

# Factores impulsores y ventajas del Edge Computing

White Paper n.º 226

Revisión 0

por Steven Carlini

## Resumen del artículo

El uso de Internet registra una tendencia hacia contenidos que hacen un uso intensivo del ancho de banda y un número creciente de “cosas” conectadas.

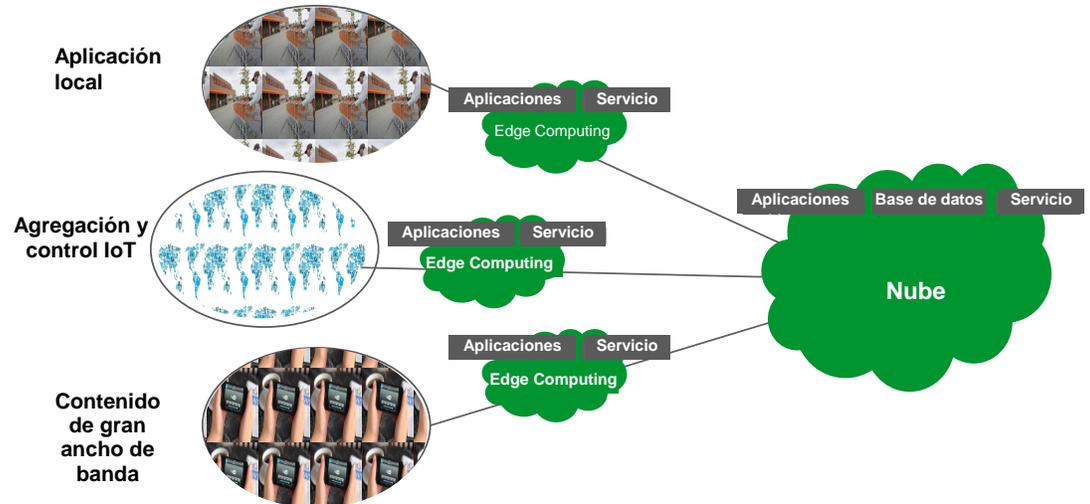
Al mismo tiempo, las redes de telecomunicaciones móviles y de datos convergen en una arquitectura de computación en nube. Para responder a las necesidades actuales y futuras, se ha comenzado a incorporar potencia de computación y almacenamiento en la periferia de la red para acortar el tiempo de transporte de los datos y aumentar la disponibilidad. El *Edge Computing* o computación en la periferia de la red acerca al usuario o a la fuente de datos a los contenidos de uso intensivo del ancho de banda y las aplicaciones sensibles a la latencia. Este White Paper explica los factores impulsores del Edge Computing y explora los diversos tipos existentes.

## Definición de Edge Computing

El Edge Computing sitúa las funciones de adquisición y control de datos, el almacenamiento de contenido de gran ancho de banda y las aplicaciones más cerca del usuario final. Se introduce como punto final lógico de una red (Internet o red privada) como parte de una arquitectura de computación en nube más extensa.

**Figura 1**

Diagrama básico de computación en nube con dispositivos en la periferia de la red



En este White Paper abordaremos tres aplicaciones primarias del Edge Computing.

1. Como herramienta para reunir grandes volúmenes de información proveniente de “objetos” locales como punto de agregación y control.
2. Como almacenamiento local y suministrador de contenidos de uso intensivo de ancho de banda formando parte de una red de distribución de contenidos.
3. Como herramienta local de aplicaciones y procesos para replicar servicios en la nube y aislar el Data Center de la nube pública.

Pero antes de tratar las aplicaciones y soluciones, vamos a definir cómo funcionan las redes e Internet.

## Cómo funciona Internet

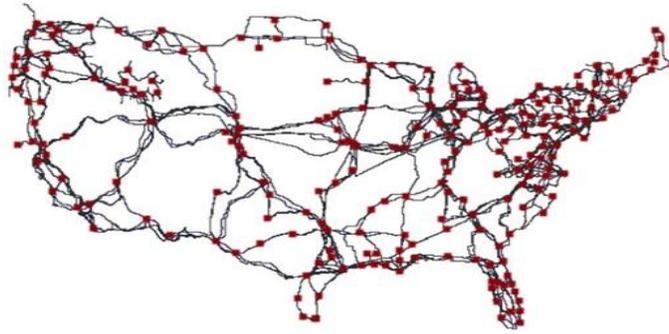
### Transmisión de datos “este-oeste”

