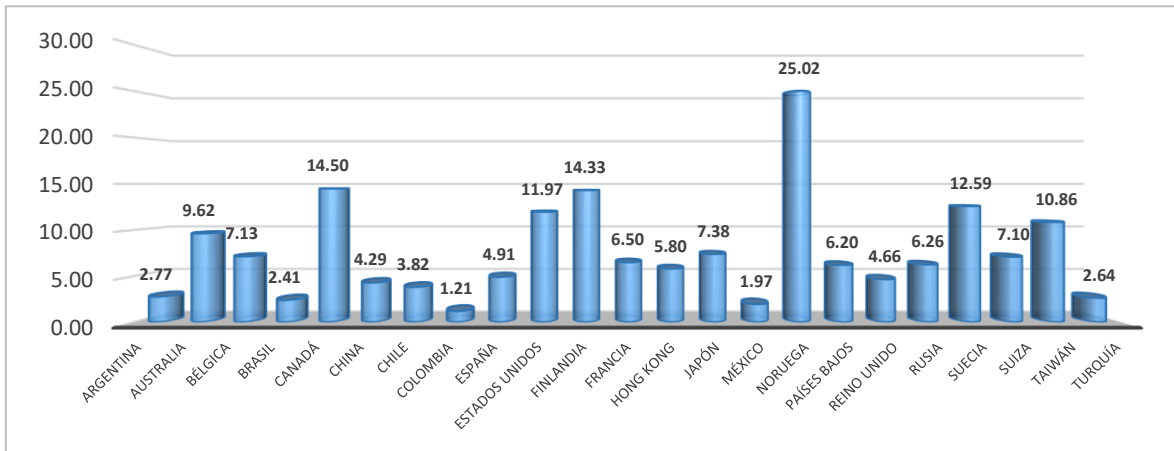
Gráfica 7: Producción (miles de millones kWh) de energía eléctrica por país 2018.



Fuente: Elaboración propia con información de IndexMundi (<https://www.indexmundi.com/g/g.aspx?v=81&c=ar&l=es>)

Esta información es comparable con el consumo per cápita de electricidad en diferentes países, donde se aprecia que México registra un nivel que solo supera a Colombia en toda la gráfica.

Gráfica 8: Consumo de electricidad per-cápita (kWh) 2018.



Fuente: Elaboración propia con información de IndexMundi (<https://www.indexmundi.com/g/g.aspx?v=81&c=ar&l=es>)

Otras fuentes indican que se requieren de 500 MW de electricidad para alimentar a todos los centros de datos en operación hoy en día⁴². Sea cual fuera la estimación más precisa el mensaje es claro respecto a que los servicios en la nube y su soporte de infraestructura en Data Centers demandan un alto suministro de energía, y por supuesto el costo que esto representa en función de las diferentes fuentes de generación accesibles en diferentes lugares. Al respecto se cuentan con indicadores complementarios que permiten determinar el desempeño de este tipo de instalaciones en el ANEXO 3: Indicadores de Eficiencia.

Podemos concluir que el consumo de energía es uno de los principales componentes de costo de los Data Centers, por lo que quienes invierten en instalación / desarrollo de estas infraestructuras, entre otros factores buscan acceso a fuentes de generación accesibles, ya que, de lo contrario, los elevados costos por el suministro de energía reducen la competitividad y atractivo para realizar inversiones de este tipo.

⁴² Allen, M. (2015). Data Center Power Costs and Requirements. Recuperado de <https://www.datacenters.com/news/data-center-power-costs-and-requirements>

5.4. Conectividad y Gestión de Redes.

El incremento de usuarios, la demanda de nuevos servicios digitales como los provistos por el Cloud Computing, los requerimientos de mayor calidad en la conectividad y el continuo aumento por parte de los usuarios para transmitir grandes volúmenes de información, impulsan fuertemente la evolución de las redes, las cuales deben cubrir los nuevos requerimientos que exige la transformación digital.

La conectividad y el Cloud Computing, hoy día no están completamente unidos, porque pueden ser empresas distintas las que proveen los servicios de conectividad (redes de telecomunicaciones), el acceso a Internet (ISP), las redes de tránsito y los Data Centers. Ahora bien, dado que los operadores de red, los proveedores de acceso a Internet y las empresas que ofrecen conectividad con el resto del mundo son parte del mismo ecosistema, es importante considerar las interdependencias entre ellos, y sí es el caso, solventar las limitantes tanto técnicas como económicas para que el servicio sea entregado al cliente de manera óptima.

La tendencia actual señala que se seguirán implementando aplicaciones sofisticadas que no podrán funcionar sobre redes rígidas o tradicionales (todas aquellas que no permiten insertar o segregarse tráfico de manera sencilla, que no soportan mecanismos para la diferenciación del tráfico, que no proveen anchos de banda bajo demanda, granulares y de alta capacidad, y que adicionalmente tengan una compleja y especializada administración), por lo contrario, estos servicios deberán contar con redes flexibles, escalables y de alta velocidad. Por ejemplo, para implementar servicios bajo el escenario de uso de la tecnología 5G (IMT 2020), que soportan la entrada del internet de las cosas (IoT), ya sean masivas (múltiples dispositivos de baja capacidad, no sensibles al retardo y comunicación esporádica), o críticas (ultra baja latencia y alta disponibilidad y confiabilidad), es esencial que las redes tengan la flexibilidad para cumplir por lo menos con las características de desempeño requeridas como se establece en modelos de referencia (ITU-T Y.4000 / Y.2060).

Desde el punto de vista del despliegue de infraestructura, para determinar la viabilidad de implementar los servicios en la nube en un país o región y por consiguiente su exitosa adopción, es necesario que la infraestructura de red (el acceso local, la red de transporte nacional y la conectividad a la red internacional), ofrezca un desempeño que sea el esperado por los proveedores de servicios en la nube. Para ello los parámetros de latencia, las velocidades de descarga y carga de información, los niveles de disponibilidad y las opciones de redundancia y enrutamiento, son indicadores esenciales para evaluar las capacidades y rendimiento de la red por donde viaja la información. Estos parámetros deberán cumplir con niveles de tal manera que se puedan soportar los servicios en la nube

de manera exitosa, otorgándole al cliente una red confiable y eficiente, lo que motivaría la disponibilidad y adopción de las soluciones que son provistas por Cloud Computing.

5.4.1. Virtualización

El primer concepto para entender cómo trabajan los equipos o servidores que conforman las unidades básicas de los Data Center que habilitan los servicios en la nube es la Virtualización. Por ejemplo, un solo disco físico en un servidor dentro de un Data Center puede contener información para varios inquilinos o clientes; el cifrado de datos y las técnicas de segregación de datos lógicos que utiliza la virtualización proporcionan más seguridad y protección contra el acceso a datos no deseados e internos, ya que es difícil apuntar a los datos de un cliente específico. Mientras que un proveedor de servicios de Cloud Computing proporcionará el mismo servicio a todos los clientes, un proveedor de servicios digitales o de TI proporcionará contratos y servicios diferentes y específicos para cada cliente.

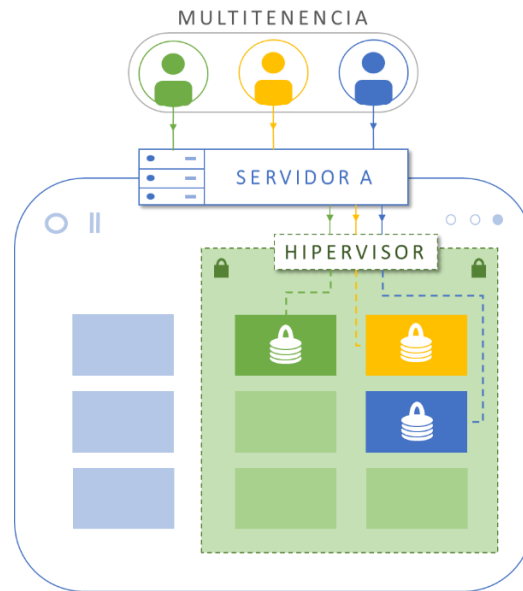
La virtualización es un proceso que separa las funciones o aplicaciones o software del hardware, y las nubes dependen de esa separación.

La forma más fácil de explicar la diferencia es describirla desde la perspectiva estricta de una infraestructura como servicio (IaaS). La base del Cloud Computing es un sistema operativo estable (como Linux⁴³). Esta es la capa que proporciona a los usuarios la independencia en los entornos públicos, privados e híbridos.

Suponiendo que el acceso a Intranet o a Internet ya está establecido, la virtualización es lo que crea las nubes. El software llamado hipervisor⁴⁴ se coloca por encima del hardware físico principal y extrae los recursos o funcionalidades de la máquina. Estos recursos pueden ser el potencial de procesamiento en bruto, el almacenamiento o las aplicaciones basadas en la nube que contienen todo el código de tiempo de ejecución y los recursos necesarios para implementarlas.

⁴³ Linux, es un sistema operativo libre tipo Unix; multiplataforma, multiusuario y multitarea. El sistema es la combinación de varios proyectos, entre los cuales destacan GNU (encabezado por Richard Stallman y la Free Software Foundation) y el núcleo Linux (encabezado por Linus Torvalds). (<https://es.wikipedia.org/wiki/GNU/Linux>).

⁴⁴ El hipervisor, también llamado monitor de máquina virtual (VMM, por sus siglas en inglés), es el núcleo central de algunas de las tecnologías de virtualización de hardware más populares y eficaces, entre las cuales se encuentran las de Microsoft: Microsoft Virtual PC, Windows Virtual PC, Microsoft Windows Server e Hyper-V. El hipervisor es la llave que habilita la virtualización. (<http://www.datakeeper.es/?p=716>).

Figura 5: Virtualización.⁴⁵

Fuente: Elaboración propia con información de Microsoft Azure

Si ese es todo el proceso, no se trata de Cloud Computing, sino de virtualización solamente. Los recursos virtuales deben asignarse a grupos centralizados antes de llamarlos nubes, y esas nubes deben coordinarse con software de gestión y automatización antes de que se les pueda considerar Cloud Computing.

Las nubes proporcionan las ventajas adicionales del acceso de autoservicio, la asignación o aumento automatizado de la infraestructura y los grupos de recursos dinámicos; esos son los beneficios que marcan más claramente la diferencia de los servicios de Cloud Computing con la virtualización tradicional.

Algunas características que permiten diferenciar los conceptos de virtualización y de Cloud Computing se muestran en la siguiente tabla:

⁴⁵ El término 'Multitenencia' se refiere a la convivencia de datos de múltiples clientes en un servidor.

Tabla 13: Características de Virtualización y Cloud Computing

	Virtualización	Cloud Computing
Definición	Tecnología	Metodología
Propósito	Crear múltiples entornos simulados desde un solo sistema de hardware físico	Agrupar y automatizar recursos virtuales para su uso sobre demanda
Uso	Entregar recursos en paquetes a usuarios específicos para un fin determinado	Entregar variedad de recursos a grupos de usuarios para varios fines
Configuración	Basado en imágenes	Basado en plantillas
Duración	Años (largo plazo)	De horas a meses (corto plazo)
Costo	Gastos de capital (CAPEX) elevados, gastos operativos (OPEX) bajos	Nube privada: CAPEX elevados, OPEX bajos Nube pública: CAPEX bajos, OPEX elevados
Escalabilidad	Escalabilidad vertical	Escalabilidad horizontal
Carga de trabajo⁴⁶	Con estado (statefull)	Sin estado (stateless)
Tenencia	Usuario único	Varios usuarios

Fuente: Elaboración Propia con información de (Geng, H. (2015). Data Center Handbook. Palo Alto, Ca. Wiley)

Una comparación práctica donde la virtualización puede hacer que un recurso cumpla la función de varios, mientras que el Cloud Computing permite que diferentes departamentos (a través de una nube privada) o empresas (a través de una nube pública) accedan a un único conjunto de recursos aprovisionados automáticamente como se muestra en la siguiente tabla.

⁴⁶ La diferencia clave entre las aplicaciones con estado y sin estado es que las aplicaciones sin estado no "almacenan" datos, como por ejemplo las aplicaciones con una sola función o servicio (Internet de las cosas y servidores de impresión); mientras que las aplicaciones con estado requieren respaldo de almacenamiento (bases de datos, transacciones bancarias y servidores de correo).

Tabla 14: Diferencias entre Virtualización y Cloud Computing

Virtualización	Cloud Computing
<p>La virtualización es una tecnología que permite crear múltiples entornos simulados o recursos dedicados desde un solo sistema de hardware físico. El software llamado "hipervisor" se conecta directamente con el hardware y permite dividir un sistema en entornos separados, distintos y seguros, conocidos como "máquinas virtuales" (VM). Estas VM dependen de la capacidad del hipervisor de separar los recursos de la máquina del hardware y distribuirlos adecuadamente.</p>	<p>Cloud Computing es un conjunto de principios que permite proporcionar recursos de infraestructura informática, de red y de almacenamiento, servicios, plataformas y aplicaciones a los usuarios, a pedido y en cualquier red. Estos recursos de infraestructura, servicios y aplicaciones se obtienen de nubes, que son conjuntos de recursos virtuales orquestados por software de gestión y automatización para que los usuarios puedan acceder a estos a pedido, a través de los portales de autoservicio a los que dan soporte, el escalado automático y la asignación dinámica de recursos.</p>

Fuente: Elaboración propia con información de Red Hat (<https://www.redhat.com/es/topics/cloud-computing/cloud-vs-virtualization>).

5.4.2. Interconexión de Data Centers

Durante la última década, los Data Centers y sus redes se han convertido en el habilitador de la tecnología para una serie de aplicaciones basadas en Internet. A partir de ese momento, la mayoría de las aplicaciones populares de Internet, desde la búsqueda tradicional, los mapas interactivos en línea, las redes sociales, hasta la transmisión de video e Internet de las cosas (IoT), se ejecutan en Data Centers. Su papel se intensificará aún más con la adopción más amplia del Cloud Computing, en el que una parte significativa del cómputo y el almacenamiento se migra a los Data Centers compartidos. Esto ya está ocurriendo a un ritmo acelerado hoy en día, con una gran cantidad de proveedores de nube que lideran el camino. Esto ha resultado en un aumento dramático en las capacidades de los centros de datos. Como ejemplo, el ancho de banda de biseción de las redes de clústeres del Data Center de Google ha aumentado en un factor de mil en la última década.

La figura: "Una vista de alto nivel de las redes de interconexión de Data Center (DC) de Google", proporciona una vista de la red de interconexión de DC de Google. Desde los puntos de vista de distancia y topología, esta red se puede dividir en cuatro segmentos:

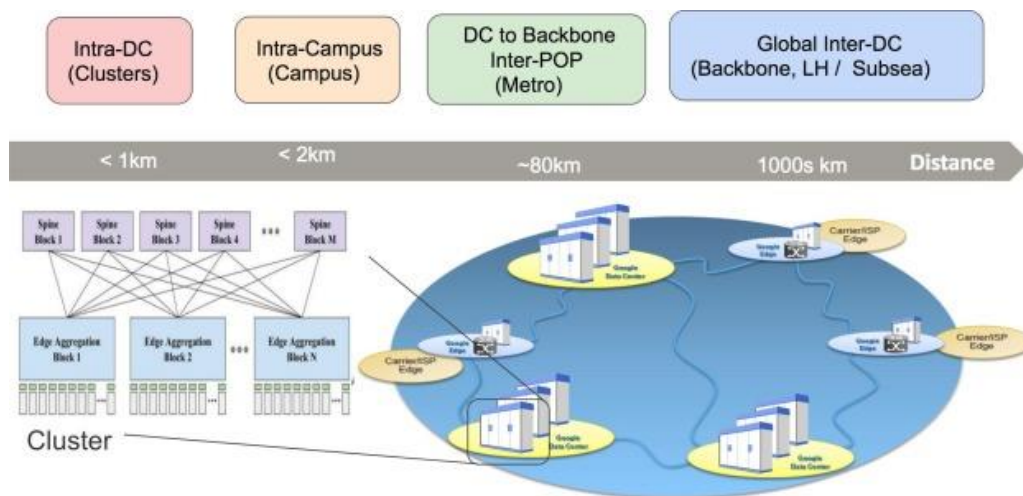
- a) La red intra-DC (Clusters), que interconecta de decenas a cientos de miles de servidores a través de una distancia de enlace de 500 m. a 1 km;

b) La red de interconexión Intra-campus (Campus), que interconecta grupos ubicados en diferentes edificios, pero dentro de un vecindario de campus de menos de 2 km;

c) La red de interconexión Inter-POP de borde (Metro), que proporciona conexiones entre los centros de datos y las redes globales de red troncal (Backbone), con una distancia de enlace generalmente inferior a 80 km; y finalmente

d) La red troncal Global Inter-DC (Backbone, LH / Subsea), que interconecta todos los centros de distribución de Google a través de tecnologías de transporte de larga distancia.

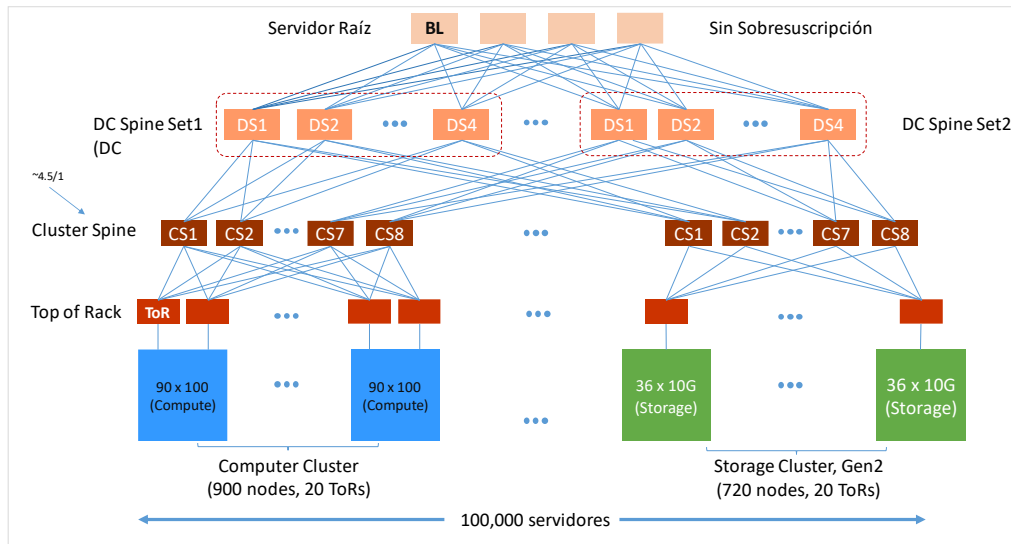
Figura 6: Una vista de alto nivel de las redes de interconexión de Data Center (DC) de Google.



Fuente: Google

Por otro lado, la evolución de la forma en que se interconectan los Data Centers para estructurar la nube, como en este caso lo ha hecho Microsoft, también conlleva grandes avances. La figura: "Redes Generación 3" nos muestra que, siendo los servidores solo computadoras, cuando se hablan entre ellos conforman una nube. En el principio los Data Centers que soportan la nube se enfocaban en el tráfico entre los usuarios y los servidores que le proveían el servicio (norte - sur). Ahora los Data Centers enfatizan una alta densidad de tráfico tanto entre servidores como con los clientes (este - oeste).

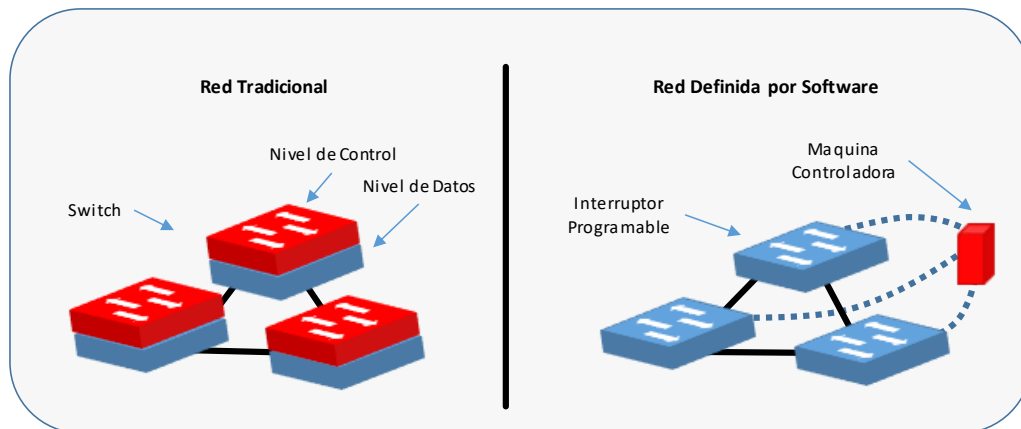
Figura 7: "Redes Generación 3"



Fuente: Elaboración propia con información de Microsoft

Esto se logra técnicamente gracias a que las redes ahora son definidas por software (SDN. Software Defined Networking), que reemplazan equipos dedicados a configurar y coordinar actividades de ruteo por programas escritos que corren en computadores genéricos como se ilustra en la Figura: "Redes definidas por software".

Figura 8: Redes definidas por software.



Fuente: Elaboración propia con información de Microsoft

Los nuevos esquemas de interconexión y topología de las redes permiten obtener beneficios que antes eran difíciles o costosos de obtener, como el cambiar o adaptar las redes en función de las demandas de los clientes, establecer fronteras

de seguridad entre clientes y servicios, modificar o expandir los Data Centers sin interrumpir los servicios y establecer mejores controles de seguridad a través de todas las capas de las redes.

