Instituto Tecnológico y de Estudios Superiores de Monterrey

Escuela de Ingeniería y Ciencias

Ingeniería en Ciencia de Datos y Matemáticas

Cloud computing — Diseño de Arquitectura en la Nube

Inteligencia Artificial Avanzada para la Ciencia de Datos II

Federico Medina García Corral A01721441

Michelle Yareni Morales Ramón A01552627

Paola Sofia Reyes Mancheno A00831314

María Fernanda Torres Alcubilla A01285041

Supervisado por:
Prof. Félix Ricardo Botello Urrutia

Monterrey, Nuevo León. Fecha, 21 de noviembre de 2023

1. Situación Problema

Eres parte del equipo Cloud en una consultoría de TI y se te ha asignado el desafío de diseñar una arquitectura en la nube para la empresa DataTech. La empresa está buscando migrar sus aplicaciones y servicios a la nube para mejorar la escalabilidad, disponibilidad y seguridad. Tu objetivo es diseñar una arquitectura en la nube que cumpla con los requisitos de DataTech. Los requerimientos del diseño son:

- Máquinas virtuales: DataTech necesita alojar sus aplicaciones en máquinas virtuales en la nube. Debes diseñar una infraestructura de máquinas virtuales que cumpla con los requisitos de rendimiento, escalabilidad y alta disponibilidad de la empresa. Considera aspectos como la selección del tamaño y tipo de las máquinas virtuales, la distribución de carga, la implementación de grupos de disponibilidad y el equilibrio de carga para garantizar la continuidad del negocio y la eficiencia operativa.

- Bases de datos IaaS y PaaS: DataTech requiere bases de datos confiables y escalables para almacenar y gestionar sus datos. Debes diseñar una combinación de bases de datos IaaS y PaaS que se ajuste a las necesidades de la empresa. Considera opciones como Azure SQL Database para bases de datos PaaS y máquinas virtuales con software de base de datos instalado para bases de datos IaaS. Asegúrate de tener en cuenta aspectos de rendimiento, seguridad, escalabilidad y opciones de respaldo y recuperación.

- Storage Account - Fileshare: DataTech necesita un almacenamiento en la nube seguro y escalable para sus archivos y datos no estructurados. Debes diseñar y configurar una o varias cuentas de almacenamiento que cumplan con los requisitos de capacidad y rendimiento de la empresa. Considera el uso de Azure Storage Accounts con opciones como File Share para compartir archivos entre aplicaciones y equipos.

- Configuración entre VNets: DataTech requiere una arquitectura en la nube con redes virtuales (VNets) bien definidas y una configuración segura entre ellas. Debes diseñar y configurar las VNets, asignar subredes adecuadas y establecer reglas de conectividad para permitir la comunicación entre las diferentes partes de la arquitectura. Asegúrate de considerar aspectos de seguridad, aislamiento de recursos y enrutamiento adecuado.

- App Service: DataTech desea implementar sus aplicaciones web utilizando servicios de App Service. Debes diseñar y configurar el entorno de App Service para permitir la implementación y escalabilidad de las aplicaciones web de la empresa. Considera aspectos como la selección del plan de App Service, la configuración de dominios personalizados, la implementación de certificados SSL, la configuración de escalabilidad automática basada en la demanda y la integración con los servicios de bases de datos.