Los datos de la fuente se convierten en paquetes que se transmiten a través de la red mediante un protocolo de red denominado IP (Internet Protocol). Otro protocolo llamado BGP (Border Gateway Protocol) se ocupa del enrutamiento en Internet. Internet se diseñó para sobrevivir a caídas masivas, sorteando los posibles problemas. BGP no tiene en cuenta la temporización del enrutamiento de datos. Solo considera el número de “saltos” (*hops*) entre las dos redes que intentan comunicarse. Esos saltos pueden estar sufriendo graves congestiones o la ruta puede seguir un trayecto físicamente más largo, con menos saltos, en lugar de un trayecto mucho más corto con múltiples saltos. La **Figura 2** muestra un mapa de los numerosos saltos de larga distancia existentes en Estados Unidos.<sup>1</sup> Aunque el funcionamiento de BGP es sumamente fiable y se trata de una tecnología fundamental sobre la cual se sustenta Internet, es muy poco eficaz desde el punto de vista de la latencia (retardos, inestabilidad y congelación de imágenes).

<sup>1</sup> <http://conferences.sigcomm.org/sigcomm/2015/pdf/papers/p5665.pdf>

**Figura 2**

Mapa de diversos saltos de red de Estados Unidos.

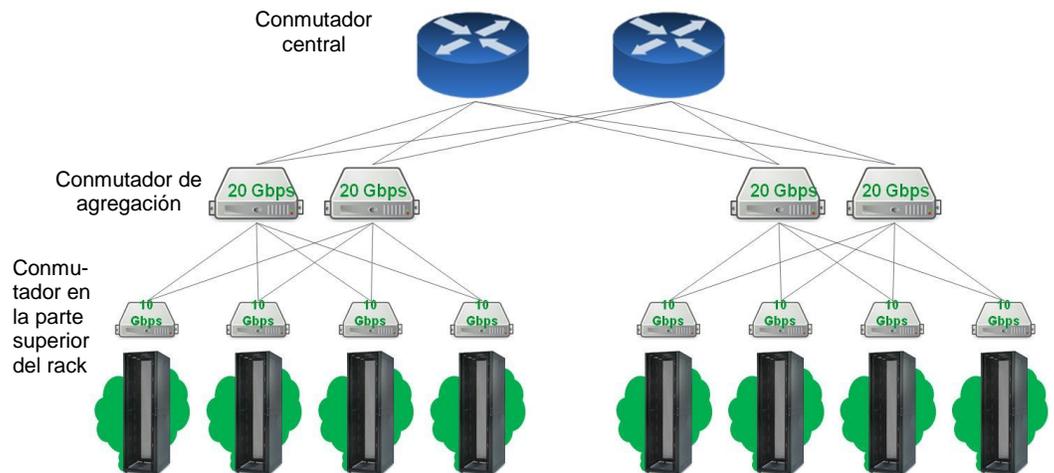


### Transmisión de datos “norte-sur”

Como refleja la **Figura 3**, los datos fluyen desde el interior de la red de un Data Center típico hacia el exterior partiendo de una interfaz de servidor físico y a través de conmutadores situados en la parte superior del rack o de conmutadores de final de fila. Desde cada conmutador situado en la parte superior del rack, los datos pasan por un conmutador de agregación, el cual dirige los datos a través de un conmutador central que es el punto principal de entrada y salida del Data Center. Cada uno de estos conmutadores transfiere datos y se considera un salto de red, con la correspondiente ralentización y posibilidad de congestionar la red. Si existe un número de abonados excesivo en cualquier capa de red (es decir, que el ancho de banda no está dimensionado para un pico de salida), aumentan las posibilidades de sufrir ralentizaciones adicionales en periodos de uso intensivo.