Además de todo lo anterior la digitalización de las actividades económicas, o transformación digital, ha promovido como se ha señalado antes en este documento, de un cambio de modelo de provisión de servicios de TI centralizado a un modelo geográficamente distribuido y regionalizado a través de la nube. Como consecuencia de este cambio, las cargas de trabajo de procesamiento y entrega de servicios digitales están migrando al borde de estas redes o nubes, conocido como el "Edge Node" o "Digital Edge", donde los datos pueden ser interconectados ahora en proximidad con los clientes, socios o empleados, lo que incrementa exponencialmente las velocidades de respuesta y capacidades de procesamiento de los negocios. Como ejemplo, con la tecnología 5G se tiene la expectativa de reducir el tiempo de latencia hasta 5 ms, en su conexión de acceso a los dispositivos (endpoint), lo que significa que las conexiones en el borde (Edge Node), serán necesarios para habilitar los servicios con esta tecnología.

Las nuevas arquitecturas distribuidas y métodos de enrutamiento requeridas por el Cloud Computing para acercar los servicios digitales y de TI a los clientes hacen imperativa una mejor interconexión que elimine el factor distancia para mantenerse competitivas. A raíz de lo anterior, en la medida que aumenta la producción y consumo de datos en los nodos en el borde de las redes o nubes, ahora los negocios digitales buscan consolidar infraestructuras de centros de datos adyacentes a nodos de interconexión o puntos de intercambio de tráfico.⁴⁷

5.4.3. Redes de Fibra Óptica

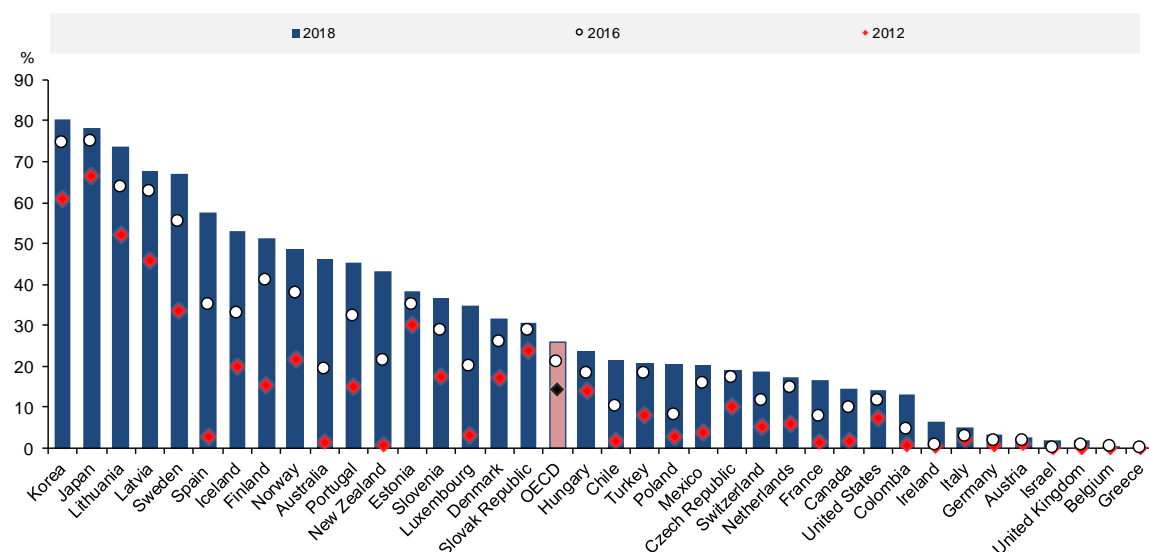
La fibra óptica de alta velocidad soporta un cuarto de Internet fijo en los países de la OCDE. La proporción de fibra de alta velocidad en las conexiones de Internet de banda ancha fija en los países de la OCDE ha creció 25% en el año 2018, 12% más respecto al año anterior, según una actualización del portal de banda ancha de la OCDE.

Sin embargo, los últimos datos muestran grandes diferencias entre países, con un porcentaje de fibra en la banda ancha total que va desde más del 70% en Corea, Japón y Lituania a menos del 10% en Grecia, Bélgica, Reino Unido, Israel, Austria,

⁴⁷ EQUINIX, (2020). Global Interconnection Index. Recuperado de https://www.equinix.com/gxi-report/?ls=Advertising%20-%20Web&lsd=19q3_cross-vertical_digital-edge+index-vol3_programs-global_equinix-run_equinix-owned-site_us-en_AMER_GXI3_awareness&utm_campaign=us-en_equinix-owned-site_GXI3_programs-global_awareness&utm_source=&utm_medium=equinix-owned-site&utm_content=digital-edge+index-vol3_Report_Further_Insights_pg44

Alemania, Italia e Irlanda. El mayor crecimiento de fibra en el año 2018 se ha visto en Irlanda, Bélgica y Australia, con suscripciones de fibra que aumentaron un 218%, 71% y 70%, respectivamente.

Gráfica 9: Porcentaje de conexiones de fibra en banda ancha fija total, a diciembre de 2018.



Notas: Definiciones: los datos de suscripciones de fibra incluyen FTTH, FTTP y FTTB y excluyen FTTC.

Australia: una nueva entidad recopila los datos informados para diciembre de 2018 en adelante utilizando una metodología diferente. Las cifras informadas a partir de diciembre de 2018 comprenden un salto de serie y son incomparables con los datos anteriores para cualquier medida de banda ancha que Australia informe a la OCDE. PbyM. Los datos de Canadá, Suiza y Estados Unidos son preliminares.

Fuente: OECD, Broadband Portal, (www.oecd.org/sti/broadband/oecd-broadband-portal.htm), información de datos sobre Israel: (<http://oe.cd/israel-disclaimer>)

El reporte de la OCDE registra que la fibra óptica participa alrededor del 25% de las conexiones de banda ancha en México. Al respecto, el Instituto recaba y analiza la información de los concesionarios que despliegan redes de fibra óptica y a finales del 2017, registra que se cuenta con una infraestructura de redes de fibra óptica de aproximadamente 433 mil km en el país, distribuida en 175 mil km de red de acceso, 189 mil km de red local o metropolitana, 61 mil km de red de transporte interestatal, 6.3 mil km de segmentos internacionales y 703 km de cable submarino⁴⁸.

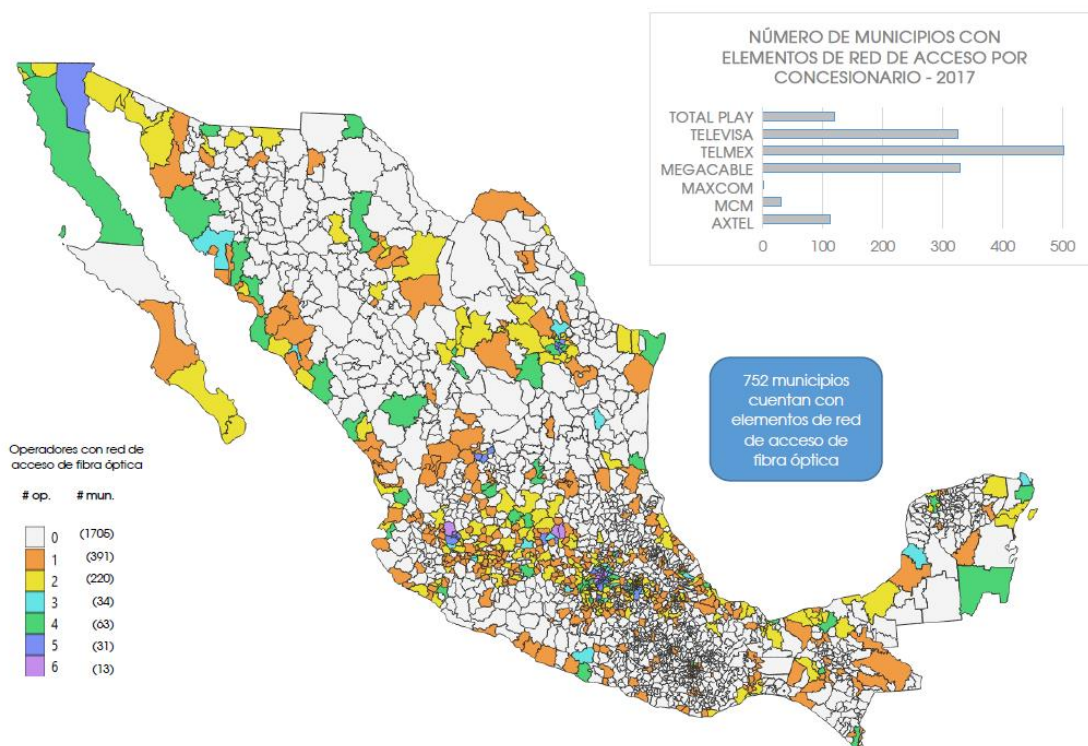
Con relación a la provisión de servicios de Cloud Computing, se considera relevante la infraestructura de la red de acceso, que conforme se ha mencionado antes, es crítica para cumplir con los parámetros de latencia o retardo y las altas capacidades que requiere la provisión de los nuevos servicios digitales, ya que es la

⁴⁸ IFT. Información de despliegues de redes de fibra óptica reportadas por concesionarios de redes públicas de telecomunicaciones. No se incluye información de la red de fibra óptica de CFE.

sección de las redes que proporcionan la conectividad de los dispositivos de los clientes con los servicios.

A continuación se muestra un mapa que muestra la disponibilidad de despliegues de fibra óptica de red de acceso en función del número de concesionarios que operan por municipios en México⁴⁹.

Mapa 1: Operadores con red de acceso en México



Fuente: Elaboración propia con información solicitada a los Operadores nacionales

Derivado de esta información, se muestra que de 2,458 municipios en el país el 30% cuentan con infraestructura de acceso de fibra óptica con al menos un proveedor de servicios de telecomunicaciones. Para mayor información sobre las coberturas de redes de fibra óptica ver el ANEXO 4: Redes de Fibra Óptica.

En el área de la OCDE, en comparación con los despliegues de fibra óptica, el xDSL (Línea digital de suscriptor, por sus siglas en inglés), sigue siendo la tecnología de banda ancha fija que prevalece, y constituye el 38% de las suscripciones de banda ancha fija, pero continúa siendo reemplazada gradualmente por fibra. Las

⁴⁹ IFT, (2018). Micro sitio de Despliegue de Infraestructura. Recuperado de http://despliegueinfra.ift.org.mx/docs/Mapa%20Operadores%20Red%20de%20acceso_0.pdf

suscripciones DSL disminuyeron un 8% en el año hasta junio de 2018, mientras que las suscripciones de fibra aumentaron en un 15%. Las suscripciones de cable, que crecieron el 4% en el año, representaron la mayor parte del resto, con un 33% de la banda ancha total.

Por otro lado, la banda ancha móvil continuó su rápida expansión con 98 millones de suscripciones adicionales, un aumento del 7,4% en el año hasta junio de 2018. Esto eleva la penetración de banda ancha móvil a más del 106% en el área de la OCDE. La penetración de banda ancha móvil es más alta en Japón, Finlandia, Estonia, Estados Unidos y Australia, con suscripciones por cada 100 habitantes del 168%, 155%, 146%, 137% y 136%, respectivamente.

Los datos sobre las comunicaciones máquina a máquina (M2M) muestran que Suecia, Nueva Zelanda, Noruega y Estados Unidos siguen siendo los líderes en el número de tarjetas SIM M2M en uso por cada 100 habitantes. Suecia cuenta con 125 tarjetas SIM M2M por cada 100 habitantes, un nivel mucho más alto que la mayoría de los países de la OCDE que proporcionaron datos debidos en parte al uso de estas tarjetas SIM en otros países por un operador sueco. En general, las suscripciones de celulares móviles M2M / integradas crecieron más del 38% en el último año en los países donde los datos estaban disponibles.

Las estadísticas de banda ancha de la OCDE ahora cubren 37 países, incluidos los nuevos miembros Colombia y Lituania. Los totales históricos, los promedios y las tasas de crecimiento de la OCDE se han ajustado para incluir a ambos países. La inclusión de Colombia y Lituania significa que los totales, promedios y tasas de crecimiento de la OCDE no son comparables con los datos de junio de 2017.⁵⁰

De lo anterior, cabe resaltar la importancia de las redes de fibra óptica en los segmentos de acceso para mejorar el desempeño de los servicios digitales y la experiencia del cliente, así como en los segmentos de transporte que serán el soporte de infraestructura de las alternativas de conectividad distribuidas y redundantes de los Data Centers y la provisión de los servicios digitales y de TI.

5.4.4. Parámetros de calidad

Se conoce como latencia al tiempo que lleva transportar una cantidad de datos como respuesta a una solicitud de información realizada por un solo usuario final

⁵⁰ OECD, (2020). Broadband Portal. Recuperado de www.oecd.org/sti/broadband/broadband-statistics/

para las condiciones en las que la red se encuentra operando en el momento de la solicitud.

De tal manera los servicios provistos por la nube al estar cursando por la red requieren valores de latencia muy cercanos a los que están definidos para su correcta operación. Ahora bien, si estos valores están fuera de su rango, implicaría que los diversos servicios presentarían retrasos o interrupciones al visualizarlos o al cargarlos, así como también se tendría una mala calidad del audio resultando en una mala experiencia de uso.

Como se ha indicado, los servicios están en constante evolución y, algunos de ellos son aún más sensibles a la latencia, entre ellos la banca en línea o la visualización de registros hospitalarios en un entorno de atención médica, para estos se requiere de latencias bajas para su correcta operación. Reducir la demora en la entrega de paquetes hacia y desde la nube es crucial para brindar los servicios avanzados que se están desarrollando (y asegurar una experiencia de alta calidad).

Otro factor que es de considerarse es el Tiempo de Ida y Vuelta (Round-Trip Time o Round Trip Delay Time, RTT por sus siglas en inglés), que representa la duración en milisegundos que tarda una solicitud de red para pasar de un punto inicial a uno de destino y volver al punto inicial, RTT es un parámetro importante comúnmente utilizada para diagnosticar la velocidad y confiabilidad de las conexiones de red.

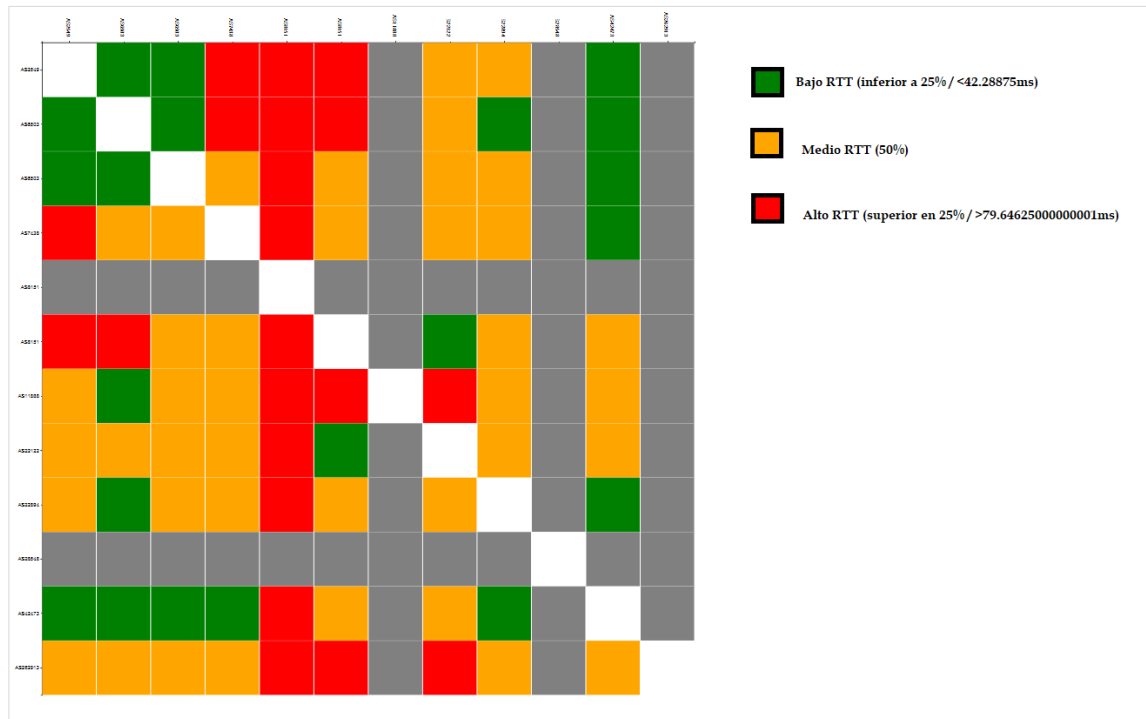
La reducción de RTT es un objetivo principal de las redes de entrega de contenido que cada vez son más populares, estas redes son las que se encargan de transmitir la mayoría de tráfico web incluido el tráfico de las principales redes sociales, servicios de *video streaming* y sitios de comercio digital.⁵¹

Para el tráfico entre redes internas en México se encontró que la mayoría de ellas cuentan con un RTT de calidad media y baja según se señala en la siguiente gráfica⁵², obtenida en enero del 2020.

⁵¹ Cloudflare, (s.f.). What is round-trip time? | RTT definition. Recuperado de <https://www.cloudflare.com/learning/cdn/glossary/round-trip-time-rtt/>

⁵² RIPE NCC, (s.f.). Recuperado de <http://sg-pub.ripe.net/emile/ixp-country-jedi/latest/MX/rftmesh/index.html?ASNS=278,2549,6503,7438,8151,11888,22122,22884,22894,28548,42473,262913,&ipv=v4>

Gráfica 10: RTT del tráfico cursado entre redes en México.



Fuente: RIPE Atlas

Consecuencia de lo anterior, los proveedores de Cloud Computing han implementado estrategias para reducir los tiempos de latencia a lo largo de grandes distancias geográficas, y lo han resuelto creando copias locales de datos o caché (cuando es posible): este es el servicio que brindan las empresas de distribución/entrega de contenido; a través de distancias más pequeñas, como en un centro de datos o un servidor, el software de almacenamiento en caché proporciona estas copias locales, y de esta manera se reducen los tiempos de latencia y se aseguran de que se disfrute una experiencia de alta calidad.

Los datos recabados de diversos proveedores de servicios en la nube presentan diversos valores de latencia en función de la estrategia de despliegue de sus redes y centros de datos.

Por otra parte, la velocidad de descarga para el usuario final es una característica importante; en el sector empresarial por ejemplo en la implementación de máquinas virtuales, *Customer Relationship Management* por sus siglas en inglés CRM⁵³, o en el uso de plataformas *Enterprise Resource Planning* por sus siglas en

⁵³ CRM: Gestión de las Relaciones con Clientes, una aplicación que permite centralizar en una única Base de Datos todas las interacciones entre una empresa y sus clientes.

inglés *ERP*⁵⁴, este indicador es crítico para medir la calidad del servicio que se está brindando; como también lo es para los usuarios masivos cuando realizan descarga de videos, juegos en línea y otros servicios provenientes de la nube. De la misma manera la velocidad de carga con la adopción cada vez mayor de máquinas virtuales, videoconferencias en empresas y usuarios masivos subiendo información, la velocidad también se convierte en un parámetro importante para la entrega de contenido a la nube.

Para ejemplificar los valores de estos indicadores en diversas aplicaciones como el correo electrónico y/o mensajería instantánea, ambos servicios requieren velocidades de descarga y carga bajas, aproximadamente menor a 1 Mbps en ambos sentidos, y una latencia media entre 100 y 159 ms. Un segundo ejemplo como lo son las aplicaciones de colaboración en servidores virtuales requieren velocidades altas de subida mayores de 1 Mbps y de bajada a 2.5 Mbps con una latencia media entre 100 y 159 ms¹³.

Las velocidades teóricas ofrecidas por los operadores fijos y móviles pueden parecer adecuadas, pero muchos factores están involucrados en el desempeño de la red, así lo muestran las diversas mediciones que se realizan, por lo que las velocidades (carga y descarga) y las latencias varían dentro de cada país y región y según el tipo de despliegue, ya sea urbano o rural, con tecnología fija y móvil, la proximidad a los centros de datos tradicionales y en la nube, sin menospreciar la calidad del equipo en las instalaciones del cliente.

El cuadro siguiente muestra un resumen por red fija o móvil las velocidades en Mb/s y latencias promedio por región para el año 2018.⁵⁵

Tabla 15: Velocidades y Latencias por región 2018.

Network	Región	Descarga Promedio	Velocidades Promedio de Cargas	Latencia (ms) Promedio
FJO	Asia Pacífico	46.2	22.1	21
	Europa Oriental y Central	32.8	18.8	33
	América Latina	11.6	3.7	41
	África Medio Oriente	8.2	41	52
	Norte América	43.2	16.6	34
	Europa Occidental	37.9	14.4	27
MOVIL	Asia Pacífico	20.8	10.1	55
	Europa Oriental y Central	14.7	7.1	49
	América Latina	10	6	70
	África Medio Oriente	7.4	3.9	80
	Norte América	34.3	15.5	67
	Europa Occidental	26.2	10.4	46

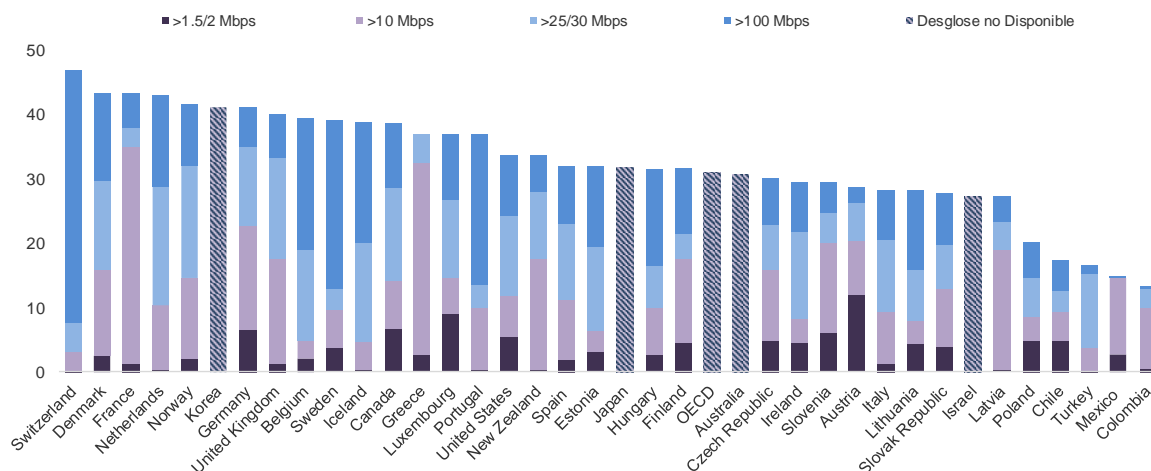
Fuente: Índice Global de Nube de Cisco, 2016-2021.