2. Diagrama de Arquitectura

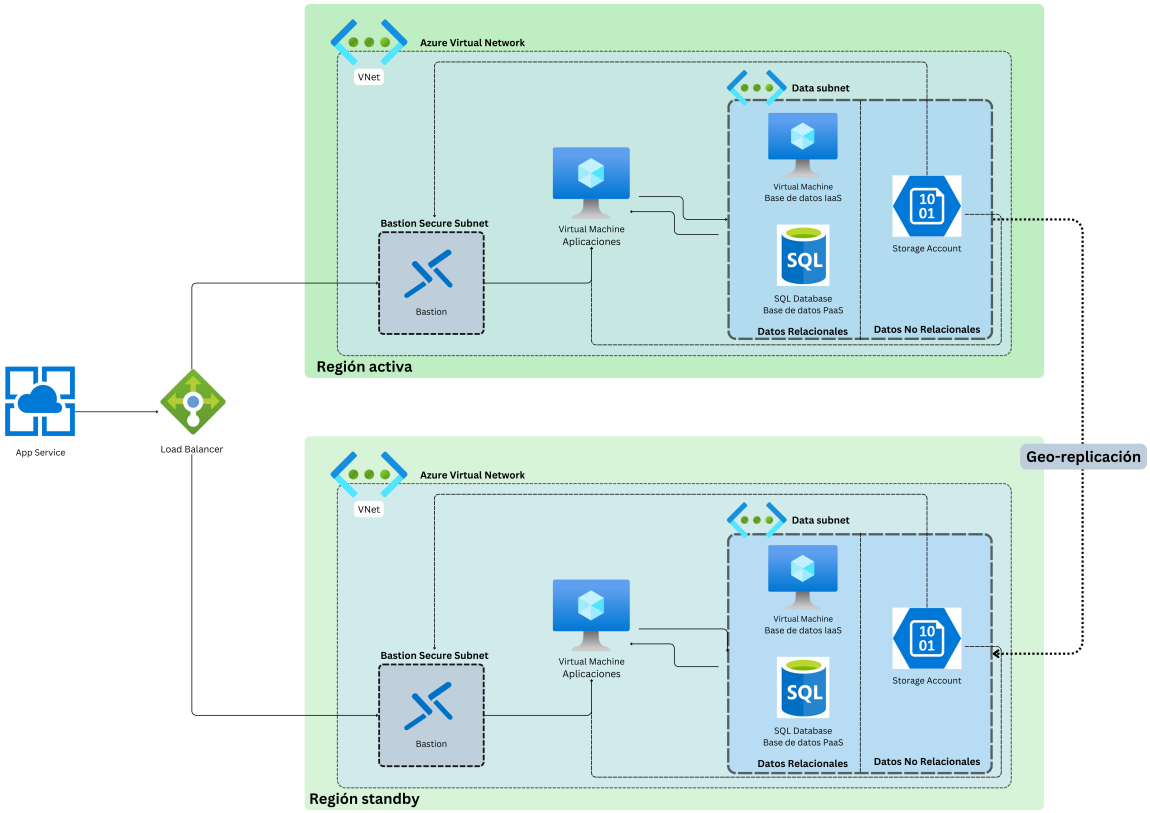

Figura 1: Diagrama de la arquitectura propuesta

La propuesta de arquitectura para la empresa Data Tech, puede evidenciarse en la figura 1, en la cual se puede observar a grandes rasgos que la misma es de alta disponibilidad de tipo 2N y activa-pasiva. Esto significa que, del número de servicios y herramientas a utilizar en la arquitectura tradicional, se hace una duplicación de esta última para asegurar una alta disponibilidad; y además una de las regiones se encuentra en standby hasta que la región principal presente algún problema, ayudando así a reducir costos de procesamiento.

Ahorabien, la arquitectura descrita de izquierda a derecha, comienza con el servicio de App Service, el cual tiene el objetivo de la implementación de sus páginas web de forma segura. Este se conecta con Load Balancer, el cual redirige el tráfico de entrada y salida entre las regiones. Posteriormente, al tener una arquitectura AD de tipo 2N, los siguientes pasos son los mismos en ambas regiones.

Cada una de las regiones tiene un VNet o Red Virtual con Azure Virtual Network, pero para tener una mejor seguridad sobre esta VNet, se agregó el servicio de Bastion Secure Network. Por esto, en el diagrama pasa primero por este servicio antes de llegar a las máquinas virtuales. Estas máquinas virtuales albergan las aplicaciones de la empresa, por lo cual las mismas tienen acceso

a las diferentes bases de datos relacionales y no relacionales. Este conjunto de bases de datos, ya sean contratadas como PaaS, IaaS para relacionales, o en Storage Account para información no relacional, están dentro de una subnet para un mejor control de seguridad de los datos que se manejan en DataTech. Específicamente Storage Account permite resguardar información, así como acceder a esta a través de las aplicaciones en las máquinas virtuales o equipos de la compañía. Finalmente, para asegurar que los 3 servicios de las Data Subnets de las 2 regiones estén al día con la información resguardada, se hace uso de la geo-replicación.