**Figura 3**

Red de un Data Center



## Aplicación n.º 1: Distribución de contenidos de gran ancho de banda

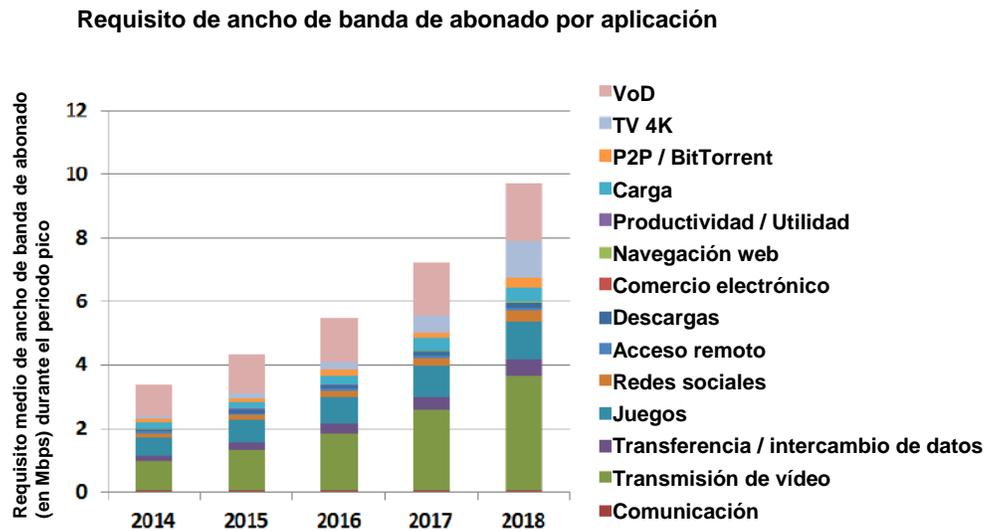
La latencia es el tiempo transcurrido entre el momento en que se transmite un paquete de datos hasta el momento en que alcanza su destino (ida) y regresa (ida y vuelta). Pese a que la mayoría de los datos solo hacen el trayecto de ida, este resulta casi imposible de medir. Por este motivo la medida de latencia más habitual es el tiempo de ida y vuelta desde un solo punto. Las latencias de ida y vuelta inferiores a 100 milisegundos (ms) son habituales, aunque lo deseable es que sean de menos de 25 ms.

El ancho de banda hace referencia a la velocidad de transmisión de los datos en la red. Cada fabricante especifica la velocidad máxima de sus equipos de red. Sin embargo, la velocidad real alcanzada en una red concreta es casi siempre inferior al valor máximo. El exceso de latencia da lugar a “atascos” que evitan que los datos aprovechen toda la capacidad de la red. El impacto de la latencia sobre el ancho de banda de la red puede ser temporal (unos pocos segundos), como un semáforo, o constante, como en el caso de un puente con un solo carril. La mayor probabilidad de congestión de la red viene dada por el contenido de vídeo de gran ancho de banda. Como se observa en la **Figura 4**, VoD, la

televisión 4K y la transmisión de vídeo son las aplicaciones de gran ancho de banda que crecen con mayor rapidez<sup>2</sup>.

**Figura 4**

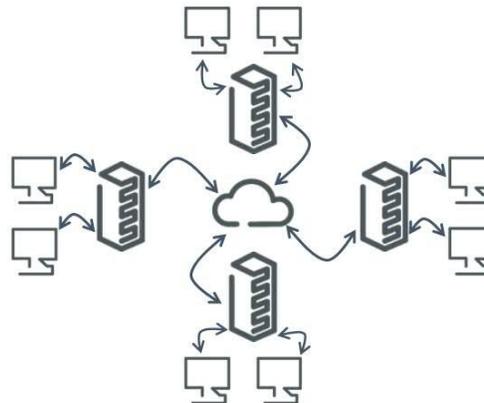
*Crecimiento de las aplicaciones de gran ancho de banda*



Para aliviar la congestión de la red y mejorar la transmisión de contenidos de gran ancho de banda, tanto ahora como en el futuro, los proveedores de servicios han comenzado a interconectar sistemas en Internet que almacenan en caché el contenido más próximo al usuario. Esto permite suministrar el contenido rápidamente a numerosos usuarios duplicándolo en múltiples servidores y dirigiéndolo a los usuarios en función de su proximidad. El almacenamiento en caché en estos sistemas es un ejemplo de Edge Computing (**Figura 5**).

**Figura 5**

*Diagrama de una red simple de distribución de contenidos (CDN)*



## Aplicación n.º 2: Edge Computing como punto de agregación y control de la IoT

Las tecnologías que dotarán de “inteligencia” a todo (ciudades, agricultura, coches, salud, etc.) en el futuro exigen una implantación masiva de sensores para establecer el Internet of Things (IoT). Un sensor IoT consiste en un nodo u objeto sin capacidad de computación con una dirección IP que se conecta a Internet.