⁵⁴ ERP: Sistema de Planificación de Recursos Empresariales: Estos programas se hacen cargo de distintas operaciones internas de una empresa, desde producción a distribución o incluso recursos humanos.

⁵⁵ Cisco, (s.f.). Cisco Annual Internet Report, Recuperado de <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/white-paper-c11-738085.html>

Según estadísticas de la OCDE, respecto a diciembre de 2018, separando la velocidad de acceso por distintos niveles, se tienen las siguientes velocidades dentro de los países que forman parte:

Gráfica 11: Suscripciones de banda ancha fija por cada 100 habitantes.



Notas: Información por niveles de velocidad (diciembre de 2018)

Fuente: OECD Broadband statistics (www.oecd.org/sti/broadband/broadband-statistics)

Como se puede ver en la tabla, de acuerdo a información de la OCDE, México cuenta con una de las velocidades más bajas entre los países que participaron en el estudio. Por cada 100 habitantes en México, existen en promedio 2.8 accesos mayores a 1.5 Mbps, 11.7 mayores a 10 Mbps y 0.3 mayores a 30 Mbps.

Derivado de que uno de los tópicos primordiales para la adopción de Cloud Computing son los parámetros de calidad con que los Servicios en la Nube se brindan, cobra mucha importancia el contar con redes que puedan operar a altas velocidades de carga y descarga, así como bajas latencias para brindar una gran experiencia de usuario.

Del reporte emitido por Ookla⁵⁶ en 2018, con más de 28 millones de pruebas realizadas en diversos proveedores de internet del país, se obtiene un promedio de 39 mili segundos (ms) de latencia, solo 2 ms por debajo del promedio señalado por Cisco para el mismo año, pero 4ms por encima de la latencia promedio para Norte América.

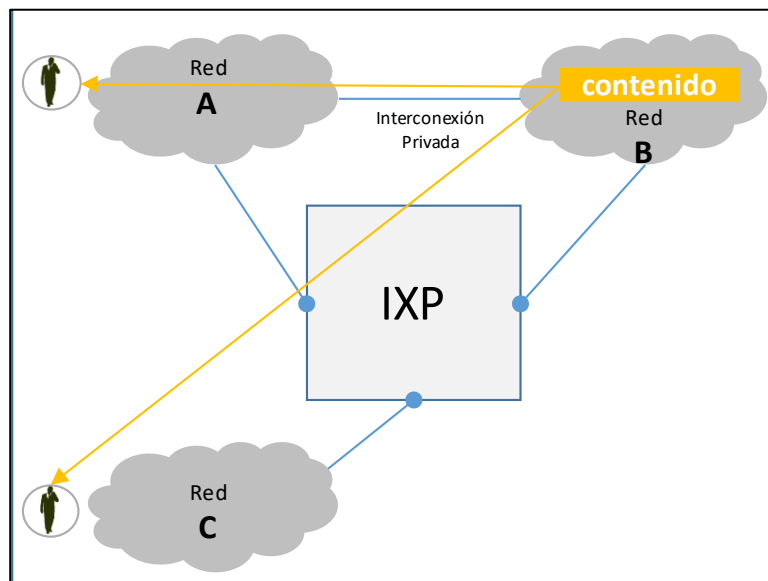
⁵⁶ Speedtest, (2018). Speedtest Awards. Recuperado de <https://www.speedtest.net/awards/mexico/>

Considerando que 1ms significa tiempo considerable si de redes y sistemas informáticos se habla, esta sección muestra que uno de los requerimientos necesarios para solventar las necesidades de los nuevos servicios como el teletrabajo, telemedicina, videoconferencias, el video streaming o el juego en línea, es contar con redes de baja latencia que conecten al país, así como la implementación y desarrollo de redes orientadas a cubrir estas necesidades, más que las redes diseñadas para brindar servicios de telecomunicaciones tradicionales que se utilizan hoy en día a lo largo del territorio mexicano.

5.4.5. Conectividad de Internet

En general las redes para cursar tráfico de internet se interconectan entre sí de dos formas; directamente o mediante un punto de intercambio de Internet (IXP), como se muestra en el siguiente diagrama.

Figura 9: Formas de interconexión de las redes.



Fuente: Elaboración propia con información de RIPE NCC

Los proveedores de acceso a Internet (ISP), recurren o dependen de los enlaces o circuitos contratados a **proveedores de tránsito**⁵⁷, que pueden ser redes de operadores establecidos o puntos de intercambio de tráfico de Internet (IXP), para poder intercambiar su tráfico a otras regiones. Bajo estos esquemas se experimentan grandes flujos transfronterizos, rutas indirectas, largos tiempos de transferencia y en

⁵⁷ Redes de nivel TIER 1, es decir, redes de alta jerarquía formada de unos cuantos miles de sistemas autónomos (AS).

general una falta de control sobre la calidad de los servicios. Esta dependencia de intermediarios resulta en costos de transporte o tránsito⁵⁸. Aun así, tradicionalmente los proveedores de Internet (ISP) encontraban más efectivo en términos de costo utilizar las conexiones internacionales para intercambiar su tráfico local o doméstico que conectarse a cada ISP en su región de forma separada, a este efecto se le conoce como *tromboning*.

Este efecto *tromboning* se compensa con los acuerdos de *peering*⁵⁹ entre ISP dentro de una misma región; sin embargo, estos acuerdos pueden ocurrir entre sistemas autónomos (AS) u operadores que también actúan como ISP, resultando en la agregación de las funciones de acceso a Internet y de transporte del tráfico de datos en sus mismas redes.

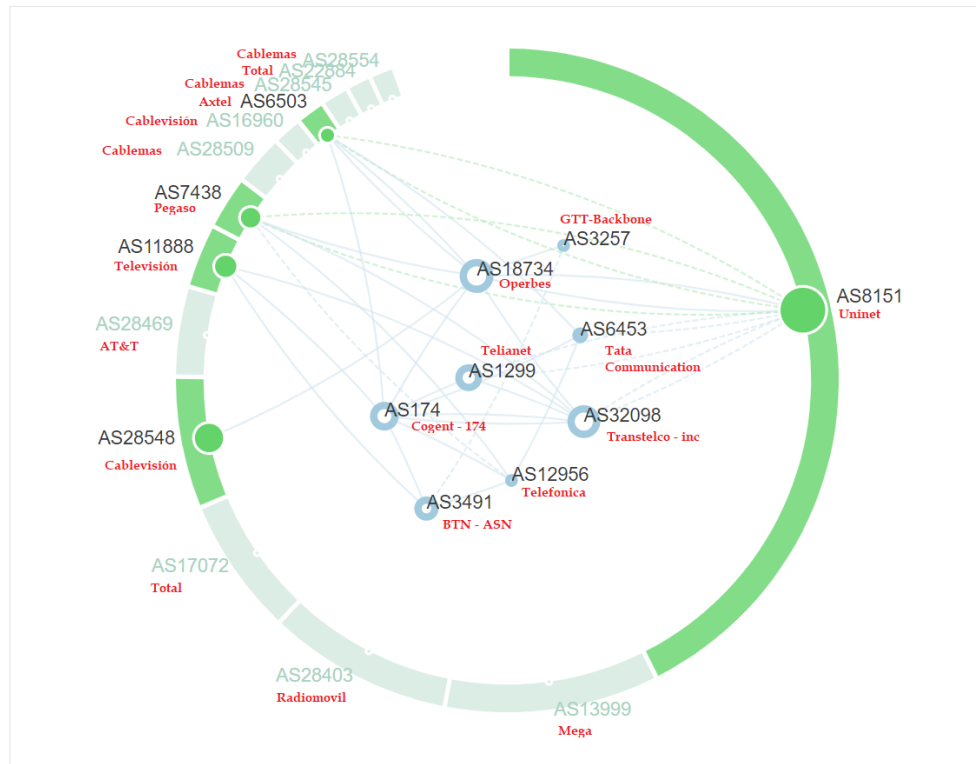
Al respecto se cuenta con información de las entidades encargadas de los registros de direcciones de Internet, como es el caso de la plataforma RIPE Atlas⁶⁰ que cuenta con miles de sensores en la red de internet en todo el mundo y colecta datos para realizar mediciones de diferentes parámetros. Para el caso que nos ocupa, consultamos la información sobre el intercambio de tráfico de internet de México (*peering*) y se reporta la siguiente gráfica.

⁵⁸ Castro, I. (2019). Shaping the Internet: History and Impact of IXP Growth. Recuperado de https://labs.ripe.net/Members/ignacio_castro/shaping-the-internet-history-and-impact-of-ixp-growth

⁵⁹ Por *peering* se refiere a la interconexión voluntaria de redes administrativamente separadas con el propósito de intercambiar tráfico entre los usuarios de diferentes redes.

⁶⁰ RIPE NCC, (s.f.). What is RIPE Atlas? Recuperado de <https://atlas.ripe.net/about/>

Gráfica 12: Esquema de red (peer to peer) en México.



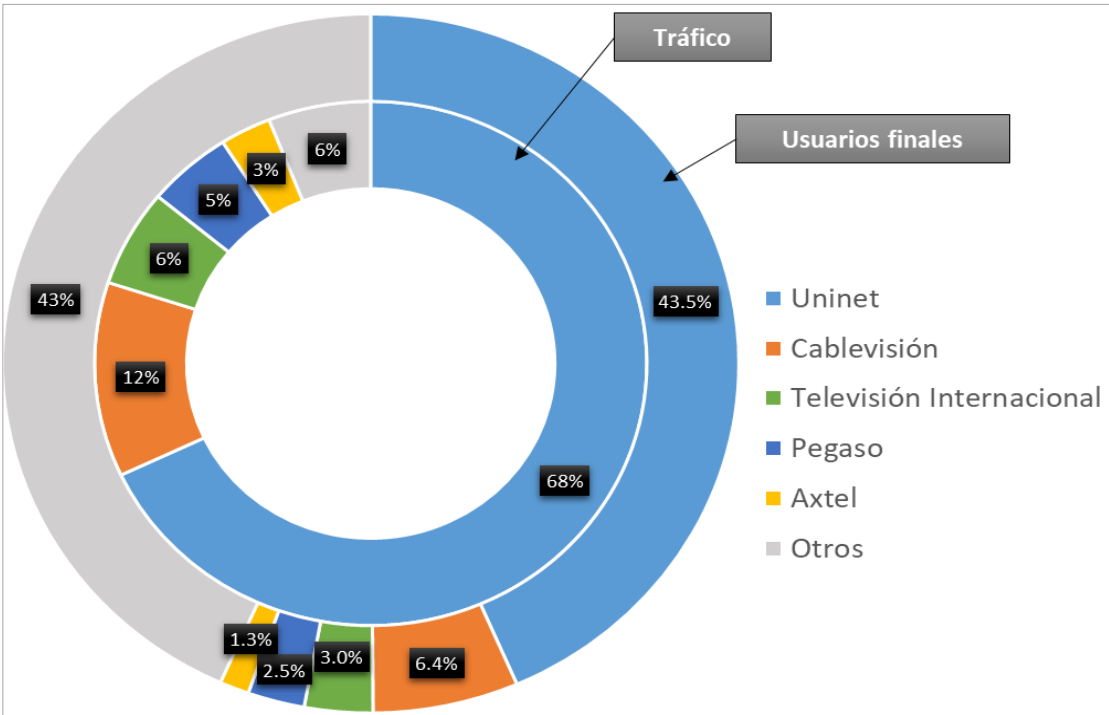
Fuente: RIPE Atlas

Esta gráfica muestra con la información del tráfico de internet una aproximación de como diferentes redes (operadores o sistemas autónomos) intervienen en la interconexión de Internet dentro del país, así como las proporciones de usuarios y de tráfico en cada red. Con la información de esta plataforma se realizó el siguiente análisis.

A continuación se muestran dos gráficas que muestran un panorama general de la estimación del intercambio de tráfico de Internet entre usuarios que se cursa a través de redes o sistemas autónomos pares (peering) dentro de México⁶¹.

⁶¹ Elaboración propia con información de la plataforma de mediciones de RIPE Atlas del 14 de enero de 2020, la cual muestra datos actualizados del tráfico cursado a través de acuerdos de peering y de proveedores de tránsito o IXP entre Sistemas Autónomos dentro de cada país. (<http://sg-pub.ripe.net/ixp-country-jedi/>)

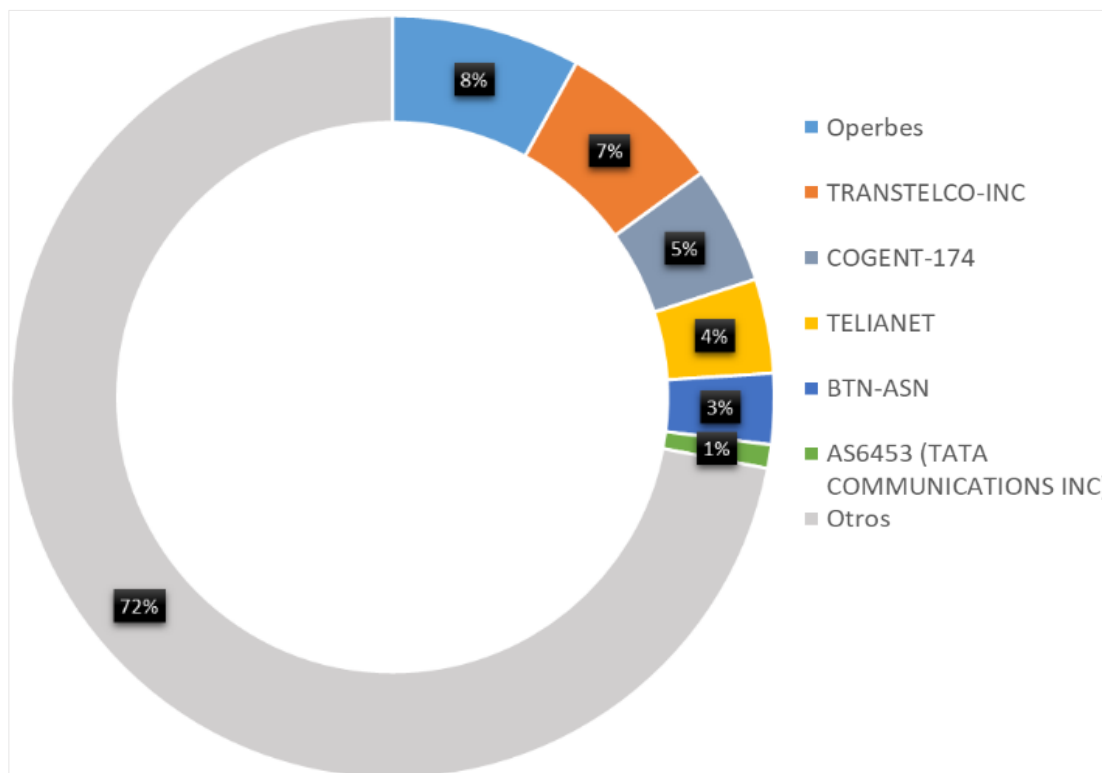
Gráfica 13: Porcentaje de usuarios finales y tráfico cursado en redes de usuarios finales en México.



Fuente: RIPE Atlas

La primera gráfica reporta la estimación de la proporción de usuarios (anillo externo) y de tráfico de Internet (anillo interno) que es intercambiado dentro del país a través de las redes que proporcionan conectividad a los usuarios finales, es decir, redes de operadores o concesionarios que proveen acceso a Internet. Se observa que una gran proporción de usuarios y de tráfico de Internet que se intercambia entre usuarios finales dentro del país se cursa a través de una sola red.

Gráfica 14: Porcentaje de tráfico cursado en redes de tránsito o IXP en México.



Fuente: RIPE Atlas

En la segunda gráfica, complementaria a la anterior, se muestra la proporción de tráfico de Internet que es intercambiado dentro del país por los operadores con ayuda o a través de redes pertenecientes a proveedores de servicios de tránsito o IXP, es decir, mediante otras redes que no son operadores o concesionarios tradicionales como en la primera gráfica. En este caso cabe señalar que la mayor parte del tráfico (72%) no se registra⁶² y en contraste las redes que reportan datos son principalmente redes de IXP que cursan tráfico del país, pero que lo gestionan en sus redes que están parcial o totalmente fuera del territorio del país.

Las gráficas y datos anteriores conllevan ciertas consideraciones con respecto a la forma en que se transporta el tráfico de Internet en México. En primer lugar, se observa que una gran proporción del tráfico de Internet agregado en una o pocas redes condiciona dicho tráfico a las políticas de gestión de tráfico y acuerdos de nivel de servicio (ASL) de dichas redes, es decir, se cuenta con pocas alternativas para disponer de diferentes o mejores parámetros de calidad (latencia, jitter, pérdidas de paquetes, disponibilidad, sobre-suscripción, etc.), así como de

⁶² Datos no registrados de redes que cursan menos del 1% del tráfico o que no cuentan con equipos sensores.

capacidades adicionales que requieran en nuestro caso los Data Centers para la provisión de servicios digitales con más exigentes características de desempeño.

En el caso de México, tenemos que los ISP son generalmente las mismas redes de acceso a Internet (operadores o concesionarios), resultando en la agregación de las funciones de acceso a Internet y de transporte del tráfico de datos en sus mismas redes, lo que posiblemente conlleve a la integración de elementos de red no relacionados con la provisión del servicio de Internet y que podría elevar los costos de este servicio restando competitividad a los usuarios o que por otra parte se apliquen estrategias comerciales para subsidiar algunos servicios.

Otra situación que resulta del bajo número de redes de acceso a Internet e interconexiones necesarias para el intercambio de datos, es la limitante que podrían tener los ISP independientes, o en su caso algunos clientes como pueden ser proveedores de servicios en la nube, para administrar y gestionar su tráfico. Esto significa que se carece de alternativas para direccionar o cursar dicho tráfico por otras rutas de salida a la red de Internet o backbone internacional de menor costo o de mayor capacidad que no sean las mismas redes de operadores establecidos.

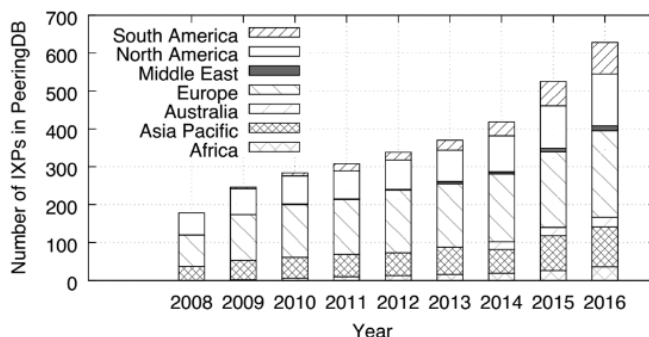
Ahora bien, como se mencionó antes, el uso de capacidad internacional para tráfico doméstico es costoso y este efecto de *tromboning* puede ser eliminado si los ISP en una región adoptan un modelo cooperativo para crear un punto de intercambio de tráfico local (IXP).

La creación de un IXP puede atraer a más proveedores de acceso a Internet (ISP) independientes y redes de entrega de contenidos, llegando a convertirse en un centro o "hub" regional para el tráfico de Internet. Como un efecto dominó, estas condiciones mejoradas atraen a más operadores internacionales y Data Centers que aprovechan la concentración de tráfico regional para expandir sus operaciones a otros países acrecentando las oportunidades de negocio e innovación y desarrollo para esas regiones.

Los puntos de intercambio de tráfico de Internet (IXP) han jugado un papel clave en el avance del ecosistema de Internet. La implementación de IXP en el mundo se ha triplicado en los últimos 10 años y su crecimiento porcentual alcanzó dos dígitos para todas las regiones, lo cual ha reducido la dependencia en general de las redes proveedoras de tránsito, de operadores tradicionales o de redes nivel TIER 1⁶³.

⁶³ Castro, I. (2019). Shaping the Internet: History and Impact of IXP Growth. Recuperado de https://labs.ripe.net/Members/ignacio_castro/shaping-the-internet-history-and-impact-of-ixp-growth

Gráfica 15: Crecimiento de IXP por región.



Fuente: RIPE NCC

En septiembre de 2019, América Latina contaba con 89 IXP⁶⁴. En general este crecimiento ha sido más representativo en Brasil con más de la mitad de los IXP en la región, sin embargo, México con el segundo mayor tráfico de América Latina, solo cuenta con dos IXP (ver ANEXO 6: Puntos de intercambio (IXP) en México), y curiosamente no registran tráfico en las gráficas de peering antes referidas. En el siguiente cuadro se puede observar un indicador a fin de comparar el número de IXP en otros países de América y ubicar a México:

Tabla 16: Número de IXP por cada millón de personas en diferentes países.

