Clase del día - 26/05/2021

En el pasado el aprovisionamiento de recursos informáticos *On-premise* (en las instalaciones de la empresa) representaba el concurso de diferentes proveedores de bienes y servicios, como eran los representantes de ventas, ingenieros de pre-venta, fabricantes de los equipos, fabricante del sistema operativo, fabricante de la base de datos, agentes aduanales, transportistas, instaladores del *site*, proveedor de energía, proveedor de comunicaciones, instaladores del hardware, instaladores del software, entre otros.

Entonces, el aprovisionamiento de recursos informáticos era un proceso complejo y tardado, el cual culminaba con el sistema en producción.

Después había que re-aprovisionar cuándo crecían las necesidades de la empresa.

**Cómputo en la nube**

En 2006 aparece en la revista Wired el artículo [**The Information Factories**](https://www.wired.com/2006/10/cloudware/) (https://www.wired.com/2006/10/cloudware/) de George Gilder que describe un nuevo modelo de arquitectura basado en una infraestructura de cómputo ofrecida como servicios virtuales a nivel masivo, a este nuevo modelo se le llamó *cloud computing* (cómputo en la nube).

El concepto clave en el cómputo en la nube es el "servicio", así, se ofrece infraestructura virtual y física como servicio (IaaS: Infrastructure as a Service), DBMS, plataformas de desarrollo y pruebas como servicio (PaaS: Platform as a Service), aplicaciones de software como servicio (SaaS: Software as a Service) y otros servicios con la terminación "as a Service", como Data as a Service (DaaS), Disaster Recovery as a Service (DRaaS), entre otros.

**La elasticidad en la nube**

Debido a que el cómputo en la nube está basado fundamentalmente en la virtualización de los recursos informáticos, este modelo de arquitectura ofrece una ventaja única, la posibilidad de hacer crecer y decrecer los recursos aprovisionados.

Supongamos un servicio de streaming bajo demanda, como es el caso de Netflix. En este tipo de servicio la demanda crece los fines de semana y decrece los días entre semana. Si el proveedor del servicio no aprovisiona los recursos suficientes para atender la demanda del fin de semana, entonces muchos usuarios se quedarán sin servicio.

Por otra parte, si el proveedor del servicio aprovisiona los recursos necesarios para atender a sus usuarios el fin de semana, estos recursos estarán sub-utilizados los días entre semana, lo cual resulta en pérdidas económicas.

Sin lugar a dudas, el éxito que han alcanzado las empresas proveedoras de streaming bajo demanda, se debe a que su modelo de negocio está basado en la posibilidad que les ofrece la nube para crecer y decrecer los recursos aprovisionados, a esta característica de la nube se le llama *elasticidad*.

El cómputo elástico es la habilidad de hacer crecer y decrecer rápidamente la capacidad de cómputo (CPUs), la memoria y el almacenamiento para adaptarse a la demanda.

Para implementar el cómputo elástico se utilizan herramientas de monitoreo, las cuales aprovisionan y des-aprovisionan recursos conforme son necesarios, sin detener la operación.

**Nube pública, nube privada y nube híbrida**

Se dice que la nube pública es un conjunto de servicios de TI ofrecidos mediante recursos (servidores, almacenamiento, red) propiedad de un proveedor de servicios en la nube, como es el caso de AWS, Azure, Softlayer, Oracle, Google, etc.

Por otra parte, se entiende como nube privada aquellos servicios ofrecidos a partir de la virtualización de los recursos propiedad de la misma empresa.

En este mismo orden de ideas, la nube híbrida sería la mezcla de servicios de nube pública y servicios de nube privada.

Ver: [**¿Qué es la nube pública, privada e híbrida?**](https://azure.microsoft.com/es-es/overview/what-are-private-public-hybrid-clouds/) (https://azure.microsoft.com/es-es/overview/what-are-private-public-hybrid-clouds/)

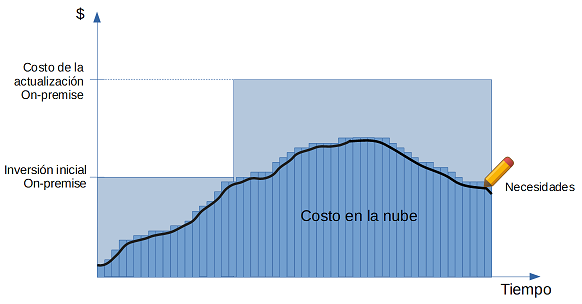
Sin embargo, hay proveedores de nube que afirman que no existe tal cosa como "nube privada", ya que el tema de elasticidad se ve acotado en un equipo "privado" debido a la limitada escalabilidad.