A medida que el precio de los sensores desciende, el número de objetos IoT se dispara. Cisco calcula que IoT contará con 50.000 millones de dispositivos conectados a Internet en 2020<sup>3</sup>. IoT puede automatizar las operaciones:

<sup>2</sup> ACG Research, [The value of content at the edge](#), 2015, p.4

<sup>3</sup> Dave Evans, [The Internet of Things: How the Next Evolution of the Internet Is Changing Everything](#), Cisco Internet Business Solutions Group, p. 3

- Recopilando automáticamente información sobre elementos físicos (máquinas, equipos, dispositivos, instalaciones o vehículos) para supervisar su estado o comportamiento.
- Utilizando esa información para proporcionar visibilidad y control con el fin de optimizar procesos y recursos.

El término Machine to Machine (M2M) hace referencia a las tecnologías que permiten que los sistemas cableados o inalámbricos se comuniquen con otros dispositivos del mismo tipo. M2M se considera parte integral de IoT y aporta varias ventajas en el ámbito industrial y comercial en general, ya que tiene una amplia gama de aplicaciones en la Ciudad Inteligente.

Industrial Internet of Things (IIoT), que incluye el aprovechamiento de los datos de los sensores, el control de la comunicación entre máquinas y las tecnologías de automatización, genera grandes cantidades de datos y tráfico de red. Los sistemas informáticos industriales privados y las tecnologías de red están migrando hacia sistemas comerciales generalizados que se comunican sobre redes IP.

La exploración de gas y petróleo es un ejemplo de aplicación IIoT. Un grupo de drones voladores denominados “robots de recogida aérea de datos” examina las zonas de trabajo durante la exploración y genera grandes cantidades de datos en forma de vídeo de alta definición. Dichas zonas de trabajo son difíciles de coordinar, con sus flotas de grandes camiones, grúas y excavadoras. Los métodos tradicionales de gestión del tráfico utilizaban helicópteros tripulados para la videovigilancia. Los drones autónomos pueden fotografiar las zonas de trabajo las 24 horas del día, proporcionando a los supervisores una vista actualizada al minuto de la distribución de sus recursos. El Edge Computing permite que los drones transmitan los datos en tiempo real y reciban instrucciones en el momento oportuno.

### Figura 6

*Exploración de petróleo y gas: Los drones recogen enormes cantidades de datos en los campos petrolíferos y utilizan el Edge Computing para transferir datos y recibir instrucciones de movimiento en tiempo real.*



## Aplicación n.º 3: Aplicaciones locales

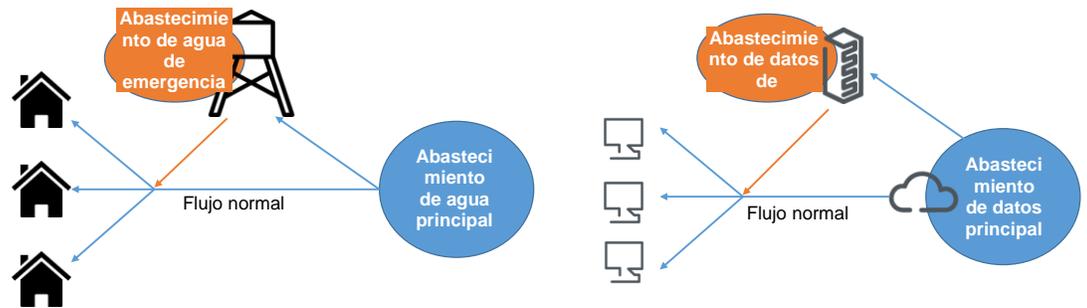
La necesidad de mantener o incrementar la disponibilidad de los recursos y las redes informáticas es casi siempre una prioridad importante. La computación en nube siempre ha sido una arquitectura centralizada. El Edge Computing transforma la computación en nube en una arquitectura de nube distribuida. La principal ventaja reside en que cualquier clase de interrupción afecta a un solo punto de la red en vez de a toda la red. Un ataque de denegación de servicio distribuido (DDoS) o un fallo de suministro prolongado, por ejemplo, afectaría únicamente al dispositivo de Edge Computing, y a las aplicaciones locales que este aloja, y no a todas las aplicaciones en ejecución en un Data Center de nube centralizada.