País		Número de IXP por cada millón de personas	País		Número de IXP por cada millón de personas
1	Panamá	0.9618	12	Nicaragua	0.1591
2	Canadá	0.8584	13	El Salvador	0.1505
3	Argentina	0.7099	14	Perú	0.1217
4	Estados Unidos	0.6348	15	Honduras	0.1110
5	Paraguay	0.4221	16	República Dominicana	0.0974
6	Chile	0.4186	17	Haití	0.0900
7	Jamaica	0.3665	18	Cuba	0.0891
8	Ecuador	0.2844	19	Colombia	0.0787
9	Bolivia	0.2580	20	México	0.0008
10	Brasil	0.2215	21	Venezuela	0.0000
11	Costa Rica	0.1999	22	Uruguay	0.0000

Fuente: Elaboración propia con información del Banco Mundial y el Directorio de Tráfico de Internet de "Packet Clearing House" (<https://www.pch.net/ixp/dir>)

⁶⁴ Hernandez, R. (2019). Confianza y neutralidad: los desafíos que enfrentan los IXP en México. Recuperado de <https://www.mdccdatacenters.com/es/interconnection-mdc/trust-neutrality-the-challenges-ixps-face-in-mexico/>

En resumen y derivado del análisis de esta sección, entendemos que las infraestructuras de las redes de telecomunicaciones en el país deben evolucionar para ofrecer características de desempeño, tanto en términos de sus arquitecturas (componentes) como en topologías (despliegues), conforme a las nuevas configuraciones de interconexión que utilizan los Data Centers y los proveedores de servicios en la nube con arquitecturas distribuidas, redundantes y con tendencia a reducir las distancias con sus socios y clientes, así como que soporten nuevos parámetros de calidad, sobre todo la reducción de las distancias para atenuar el factor de Latencia o RTT que es lo que más afecta la provisión de servicios digitales.

Derivado del razonamiento respecto al desempeño de las redes, se hace necesaria la promoción de despliegues de proyectos de fibra óptica, tanto en segmentos de acceso como de transporte, ya que éstas serán el soporte de infraestructura de las alternativas de conectividad distribuidas y redundantes, y que sea aprovechada por más y diferentes proveedores de servicio de acceso a internet (ISP) independientes, que a su vez demanden o propicien la creación y utilización de puntos de intercambio de tráfico de internet (IXP).

Finalmente, de manera preliminar se identifica que la provisión del acceso y el tránsito hacia otras redes, tanto dentro como fuera del país, es manejado por pocos agentes que integran estas funciones y que recurren a los acuerdos de peering o rutas con puntos de intercambio extranjeros, lo que resulta en pocas alternativas para que otros proveedores (ISP) o proveedores de Cloud Computing, que quieran operar de forma diferenciada, tengan la posibilidad de acceder directamente a la red mundial de internet y puedan administrar y gestionar su tráfico sin depender de algún grado de la intermediación de las grandes redes tradicionales o puntos de intercambio (IXP) fuera del país.

6. Mercados y Regulación

El papel que juega el sector de las tecnologías de la información y comunicación (TIC) en la economía no ha pasado desapercibido en México, ya que ha mostrado un crecimiento sostenido, adquiriendo cada vez una mayor importancia en la economía del país. De acuerdo con datos del Instituto Nacional de Estadística y Geografía (INEGI), el porcentaje del PIB atribuible al sector TIC pasó de 3.2% en el 2000 al 5.6% en el 2010, lo que representa un crecimiento del 75% en 10 años.

Por otra parte, asociado al crecimiento de las TIC se encuentra el componente de su adopción por parte de la población. En la siguiente tabla se muestra un resumen de los resultados de la Encuesta Nacional sobre Disponibilidad y Uso de Tecnologías

de la Información en los Hogares, 2019 (ENDUTIH 2019)⁶⁵, donde se muestran algunos de los principales datos relativos a la adopción de las TIC en México, con el propósito de comprender el contexto actual del uso y aprovechamiento de estas tecnologías, así como contar con una referencia sobre su penetración en distintos grupos de la sociedad.

Tabla 17: Encuesta TIC 2019.

Proporción de usuarios de computadora	Proporción de usuarios de Internet	Proporción de usuarios de telefonía celular
43.0%	70.1%	75.1%

Proporción de hogares con computadora	Proporción de hogares con Internet	Proporción de hogares con televisor digital
44.3%	56.4%	76.5%

Proporción de usuarios de TIC por sexo	Mujeres	Hombres
Computadora	50.1%	49.9%
Internet	51.6%	48.4%
Celular	51.7%	48.3%

Edad	Usuarios de Internet por edad	Mujeres	Hombres
Total	100%	51.6%	48.4%
55 años o más	9.8%	53.3%	46.7%
45 a 54 años	13.0%	54.8%	45.2%
35 a 44 años	17.3%	52.1%	47.9%
25 a 34 años	19.2%	53.1%	46.9%
18 a 24 años	15.8%	49.9%	50.1%
12 a 17 años	14.9%	50.6%	49.4%
6 a 11 años	10.0%	46.8%	53.2%

Nota: Los valores para mujeres y hombres están calculados respecto del total de usuarios de Internet de seis años o más por grupos de edad y sexo.

Usuarios de Internet por tipo de uso	
Actividad	Proporción
Para entretenimiento	91.5%
Para obtener información	90.7%
Para comunicarse	90.6%
Para acceder a redes sociales	87.8%
Para apoyar la educación/capacitación	83.8%
Para acceder a contenidos audiovisuales ¹	80.5%
Para descargar software	48.0%
Para leer periódicos, revistas o libros ¹	47.3%
Para interactuar con el gobierno	35.6%
Para ordenar o comprar productos	22.1%
Utilizar servicios en la nube	19.4%
Para operaciones bancarias en línea	16.8%
Para ventas por Internet	9.3%

Estas opciones de respuesta están consideradas dentro de la opción de entretenimiento.
Para más información consulte el portal del INEGI en: <https://www.inegi.org.mx/datos/>

Fuente: ENDUTIH 2019

De los resultados del ENDUTIH destaca, por una parte, que en México hay 80.6 millones de usuarios de Internet de seis años o más, que representan el 70.1% de la población en ese rango de edad. Y, por otra parte, que las tres principales actividades de los usuarios de Internet en 2019 fueron: entretenimiento (91.5%), obtención de información (90.7%) y comunicación (90.6%). Así como que las tres

⁶⁵ https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2020/OtrTemEcon/ENDUTIH_2019.pdf

actividades menos realizadas fueron: utilizar servicios en la nube (19.4%), operaciones bancarias en línea (16.8%) y ventas por Internet (9.3%).

Es así que, de acuerdo a las tendencias del uso de Internet de los usuarios mexicanos, se encuentra que aún existen áreas de oportunidad principalmente en actividades relacionadas con el Cloud Computing y con la economía digital, tales como el comercio y el gobierno electrónicos. Por ejemplo, en la Encuesta Nacional de Calidad e Impacto Gubernamental 2017 del INEGI, se encontró que de la población de 18 años y más, solo el 5.7% ha realizado trámites, pagos o servicios del Gobierno a través de Internet.

En relación al Cloud Computing, no se cuenta con información específica de las actividades que utilizan esta nueva plataforma de servicios y en consecuencia su impacto en el desempeño de las redes de telecomunicaciones. Por ejemplo, el aumento de utilización de capacidad de las redes como consecuencia de las actividades de minería de datos que soportan el procesamiento de información de nuevas aplicaciones basadas en la nube.

Al respecto se hace necesario recabar información que permita dimensionar si existe la capacidad tanto en infraestructura como en procesamiento que permita el despliegue y crecimiento de los servicios de forma sostenida para soportar los nuevos modelos de operación para las soluciones en la nube.

Relación de Empresas ver ANEXO 1: Empresas Cloud Computing en México.

En esta sección se comprenden los aspectos que han permitido la adopción de servicios de Cloud Computing en otros países. Es decir, se pretende reconocer o clasificar los factores que pueden incentivar el desarrollo de las tecnologías de Cloud Computing en el país y de acuerdo a su naturaleza identificar la mejor forma para su implementación.

6.1. Regulación del Tráfico de datos fronterizos

Algunos países imponen políticas de datos que restringen tanto el uso doméstico de datos como el flujo de datos a través de las fronteras. Esta tendencia ha dado como resultado un nivel creciente de restricción en los datos en muchos países y conlleva costos significativos tanto para las empresas nacionales como extranjeras y, en última instancia, para los consumidores. Estos países estudian o adoptan medidas para desalentar u obstaculizar los flujos transfronterizos de datos, por motivos diversos. Por ejemplo, garantizar la seguridad nacional, proteger los datos personales y la privacidad, asegurar el acceso a información sobre la aplicación de la legislación, evitar flujos que puedan perturbar el orden público nacional, o

proteger y promover la actividad económica dentro del territorio nacional. En algunos países, estas políticas pueden formar parte de una estrategia gubernamental más amplia orientada al control cyber-soberano de la economía digital y la sociedad. En esos casos, los obstáculos a los flujos transfronterizos de datos han ido a veces acompañados de políticas de localización, que exigen que los datos permanezcan en una jurisdicción determinada y sean procesados allí.

Dentro de los países que imponen políticas de datos más estrictas, se encuentran en Rusia, China y Turquía que restringen el entorno regulatorio para mover y usar datos. A estos países les siguen dos grandes economías europeas, Francia y Alemania. Curiosamente, los cinco países son relativamente grandes, a menudo con una sólida base de fabricación.

Permitir el flujo libre de datos a través de las fronteras tiene un impacto positivo neto significativo en la economía global. Un informe de febrero de 2016 del McKinsey Global Institute, estimó que los flujos de datos transfronterizos contribuyeron con casi US \$ 2.8 billones a la economía global en 2014 al permitir el flujo de bienes, servicios y otros recursos (McKinsey Global Institute, marzo 2016). El informe estima que esta cifra podría alcanzar los US \$ 11 billones para 2025.

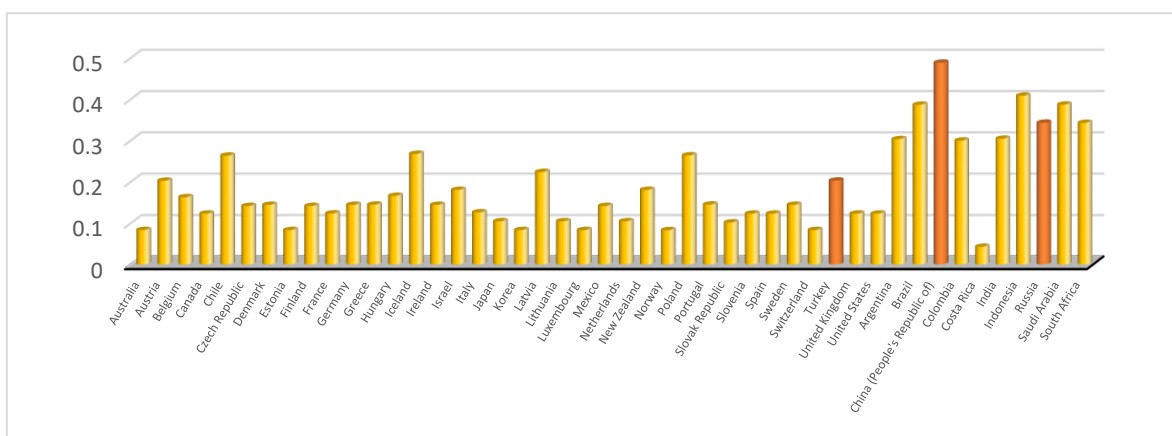
Por su parte la Conferencia de las Naciones Unidas sobre Comercio y Desarrollo (UNCTAD, por sus siglas en inglés), en su nota presentada ante la junta de comercio y desarrollo 2019, señala que el tráfico mundial sobre el Protocolo de Internet (tráfico IP), que es un indicador de los flujos de datos, ha crecido exponencialmente en los dos últimos decenios. En 1992, las redes de Internet transportaban en todo el mundo 100 gigabytes (GB) al día, aproximadamente. Diez años más tarde, el tráfico global en Internet ascendía a 100 GB por segundo. En 2017, era de más de 45,000 GB por segundo, debido a los cambios cualitativos y cuantitativos de los contenidos que circulan por la Red. En 2022, con 150,700 GB por segundo, se espera que el tráfico IP mundial sea 75 veces mayor que en 2007.

La transición hacia el Cloud Computing puede considerarse un cambio radical en la relación entre las telecomunicaciones, las empresas y la sociedad, resultado del enorme aumento de la capacidad de procesamiento y almacenamiento de datos, así como de su velocidad de transmisión, acompañados de una drástica reducción de los precios. Por ejemplo, el costo medio de 1 GB de capacidad de almacenamiento pasó de más de 400.000 dólares en 1980 a 0,02 dólares en 2016.

El Centro Europeo para la Economía Política Internacional (ECIPE), considera que las restricciones unilaterales sobre el flujo de datos transfronterizos y el acceso a los mercados extranjeros puede frenar el crecimiento económico y la recuperación porque limitan el acceso a precios competitivos, el crecimiento del empleo en muchos sectores y las oportunidades de inversión.

El Índice de Restricción del Comercio de Servicios Digitales de la OCDE (Digital STRI, por sus siglas en inglés), es una herramienta que identifica, cataloga y cuantifica las barreras transversales que afectan los servicios comercializados digitalmente. Consta de dos componentes, la base de datos reguladora y los índices, que reúnen información comparable de 44 países. El Digital STRI muestra un entorno regulatorio global diverso y complejo que afecta el comercio de servicios habilitados digitalmente. Además, en los últimos años, los índices muestran un entorno regulatorio cada vez más estricto, destacando la necesidad de cooperación y diálogo internacional para maximizar los beneficios de la digitalización.

Gráfica 16: Índice de restricción del comercio de servicios digitales.



Fuente: Elaboración propia con información estadística de la OECD. Stat. (https://stats.oecd.org/Index.aspx?DataSetCode=STRI_DIGITAL)

El indicador de restricción del comercio de servicios digitales ubica a México en la mitad de la tabla, al nivel de países como Israel, Hungría, Bélgica y Turquía, pero mejor que países como Brasil, China e Indonesia. Lo anterior refleja que existe un área de oportunidad y se podría encontrar en una revisión de la regulación una palanca en los servicios en la nube que facilite o promueva que la economía se mueva hacia un mayor grado de digitalización que beneficie al país como la política de “Cloud First” que se explicará más adelante.

6.2. Seguridad

Hoy en día las organizaciones del sector público siguen teniendo una preocupación legítima por la seguridad de sus datos, por lo que algunos gobiernos han llegado a establecer a los Proveedores de Servicios de Comunicaciones (PSC) como requisito de que todo contenido de los clientes procesado y almacenado en un sistema de

TI, permanezca dentro de las fronteras del país, con el afán de brindar una capa adicional de seguridad.

La residencia de los datos es una combinación de problemas asociados principalmente a riesgos de seguridad percibidos (y en algunos casos reales), relacionados con el **acceso de terceros** a la información, incluidas las autoridades competentes extranjeras. Los clientes del sector público desean tener la certeza de que sus datos están protegidos frente al acceso no deseado de atacantes maliciosos o de otros gobiernos.

Tomar una postura rigurosa sobre la residencia de los datos a veces restringe el uso del Cloud Computing. La preocupación general en torno a la ciberseguridad, así como la posible extralimitación de la vigilancia gubernamental por parte de algunos países ha contribuido a que el **debate se centre continuamente en mantener los datos en el país**. No obstante, esta restricción puede ser contraproducente para el objetivo de proteger de manera efectiva datos sensibles respecto a diversos ámbitos del sector público. Los niveles más altos de protección sólo se pueden garantizar en el entorno y hábitat donde se ubica la plataforma del proveedor de Cloud Computing ya que se puede mantener la soberanía reglamentaria de nación-estado.

Los proveedores de Cloud Computing diseñan, operan y mantienen ofrecimientos que permiten a clientes de diferentes sectores (comercial, público, regulado), abordar algunos de los riesgos de seguridad y vulnerabilidades predominantes. Los clientes confían en los ofrecimientos, para que apliquen prácticas de seguridad que sean dinámicas y respondan ante amenazas en tiempo real, mejorando así drásticamente la seguridad de cada cliente. Por su parte, los proveedores tienen los incentivos adecuados para mantener una ciberseguridad de clase mundial dado que se enfrentarían a consecuencias sustanciales a largo plazo, tales como las repercusiones asociadas a un sistema en peligro, la pérdida de confianza de los clientes y el daño a la marca. En otras palabras, la seguridad de primer nivel es requerida para el éxito de un PSC. La seguridad tiene que estar totalmente integrada en el diseño, el desarrollo y las operaciones de los Servicios en la Nube.

La protección a la seguridad de los datos del gobierno y la infraestructura de TI y la privacidad de los ciudadanos deben abordarse con planes nacionales de ciberseguridad sólidos y un marco legal y normativo moderno para la economía digital. El Cloud Computing es un recurso adicional a disposición de los gobiernos para lograr este objetivo de política.

La transición hacia el Cloud Computing debe ser estable, confiable y segura, que genere confianza a los usuarios de los distintos servicios, es por esto que los gobiernos a nivel internacional están considerando a la ciberseguridad como parte de su

estrategia nacional, con la intención de generar políticas públicas que ayuden a generar un entorno de adopción fiable y seguro.

Diferentes aspectos de seguridad han sido cubiertos por estándares de distintas organizaciones y en el caso específico de los servicios en la nube tenemos el marco de seguridad ITU-T X.1601, que permite analizar los riesgos y amenazas en el entorno de Cloud Computing, así como describe medidas de seguridad que pueden atender dichas amenazas y resolver retos de seguridad y privacidad.

Derivado de que gran parte de la información y los servicios que son brindados en los países a través **de infraestructuras críticas** están adoptando Cloud Computing, la creación de legislación para la protección de estas infraestructuras se vuelve prioritario, sin embargo, un tema que presenta un área de oportunidad es la armonización de esta legislación, ya que si bien se están generando políticas para atender la ciberseguridad, estas tienden a tipificar los distintos ataques, ataques que pueden o no ser considerados graves en las distintas legislaciones, lo que genera incertidumbre sobre las acciones a generar para minimizar estas incidencias.

Para lograr una armonización de estas políticas públicas, se requiere de gran participación de los países, así como de reportar y compartir la información sobre los ataques recibidos de manera oportuna. Asimismo, para una efectiva implementación de la legislación se requiere del cumplimiento, certificación y verificación de la misma.

6.3 Política de Cloud First

Una política de Cloud-First requiere que las agencias gubernamentales utilicen los servicios comerciales en la nube como el habilitador principal para la modernización de TI. Los Proveedores de Servicios en la Nube ahora facilitan que los clientes gubernamentales se alejen del modelo de gastos de capital (CapEx) de comprar y poseer activos físicos de TI que se deprecian.

En cambio, las agencias pueden asignar sus presupuestos de TI para gastos operativos (OpEx) para cubrir solo los servicios que utilizan, al tiempo que evitan la *'deuda tecnológica'* tradicional relacionada con CapEx al tener acceso a pedido a productos y servicios de TI de vanguardia. Una política de Cloud-First bien diseñada conduce a ahorros de costos, brinda mayor seguridad que las soluciones locales, permite la flexibilidad para que las entidades gubernamentales ajusten el uso y aumenta los esfuerzos para el desarrollo y la transparencia de la fuerza laboral.

Tabla 18: Marco Internacional del Cloud First

Australia	Emitió su Estrategia de Nube Segura, reemplazando la Política de Computación en la Nube del Gobierno que se lanzó en 2014. La nueva estrategia se centra en ayudar a las agencias gubernamentales a usar la nube más fácilmente, con el principio de "usar los servicios de nube pública como el predeterminado".
Canadá	Emitió su actualización de la Estrategia de Adopción de la Nube en 2018, destinada como una directiva de política para que los departamentos y agencias se centren en la adopción de "la nube pública primero".
Chile	Emitió una Orden Ejecutiva "Instructivo Presidencial sobre servicios en la nube" en febrero de 2018.
Filipinas	Anunció su Política de Cloud First en enero de 2017.
Singapur.	Ha sido un líder en la política de la nube. Su folleto de computación en la nube de 2016 proporciona una visión general de las estrategias de uso de la nube.
Reino Unido.	Introdujo una Política Cloud First y ha creado la red de adquisición y proveedores necesaria para implementar esta estrategia.
Estados Unidos.	Lanzó una Política de Cloud First en 2011. En septiembre de 2018, el Gobierno Federal emitió un Borrador actualizado de Estrategia Federal de Computación en la Nube: De Cloud First a Cloud Smart.

Gracias al marco internacional se pueden identificar un conjunto de mejores prácticas para respaldar la transición al entorno del Cloud First.

- **Promoción de las políticas Cloud First:** Los gobiernos emiten una declaración de política con una directiva procesable, incluidos los horarios, que crea un marco para la implementación de tecnologías en la nube, asigna las funciones y responsabilidades de las entidades gubernamentales y los PSC y establece un canal de adquisición diseñado para obtener todos los beneficios de las tecnologías en la nube. Una vez que los gobiernos han establecido políticas e infraestructura para Cloud First, sus profesionales de TI pueden diseñar soluciones "nacidas en la nube" o "nativas de la nube".
- **Acreditación, cumplimiento y seguridad en la nube:** Los gobiernos utilizan los sistemas de acreditación nacionales e internacionales enfocados en la nube para evaluar a los PSC (en lugar de crear sus propios programas de certificación únicos) y aprovechar el modelo de responsabilidad compartida para la seguridad en la nube.
- **Clasificación de datos:** Los gobiernos clasifican sus datos en función de su nivel de sensibilidad y luego gestionan cada segmento de manera congruente con su nivel de sensibilidad.
- **Privacidad y control de datos:** Los gobiernos establecen y/o adoptan políticas complementarias de seguridad y procesamiento de datos y privacidad para respaldar una transición exitosa a la nube.

- **Contratación y adquisición en la nube:** Los gobiernos diseñan e instituyen un canal de contratación en la nube que las agencias pueden utilizar para obtener todos los beneficios de la nube, con términos y condiciones centrados en la nube, precios, gobernanza y seguridad. Los contratos de adquisición reconocen el modelo de pago por uso de la computación en la nube comercial y el papel de los socios y revendedores de los PSC en el modelo de entrega en la nube.

México se encuentra rezagado respecto a otros países del mundo y América Latina al no haber implementado aún políticas enfocadas a políticas a Cloud First. A pesar de que tiene un tráfico de datos mayor al de países que han comenzado a implementar dichas políticas, como Chile.