2.1. Configuración de la arquitectura

En base a la arquitectura creada, hay algunas configuraciones de red y seguridad que hemos considerado que son importantes para que DataTech las tenga en cuenta. Estas tienen la intención de optimizar el funcionamiento de las operaciones dentro de la nube al mantenerse privada y protegida de ataques o vulnerabilidades que pudieran encontrarse dentro. Primeramente se decidió tener dos VNets, una en cada región, segmentadas en subredes lógicas para tener una mejor organización de los recursos, al igual que mantenerlas conectadas para que haya una comunicación segura entre ellas dentro de las diferentes regiones [1].

Para la parte del Load Balancer se recomienda que se configure para distribuir el tráfico entre las instancias del App Service en las regiones activa y standby. Igualmente, como recomendación extra, se recomienda que se utilice la opción de "Backend Pools" para así especificar las Máquinas Virtuales (VMs) o las instancias del App Service como destinos [2].

Ahora, para poder configurar el Bastion Secure Subnet y las Máquinas Virtuales de una manera segura, se recomienda seguir las reglas del Network Security Group (NSG) para permitir solamente el tráfico necesario desde y hacia las VMs a través de este componente. Para este, Azure tiene un servicio llamado Azure Bastion con el cual se puede implementar como servicio de administración seguro para las VMs [**Morris˙2023**]. Igualmente, se recomienda utilizar Azure Security Center para monitorear y mejorar la postura de seguridad de las VMs.

Por otro lado, tenemos la SQL Database y las VMs dentro de los Datos Relacionales, la cual es de Azure. Para este simplemente se pueden crear reglas de firewall y de acceso para únicamente permitir la entrada a esta de ciertas ubicaciones específicas, lo cual eliminaría considerablemente la vulnerabilidad de ser infiltrado en las zonas no especificadas [**Mas˙2023**]. Por otro lado, para el Storage Account se puede hacer algo similar; se puede utilizar un SAS (Shared Access Signatures) para controlar el acceso a los datos que están almacenados aquí.

Finalmente podemos implementar Certificados SSL en donde se puede configurar nuevamente reglas de NSG para garantizar la seguridad de las conexiones cifradas. cone ste, se puede utilizar Azure Key Vault para gestionar dichas claves secretas de manera segura. Igualmente, es importante mencionar que se debe de estar monitoreando la arquitectura, con lo cual se recomienda utilizar Azure Monitor para mantener alerta según situaciones específicas que se establezcan.

3. Estimación de Costos

Para poder calcular con precisión los costos en la nube se debe de considerar que esto es un desafío debido a la variabilidad de los precios según la región, el tipo de recurso, aplicaciones de este y otros factores adicionales. Sin embargo, se va a hacer un plan de costos con suposiciones sobre sus usos y restringido a situaciones "ideales" en donde no haya problemas ni costos adicionales a

los servicios mencionados en la arquitectura. A continuación se muestran las especificaciones que se tomaron en cuenta para la estimación de costos y sus respectivos precios:

- **App Service**: Para estimar su costo, se debe de considerar el número de instancias y el nivel de rendimiento que se necesita. Para este ejercicio, se asumió el plan Premium con dos instancias, el cual tiene un costo aproximado de $166.66 por mes.
- **Load Balancer**: El costo de este dependerá del tipo y la cantidad de tráfico que se vaya a manejar. El Standard SKU tiene un costo de $41.66 mensuales.
- **Azure Virtual Network**: Para este se debe de tener en cuenta el tráfico de datos entre regiones y subredes. Podemos suponer que el costo de red virtual y de tráfico entre regiones es de $16.66 y $25 mensuales, respectivamente.
- **Bastion Security Subnet**: Para este se deben de considerar los usuarios que tendrán acceso a la máquina virtual. Tomaremos como ejemplo 10 usuarios, el cual tiene un costo aproximado de $33.33 mensuales.
- **Máquinas Virtuales (VMs)**: Para calcular el costo de las VMs, se debe de calcular en base a su tamaño, el tipo y la cantidad de instancias en ejecución. Para esto, se asumen dos instancias de tamaño estándar, el cual tiene un costo de $416.66 al mes.
- **Azure SQL Database**: Para calcular este se necesita saber la cantidad de almacenamiento que se va a contratar y de las DTUs. Se asumirá que se utilizarán 100 DTUs, el cual tiene un costo aproximado de $250 menusales.
- **Máquina Virtual en Datos Relacionales**: Para este es lo mismo que la máquina virtual anteriormente mencionada pero con solo una instancia, el cual tiene un costo de $208.33 al mes.
- **Azure Storage Account**: Para este se ocupa el almacenamiento y las operaciones de lectura/escritura. Para este, se estima un costo de $83.33 y de $25 mensuales respectivamente.
- **Georeplicación**: Se asume que el costo adicional para replicar entre regiones es de $66.66 mensuales.
- **Servicios Adicionales**: El costo de los certificados SSL es de $16.66 y el costo de monitorización, en este caso Azure Monitor, es de $41.66 mensuales.

En base a estas suposiciones, las cuales son datos extraídos de diversas páginas de internet y de blogs que ayudan a estimar costos para trabajar con la nube, encontramos que el costo para nuestra arquitectura es de $1,375 mensuales.

4. Configuraciones de red y seguridad

La migración y el uso de un servicio en la nube se debe realizar con un enfoque que garantice la seguridad y proteja la información, con el objetivo de que esta se mantenga confidencial, integra y disponible tanto en reposo como en tránsito. Para esto, se sugieren realizar acciones como:

- **Autenticación y control de acceso.** Las personas que ingresen a la infraestructura deben contar con las credenciales adecuadas y se debe controlar el acceso a cierta información con la otorgación de privilegios.
- **Cifrado de datos.** Ya sea en tránsito o reposo, esto se puede a través de la configuración de las distintas plataformas utilizadas.
- **Firewalls y grupos de seguridad.** Es necesario configurar firewalls para restringir el tráfico de red no autorizado y controlar el acceso a los recursos.
- **Estrategias de respaldo y recuperación.** Son necesarias para garantizar la disponibilidad de los datos en caso de pérdida o corrupción.
- **Protección de la red.** Aunque nuestros servicios estén en la nube, se debe proteger la red con la cual se accede ya que una violación a su seguridad puede afectar a nuestra información.
- **Monitoreo y detección de amenazas.** La mejor manera de evitar una violación, es la detección previa de las amenazas y riesgos que presenta nuestro modelo, así como su solución.
- **Fomentación de la cultura de la ciberseguridad.** Aunque nuestros datos y servicios estén protegidos, el factor humano siempre debe ser considerado ya que es quien hace uso de las plataformas, por lo que la concientización del personal es esencial para la prevención.
- **Implementación de las normas.** Cada país o región cuenta con sus normas de cómo se debe tratar la información (sensible o no) y cómo terceras partes pueden verse involucradas, por lo que se sugiere contar con personas encargadas de verificar el cumplimiento de las normas para prevenir asuntos legales en un futuro.
- **Actualización y mantenimiento.** Al ser un modelo serverless en la nube el mantenimiento se ve reducido, ya que las actualizaciones de sistemas operativos y parches de seguridad son responsabilidad del proveedor, por lo mismo, se deben contratar empresas de confianza.
- **Eliminación y destrucción de la información.** Cuando esta ya no presenta importancia y no se usará en algún futuro, su destrucción debe ser segura para evitar filtraciones.

