



7. Estrategias y modelos de negocio cloud

7.3.1. El Cloud y el Big Data (I).

El proceso de digitalización de la sociedad que venimos experimentando en los últimos años ha multiplicado exponencialmente la cantidad de datos que debemos almacenar. Cada vez generamos más información, y es más compleja de procesar.

Hace unos años hablábamos en términos de terabytes o petabytes de información. Según la Wikipedia, cada día de 2012 se generaron 2,5 trillones de bytes de información de contenido de todo tipo.

Esto es debido en gran parte a la proliferación de dispositivos inteligentes como smartphones y tabletas, las redes sociales y los avances tecnológicos de los últimos años.

Los sets de datos que las aplicaciones informáticas tenían que procesar son ahora mucho más complejos, crecen de forma constante y son mucho más difíciles de gestionar. La información relacionada al dato en sí mismo es también mucho más compleja, lo que provoca que la gestión del dato sea realmente el reto que nuestros sistemas han debido afrontar.

Llamamos **Big Data** al conjunto de sets de datos grandes y complejos que debemos gestionar.

Su integración con el cloud ha permitido abstraer el dato de la capa hardware y software que tiene por debajo. Partiendo del concepto de DaaS (Data as a Service) han proliferado servicios basados en poner a disposición de las organizaciones el dato de la forma más eficiente posible y con sus datos relacionales asociados sin preocuparse de las capas que tiene por debajo a nivel tecnológico.

¿Cómo impacta esto en nuestro día a día? El cloud es la capa de mediación que permite gestionar ese gran volumen de datos que no para de crecer, y que lo hace de forma exponencial, y disfrutar del mismo en modo servicio y sin que su gestión suponga un problema.

A modo de ejemplo, recordemos la tecnología que utilizábamos en el año 2000 y nuestras preocupaciones: el temido efecto 2000 y su efecto sobre la banca, los teléfonos móviles continuaban con su proceso de generalización iniciado a finales de los años 90, Apple "revivía" gracias al lanzamiento de su revolucionario iPod y nuestros ordenadores PC vivían en los tiempos de Windows XP.

La cantidad de datos que cada uno generaba al día en ese momento era muy inferior a la que generamos ahora mismo. Sólo con nuestro smartphones, una tableta y las difusión de las soluciones 3G/4G a usuario final somos capaces de crear una foto de alta resolución y compartirla en Internet con nuestros amigos y familiares en segundos.

Asimismo, las empresas han evolucionado de forma acorde y a día de hoy muchas de ellas disponen de soluciones integradas de compartición de documentación que permiten compartir un documento en tiempo real entre dos personas ubicadas a miles de kilómetros de distancia y trabajar sobre el mismo de forma coordinada mientras celebran una videoconferencia con notas, que además, están grabando para poder consultar cuando precisen en el futuro.

Desde una perspectiva más técnica, el Big Data supone que debemos crear enormes entornos de servidores, aceleradores, balanceadores y cabinas de almacenamiento que deben convivir en la misma plataforma para garantizar un rápido y ágil acceso al dato.

El problema ya no lo encontramos en la gestión de la información, sino en la cantidad de información que debemos gestionar, debido a la integración de tecnologías convergentes y los cambios de modelo relacional y de trabajo de las propias empresas y organizaciones, que son capaces de generar mucha más información en el mismo periodo de tiempo que hace unos años.

Esto complica de forma drástica la gestión del dato, y los procesos de análisis y data mining que una empresa debe realizar sobre los mismos: un ERP corporativo a día de hoy debe correlar información de distintos orígenes: estructurados, como Bases de Datos; y desestructurados, como documentos, vídeos o correos electrónicos. También debe ser capaz de entregar los mismos niveles de operatividad y agilidad que hace unos años, si bien la capacidad de datos se ha multiplicado de forma exponencial, así como el tamaño de los mismos.

Tiene toda la lógica del mundo que la computación distribuida, y de forma concreta, el cloud, sea la capa de mediación sobre la que se desarrollen todos los modelos de Big Data.

Referencias:

- [Mapping Big Data with Google's Cloud Platform](#)