7. Conclusiones y recomendaciones

El fenómeno de la transformación digital está cambiando todo a su paso, redefiniendo nuevos esquemas comerciales, educativos, médicos, económicos, sociales, políticos y de entretenimiento, entre otros más, que impactan a todos los sectores de la población. Los datos de incremento de tráfico auguran el crecimiento intensivo de nuevos servicios basados en la nube y las infraestructuras subyacentes que los soportan, como lo son los Data Centers y redes de nueva generación. Es así que nos encontramos ante una coyuntura en la que se deben anticipar los planes de despliegues e inversiones de infraestructura en las redes y servicios digitales para poder sostener estas nuevas demandas de capacidad, aún más considerando que estos cambios en el uso intensivo de los servicios digitales pueden resultar permanentes, a raíz del aislamiento generado por la pandemia COVID-19.

En este documento se mencionaron los principales conceptos que describen al Cloud Computing, así como sus modelos de despliegue, características y tipos de servicios bajo los que se ofrece con el fin de comprender los fundamentos de su desarrollo y operación. Además del cambio de paradigma respecto a la consideración de los costos de capital y operación que representa un avance sin precedentes en la manera de prestar servicios digitales en la actualidad y en un futuro.

En el caso de México, los Data Centers se ubican principalmente en la Ciudad de México, Estado de México, Querétaro, Nuevo León, Jalisco y Aguascalientes, siendo el país el segundo mercado de Data Centers más grande de Latinoamérica. Además de que se prevé un crecimiento significativo de los Data Center en el país durante los próximos años, una de las principales ventajas para la construcción de Data Centers en México es que posee una ubicación geográfica estratégica, ya que el país es el punto intermedio para ofrecer conectividad entre Estados Unidos y

el resto de América Latina, por lo que existe un gran potencial de desarrollo en este mercado para generar una mayor inversión en los próximos años.

Por eso los Data Centers son en la actualidad el corazón del ecosistema digital, y ante la creciente demanda por servicios de calidad ofrecidos a través de la Nube, las empresas de telecomunicaciones ahora deben realizar a su vez importantes inversiones en la reconversión de su infraestructura tradicional para soportar los parámetros de desempeño que requieren los nuevos servicios digitales.

Asimismo, se da cuenta que el suministro de energía eléctrica es uno de los principales componentes de costo de los Data Centers y que aun resultan comparativamente mayores en aproximadamente 30% con respecto a los Estados Unidos, costos que reducen la competitividad y atractivo para atraer y realizar inversiones de este tipo en nuestro país.

Por otra parte, en este documento se explicaron las diferencias subyacentes en las infraestructuras necesarias para las telecomunicaciones tradicionales y para los Data Centers, que soportan los nuevos servicios digitales, ya que la configuración de las redes actuales aporta factores de latencia al tráfico de datos que pudieran afectar los servicios de Cloud Computing. Esto es crítico en el sentido de que actualmente gran parte de las soluciones de las tecnologías de información y comunicaciones están en la Nube y su crecimiento es a gran escala gracias a la adopción de nuevas tecnologías como el acceso 5G, el Internet de las cosas (IoT) y el empleo de soluciones de inteligencia artificial (AI), que aceleran las cargas de trabajo (workloads) de misión crítica (control de tráfico, drones, vehículos autónomos, etc.) en la nube.

Con relación a lo anterior y en la medida que aumenta la producción y consumo de datos, las nuevas arquitecturas distribuidas y métodos de enrutamiento requeridas por el Cloud Computing hacen imperativa una mejor interconexión que elimine el factor distancia, por lo que los negocios digitales buscan consolidar infraestructuras de centros de datos adyacentes a nodos de interconexión o puntos de intercambio de tráfico para mantenerse competitivos.

Se reconoce que los factores de interconexión y las estrategias de despliegue de las redes, principalmente las de acceso de fibra óptica, inciden en los parámetros de calidad con que los Servicios en la Nube se brindan, por lo que cobra especial importancia el contar con redes que puedan operar a altas velocidades de carga y descarga, así como bajas latencias para brindar una experiencia óptima al usuario final.

Los servicios digitales van a requerir cada vez con más prioridad, mejores estándares de desempeño de las redes con las que cuenta el país, originalmente diseñadas

para brindar servicios de telecomunicaciones tradicionales las cuales pudieran no estar completamente adaptadas a estas nuevas necesidades. Actualmente, conforme a información del Instituto, apenas aproximadamente en 30% de los municipios del país operan redes de acceso de fibra óptica, por lo que se identifica la necesidad de mayores despliegues para garantizar mejores niveles de calidad, ya que, como se muestra en la sección 5.4.4., en promedio en México se registran parámetros de latencias o retardos para el tráfico cursado entre redes alejados de los requeridos para el funcionamiento óptimo de los servicios provistos por el Cloud Computing.

Es por ello que se hace necesaria la reducción de los costos de transacción que enfrentan concesionarios o terceros para instalar redes de fibra óptica, en ocasiones impuestos por autoridades locales a través de trámites y cobros innecesarios. Inclusive, ante la importancia de que los gobiernos desarrollen e implementen políticas que favorezcan el uso de Servicios en la Nube, se podrían establecer acuerdos que reserven fibras o capacidad para uso compartido o asignado al gobierno para que soporte sus diferentes trámites ante la ciudadanía, es decir, para evolucionar más aceleradamente hacia una efectiva práctica de e-Gobierno.

Además de la promoción de despliegues de proyectos de fibra óptica, tanto en segmentos de acceso como de transporte, resulta relevante que sea aprovechada junto con las redes actuales por más y diferentes proveedores de servicio de acceso a Internet (ISP) independientes, que a su vez demandan o propician la creación y utilización de más puntos de intercambio de tráfico de internet (IXP) para optimizar la provisión de servicios.

Es decir, se identifica la necesidad de evaluar los esquemas de intercambio de tráfico de datos en nuestro país, asociados a la interconexión de Internet, ya que actualmente la provisión del acceso y el tránsito hacia otras redes de Internet está bajo control de pocos agentes que integran estas funciones, que si bien, soportan por ahora la infraestructura para Data Centers, no son necesariamente las más eficientes para el óptimo funcionamiento de Internet, ya que podrían estar incorporando otros costos no relacionados con la provisión del servicio de acceso a Internet.

La estructura del mercado actual donde predominan pocas redes de telecomunicaciones con un alcance importante, que a su vez son proveedores de su propio servicio de Internet, al concentrar un importante tráfico de sus propios usuarios pudiera no generar incentivos a otros esquemas de intercambio de tráfico, ya que cada una los gestiona de manera individual, limitando a pequeños ISPs de la posibilidad de obtener mejores condiciones para la gestión de su tráfico.

Esto resulta en pocas alternativas para que proveedores de Internet independientes o de menor escala o proveedores de Cloud Computing operen de forma diferenciada mediante el acceso directo a la red mundial de Internet y puedan administrar y gestionar su tráfico sin recurrir a la intermediación o arbitraje de las grandes redes tradicionales o puntos de intercambio (IXP) fuera del país.

Asimismo, se debe promover la eliminación de posibles barreras para un mayor despliegue de Data Centers en el país con el fin de incrementar su oferta existente y hacer crecer dicho mercado, promoviendo sus aplicaciones en los diversos sectores e industrias, lo que generaría más inversión y un entorno de Servicios de Cloud Computing competitivo en México. Lo anterior, aprovechando la ubicación geográfica del país que lo posiciona como un centro estratégico para proveer conectividad y servicios digitales a la región de América Latina. Además, es de gran relevancia contar con esquemas eficientes de provisión de energía eléctrica para soportar la infraestructura, tanto de redes de telecomunicaciones como de Data Centers, que son consumidores intensivos y constantes de energía. En este sentido, y en complemento a lo anterior, también es importante que en el país se continúe desarrollando y capacitando el capital humano para instalar, mantener y operar la infraestructura de Cloud Computing.

Acorde a los resultados de la ENDUTIH 2019, de los 80.6 millones de usuarios de Internet en el país se aprecia que las tres principales actividades en 2019 fueron: entretenimiento (91.5%), obtención de información (90.7%) y comunicación (90.6%). Así como que las tres actividades menos realizadas fueron: utilizar servicios en la nube (19.4%), operaciones bancarias en línea (16.8%) y ventas por Internet (9.3%), lo cual revela áreas de oportunidad principalmente en actividades relacionadas con el Cloud Computing y con la economía digital, tales como el comercio y el gobierno electrónicos. Al respecto se deben revisar políticas como *-Cloud First-* por parte de los agentes reguladores y promotores de los servicios digitales dentro del país.

En suma, para materializar la promesa de la transformación digital y sus servicios se identifican tres factores indispensables: las redes de fibra óptica, las tecnologías que habiliten los servicios de alta velocidad y sobretodo los servicios en la nube o **Cloud Computing**. Estos factores o elementos son los que permitirán el efectivo despliegue de las nuevas tecnologías como 5G y su consecuente aplicación al Internet de las cosas (IoT), así como la habilitación de la inteligencia artificial (AI), y de todas aquellas tecnologías y soluciones que necesita ahora más que nunca nuestra sociedad.

ANEXO 1: Empresas Cloud Computing en México

AWS (Amazon Web Services)

Descripción: Empresa de Amazon que ofrece servicios de computación en la nube, integrando una plataforma de computación en la nube, orientado a empresas, emprendedores y agencias gubernamentales.

Cobertura: La nube AWS incluye 61 zonas de disponibilidad en 20 regiones geográficas de todo el mundo, esto permite varias zonas de disponibilidad físicamente independientes y aisladas, conectadas mediante redes con un alto nivel de desempeño, redundancia, y baja latencia.

Servicios: Plataforma en la nube que ofrece más de 175 servicios integrales de centros de datos a nivel global, ofreciendo desde tecnologías de infraestructura como cómputo, almacenamiento y bases de datos hasta tecnologías emergentes como aprendizaje automático e inteligencia artificial, lagos de datos y análisis e Internet de las cosas.

Mapa 2: AWS



Fuente: (<https://aws.amazon.com/es/about-aws/global-infrastructure/?hp=tile&tile=map>)

Axity

Descripción: Compañía conformada por el fondo de capital privado, Southern Cross Group, con amplia trayectoria en América latina. Actualmente cuenta con 4,000 colaboradores y 400 clientes, en asociación tecnológica con Forcepoint, IBM,

McAfee, Microsoft, Palo Alto Networks, Google, BMC Software, Now Platform, Splunk Technology, Amazon Web Services, Dell EMC, y CISCO.

Ubicación: Presencia regional en México, Colombia, Perú, Chile, Argentina y Estados Unidos.

Servicios: Integra un portafolio de servicios con soluciones en ciberseguridad, gestión de TI, aplicaciones, transformación digital y tecnología en la que se destacan Platform as a Service (PaaS) e Infrastructure as a Service (IaaS).

Fuente: <https://www.axity.com/es/>

Axtel/Alestra

Ubicación: Los Centros de Datos se encuentran ubicados en:

- Monterrey
- Querétaro
- Guadalajara
- Ciudad de México

Descripción: Más de 7 mil 200 m² en 6 DC. Diseñados bajo los más altos estándares ICREA, Uptime Institute, CEEDA y CMMI. Redundancia a nivel metropolitano y nacional. Centros de Datos con certificaciones, energía, climatización, seguridad, conectividad, escalabilidad y personal especializado. El Data Center más grande está ubicado en la ciudad de Querétaro, el cual se distingue por ser Multi Tier y Multi Densidad.

Servicios: Ofrece servicios de SaaS, PaaS e IaaS.

Fuente: <http://www.axtel.mx/> (Centros de Datos ahora Equinix)

Bestel

Descripción: Empresa de Telecomunicaciones y Tecnologías de Información, en 2015 integraron las capacidades de Bestel, Metrored y Quantumlink.

Cobertura: Con su actual red de fibra, tiene una cobertura en 225 ciudades del país.

Servicios: Ofrece servicios SaaS en la nube con CenturyLink, Rackspace, Limelight; nube pública con Century Link; IaaS con RackSpace y Servicios de Data Center (Data Center Design Awareness).

Mapa 3: Bestel



Fuente: (<https://www.bestel.com.mx/mapa-cobertura>)

C3ntro Telecom

Descripción: Operador mexicano con 22 años de experiencia en Telecomunicaciones constituido en 2011 con la razón social Bicentel SA de CV, parte del grupo de IPBtel.

Servicios: Dispone de 20 Data Center POP en México y USA con los que ofrece Servicios en la Nube de SaaS, PaaS e IaaS, con un ingreso anual de ventas de más de USD\$150,000.00.

Mapa 4: C3ntro Telecom



Fuente: (<https://www.c3ntro.mx/proveedor-de-telecomunicaciones-en-mexico>)

Century Link

Descripción: Empresa que se fusionó con Level 3 en 2017, entra a México a los negocios mayorista y de corporativos en cuanto a provisión de telefonía, Internet y capacidad de transporte; también en aquel sobre enlaces dedicados, redes privadas virtuales, administración de información en Data Centers y conectividad internacional a través de cables submarinos, con ayuda de la infraestructura ya instalada de Level 3.

Servicios: Ofrece servicios de SaaS, PaaS e IaaS.

Mapa 5: Century Link



Fuente: (<https://www.centurylink.com.mx/resources/network-maps.html>)

EMC/Dell

Descripción: Dell y EMC son dos empresas que se fusionaron para convertirse en Dell Technologies. Dell Technologies está conformada por siete líderes en tecnología (Dell, Dell EMC, Pivotal, RSA, Secureworks, Virtustream y VMware) que juntos forman una empresa con el objetivo de impulsar la transformación digital y lograr resultados reales todos los días para los clientes y las personas que se asocian con ella.

Servicios: Proveedor de infraestructura de nube, ayuda a las organizaciones a ampliar un modelo operacional coherente en nubes públicas y privadas, para una experiencia de nube híbrida. Permite crear una estrategia de múltiples nubes duradera que unifique los ambientes y reduzca el riesgo en las nubes.

Equinix

Descripción: Empresa multinacional con sede en Estados Unidos, especializada en conexión de Internet y centros de datos.

Cobertura: Centros de datos distribuidos en 24 países de los 5 continentes.

Servicios: Soluciones de infraestructura en la nube privada, pública e híbrida. Así como IaaS, PaaS y SaaS.

Mapa 6: Equinix



Fuente: (<https://www.equinix.com/data-centers/>)

Fujitsu

Descripción: Fujitsu es una empresa japonesa en el sector de tecnologías de la información y las comunicaciones (TIC) que ofrece una gama de productos, soluciones y servicios de tecnología. Aproximadamente 159,000 empleados de Fujitsu brindan soporte a los clientes que están distribuidos en más de 100 países.

Servicios: Soluciones en cloud tales como IaaS (servicios en la nube públicos, privados y locales) y SaaS (servicios de suscripción, que incluyen aplicaciones para mejora de la productividad en la oficina, gestión de relaciones con los clientes, gestión de TI y otras soluciones empresariales clave).

Gold Data

Descripción: Empresa proveedora de telecomunicaciones con presencia en América.

Cobertura: Ofrece conectividad internacional desde México DF, Monterrey y Guadalajara al resto de la región.

Servicios: Cuenta con puntos de presencia estratégicos, capacidad submarina y acceso a últimas millas para ofrecer soluciones completas a sus clientes. Soluciones de conectividad tales como Cloud Connection, Soluciones de Broadcasting y Seguridad de Red.

Mapa 7: Gold Data



Fuente: (<https://golddata.net/our-network/>)

Google Cloud

Descripción: Plataforma de la empresa Google que proporciona soluciones tecnológicas en la nube.

Servicios: Proporciona opciones con nubes híbridas y múltiples, modernización de la infraestructura, administración de datos, desarrollo de aplicaciones, IA y estadísticas de negocios inteligentes.

Ho1a

Descripción: Compañía de telecomunicaciones y TI, originaria de Guadalajara, Jalisco; desde 2013 forma parte de Megacable Holdings. Ofrece un modelo de integración de tecnología orientada a resolver problemas de negocio.

Servicios: Ofrece servicios de Data Center, así como servicios de IaaS, SaaS y PaaS.

HPE (Hewlett Packard Enterprise)

Descripción: Empresa de Estados Unidos en tecnologías de la información que brinda servicios de consultoría y de soporte en diversas áreas.

Servicios: Nube privada, nube pública e híbrida. Consultoría de informática y gestión en la nube.

Huawei Cloud

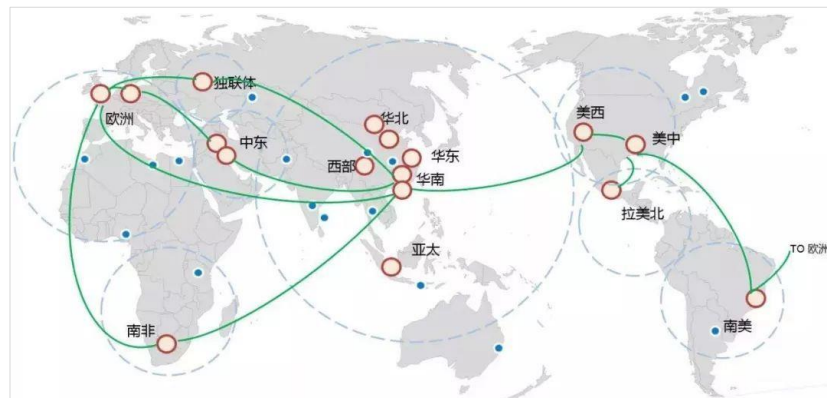
Descripción: Huawei Cloud despliega múltiples regiones geográficas y zonas de disponibilidad en todo el mundo, proporciona redes de conexión a la nube globales de alta velocidad y estables, y servicios localizados que están cerca de los clientes.

Beneficiándose del diseño comercial global de Huawei, el centro de datos de Huawei ha formado ocho redes troncales completas y de alta velocidad de 100 ms en el mundo, que cubre 8 regiones y 180 países de todo el mundo. Los usuarios pueden acceder al centro de datos más cercano dentro de los 100 ms en cualquier parte del mundo.

Los Data Centers de Huawei se encuentran operando en Pekín, Dalián, Guiyang, Shanghái, Cantón, Shenzhen, Hong Kong, Bangkok, Singapur, Johannesburgo, Sao Paulo, Buenos Aires, Lima, Santiago de Chile y **Ciudad de México** (<https://www.huaweicloud.com/intl/es-us/>).

Servicios: Huawei Cloud es la marca de servicios en la nube de Huawei. Utiliza métodos en línea para abrir a los clientes los más de 30 años de acumulación de tecnología y soluciones de productos en el campo de la infraestructura de las TIC para los clientes. Se compromete a proporcionar servicios en la nube estables, confiables, seguros, sostenibles e innovadores.

Mapa 8: Data Centers Huawei Cloud



Fuente: (https://www.sohu.com/a/256573421_464033)

IBM

Descripción: Compañía multinacional estadounidense de tecnología y consultoría, que ofrece diversos servicios de consultoría, Internet e infraestructura tecnológica.

Servicios: IaaS, PaaS y SaaS. Cuenta con un producto denominado IBM Cloud, el cual lo define como una robusta suite de herramientas de datos y de IA avanzadas, para ayudar en la transición de la industria hacia la nube, con recursos que pueden ser compartidos o de uso exclusivo y privado, al crear aplicaciones y piezas de software en un entorno virtual dentro de una infraestructura que se encuentra en la nube.

Infotec

Descripción: Centro de Investigación e Innovación en Tecnologías de la Información y Comunicación, tiene como propósito mejorar la competitividad, transparencia y eficiencia de las empresas y organizaciones de los sectores público, académico, social y privado.

Lo anterior, mediante actividades de investigación, innovación, desarrollo, consultoría, difusión, formación, capacitación, actualización de recursos humanos, servicios especializados, así como intercambio y alianzas tecnológicas, tanto a nivel nacional como internacional.

Servicios: Cuenta con dos Data Center, con espacio físico en Aguascalientes y la Ciudad de México para la gestión de servicios (certificación ISO/IEC 20000-1:2011) de infraestructura, co-ubicación y bóveda de medios con los ecosistemas de TIC de sus clientes. En la ciudad de Aguascalientes cuenta con alta seguridad y disponibilidad (diseño certificado bajo TIER III por el Uptime Institute), lo que lo convierte en el primer Data Center certificado del gobierno mexicano. En la Ciudad de México, cuenta con un Data Center donde se brinda consultoría especializada para migrar a procesamiento físico de última generación y/o inicios de virtualización. Ambos centros de datos están disponibles para el uso del sector público y privado.

IPXON

Descripción: Con más de 20 centros de datos distribuidos geográficamente en ciudades clave, IPXON Networks opera una red cloud extensa, en asociación con CISCO, Toshiba, MikroTik, Supermicro e Intel.

Cobertura: Ciudad de México, Guadalajara y diversos países de América Latina.

Servicios: Proveedor de Infrastructure as a Service (IaaS).

KIO Networks

Ubicación: 40 Centros de Datos (Cd. México, Puebla, Veracruz, San Luis Potosí, Xalapa, Mérida (1&2), Villahermosa, Cuernavaca, Tuxtla Gutierrez, Guadalajara, Hermosillo).

Servicios: Ofrecen servicios de SaaS, PaaS e IaaS, Data Analytics, Ciber Security.

Descripción de sus centros de datos:

KIO MEX 1. Localizado en Cuajimalpa, Ciudad de México. Cuenta con 4 Salas de Cómputo de 650 m². El Data Center fue diseñado basado en Estándar del Uptime Institute, específicamente bajo una topología TIER III, Certificaciones: ICREA Nivel 4, ISO 14001, ISO 27001, ISO 9001, PCI "Payment Card Industry Data Security Standard" y SSAE16 Type II. Carrier acometidos: redIT, MetroRed, Bestel, Telefónica, Telmex, Axtel-Alestra, Totalplay, Verizon, Transtelco y MetroCarrier.