En cambio, en la nube pública la escalabilidad es casi ilimitada, ya que si se agotan los recursos de un data center, sin ningún problema se puede escalar a otro data center, ya sea del mismo proveedor o de otros proveedores.

Una característica importante de la nube (pública) es "pagar por lo que se usa", esto significa que solo se paga por los recursos aprovisionados.

En el caso de un centro de cómputo tradicional (también es el caso de la "nube privada") la empresa paga por todo el equipo, lo use a toda su capacidad o no.

Consideremos la siguiente gráfica:



**Escenario On-premise**

Al inicio los requerimientos informáticos de la empresa son pocos, sin embargo ésta debe hacer una inversión inicial grande ya que debe adquirir el equipo que le permita operar durante un periodo de tiempo determinado. Para saber cuál será la inversión inicial  se deberá hacer una *planeación de capacidad*.

Suponiendo que la empresa crece, entonces en determinado momento deberá realizar la actualización de su equipo informático. Entonces se deberá hacer una planeación de capacidad para establecer el tamaño de la inversión, estimando un tiempo determinado de operación.

No obstante las previsiones, la empresa podría decrecer, por tanto los requerimientos informáticos también decrecen. El costo de la actualización no se recupera, incluso, en poco tiempo el equipo se hará obsoleto.

**Escenario nube**

Ahora supongamos que la empresa utiliza servicios en la nube. Entonces la inversión inicial es mínima, ya que la empresa contratará los servicios indispensables para operar. Conforme la empresa crece, la elasticidad de la nube permite adecuar el tamaño de la infraestructura informática de acuerdo a las necesidades.

Si la empresa decrece, también decrecen sus necesidades y el costo de la infraestructura informática. En la gráfica, las barras corresponden al costo de los servicios de nube. Podemos ver que este costo se ajusta a las necesidades de la empresa.

Podemos concluir que al utilizar los servicios de nube se optimiza la relación costo/beneficio, ya que solo pagamos lo que realmente usamos.

Veamos algunos ejemplos de elasticidad en Azure.

**Nota importante**. Es recomendable detener la máquina virtual antes de cambiar la configuración ya que es posible que se experimenten problemas de conexión mediante SSH y problemas de conexión a los servicios que ofrece la máquina virtual (HTTP, HTTPS, etc.). Si es el caso, el problema de conexión puede resolverse deteniendo la máquina virtual y volviendo a encenderla para cambiar la IP pública. Así mismo, se puede re-establecer la contraseña del usuario seleccionando la opción "Restablecer contraseña" en la sección "Soporte y solución de problemas".

**Cambio del tamaño de una máquina virtual**

Para cambiar el número CPUs virtuales y/o el tamaño de la memoria RAM de una máquina virtual:

1. Ejecutar el portal de Azure.
2. Seleccionar la opción "Maquinas virtuales".
3. Seleccionar la máquina virtual a modificar.
4. En el menú "Configuración" seleccionar la opción "Tamaño".
5. Seleccionar el tamaño requerido (vCPU, RAM)
6. Dar click al botón "Cambiar tamaño"
7. Nota: si la máquina virtual está en ejecución podría reiniciarse al cambiar el tamaño
8. Dar click en la ca**m**pana de notificaciones para verificar que el cambio se haya realizado con éxito.

**Cambio del tamaño del disco de S.O.**

Para cambiar el tamaño del disco de sistema operativo de una máquina virtual:

1. Seleccionar la máquina virtual.
2. Detener la máquina virtual.
3. Seleccionar "Discos".
4. Seleccionar el disco a cambiar de tamaño.
5. Seleccionar "tamaño y rendimiento" en el menú Configuración.
6. Seleccionar el nuevo tamaño del disco.
7. Dar clic en el botón "Cambiar tamaño".
8. Para encender la máquina virtual seleccionar la máquina virtual en el menú Inicio -> Máquinas virtuales y seleccionar Iniciar.

Actividades individuales a realizar

Ver el artículo: [What is elastic computing or cloud elasticity?](https://azure.microsoft.com/en-us/overview/what-is-elastic-computing/)