

Cloud Computing - DBaaS



Marta Zorrilla - Diego García-Saiz
Enero 2017

Este material se ofrece con licencia: [Creative Commons BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/)



Tabla de contenidos

- Cloud computing
 - Definición
 - Taxonomía
 - Ventajas y preocupaciones
 - Del gestor relacional al DBaaS

Bibliografía

- Libros
 - Lee Chao. Cloud Database Development and Management. 2013. Auerbach Publications
 - Ozsu, M. Tamer. Principles of distributed database systems. 3rd ed. Prentice Hall, cop. 2011. Springer
- Artículos:
 - Waleed Al Shehri. Cloud Database: Database as a Service. International Journal of Database Management Systems Vol.5, No.2, April 2013
 - M. Stonebraker, A. Pavlo, R. Taft and M. L. Brodie, "Enterprise Database Applications and the Cloud: A Difficult Road Ahead," *2014 IEEE International Conference on Cloud Engineering*, Boston, MA, 2014, pp. 1-6.
 - Katarina Grolinger et al. Data management in cloud environments: NoSQL and NewSQL data stores. *J. Cloud Comput.* 2 (1) 1-24, 2013

URLs interesantes

- Vídeo: Cloud wars
 - <https://youtu.be/342KEaxFVjM>
- AWS, AZURE y GOOGLE comparison
<https://www.sitepoint.com/a-side-by-side-comparison-of-aws-google-cloud-and-azure/>
- Multi-tenant vs multi-instance
 - <https://servicematters.servicenow.com/why-cloud-architecture-matters-the-multi-instance-advantage-over-multi-tenant/>
- DBaaS: pros and cons
 - <http://searchitchannel.techtarget.com/tip/DBaaS-pros-and-cons-for-solution-providers>

Cloud Computing

- Visión
 - Servicios de computación, de almacenamiento y red proporcionados a través de Internet (la "nube") bajo petición.
- Nuevo modelo de negocio
 - Usando interfaces sencillas, los usuarios pueden externalizar la gestión de datos, desarrollo de aplicaciones o la admón. de sistemas.
 - El proveedor del servicio se encarga de gestionar la infraestructura.
- Se apoya en modelos de computación previos tales como *web services, cluster computing, virtualization, grid computing,...*