KIO MEX 2. Localizado en Cuajimalpa, Ciudad de México. Cuenta con 2 Salas de cómputo de 600 m² cada una. TIER IV y certificaciones: ICREA Nivel 5, CEEDA Gold, ISO 14001, ISO 27001, ISO 9001, PCI "Payment Card Industry Data Security Standard". SSAE16 Type II. Carrier acometidos: redIT, Marcatel, Level 3, Metrored, GTT- Perseus Telecom, Totalplay, Cogent, Transtelco, Telmex, Bursatec, Bestel, MCM y Axtel-Alestra.

KIO MEX 3. Localizado en Ciudad de México. Cuenta con una Fase de 900 m², con estructura categoría 4. Certificaciones: ICREA Nivel 3, ISO 14001, ISO 27001, ISO 9001, PCI "Payment Card Industry Data Security Standard" y SSAE16 Type II. Carriers acometidos: Metro Red, Bestel, Telefónica, Telmex, Axtel-Alestra, Totalplay, Verizon, Transtelco y Metro Carrier.

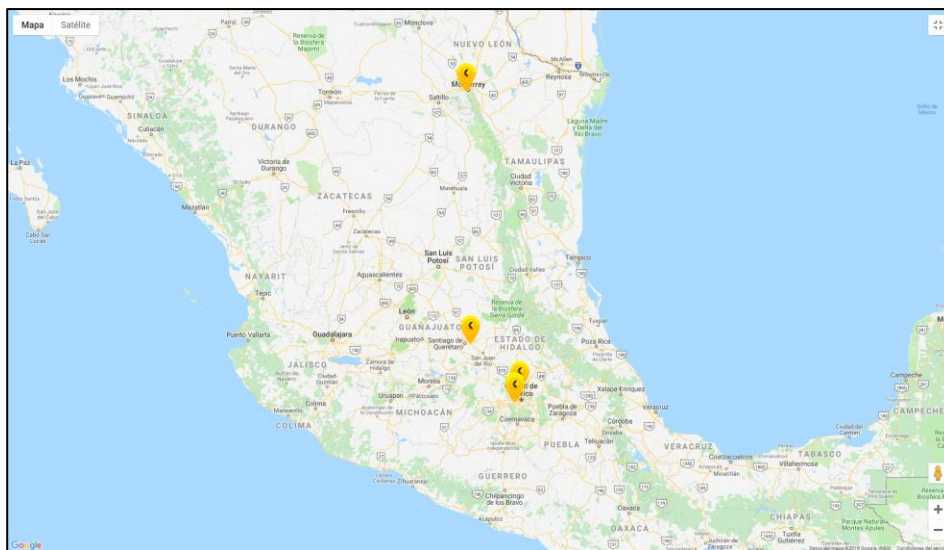
KIO MEX 4. Localizado en Huixquilucan, Estado de México. 3,450 m² de piso blanco dividido en tres Salas de Cómputo. TIER III y certificaciones: ICREA Nivel 3, ISO 14001, ISO 27001, ISO 9001, PCI "Payment Card Industry Data Security Standard" y SSAE16 Type II. Carriers acometidos: Axtel-Alestra, redIT, Marcatel, Level 3, MetroRed, GTT-Perseus Telecom, Totalpay, Cogent, Verizon, Telefónica, Transtelco, Telmex, Bestel y MCM.

KIO MEX 5. Localizado en Tultitlán, Estado de México. Cuenta con 8 Salas de Cómputo de 1,000 m² cada una. Certificaciones: Uptime Institute Tier III Facility, ICREA Nivel 4, CEEDA Gold, LEED Silver, ISO 14001, ISO 27001, ISO 9001, PCI "Payment Card Industry Data Security Standard" y SSAE16 Type II. Carriers acometidos: redIT, Metrored, Bestel, Telefónica, Telmex, Axtel-Alestra, Marcatel, MCM y Totalplay.

KIO QRO 1. Localizado en Parque Industrial El Marqués, Querétaro. Cuenta con 5 Salas de Cómputo de 1,000 m2 cada una. Certificaciones: Uptime Institute Diseño TIER IV, ICREA Nivel 5, CEEDA Gold, ISO 14001, ISO 27001, ISO 9001, PCI "Payment Card Industry Data Security Standard". SSAE16 Type II. Carrier acometidos: redIT, Metrored, Bestel, MCM, Axtel-Alestra, Marcatel, Telmex, Totalplay, Transtelco y MetroCarrier.

Además, KIO cuenta con la red de interconexión de alta velocidad entre sus Centros de Datos.

Mapa 9: KIO Networks



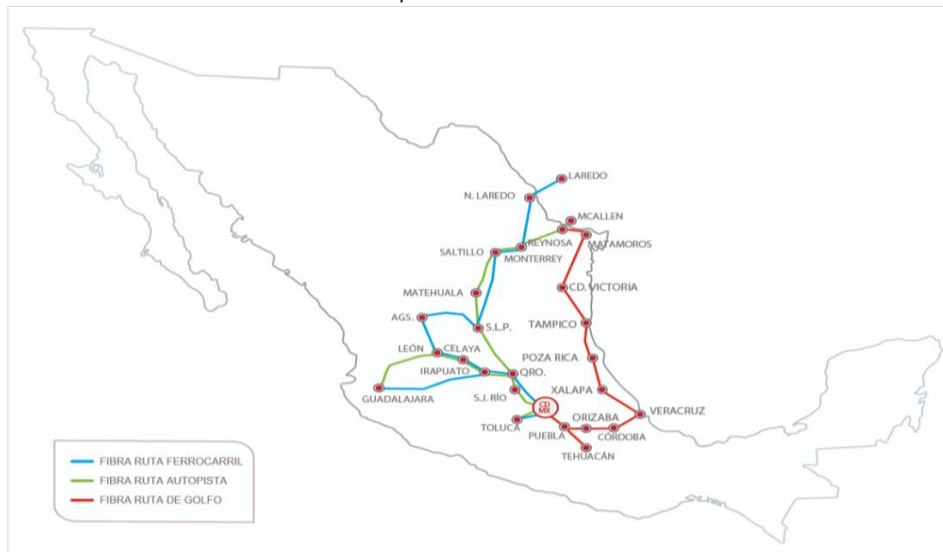
Fuente: (<https://www.kionetworks.com/data-center-services>)

Marcatel

Descripción: Empresa mexicana de telecomunicaciones con más de 50 años de experiencia y operando actualmente en más de 100 países.

Servicios: Por medio de su red de fibra óptica global incorpora servicios de datos y de nube, Internet y transporte de alta calidad, cuenta con las certificaciones CE 2.0 MEF, ESR, e ISO 9001:2015.

Mapa 11: Maxcom



Fuente: (<https://www.maxcom.com.mx/>)

MCM

Ubicación: Ciudad de México, Monterrey y Guadalajara.

Descripción: Empresa parte de Megacable que cuenta con un Data Center con las siguientes características: alta seguridad, monitoreo 24/7, control de temperatura y humedad, respaldo de energía, detección y control de incendios FM200, video vigilancia, gabinetes y racks de servidores anti-estática, certificado ISO/IEC 27001:2013 en seguridad de la información y servicios de co-ubicación.

Servicios: Enlaces directos a la nube.

MetroCarrier

Descripción: Busca mejorar y optimizar los procesos de comunicación en México para Empresas, Corporativos, Operadores de Telecomunicaciones, Sector Público y Hospitalidad. Con una trayectoria de 18 años, brinda conectividad a través de red nacional de fibra óptica para operadores en 77,000 kilómetros distribuidos en más de 25 estados.

Servicios: En sus servicios incluye soluciones en cloud, Internet, e infraestructura.

Mapa 12: Metrocarrier



Fuente: (<https://www.metrocarrier.com.mx/acerca-de-metrocarrier/>)

Microsoft (Azure)

Descripción: Azure, de Microsoft, es un conjunto completo y en expansión constante de servicios de informática en la nube que ayudan a las organizaciones a afrontar sus desafíos empresariales. Azure ofrece la flexibilidad de crear, administrar e implementar aplicaciones en una red mundial con las herramientas y las plataformas que se prefieran.

Servicios: En general, servicios de informática en la nube.

Mapa 13: Microsoft Azure
58 regiones, más que cualquier otro proveedor en la nube



Fuente: (<https://azure.microsoft.com/es-mx/>)

Neutrona Networks International

Descripción: Operador independiente que proporciona soluciones en telecomunicaciones en América.

Cobertura: Presencia en Latinoamérica, con presencia directa en México.

Servicios: Ofrece servicios de MPLS, VPN, enlaces punto a punto, seguridad, y conexión a la nube.

Oracle

Descripción: Compañía multinacional especializada en el desarrollo de soluciones de nube, bases de datos, aplicaciones comerciales y desarrollo de aplicaciones.

Servicios: Oracle Cloud Infrastructure es una nube de IaaS y PaaS de segunda generación que permite a las empresas ejecutar de manera segura cargas de trabajo de aplicaciones y bases de datos de misión crítica, así como acuerdos de nivel de servicio (SLA) sólidos en Cloud Computing.

Rackspace

Descripción: Empresa de Estados Unidos que proporciona servicios de computación en la nube, aplicaciones, datos, seguridad y de infraestructura.

Servicios: Soluciones en nube pública, privada e híbrida, así como de nubes múltiples.

Red Hat

Descripción: Empresa estadounidense que en la actualidad cuenta con una amplia cartera de productos, que incluye infraestructuras de nube híbrida, middleware, integración ágil, desarrollo de aplicaciones nativas de la nube y soluciones de automatización y gestión. Permite a las empresas adaptarse a los negocios cada vez más digitales e interconectados.

Servicios: Diseño e implementación de nube privada, uso de la nube pública, herramientas de gestión y automatización, almacenamiento, middleware y una plataforma de contenedores que empaqueta y traslada las aplicaciones de una nube a otra.

SAP

Descripción: Compañía con sede en Estados Unidos y con presencia en todos los países de la región, que se desempeña en el mercado de software para aplicaciones empresariales, soluciones analíticas, soluciones cloud y bases de datos.

Servicios: Permite simplificar la integración empresarial para entornos híbridos y heterogéneos, optimización de procesos de negocios en la nube, gestión y procesamiento de datos.

SDRMéxico (Servicios Disaster Recovery)

Ubicación: Metepec, Estado de México.

Descripción: Empresa mexicana al servicio de más del 65% de los bancos instalados en México, incluyendo a otras empresas de servicios, comunicaciones y diversas industrias en todo el continente.

Servicios: Cuenta con 6,000 m² de infraestructura diseñada y construida para ser un sitio alternativo, permitiendo el alojamiento de equipo de cómputo y comunicaciones en su Data Center, el cual cumple con los más altos estándares mundiales (ICREA Nivel III). Así mismo ofrece servicios administrados en su nube privada.

Serveris

Descripción: Data Center que se encuentra en la ciudad de Monterrey.

Características:

- 1,000 m² totales
- Disponibilidad 99.982%
- Categoría Tier III
- Enlaces dedicados de alta velocidad, por lo que puede ofrecer una conexión de hasta 10 Gigabits para sus sistemas.

Servicios: En este Data Center se brindan servicios de Infraestructura de Cómputo como Servicio (IaaS) con el cual brindan servidores dedicados, nubes privadas, servicios para almacenar y respaldo de datos.

Servnet Telecomunicaciones

Descripción: Empresa de telecomunicaciones que cuenta con un Data Center que cumple con estándares de calidad en sus instalaciones, tales como energía ininterrumpible y regulada, sistemas contra incendios, centro de operación de red, pisos y techos falsos, aire acondicionado de precisión, planta de energía y sistemas

de cableado de primer orden, así como un equipo humano integrado por ingenieros expertos en sistemas operativos, servicios web, correo y bases de datos.

Cobertura: Tiene presencia en 14 ciudades del país.

Servicios: Ofrece servicios de voz, datos, cloud y de Data Center.

Telefónica

Descripción: Compañía española de telecomunicaciones integradas y operador de telefonía móvil. Posee y opera su propia red IP global en más de 40 mercados. Y a través de acuerdos con partners también ofrece servicio en más de 170 países.

Servicios: Proporciona soluciones escalables en la nube, permite desplegar y publicar aplicaciones en Internet de manera automática y escalable mediante servidores cloud, así como un centro de datos virtual.

Telmex (Triara)

Descripción: Triara es la red de Data Center de Telmex. Brinda Soluciones Integrales de Nube, proporcionando infraestructura, conectividad, almacenamiento, servicios administrados de TI y gestión de sus aplicaciones.

Tanto Triara Monterrey como Triara Querétaro fueron galardonados con la certificación ICREA Nivel 5 HSHA-WCQA (High Security High Available World Class Quality Assurance).

Cobertura: Presencia en México (Querétaro y Monterrey), Colombia (Bogotá) y Brasil (Río de Janeiro).

Servicios: Ofrece servicios de SaaS, PaaS, IaaS y Big Data. 38.4 Km² repartidos en 44 Data Center.

Total Play

Descripción: Data Center ubicado en territorio nacional. Cuenta con certificaciones que garantizan calidad en procesos, seguridad y máxima protección de la información y continuidad del negocio.

Servicios: Capacidad de crecimiento inmediato, máxima disponibilidad, alto desempeño, escritorios virtuales, co-ubicación, DRP (Plan de Recuperación ante Desastres), consultores especializados. Certificación ICREA Nivel IV: Disponibilidad 99.99%, ISO 22301: Gestión de la continuidad de negocio, ISO 27001: Procesos y seguridad de la información, ISO 20000: Sistema de Gestión de Servicios de TI.

TOTVS

Descripción: TOTVS es una empresa brasileña de tecnología, especialista en el desarrollo de soluciones de negocios para reproductores de todos los tamaños. Sus herramientas son integradas y atienden tanto el core business como el back office de sus clientes, alcanzando toda su cadena de valor, de extremo a extremo. Se destaca la oferta de servicios en la nube.

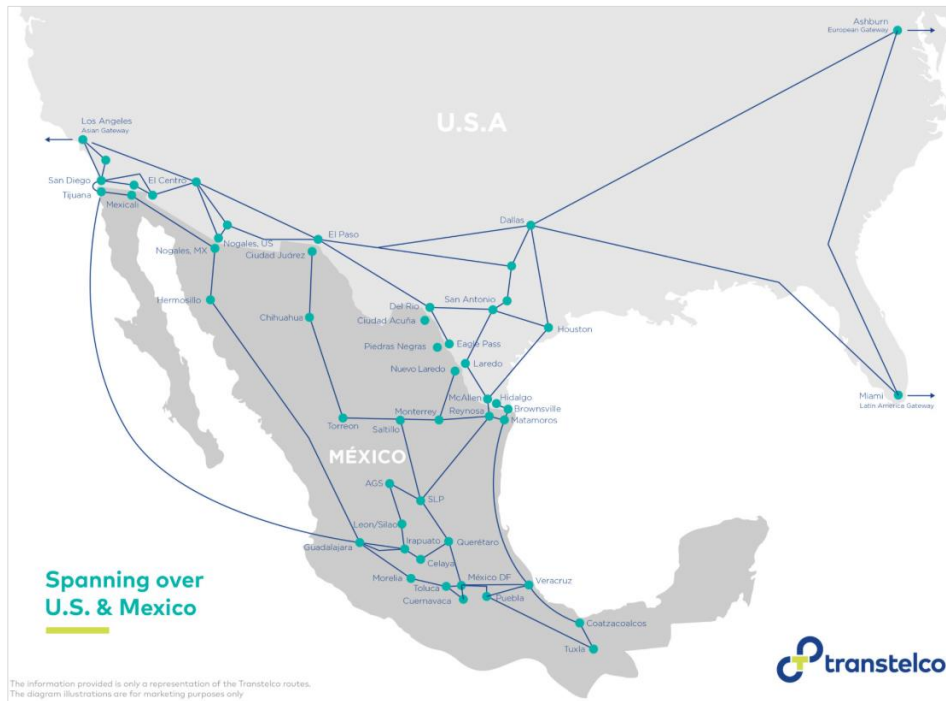
Servicios: Ofrece su servicio TOTVS Cloud, el cual es una solución de software de gestión integral, para administrar una empresa a través de la web, delegando en TOTVS el mantenimiento y soporte de la solución, lo anterior con flexibilidad, seguridad, accesibilidad, monitorización y atención al cliente.

Transtelco

Descripción: Fundada en 2001, se ha centrado en desarrollar progresivamente su infraestructura de red de alta velocidad y ofertas de servicios diseñadas para satisfacer las necesidades de las compañías.

Servicios: El software TCloud Connect de Transtelco permite a las organizaciones acceder y administrar de manera transparente los principales datos de Cloud y ofertas SaaS en México. Las plataformas soportadas actualmente incluyen: Amazon AWS (Direct Connect), Microsoft Azure (ruta exprés) y Google Cloud. También ofrecen servicios de IaaS en los DC ubicados en Dallas y El Paso, Texas.

Mapa 15: Transtelco



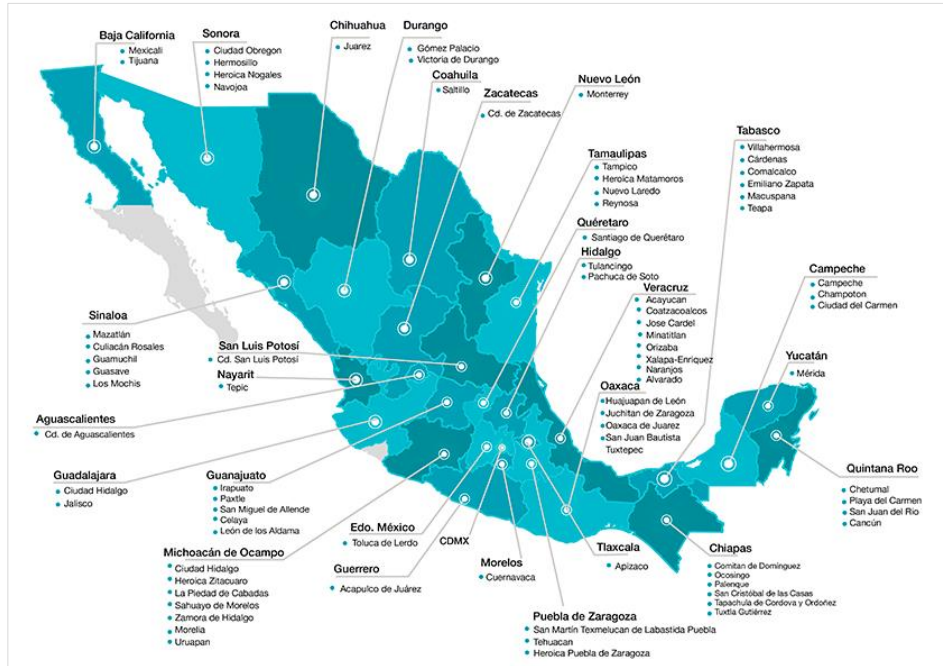
Fuente: (<https://transtelco.net/company/#network-maps>)

UC Telecom

Descripción: Operador mexicano de telecomunicaciones que brinda soluciones IP empresariales de voz y datos administradas desde la nube.

Servicios: Ofrece soluciones de conectividad inteligente a la nube a través de su red de fibra óptica, así como Internet dedicado de alta capacidad.

Mapa 16: UC Telecom



Fuente: (<http://www.uc-tel.com/>)

Verizon

Descripción: Operador de telefonía móvil de Estados Unidos.

Servicios: En general, servicios en la nube y red privada con interconexión de nube segura.

ANEXO 2: Clasificaciones para Data Centers

Uptime Institute⁶⁶; fue el primero en definir las clasificaciones como TIER I a IV a mediados de los 90. Uptime Institute certifica centros de datos líderes en todo el mundo para el diseño, la construcción, la administración y las operaciones en función de los estándares TIER.

Como podemos apreciar, la clasificación de niveles es importante para los planificadores de Data Centers en preparación para el presupuesto de construcción. Uptime Institute definió originalmente los cuatro niveles de infraestructura del centro de datos en el documento técnico titulado Clasificaciones de nivel Definir el rendimiento de la infraestructura del sitio (Chen, Gao, & Chen, 2016). La Norma de infraestructura de telecomunicaciones para centros de datos (ANSI/TIA, 2012), también adopta las definiciones de este libro blanco (Turner IV, 2006). En general, el nivel de un controlador de dominio depende del sistema más débil. Por ejemplo, si el sistema de energía tiene una clasificación de nivel 3, mientras que el sistema de refrigeración tiene una calificación de nivel 2, el controlador de potencia tiene una calificación de nivel 2. En la siguiente Tabla, se resumen los requisitos de nivel y los atributos comunes.

Tabla 19: Requerimientos TIER y atributos comunes.

	Tier I	Tier II	Tier III	Tier IV
Tipo de construcción	Inquilino	Inquilino	Solo	Solo
Turnos de personal Personal/Turno	Ninguno Ninguno	1 turno 1/Turno	Más de 1 turno 1-2/turno	24 por siempre 2+/Turno
Utilizable para carga crítica	100% N	100% N	90% N	90% N
Límite inicial de KW por gabinete (típico)	<1kW	1-2 kW	1-2 kW	1-3 kW
Último kW por gabinete (típico)	<1 kW	1-2 kW	>3 kW ^{2,3}	>4 kW ^{1,2}
Espacio de apoyo elevado- Proporción de piso	20%	30%	80-90+%	100+%
Altura del piso elevado (típico)	12"	18"	30-36"	30-42"
Carga de suelo lbs/ft (típico)	85	100	150	150+
Voltaje de utilidad (típica)	208, 480	280, 480	12 - 15 kV	12 - 15 kV
Puntos únicos de falla	Muchos + errores humanos	Muchos + errores humanos	Algunos + errores humanos	Fuego, EPO + un error humano

⁶⁶ Uptime Institute, (2019). Certificación Tier. Recuperado de <https://es.uptimeinstitute.com/tier-certification>

Desconexiones Representativas de Mantenimiento Planificado	2 eventos anuales de 12 Horas c/u	3 eventos durante 2 años de 12 horas c/u	Ninguno requerido	Ninguno requerido
Fallas de sitios representativos	6 fallos en 5 años	1 falla cada año	1 fracaso cada 2,5 años	1 fracaso cada 5 años
Período de inactividad anual del usuario final causado por el sitio (basado en datos de campo)	28.8 horas	22.0 horas	1.6 horas	0.8 horas
Disponibilidad del usuario final resultante según el tiempo de inactividad causado por el sitio	99.67%	99.75%	99.98%	99.99%
Meses típicos para planificar y construir	3	3-6	15-20	15-20
Año del primer despliegue	1965	1970	1985	1995
Costo de construcción	220	220	220	220
Salida de UPS utilizable en piso elevado	\$10,000/kW	\$11,000/kW	\$20,000/kW	\$22,000/kW

Notas:

¹100 W / ft² máximo para enfriamiento por aire en áreas extensas, agua o métodos alternativos de enfriamiento superiores a 100 W / ft² (sin costo adicional).