Las empresas que han migrado a la computación en nube deslocalizada pueden beneficiarse de la mayor redundancia y disponibilidad del Edge Computing. Las aplicaciones de negocio críticas o aquellas necesarias para las funciones básicas del negocio pueden duplicarse in situ. Podría compararse a una pequeña localidad que disfruta de un abastecimiento de agua abundante compartido como fuente principal, como recoge la **Figura 7**. En caso de inte-

rrumpirse el abastecimiento debido a un problema en el suministro principal o la red de distribución, la localidad dispone de un depósito de agua de emergencia.

**Figura 7**

El abastecimiento de agua de una localidad como representación del Edge Computing.

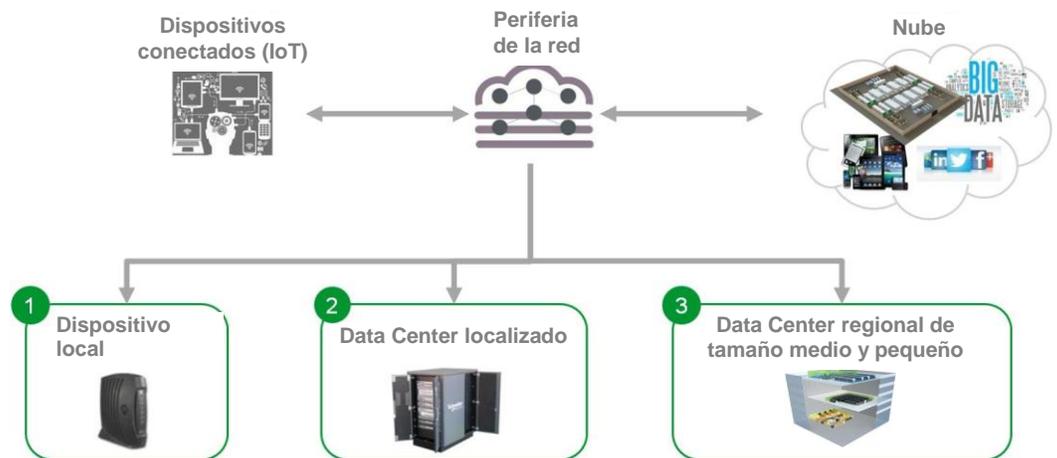


## Tipos de Edge Computing

En general, existen tres tipos de Edge Computing, como muestra la **Figura 8**.

**Figura 8**

Tipos de Edge Computing



### Dispositivos locales:

Dispositivos dimensionados para una finalidad definida específica. La implantación es "inmediata" y son adecuados para aplicaciones domésticas o en pequeñas oficinas. Algunos ejemplos serían el sistema de seguridad de un edificio (dispositivo Intel SOC) o el almacenamiento de contenido de vídeo local en un DVR. Otro ejemplo es una pasarela de almacenamiento en la nube compuesta por un dispositivo local, que normalmente es un *appliance* o servidor con las APIs necesarias, como SOAP o REST. Las pasarelas de almacenamiento en la nube permiten que los usuarios integren dicho almacenamiento en las aplicaciones sin trasladar las aplicaciones a la nube misma.

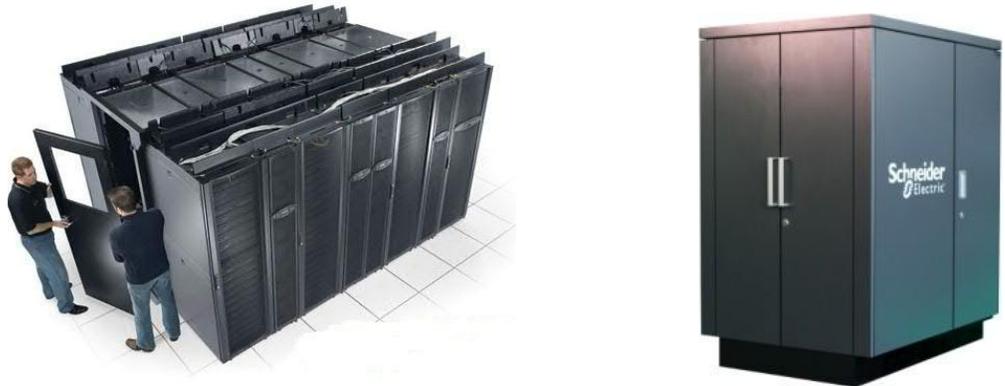
### Data Centers de 1-10 racks localizados:

Estos Data Centers ofrecen capacidades de procesamiento y almacenamiento importantes y se instalan con rapidez en entornos ya existentes. Es frecuente que estén disponibles como sistemas configurables a la carta prediseñados que posteriormente se ensamblan in situ, como muestra la **Figura 9** (izquierda). Otra clase de Data Center localizado son los micro Data Centers prefabricados, montados en fábrica y depositados directamente en su centro de trabajo, como los de la **Figura 9** (derecha). Estos sistemas, compuestos por un solo envoltorio, pueden utilizar distintas clases de envoltorios (impermeables, a prueba de corrosión, antiincendios, etc.) o envoltorios estándar para oficinas. Las versiones de rack individual pueden aprovechar el edificio, la refrigeración y la alimentación ya disponibles, ahorrando gastos de capital, puesto que no es necesario construir un centro específico. La instalación requiere elegir un lugar próximo a la acometida de alimentación y de fibra del inmueble. Las versiones con múltiples racks ofrecen más capacidad y versatilidad debido a

su tamaño, pero exigen más tiempo de planificación e instalación y necesitan refrigeración propia. Estos sistemas de 1 a 10 racks son adecuados para una amplia gama de aplicaciones que necesitan baja latencia o gran ancho de banda, o bien más seguridad o disponibilidad.

**Figura 9**

*Un ejemplo de Data Center configurado a la carta (izquierda) y de micro Data Center (derecha)*



#### **Data Centers regionales:**

Los Data Centers con más de diez racks y situados más próximos al usuario y a la fuente de datos que los Data Centers de nube centralizada se denominan Data Centers regionales. Debido a su tamaño, su capacidad de procesamiento y almacenamiento es superior a la de los Data Centers localizados de 1 a 10 racks. Incluso cuando son prefabricados, su construcción tardará más tiempo que la de los Data Centers localizados debido a la posible necesidad de realizar obra, de obtener permisos y de cumplir los requisitos legales locales. También requieren fuentes de alimentación y refrigeración propias. La latencia dependerá de la cercanía física a los usuarios y los datos, así como del número de saltos que los separen.

## Conclusión

El Edge Computing puede solucionar los problemas de latencia y permitir que las empresas aprovechen las oportunidades asociadas al uso de una arquitectura de computación en nube. Las cargas de trabajo generadas por la transmisión de vídeo de gran ancho de banda causan congestiones y latencia en la red. Los Data Centers de la periferia de la red acercan los contenidos que hacen uso intensivo del ancho de banda al usuario final y los datos a las aplicaciones sensibles a la latencia. Se introduce potencia de computación y capacidades de almacenamiento directamente en la periferia de la red para reducir el tiempo de transporte y aumentar la disponibilidad. El Edge Computing comprende dispositivos locales, Data Centers localizados y Data Centers regionales. Las versiones localizadas de 1-10 racks son las que proporcionan una rapidez de instalación y una capacidad más acordes con las futuras demandas de las aplicaciones IoT. Pueden diseñarse e instalarse de manera rápida y sencilla en versiones configuradas a la carta o prefabricadas.



### Acerca del autor

**Steven Carlini** es Director de Marketing para Soluciones de Data Centers de Schneider Electric. A lo largo de su carrera profesional ha sido responsable de algunas de las soluciones más innovadoras que han transformado el panorama y la arquitectura de los Data Centers. Es titulado en Ingeniería Eléctrica por la Universidad de Oklahoma y ha cursado un MBA en Comercio Internacional por la Universidad de Houston. Es un experto reconocido en este campo y ponente y conferenciante habitual en eventos relacionados con el sector de los Data Centers.



**[Cost Advantages of Using Single-Rack Micro Data Centers](#)**

White Paper n.º 223



**[Opciones prácticas para la instalación de equipos de IT en sucursales y salas de servidores pequeñas](#)**

White Paper n.º 174



**[Acceda a todos los White Papers](#)**

[whitepapers.apc.com](http://whitepapers.apc.com)



**[Acceda a todas las TradeOff Tools™](#)**

[tools.apc.com](http://tools.apc.com)



## Contacte con nosotros

Para enviar sus comentarios y observaciones acerca del contenido de este White Paper:

Data Center Science Center  
[dcsc@schneider-electric.com](mailto:dcsc@schneider-electric.com)

Si es usted cliente y tiene una pregunta concreta acerca de su proyecto de Data Center:

Póngase en contacto con su representante de Schneider Electric en:  
[www.apc.com/support/contact/index.cfm](http://www.apc.com/support/contact/index.cfm)