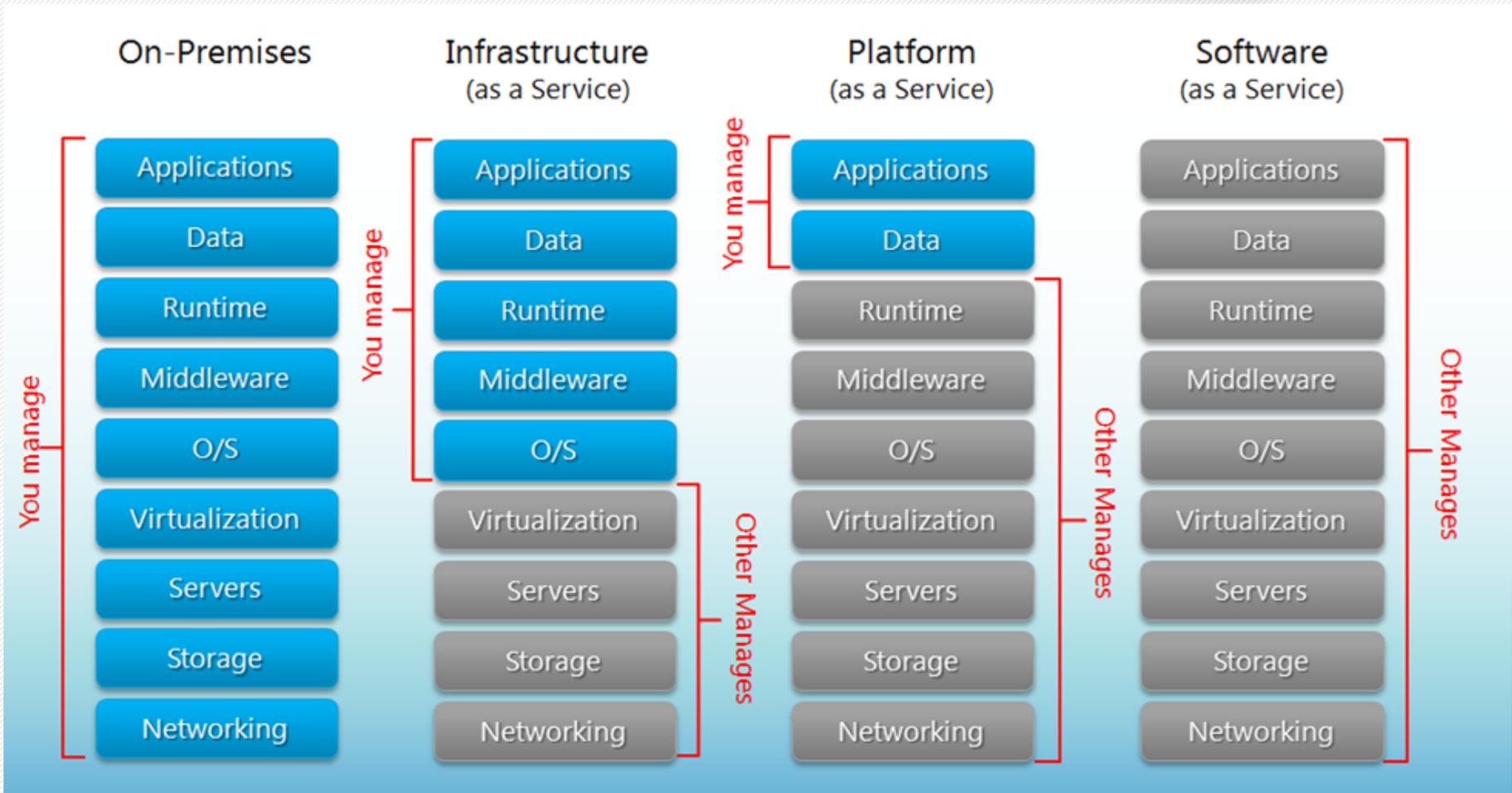
Cloud Computing: Definición

- Modelo de negocio que proporciona servicios y recursos bajo demanda, generalmente sobre Internet, con escalabilidad y con la fiabilidad de un centro de datos
- Ha sido desarrollada por gigantes de la industria web, como Amazon, Google, Microsoft y Yahoo.
- Se contrata como “servicios”
 - Se paga por los recursos que consumes “*Pay-as-you-go pricing model*”
 - Se acuerda el “*Service Level Agreement (SLA)*” que establece el uso y calidad de los servicios y política de precios

Cloud Computing: Taxonomía

- Infrastructure-as-a-Service (**IaaS**)
 - Computación, redes y recursos de almacenamiento, como un servicio
 - Proporciona elasticidad: capacidad para incrementar o liberar recursos según sea necesario
 - P.ej. Servicios Web de Amazon
- Platform-as-a-Service (**PaaS**)
 - Plataforma informática con IDEs y APIs como servicio
 - Permite a los desarrolladores crear e implementar aplicaciones directamente en la infraestructura de la nube e integrarlas con las aplicaciones proporcionadas como SaaS
 - P. ej. Google Apps
- Software-as-a-Service (**SaaS**)
 - Software de aplicación como servicio
 - Aloja aplicaciones: desde simple (correo electrónico, calendario) hasta complejo (CRM, análisis de datos o red social)
 - P.ej. Sistema Salesforce CRM

Cloud Computing: Taxonomía



Cloud Computing: Ventajas

- Costes más bajos
 - Cliente: ni es propietario ni gestiona la infraestructura IT, paga por lo que usa
 - Proveedor: los gastos de infraestructura los comparte entre muchos clientes
- Fácil acceso y uso
 - El acceso a los servicios IT es vía Internet, accesible desde cualquier dispositivo conectado
- Elasticidad
 - Flexibilidad del cliente para incrementar/decrementar el uso de más máquinas virtuales (VMs)
- Calidad de Servicio (QoS)
 - Al ser el proveedor especializado en este tipo de servicios, aumenta la QoS

Cloud Computing: «preocupaciones»

- Empresas ceden el control al proveedor del servicio y éste debe velar por la **privacidad**, **seguridad** e **integridad** de los datos.
 - La solución pasa por crear **nubes privadas**, pero esto está en contradicción con la compartición de la infraestructura entre múltiples clientes
- El servicio depende de la velocidad de la red y su disponibilidad. Al confiarse en Internet, esta es inferior y con menos QoS que una línea dedicada.
- Rendimiento: transacciones por segundo y tiempos de respuesta no acotados.
- Al usarse HTTP, protocolo sin estado, toda petición debe ser una transacción para poder realizar commits y rollbacks.

Cloud Computing: privacidad

- Nubes públicas
 - Los usuarios acceden a los servicios de manera compartida sin que exista un exhaustivo control sobre la ubicación de la información que reside en los servidores del proveedor. El hecho de sean públicas no es un sinónimo de que sean inseguras.
- Nubes privadas
 - Adecuadas para aquéllos que necesiten, por la criticidad de la información que manejen, una infraestructura para su uso exclusivo.
- Nubes híbridas
 - Combinan características de las dos anteriores, de manera que parte del servicio se puede ofrecer de manera privada (por ejemplo, la infraestructura) y otra parte de manera compartida (por ejemplo, las herramientas de desarrollo).