² Las densidades W / ft² mayores requieren mayor espacio de soporte (100% a 100 W / ft² y hasta 2 o más veces a densidades mayores), pisos elevados más altos y, si se requiere en áreas grandes, entrada de servicio de voltaje medio.

³ Excluye tierra; Requisitos arquitectónicos únicos, permisos y otras tarifas; interesar; y costos civiles anormales. Estos pueden ser de varios millones de dólares. Asume un mínimo de 15,000 pies cuadrados de piso elevado, arquitectónicamente sencillo, edificio de una sola planta, con una red troncal eléctrica dimensionada para lograr la máxima capacidad con la instalación de componentes o sistemas adicionales. Realice ajustes para NYC, Chicago y otras áreas de alto costo.

⁴ Los costos se basan en datos de 2005. Los costos del año futuro deben ajustarse utilizando los índices ENR.

⁵ Consulte el Libro Blanco del Instituto titulado Dólares por kW más dólares por pie cuadrado. Un modelo de costo de centro de datos mejor que los dólares por pie cuadrado solo para obtener información adicional sobre este modelo de costo.

BICSI⁶⁷: es la autoridad líder en programas de diseño de centros de datos, diseñada para educar a los profesionales de las TIC en los métodos adecuados para diseñar los centros de datos más grandes y más rápidos del mundo. Establece clasificaciones de los centros de datos a modo de "Clase" con niveles que van desde el 0 (carente de alternativas eléctricas, UPS o puesta a tierra) al 4 y que sirven de base para otorgar certificaciones.

⁶⁷ Bicsi, (2019). Data Center Design Consultant (DCDC). Recuperado de <https://www.bicsi.org/education-certification/certification/dcdc>

UL (Underwriter Laboratories): es una consultoría de seguridad y certificación de la empresa con sede en Illinois de E.U.A., la cual está preparando su estándar UL3223 como base para un programa de certificación de Data Center.

Certificaciones TSI/EN 50600 (Trusted Site Infrastructure): es un método para la evaluación y certificación de seguridad física y disponibilidad de centros de datos establecidos desde 2002. Con un esquema privado de la entidad de certificación TÜVIT (TÜV NORD GROUP) que cubre, entre otras adicionales, las especificaciones de la norma europea EN 50600 en relación a criterios de disponibilidad considerando aspectos que van más allá de la capa específica de las instalaciones y en términos de “Niveles” de 1 a 4.

The Open Standard for Data Center Availability (OSDA): es un sistema de clasificación y categorías de disponibilidad que tiene como objetivo promover la innovación en eficiencia energética y diseños sostenibles de centros de datos.

Open-IX Association (OIX)⁶⁸: es una comunidad autorregulada que se forma para fomentar el desarrollo de centros de datos críticos e IXP. Dispone de dos certificaciones para Data Centers orientados a la actividad como IXP - Internet Exchange Provider - en base los estándares OIX-1 y OIX-2 respectivamente y enfocadas al rendimiento, la resistencia y la fiabilidad de la infraestructura de red que intervienen en los servicios de interconexión masivos.

ICREA⁶⁹: “International Computer Room Experts Association” es una asociación Internacional sin fines de lucro formada por ingenieros especializados en el diseño, construcción, operación, administración, mantenimiento, adquisición, instalación y auditoría de centros de cómputo. Fue fundada en 1999 en la Ciudad de México, tiene presencia internacional en 23 países, actualmente establece la norma ICREA-Std-131-2017 que marca niveles (I a V).

La norma ICREA Std-131-2017⁷⁰ establece los niveles en función de la disponibilidad esperada:

NIVEL I: Sala de cómputo en ambiente Certificado QADC (Quality Assurance Data Center). Esta topología aporta un 95% de disponibilidad y es una configuración básica.

⁶⁸ OPEN-IX, (2019). Building Community and Consensus to Foster Data Center and Interconnection Standards. Recuperado de <https://www.open-ix.org/cpages/home>

⁶⁹ ICREA, (2019). Acerca de ICREA. Recuperado de <https://icrea-international.org/somos/>

⁷⁰ ICREA, (2017). NOTA TÉCNICA. Criterios generales de Certificación ICREA 2017. Recuperado de https://icrea-international.org/wp-content/uploads/2019/08/Criterios_Generales_de_Certificacion.pdf

- NIVEL II: Sala de cómputo en ambiente Certificado de clase mundial WCQA (World Class Quality Assurance). Esta topología aporta un 99% de disponibilidad y es una configuración con redundancia básica.
- NIVEL III: Sala de cómputo confiable con Ambiente Certificado de clase mundial S-WCQA (Safety World Class Quality Assurance). Esta topología aporta un 99.9% de disponibilidad y es una configuración con redundancia que permite darle mantenimiento sin suspender la operación.
- NIVEL IV: Sala de cómputo de alta seguridad con certificación HS-WCQA (High Security World Class Quality Assurance). Esta topología aporta un 99.99% de disponibilidad y es una configuración con redundancia, que permite darle mantenimiento con elementos propios y fijos sin suspender la operación y tolerante a fallos.
- NIVEL V: Sala de cómputo de alta seguridad y alta disponibilidad con certificación de clase mundial HSHA-WCQA (High Security, High Available World Class Quality Assurance). Esta topología aporta un 99.999% de disponibilidad y es una configuración con redundancia sin puntos únicos de falla (PUF), que permite darle mantenimiento con elementos propios y fijos sin suspender la operación, tolerante a fallas.

ANEXO 3: Indicadores de Eficiencia

Indicadores de Eficiencia

Para poder comprender adecuadamente los problemas y brindar mejoras, es esencial que se acuerde un conjunto inicial de mediciones que realicen recopilación de datos a gran escala. En este sentido, las métricas son empleadas para que los centros de datos puedan medir el rendimiento del equipamiento de Tecnología de la Información y puedan mejorar las áreas problemáticas.

PUE (Power Usage Effectiveness)⁷¹: mide el valor de la eficiencia eléctrica en relación al consumo eléctrico total, con la finalidad de identificar qué tan eficiente es el consumo actual de los equipos.

Mientras menor sea el valor PUE, mejor será el aprovechamiento eléctrico, lo que se traduce en menores costos y menores emisiones de CO_2 .

$$PUE: \frac{\text{Total de energía consumida por la instalación}}{\text{Energía consumida por el equipamiento TI}}$$

La “energía consumida por el equipamiento TI” incluye toda la energía asociada a todos los equipos de TI (por ejemplo, equipos de computación, almacenamiento y red) junto con el equipo adicional (por ejemplo, conmutadores, monitores y estaciones de trabajo, así como computadoras portátiles utilizadas para monitorear o controlar el Data Center).

La “energía total consumida por la instalación” incluye toda la energía de los equipos de TI, más todo lo que sirve como sostén de los equipos de TI que usan energía (por ejemplo, componentes de suministro de energía, UPS, baterías y componentes del sistema de enfriamiento, bombas, torres de enfriamiento).

Cuando el PUE es perfecto (igual a 1), significa que cada kilowatt es usado en la energía de los servidores, no existe energía desperdiciada. Por lo que una relación 1.5/1 significa que existe un 0.5 que se está desperdiciando en la refrigeración del Data Center.

⁷¹ Avelar, V. Azevedo, D. French, A. The Green Grid, (2012). White Paper #49. PUE: A Comprehensive Examination of The Metric. Recuperado de https://datacenters.lbl.gov/sites/default/files/WP49-PUE%20A%20Comprehensive%20Examination%20of%20the%20Metric_v6.pdf

DCiE (Data Center Infrastructure Efficiency)⁷²: Es la potencia del equipamiento TI, multiplicada por 100 y dividida por la potencia total de la instalación. Cuanto mayor sea el número, mejor. El DCiE de un Data Center no debería ser mayor de 1.

$$DCiE = \frac{\text{Potencia del equipamiento TI}}{\text{Potencia total de la instalación}}$$

La “potencia total de la instalación” se define como la potencia medida en el medidor de luz, es la potencia dedicada exclusivamente al Data Center.

La potencia del equipamiento TI se define como la potencia consumida por el equipo que se utiliza para administrar, procesar, almacenar o enrutar datos dentro del espacio de cómputo.

WUE (Water Usage Effectiveness)⁷³: se utiliza para evaluar la eficiencia del consumo de agua en los equipos de refrigeración en relación a la cantidad de kW/h.

Existen dos definiciones de la métrica:

WUE: es una métrica basada en el sitio, donde se realiza una evaluación del agua utilizada en el sitio para la operación del Data Center. Esto incluye el agua utilizada para la humidificación y el agua evaporada en el lugar para la producción de energía o el enfriamiento del Data Center y sus sistemas de soporte.

$$WUE = \frac{\text{Uso anual de agua utilizada en el sitio}}{\text{Energía consumida por el equipamiento TI}} \left[\frac{L}{kWh} \right]$$

WUE source: es una métrica basada en la fuente, la cual incluye el agua utilizada en el sitio y el agua utilizada fuera del sitio en la producción de la energía utilizada en el sitio. Por lo general, esto añade el agua utilizada en la fuente de generación de energía al agua utilizada en el sitio.

$$WUE_{source} = \frac{\text{Uso anual de agua utilizada en la fuente de energía} + \text{Uso anual de agua utilizada en el sitio}}{\text{Energía consumida por el equipamiento TI}} \left[\frac{L}{kWh} \right]$$

⁷² Verdun, G. (2008). White Paper #14. THE GREEN GRID METRICS: DATA CENTER INFRASTRUCTURE EFFICIENCY (DCiE) DETAILED ANALYSIS. Recuperado de <https://leonardo-energy.pl/wp-content/uploads/2017/08/the-green-grid-metrics-datacenter-infrastructure-efficiency-DCiE-detailed-analysis.pdf>

⁷³ Patterson, M. (2011). The Green Grid (2011). White Paper #35. WATER USAGE EFFECTIVENESS (WUE™): A GREEN GRID DATA CENTER SUSTAINABILITY METRIC. Recuperado de <https://airatwork.com/wp-content/uploads/The-Green-Grid-White-Paper-35-WUE-Usage-Guidelines.pdf>

CUE (Carbon Usage Effectiveness)⁷⁴: Para los centros de datos que obtienen toda su fuente de energía de la red de energía y no generan CO_2 local, CUE se define de la siguiente manera:

$$CUE = \frac{\text{Total de emisiones } CO_2 \text{ causadas por el total de energía consumida en el centro de datos [kgCO}_2\text{eq]}}{\text{Energía consumida por el equipamiento TI}} \left[\frac{\text{kgCO}_2\text{eq}}{\text{kWh}} \right]$$

En esta ecuación, el "Total de energía del Data Center" tiene el mismo valor que el numerador de la métrica PUE. Las unidades de la métrica CUE son kilogramos de dióxido de carbono ($kgCO_2eq$) por kilovatio-hora (kWh).

Un enfoque alternativo para calcular la CUE es multiplicar el factor de emisión de carbono (CEF) por la PUE anual del Data Center:

$$CUE = CEF \times PUE$$

$$CUE = \frac{CO_2 \text{ emitida (kgCO}_2\text{eq)}}{\text{unidad de energía (kWh)}} \times \frac{\text{Total de energía consumida por el centro de datos}}{\text{Energía consumida por el equipamiento TI}}$$

Total de emisiones de CO_2 : Este componente incluye las emisiones de CO_2 de las fuentes de energía locales y de la red de energía. Idealmente, las emisiones de CO_2 se determinarán para la mezcla real de energía entregada en el sitio.

ERE (Energy Reuse Effectiveness)⁷⁵: se utiliza para calcular la eficiencia de la reutilización de la energía, como por ejemplo podría ser la reutilización del calor generado por los equipos para aclimatar las instalaciones.

Por lo tanto, si la energía del Data Center se usa en otra ubicación y reduce la energía que el sitio necesitaría comprar, obtener o generar, puede contabilizarse en la métrica:

$$ERE = \frac{\text{Enfriamiento} + \text{Energía} + \text{Iluminación} + \text{TI} + \text{Reuso}}{\text{TI}}$$

Donde el "enfriamiento" representa la energía utilizada por todo el sistema de enfriamiento atribuible al Data Center, la "energía" es la energía perdida en el sistema de distribución de energía a través de la pérdida de línea y otras ineficiencias de la infraestructura (por ejemplo, UPS o PDU), la "iluminación" representa la energía utilizada para la iluminación, el Data Center y espacios de soporte, la "TI" es la energía utilizada por todos los equipos de TI (servidores, red, almacenamiento) en el Data Center y el "reuso" se refiere a la energía del Data

⁷⁴ Belady, C. (2010). White Paper #32. Carbon Usage Effectiveness (CUE): A Green Grid Data Center Sustainability Metric. Recuperado de <https://airatwork.com/wp-content/uploads/The-Green-Grid-White-Paper-32-CUE-Usage-Guidelines.pdf>

⁷⁵ Patterson, M. (2010). White Paper #29. ERE: A METRIC FOR MEASURING THE BENEFIT OF REUSE ENERGY FROM A DATA CENTER. Recuperado de https://eehpcwg.llnl.gov/documents/infra/06_energyreuseefficiencymetric.pdf

Center usado en otra ubicación, por lo que se reduce la energía que el sitio necesitaría comprar, obtener o generar.

CCF (Cooling Capacity Factor) ⁷⁶: se utiliza para gestionar la eficiencia de refrigeración en el Data Center. Se calcula dividiendo la capacidad total de enfriamiento declarada por el fabricante en funcionamiento (kW) por el 110% de la carga crítica de TI (kW).

$$CCF = \frac{\textit{Capacidad total de enfriamiento en funcionamiento}}{\textit{Salida de UPS} \times 1.1}$$

La capacidad total de enfriamiento en funcionamiento es la suma de las capacidades nominales de las unidades de refrigeración en funcionamiento.

La carga crítica de TI en la sala es igual a la(s) salida(s) de UPS para la sala.

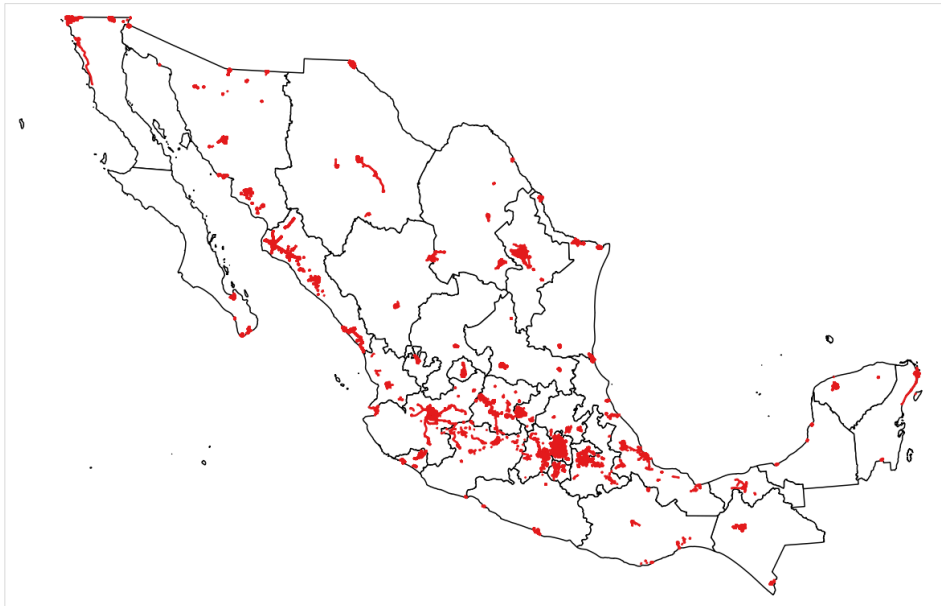
UUR (Utilización Unitaria de Rack): consiste en evaluar el porcentaje de utilización de cada Rack medido en unidades de rack. Una unidad de rack equivale a 4.445 (cm) de alto.

⁷⁶ Brill, K. Strong, L. (2013). WHITE PAPER. Cooling Capacity Factor (CCF) Reveals Stranded Capacity and Data Center Cost Savings. Recuperado de <https://www.anixter.com/content/dam/Suppliers/Upsite/Cooling%20Capacity%20Factor%20White%20Paper.pdf>

ANEXO 4: Redes de Fibra Óptica en México.

Con lo que respecta al tema de infraestructura de Red de Fibra Óptica en México a continuación se muestran trayectorias de despliegues de fibra óptica tanto en la red de acceso como en la red de transporte identificables al cierre de 2017.

Mapa 8 Red de Acceso en México

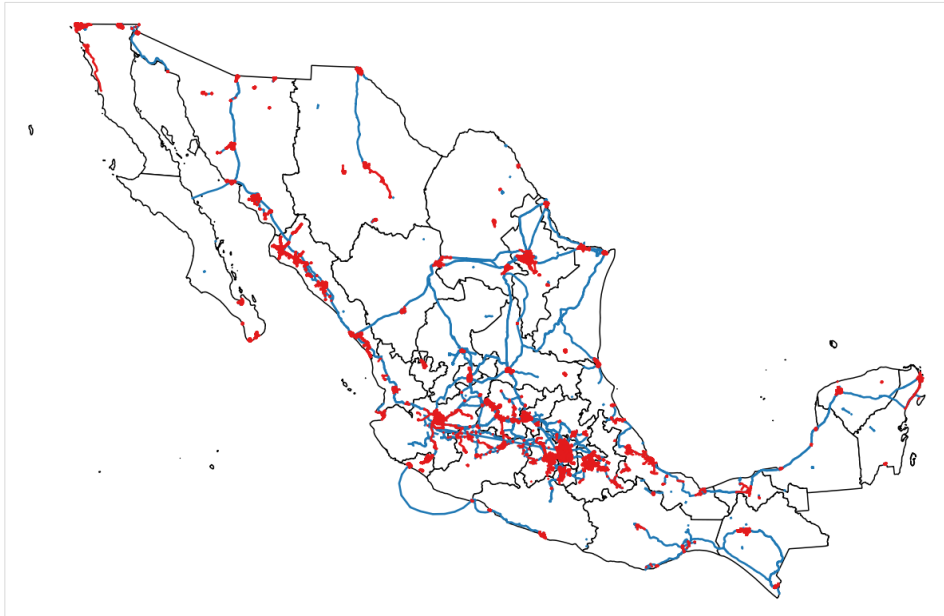


Fuente: Elaboración propia con información de operadores.

Mapa 9: Red de Transporte por Entidad Federativa de México.



Mapa 10: Trayectorias de la Red de fibra óptica de México.



Mapa 11: Red de acceso para la Ciudad de México.

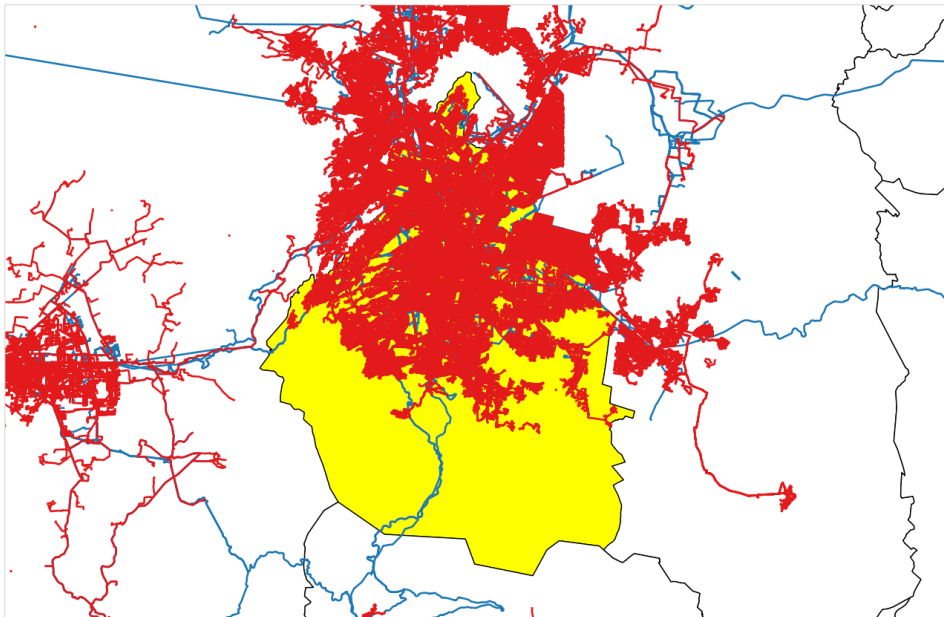
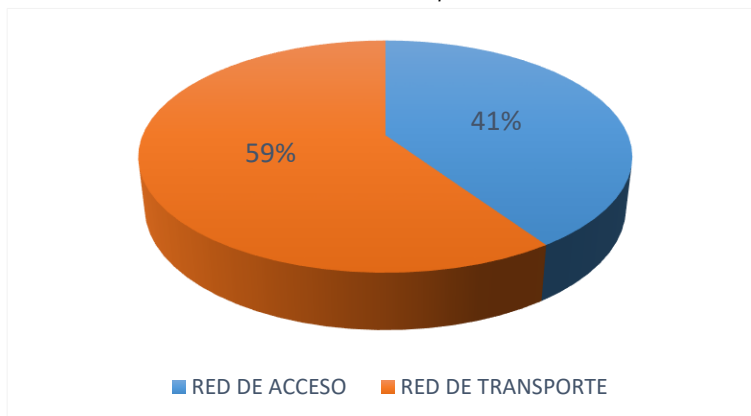


Figura. Red de acceso para la Ciudad de México.