5. Conclusiones Individuales

- **Sofía Reyes**: Teniendo cada vez más empresas que desean migrar sus sistemas a la nube, es importante poder identificar ciertas arquitecturas base que sean sencillas y muy generales que permitan tener una idea de cómo se manejaría esta migración en la nube. Teniendo estas arquitecturas básicas de guía se podrían ir ajustando a las necesidades y actividades de cada una de las empresas, pues es indispensable saber en su totalidad el manejo de los datos tanto vía consumidor como con empleadxs, para poder diseñar una arquitectura robusta. En este caso, al tener una idea general de lo que realiza la empresa, y de cómo se maneja el acceso a información para empleadxs, significó un reto en cómo definir la mejor arquitectura de datos, así como la configuración de los diferentes servicios. No obstante, la arquitectura creada, a pesar de ser bastante simple, me ha dado una idea más clara de la construcción de las mismas desde cero, y la importancia de nunca dejar de lado la alta disponibilidad dentro de ellas.

- **Fernanda Torres**: En los últimos años una gran cantidad de empresas han realizado un proceso de migración de servicios a la nube, esto debido a la seguridad, disponibilidad, escalabilidad y en ciertos escenarios rentabilidad que genera, sin embargo, lo que normalmente no es tomado en cuenta es la diversidad de soluciones de arquitecturas para cada empresa y cómo esta puede verse afectada por el giro de la empresa. En nuestro caso, para la empresa DataTech propusimos un escenario base pero nuestras estimaciones de precios y arquitectura puede verse fácilmente afectadas por los pequeños detalles de la empresa, por lo que para casos reales y propuestas más detalladas y aterrizables sugiero consultar con personas expertas en el tema y un acercamiento con personal de la empresa para comprender sus necesidades, sin embargo, para fines de la actividad y aplicación del aprendizaje obtenido en clase, considero que el escenario es el adecuado.
- **Federico Medina**: Observamos que el crear arquitecturas y establecer un presupuesto es muy difícil si no se sabe la idea completa de cómo es que servirá dicha herramienta en las operaciones diarias de la empresa al igual que sin tener el contexto completo del negocio. Dicha arquitectura creada fue hecha con la suposición de una situación muy específica, lo cual significa que no necesariamente sea la arquitectura que DataTech deba utilizar. Esto debido a que hay muchas variables que peuden afectar drásticamente la arquitectura como las regiones específicas, la cantidad de tráfico que reciben, el presupuesto que la empresa pone para migrar a la nube, los servicios específicos que buscan, etc. Por la misma cuestión, el estimar los costos fue difícil y tuvimos que "inventar" ciertos precios ya que los servicios te ofrecen sus precios en base a diferentes cuestiones que tú como empresa les presentas a estos. Sin embargo, el tener una idea o una visión sobre cómo diseñar una arquitectura y cómo calcular sus costos es muy importante porque ya sabiendo la parte general de este proceso, el buscar migrar a la parte específica es mucho más fácil y se podrá ayudar de una mejor manera a diferentes empresas como lo es DataTech.
- **Michelle Morales**: Diseñar una arquitectura en la nube para DataTech no es tan sencillo si no conoces a detalle los recursos y limitaciones con las que te enfrentas, sin embargo después de analizar e investigar un poco más acerca de los requerimientos se logró hacer un diseño adecuado. La estimación de costos fue tal vez la parte más complicada, ya que existe un rango de precios muy variados que dependen completamente de las necesidades y preferencias de la empresa, pero consideramos los precios que ya conocíamos previamente de algunos servicios. Las arquitecturas en la nube pueden brindar una mayor eficiencia operativa, flexibilidad estratégica y capacidad para innovar, lo cual es de suma importancia para las empresas, sin embargo, para obtener los mejores resultados es importante conocer bien los recursos y necesidades de la empresa y adaptar las arquitecturas a ello, para así mismo hacer una estimación de costos más acertada y brindar un diseño realmente eficiente.

Referencias

[1] Terry Lanfear. *Best practices for network security - Microsoft Azure*. 2023. URL: `https://learn.microsoft.com/en-us/azure/security/fundamentals/network-best-practices`.

[2] Shuhei Uda. *Backend Pool Management - Azure Load Balancer*. 2023. URL: `https://learn.microsoft.com/en-us/azure/load-balancer/backend-pool-management`.