Cloud Computing: seguridad

- Datos deben asegurarse en los tres estados: “*at rest*”, “*in transit*” y “*in use*”
 - El cliente puede proteger los datos “*in transit*” encriptándolos, lo que conlleva que el proveedor del servicio gestione política de claves antes de su almacenamiento.
 - El cliente puede reforzar las técnicas de control de acceso (*data in use*) pero depende de las posibilidades que ofrezca el modelo de servicio contratado.
 - Para datos “*at rest*”, la mejor política sería que el cliente guardara todos los datos encriptados, incluso la BD sin que el proveedor tenga acceso a las claves. Se evita así que haga un mal uso de los datos, aunque siguen vigentes las amenazas de ataque para denegación del servicio y corromper los datos.

Cloud Computing: aplicaciones empresariales

- **OLTP:** aplicaciones de escritura intensiva que requieren garantías transaccionales ACID, protección de datos sólida y respuesta precisa. BD de tamaño medio (TB)
 - **La solución Cloud no se adapta bien** ya que requiere multiprocesadores de disco compartido (*shared-disk*) y los datos corporativos se almacenan en un host no confiable. No válido para aplicaciones de misión crítica
- **OLAP:** aplicaciones analíticas que manejan volúmenes de datos grandes (PB), de lectura intensiva y que pueden aceptar propiedades ACID relajadas
 - **Adecuadas para la nube** ya que el modelo es fácilmente particionado y las consultas pueden ser paralelizadas por lo que la arquitectura *shared-nothing clusters* es apta para este propósito
 - Además, los datos sensibles pueden ser anonimizados en la nube (datos generalmente agregados).
- **Otras:** contexto muy variado. Adecuado para servicios (ej: blogs, servicios de redes sociales, etc.) que utilicen tecnología NoSQL.

Del modelo relacional a DBaaS

1970-2000:
Soluciones
relacionales

2000-2005:
PuntoCom,
aparición NoSQL

2005-2010
Open source

2010 +:
Adopción
cloud



DBaaS

- DataBase as a Service:
 - Servicio gestionado que ofrece acceso a una base de datos.
 - El control se realiza a través de una API.
 - Esta API es accesible al usuario a través de una consola de administración, usualmente una aplicación Web, que el usuario puede utilizar para administrar y configurar la base de datos e incluso añadir o reducir instancias de base de datos.
- DBMS en la nube (**DBaaS**), los más conocidos:
 - Amazon Web Services (AWS)
 - Google Cloud SQL
 - Microsoft Azure

Amazon Web Services (AWS)

<https://aws.amazon.com/es/>

- Servicios que ofrece en relación a datos
 - Relacional
 - MySQL, Maria, Postgresql
 - Oracle, SQL Server, Aurora
 - In-memory
 - ElasticCache, Redshift
 - NoSQL
 - Amazon SimpleDB: key-value store
 - Amazon DynamoDB: BD NoSQL sobre disco de estado sólido. Adecuado para datos almacenados con bajo nº de consultas.

Google Cloud SQL

<https://cloud.google.com/>

- Tres productos:
 - Google Cloud SQL:
 - Bases de datos MySQL alojadas en la nube de Google. Completamente relacional. Google gestiona replicación, encriptación, backups y gestión de parches
 - Cloud Datastore:
 - almacenamiento NoSQL
 - BigQuery:
 - herramienta analítica

Google ofrece la posibilidad de alojar otras BD en su plataforma.

Microsoft Azure

<https://azure.microsoft.com/es-es/>

- Algunos de los productos:
 - Relacional:
 - Azure SQL Database.
 - NoSQL:
 - Document Stores: DocumentDB
 - Key/Value Stores: Redis Caché
- Azure ofrece MongoDB y otras BD conocidas (MemSQL, Riak, María, ...), disponibles en Azure Marketplace.

DBaaS: Modelos de servicio

- Modelo Multi-instancia, un DBMS en una VM para cada suscriptor dándole control al cliente en las tareas administrativas
- Modelo Multi-tenant, el suscriptor comparte el entorno para lo cual se etiquetan los datos con un ID del suscriptor para su identificación. El proveedor se encarga de mantener el entorno seguro
 - Datos aislados “lógicamente” de tus competidores
 - Arquitecturas grandes, complejas y centralizadas
 - Las paradas afectan a todos los clientes, aunque la actualización se requiera por uno
- Las características comunes de un RDBMS “on-premise” no están siempre disponibles en la “cloud”

Características a valorar a la hora de seleccionar un DBaaS

- Capacidad de almacenamiento en una base de datos
- Portabilidad a otro proveedor de servicios
- Transacciones: soporta ACID o no
- Grado de configuración/administración que te permite el proveedor realizar
- Protocolos que soporta para acceder a la BD (JDBC, SOAP,...)
- Protocolos de encriptación que soporta
- Localizaciones que dispones para replicar/asegurar datos