Asimismo, la proporción de kilómetros de fibra óptica desplegada entre la red de acceso y la red de transporte se muestra en la siguiente gráfica.

Gráfica 17: Kilómetros de Fibra Óptica en México al 2017.



Fuente: Elaboración del Instituto con información al cierre de 2017.

Análisis adicional a nivel Entidad Federativa sobre accesos de banda ancha fija y cobertura de fibra óptica se pueden consultar en el Micrositio de Despliegue de Infraestructura (<http://despliegueinfra.ift.org.mx/estudios.php>).

ANEXO 5: Data Centers Internacionales.

El software juega un papel relevante en el manejo de las eficiencias operativas. Compañías como Facebook, Amazon, Google y Microsoft han sido líderes en la definición y evolución del movimiento de *Hyperscale*, sin embargo, no ha sido un fenómeno exclusivo de los Data Centers estadounidenses; es global y es liderado por empresas en todo el mundo. Por ejemplo, en China, los tres gigantes digitales (Baidu, Alibaba y Tencent), están impulsando innovaciones significativas por toda Asia.

Por un lado, Baidu, la herramienta de búsqueda China, controla la mayoría del mercado de búsquedas. Mientras que Alibaba, una herramienta de comercio electrónico, sirve de intermediario entre compradores y vendedores en línea y facilita la venta de productos entre las dos partes a través de su extensa red de sitios web. Y por último Tencent, tiene servicios que van desde una aplicación de redes sociales hasta juegos en línea para múltiples jugadores. Estas empresas se suman a empresas de todo el mundo que están impulsando la innovación a nivel *Hyperscale*.

Por ejemplo, *Emerson Network Power* decía en que en el año 2011 había 509,147 Data Centers en el mundo, el espacio total de Data Centers equivalía a 5,955 campos de fútbol.⁷⁷ Según un estudio de Data Centers realizado por Ovum, existen en el año XXX, más de 10 mil Data Centers en el mundo, que se encuentran distribuidos de la siguiente manera:

Tabla 20: Distribución de Data Centers en el mundo.

Región	Número de Data Centers	Porcentaje por área geográfica.
América del Norte	4,685	44.8%
Europa Occidental	2,779	26.6%
Asia y Oceanía	1,756	16.8%
Europa Oriental	501	4.8%
Medio Oriente y África	377	3.6%
América Latina y el Caribe	363	3.5%

Fuente: OVUM Ltd.

Data Centers (DC) de producción representativos.

Las grandes empresas de TI construyeron varios centros de producción para respaldar sus negocios. Otros se alquilan para brindar servicios a medianas y pequeñas empresas que no pueden pagar sus propios centros de distribución.

⁷⁷ Emerson Network Power (<https://www.datacenterknowledge.com/archives/2011/12/14/how-many-data-centers-emerson-says-500000>)

Google posee 36 DC de producción a nivel mundial, 19 de los cuales están en América, 12 en Europa, 3 en Asia, 1 en Rusia y 1 en Sudamérica, como se muestra en la siguiente Figura. Estos DC son compatibles con los servicios de Google, como búsqueda, correo y mapas.

Mapa 12: Data Centers Google



Fuente: (<https://www.google.com/about/datacenters/locations/>)

El prototipo del primer DC de Google, BackRub se ubicó en el dormitorio de Larry Page (uno de los fundadores de Google). Aunque era simple, BackRub había cumplido con los requisitos básicos de búsqueda de Google en ese momento.

A Google le costó casi \$ 600 millones de dólares construir el primer DC en 2006, por ejemplo, el Data Center de Portland Dalles. Es un par de DC de 94,000 pies cuadrados que se encuentran cerca de la ciudad de Portland en el Estado de Oregón en Estados Unidos. Google anunció otros \$600 millones para construir un nuevo CD con 49,987.2 metros cuadrados en Dalles en 2013, y lo abrió en 2015. Otro DC de Google es el Data Center del condado de Douglas, en el estado de Georgia, éste proporciona servicios para los negocios clave como búsqueda, correo y mapas.

El Data Center de Hamina, en Finlandia, se reconstruye a partir de una fábrica de papel, que utiliza agua de mar a lo largo de las tuberías de la fábrica de papel para controlar la temperatura del Data Center.

A Google le costó más de 2 años y 250 millones de euros construir el Data Center de Saint-Ghislain en Bélgica, que se inauguró en 2010. Es el primer Google DC en todo el mundo que opera completamente sin refrigeración. En su lugar, utiliza un avanzado sistema de enfriamiento por evaporación, que extrae aguas grises (aguas residuales relativamente limpias) de un canal industrial cercano. Google invirtió en 2013, \$300 millones de dólares para mejorar las instalaciones para satisfacer la creciente demanda de servicios en línea. El Data Center del condado de Oklahoma

Mayers, tiene dos edificios de CC con una inversión de más de \$700 millones de dólares, donde las unidades de refrigeración modulares controlan la temperatura.

Microsoft también posee Data Centers de producción en América, Europa y Asia. Construyó el Washington Quincy Data Center con un área de 75 acres en 2007. Quincy Modular Data Center empezó operaciones en 2011, que cubre 28,353 metros cuadrados y utiliza tecnologías verdes. A fines de 2013, Microsoft aprobó un presupuesto corporativo de \$11 millones de dólares para comprar un terreno de 200 acres donde construyó un centro de distribución a gran escala a principios de 2015.

Microsoft estableció el Data Center de San Antonio en 2008, que ocupa aproximadamente medio millón de pies cuadrados y costó \$550 millones de dólares. Cuesta 8 millones de galones de agua reciclada cada mes como parte del sistema de enfriamiento. El Illinois Chicago Data Center fue uno de los centros de distribución de CD más grandes del mundo, con más de 213,360 metros cuadrados y costo aproximadamente \$500 millones de dólares. 56 contenedores de envío de 12 metros (cada uno contiene 1800–2500 servidores) están ubicados en el primer piso, y el número crecerá con demandas adicionales. En el segundo piso, los servidores se ubican en cuatro habitaciones de pisos elevados tradicionales (cada una con 3,657 metros cuadrados). El agua de enfriamiento a lo largo de unas 11.26 kilómetros de tuberías mantiene el CD en una temperatura baja.

El Data Center de Dublín, es el mayor centro de distribución de ultramar de Microsoft. Cubre 92,354.4 metros cuadrados y se enfría por medio de viento natural para ahorrar energía.

Otras grandes empresas de TI también poseen CD de producción. IBM, por ejemplo, siempre se ha dedicado a construir DC más inteligentes. IBM administra más de 430 DC en todo el mundo, con un tamaño total de hasta 8 millones de pies cuadrados. En Canadá, IBM construyó un DC en colaboración con el gobierno y las universidades, como *Barrie Cloud Data Center*, que cubre hasta 30,480 metros cuadrados y mejora significativamente la eficacia del uso de energía (PUE) por tecnologías innovadoras. En 2014, IBM anunció un compromiso de \$1.200 millones de dólares para construir 15 nuevos centros de distribución en 15 países en los cinco continentes, excepto en África y la Antártida.

AMAZON

Empresa valuada en 880 millones de dólares, dirigida por Jeff Bezos lideró el consumo durante el tercer trimestre de 2019⁷⁸.

Como hemos señalado, Amazon es una empresa exitosa que como todas aquellas que se fundan en el amparo de las tecnologías disruptivas, no tiene activos dentro del ramo que representa. Sus ingresos se incrementaron 24%, a pesar de hacer

⁷⁸ <http://www.globaltimes.cn/content/1168204.shtml>

fuertes inversiones en envíos rápidos, en publicidad y nuevos mercados como el de la Salud, además de reportar una ligera caída del 26% en sus acciones (ver gráfica del comportamiento accionario).

Gráfica 18: Comportamiento accionario Amazon Oct 2018 - Oct 2019.



Fuente: TradingView.

Para los inversionistas de Amazon lo más preocupante está en la desaceleración del crecimiento en los servicios que ofrece como AWS, El crecimiento de los ingresos en AWS se desaceleró del 46% el año pasado al 35% en el último trimestre. Si bien el crecimiento de AWS es más rápido que el aumento de las ventas corporativas del 24% de Amazon durante ese período. Los servicios en la nube se están convirtiendo en un impulsor clave de ganancias en Amazon.

AMAZON en números			
	Q3 2019 (Est.)	Q3 2018	Q3 2017
Ganancias por acción (en dólares)	4.57	5,75	0,34
Ingresos (en miles de millones de dólares)	68,7	56,6	43,7
Ingresos de AWS (en miles de millones)	N / A	4.6	3.2

Fuentes: YCharts, Amazon Investor Relations

Sin embargo, y a pesar de que Amazon no actúa como un monopolio tradicional, arrasando contra los consumidores y maximizando sus ganancias, más bien mantiene los precios bajos, subvenciona las entregas y se expande.

Pero el creciente poder de la firma, y especialmente su capacidad de utilizar el dominio digital para superar a sus rivales, hace que los reguladores europeos y los legisladores estadounidenses se sientan cada vez más incómodos. Por ejemplo, la publicidad en Internet. Los artículos que no están en la primera página podrían perfectamente ni existir para muchos compradores, lo cual obliga a los vendedores a pagar para que se destaque su producto. Amazon no detalla lo que ingresa por publicidad, pero representa la mayor parte de la línea "otros" de ventas de la empresa, que creció más del 40%.



En 2018, Amazon recibió \$ 258.22 mil millones en ventas, la mitad del mercado en línea y el 5% de las ventas minoristas totales de EE. UU. Parte del crecimiento se produjo orgánicamente, y algunos de las principales adquisiciones de grandes minoristas en línea, como Zappos, sitios de revisión, como Goodreads, y el minorista en persona, Whole Foods. Aunque sus números de ventas todavía están detrás de Walmart, \$ 331.67 mil millones (incluye ventas de ladrillo y mortero). Al ritmo que está creciendo, Amazon y el comercio minorista en línea pueden algún día superar las tiendas físicas. En solo 25 años, Amazon se convirtió en uno de los minoristas más grandes del mundo.

Pero no sólo Amazon se encuentra en la lista de Congreso norteamericano para rendir cuentas, Facebook ya compareció dos veces ante el Congreso y una, Google y Twitter.

Habrà que esperar cual será el argumento del congreso sobre la responsabilidad a la que tendrá que enfrentar Amazon, mientras se puede especular que no son exclusivamente temas que tienen que ver con los 1650 millones de dólares invertidos

en 2018 para la compra, adquisición o fusiones de startups digitales⁷⁹. Amazon hizo más que facilitar y agilizar la compra de cosas; alteró la naturaleza del comercio y con ello la economía de los Estados Unidos. Cambió cómo y dónde trabaja la gente e, igualmente importante, cambió cómo se establecen los precios y con qué frecuencia cambian.

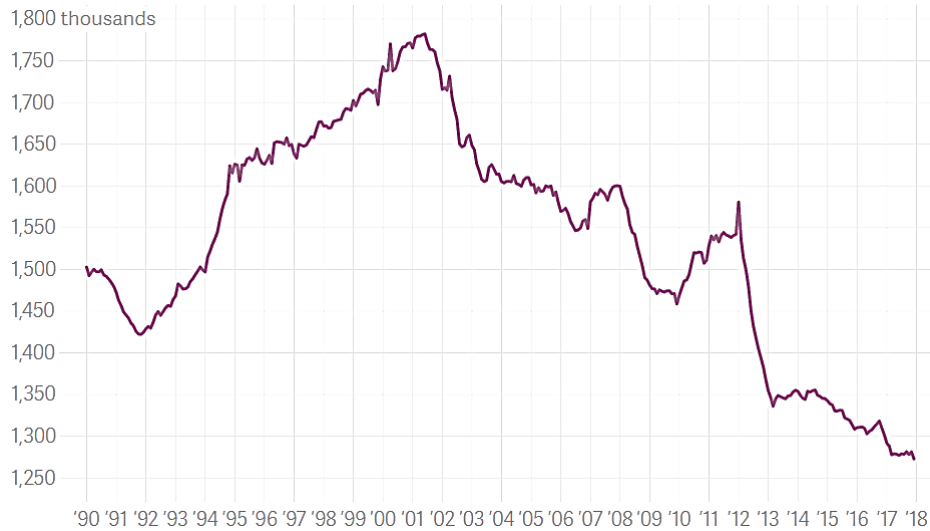
Antes de comprar en línea, los comerciantes solo podían aumentar los precios una o dos veces al año en respuesta a la inflación o la competencia. Ahora, para mantenerse competitivos, los minoristas en línea usan algoritmos, que cambian los precios con mayor frecuencia para satisfacer la demanda. Algunos minoristas en línea, como Walmart, obtienen datos de sus competidores para mantenerse actualizados sobre el último precio del mercado. Se estima que los cambios de precios han llegado a ser casi dos veces más frecuente a medida que las transacciones en línea se hicieron más comunes. También se estima que las rebajas se han vuelto más comunes. Se atribuye al incremento de la competencia como la razón por la cual el precio de los productos vendidos en Amazon cambia con más frecuencia.

Antes, los productos podían ser más baratos en las zonas más pobres o más caros en ciertos países. Ahora, los precios tienden a ser los mismos en todas partes, tanto en línea como en tiendas.

Principalmente, esto ha significado el acceso a bienes más baratos, lo que beneficia a todos los consumidores, especialmente a los estadounidenses de bajos ingresos. Pero los precios más bajos en todas las geografías también exprimen a los minoristas en áreas de alto costo cuando no pueden cobrar mucho más por sus productos, pero aun así deben pagar alquileres y salarios más altos. La presión sobre los precios y la competencia del comercio electrónico son grandes razones por las cuales muchos escaparates están vacíos y los minoristas como la cadena de librerías Borders ya no existen. Incluso a los grandes minoristas con múltiples cadenas les resulta difícil sobrevivir en un entorno de precios más feroz. La siguiente figura es el número de empleados minoristas en los grandes almacenes estadounidenses. Después de 2000, los trabajos comenzaron a disminuir, incluso cuando la economía estaba en auge.

⁷⁹ Las adquisiciones más ambiciosas de Amazon no llegaron a los 1,000 millones de dólares en valoración. Al menos eso ocurrió en las adquisiciones del fabricante de timbres inteligentes Ring (comprada en efectivo por 839 millones) y la compañía farmacéutica online PillPack (trato cerrado en 753 millones). (<https://www.businessinsider.es/amazon-gasto-1650-millones-dolares-adquisiciones-durante-2018-370573>)

Gráfica 19: Número de empleados en tiendas departamentales en EU



Fuente: (<https://theatlant.com/charts/wB-U18pZg>)

Esto se ha convertido en una fuente de preocupación para los responsables políticos. El secretario del Tesoro, Steven Mnuchin, dijo que Amazon "destruyó el comercio minorista" y quiere investigar su ruptura. Amazon no es técnicamente un monopolio; sigue siendo solo el 5% de las ventas minoristas y la mitad de las ventas en línea, aunque domina e impacta los precios para todo el mercado. En el pasado, los monopolios se rompieron porque sobrecargaban a los clientes; Se culpa a Amazon por cobrar muy poco y exprimir a la competencia. El temor es que eventualmente Amazon podría ganar aún más poder de mercado y ser libre de aumentar sus precios. Sin embargo, ese argumento es especulativo porque la historia del comercio minorista en línea es demasiado corta para saber cómo evolucionará la competencia.

Amazon también es propietaria de un DC a nivel mundial, que no solo respalda el negocio de comercio electrónico, sino también los servicios para empresas, gobiernos y empresas emergentes de todo el mundo por parte de *Amazon Web Service (AWS)*.

El Data Center de Oklahoma Tulsa de HP, cubre 404,000 pies cuadrados con 4 salas de datos (cada 40,000 pies cuadrados). Instala un sistema de techo reflectante para evitar que la luz solar aumente la temperatura, y un sistema de enfriamiento innovador para mantener la corriente continua, funcionando de manera estable. Puede soportar un tornado Force 5.

El Quincy Cloud Data Center de Dell, cubre un área de 12,192 metros cuadrados y costó \$3.6 millones de dólares en la primera fase. Dell también posee otros CD en India y China.

APPLE

Apple tiene todos sus Data Centers alimentados 100% de fuentes renovables. Es propietario de DC en Maiden (Carolina del Norte), Austin (Texas), Prineville (Oregón), Newark (California), Reno (Nevada), Cork (Irlanda) y Munich (Alemania). El Data Center *iCloud Maiden* de Carolina del Norte, cubre 152400 metros cuadrados, que suministra 20 megavatios de potencia con una granja solar de 100 acres.

ANEXO 6: Puntos de intercambio (IXP) en México.

México registra a finales del 2018 el segundo mayor número de usuarios de Internet en América Latina con 85 millones, después de Brasil con 149 millones⁸⁰, pero respecto a la adopción de puntos de intercambio de tráfico de Internet (IXP), solo cuenta con dos.

En la comparación de los cargos de tránsito de tráfico IP entre los países de la OCDE, México se encuentra dentro de los más costosos con un promedio de 60 USD/Mbps mensuales por Puerto (STM-1/OC-3), que equivale en capacidad a 155 Mbps por Puerto de tránsito IP⁸¹, que puede considerarse como un reflejo del costo de la conectividad de Internet con el resto del mundo.

IXP en México.

El 20 de noviembre de 2012 se constituyó el Consorcio para el Intercambio de Tráfico de Internet, Asociación Civil (CITI) sin fines de lucro que manejará el IXP. Los seis socios fundadores del IXP son: CUDI, Kio Networks, Megacable, Nextel, Redit y Transtelco⁸². El 30 de abril de 2014, CITI inauguró el primer IXP mexicano cuando en ese momento existían alrededor de 350 puntos de intercambio de tráfico de Internet alrededor del mundo y México era el único país de la OECD que no contaba con un IXP⁸³. Cualquier empresa, organismo o dependencia gubernamental puede ser parte del IXP mediante la adquisición de una membresía a la Asociación Civil. Los ingresos generados por las membresías se utilizarán para reinvertir en el robustecimiento de este primer IXP mexicano.

En abril de 2016 se creó el Nodo IXP en Yucatán que continúa el proceso de digitalización para una mejor conectividad a la red y su adhesión a la política inmobiliaria federal en materia de telecomunicaciones. La agrupación Internet Exchange Services Yucatán, A. C., funciona con el gobierno, la iniciativa privada y la academia, dispone de 12 puertos de interconexión configurables de uno a 10 gigabytes, así como banda ancha. Se trata de infraestructura física que funge como punto de intercambio de Internet operado por algún proveedor autorizado⁸⁴.

De lo anterior se puede decir que los IXP son elementos que inciden en la evolución de los servicios implementados en la nube, ya que procuran eficiencias en el

⁸⁰ Internet World Stats, (s.f.). Latin American Internet Usage Statistics. Recuperado de <https://www.internetworldstats.com/stats10.htm>

⁸¹ OCDE (2012), Estudio de la OCDE sobre políticas y regulación de telecomunicaciones en México, OECD Publishing. <http://dx.doi.org/10.1787/9789264166790-es>

⁸² Consorcio para el intercambio de tráfico en internet AC. (<http://www.ixp.mx/>)

⁸³ Según nota informativa del consorcio IXP. (<http://ixp.mx/2018/06/>)

⁸⁴ Instituto de Administración y Avalúos de Bienes Nacionales. (2018). Por un Yucatán mejor conectado. Recuperado de <https://www.gob.mx/indaabin/prensa/por-un-yucatan-mejor-conectado?idiom=es>

funcionamiento de los Data Centers, sobre todo en los costos de conectividad y factores de desempeño como el retardo o latencia, así como capacidades necesarias para los servicios de Cloud Computing. El despliegue de IXP en el territorio mexicano tiene impactos económicos y sociales positivos ya que permite intercambiar tráfico entre diversos agentes en un punto central, reduciendo así costos de tránsito y transporte, disminuye latencias, elimina la necesidad de que el tráfico intercambiado tenga que salir del país para llegar a su destino, da incentivos a contar con contenidos ubicados en el territorio nacional y facilita la adopción de protocolos más modernos, como es IPv6.

COLABORADORES QUE PARTICIPARON EN LA ELABORACIÓN DE ESTE ESTUDIO.

Rey Jiménez Luis Raúl

Dirección General de Compartición de Infraestructura

Rojas Castañeda Fernando

Dirección de Análisis Técnico de Acceso y Compartición de Infraestructura

Ortiz Ángeles Sonia Lizde

Dirección de Análisis Prospectivo de Infraestructura

Cázares Preciado José Pablo

Dirección de Área

Rodríguez Chavarría Carlos

Subdirección de Análisis Prospectivo de Infraestructura

Justo Arteaga Alin

Subdirección de Análisis Técnico de Acceso y Compartición de Infraestructura 1

Belli Gómez Giovanni

Subdirección de Análisis Técnico de Acceso y Compartición de Infraestructura 2

Briceño Díaz Jorge Antonio

Jefatura de Departamento de Análisis Técnico de Acceso y Compartición de Infraestructura 1

Reglero Crespo Alan Benjamín

Jefatura de Departamento de Análisis Técnico de Acceso y Compartición de Infraestructura 2

Botello Martínez Luis Javier

Jefatura de Departamento de Análisis Técnico de Acceso y Compartición de Infraestructura 3



INSTITUTO FEDERAL DE
TELECOMUNICACIONES

INSTITUTO FEDERAL DE TELECOMUNICACIONES

Insurgentes Sur 1143, Col. Nochebuena,
Demarcación Territorial Benito Juárez, C.P. 03720
Ciudad de México, Tel: 55 5015 4000 / 800 2000 120

www.ift.org.mx



[/IFT.mx](https://www.facebook.com/IFT.mx)



[@IFT_MX](https://twitter.com/IFT_MX)



[IFTmx](https://www.youtube.com/IFTmx)