Campus Monterrey

Inteligencia artificial avanzada para la ciencia de datos II (Gpo 501)

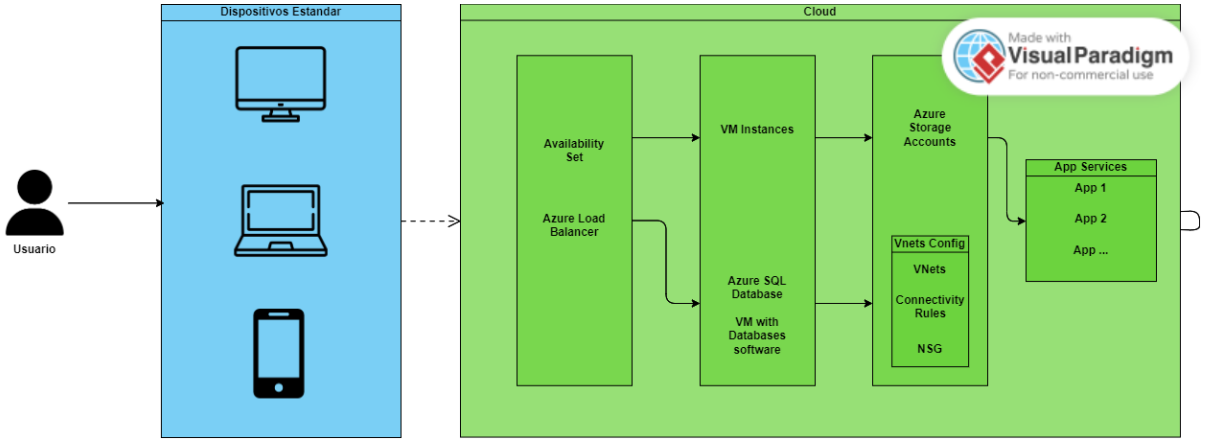
Diseño de Arquitectura en la Nube

Juan Pablo Castañeda Serrano	A01752030
Aldo Daniel Villaseñor Fierro	A01637907
José Alfredo García Rodriguez	A00830952
Francisco Castorena Salazar	A00827756

Actividad: Diseño de Arquitectura en la Nube

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.

Diagrama de la arquitectura propuesta



- Aquí el flujo comienza desde el usuario y pasa a traves de algun servicio DaaS (Desktop as a Service), en algún tipo de dispositivo electrónico conectado a internet, luego se distribuye a instancias de maquinas virtuales con un equilibrio de carga.
- Para garantizar alta disponibilidad las maquinas virtuales están conectadas a grupos de disponibilidad.
- Se accede a bases de datos a través de la configuración entre VNets, y las reglas de conectividad y NSG se aplican para garantizar la seguridad.
- El almacenamiento en la nube se utiliza para archivos y datos no estructurados, y los servicios de App Service alojan las aplicaciones web de DataTech.

Este es un diagrama de alto nivel, profundizaremos acerca del funcionamiento de cada una de estas partes en la siguiente sección del documento.

Sobre el diseño de la arquitectura

Para este escenario de migración y creación de una arquitectura completa para DataTech, se decidió utilizar los servicios de Microsoft Azure, esto al considerar distintos aspectos como diversificar la experiencia en la nube, habiendo trabajado previamente en nuestras actividades de consultoría principalmente con AWS. Nuestro equipo buscará explorar y comprender otros proveedores de servicios en la nube, para ampliar conocimientos, ademas de esto, consideramos que Azure ofrece una gama completa de servicios con precios competitivos, como por ejemplo en instancias de maquinas virtuales de serie B, Azure SQL Database y opciones de almacenamiento. La flexibilidad y escalabilidad de los servicios de Azure así

como las características de seguridad y sus políticas de seguridad (https://learn.microsoft.com/en-us/azure/governance/policy/overview) proporcionan un entorno confiable y protegido para montar nuestras propias arquitecturas Cloud.

Maquinas Virtuales

- Las instancias de máquinas virtuales de la serie Bs en Azure son una elección sólida debido a su atractiva relación precio-rendimiento. Para optimizar aún más los costos, es fundamental considerar cuidadosamente el tipo de instancia según la carga de trabajo específica. La instancia B20ms se destaca por sus impresionantes recursos, como 20 núcleos, 80 GiB de RAM y 160 GiB de almacenamiento temporal, lo que la convierte en una opción eficaz para momentos de alta demanda. Por otro lado, la instancia B1, con 1 GiB de RAM y 4 GiB de almacenamiento temporal, es una opción más económica para flujos de datos bajos.
- Distribución de carga: Implementaremos Azure Load Balancer ya que garantiza una distribución eficiente y rentable de la carga entre las instancias de máquinas virtuales. Este servicio distribuye el tráfico de red de manera equitativa, mejorando la disponibilidad y evitando la congestión en una instancia específica.
- Grupos de Disponibilidad: Utilizaremos zonas de disponibilidad para asegurar la alta disponibilidad a un costo competitivo, Azure divide las regiones en varias zonas, cada una con su propia fuente de energía, refrigeración y red independiente. Esto asegura la continuidad del servicio incluso en el caso de fallas en una zona específica.

La elección consciente de instancias de máquinas virtuales, el uso eficiente de Azure Load Balancer y la implementación de zonas de disponibilidad permite optimizar los costos sin comprometer capacidades de respuesta y disponibilidad.

Bases de Datos IaaS y PaaS

- Base de datos PaaS: Se escogerá Azure SQL Database para bases de datos PaaS debido a su capacidad de escalabilidad automática. Este servicio permite que la escalabilidad sea dinámica según la demanda. Es esencial destacar que Azure SQL Database proporciona características integrales de seguridad, como la autenticación multifactor y la encriptación de datos en reposo y en tránsito.
- Base de datos IaaS: En caso de ser necesario se implementará el uso de instancias de maquinas virtuales serie B con software de base de datos instalado para un costo más bajo y mayor control sobre los datos que se manejen, estas VM solo serán utilizadas cuando la cantidad de información sea baja, o si se requiere algún otro tipo de software del almacenamiento de datos del tipo Non SQL.
- Backup files y recovery systems: Se configurará Azure Backup para las bases de datos Iaas, por parte de las bases de datos PaaS en Azure SQL Database crean copias de seguridad de forma automatica y periodica en diferentes frecuencias de tiempo, por ejemplo full backup cada semana, o cada 10 minutos aproximadamente ante entry logs (https://learn.microsoft.com/en-us/azure/azure-sql/database/automated-backups-overview?view=azuresql)

Storage Accounts

- Configuraremos Azure Blob Storage para File Share, ya que es un servicio de almacenamiento escalable diseñado para manejar grandes volúmenes de datos no estructurados, como archivos y objetos. Al optar por este servicio, DataTech se beneficia de un modelo de precios flexible.
- Azure Blob Storage proporciona la funcionalidad de File Share a través de su característica de Blob Storage con la opción de acceso mediante el protocolo SMB (Server Message Block). Esto permite compartir archivos entre aplicaciones y equipos de manera eficiente, estableciendo un entorno colaborativo y facilitando el acceso a datos no estructurados.
- El almacenar los datos en varias cuentas de almacenamiento puede beneficiar a DataTech ya que puede ayudar a distribuir la carga y mejorar rendimientos, por otro lado, DataTech puede aplicar mejores políticas de acceso y de seguridad de manera más controlada dependiendo de los tipos de datos almacenados en distintas cuentas , lo que mejora la gobernanza.

Configuración entre VNets

- Se utilizará Azure Virtual Network para configurar las VNets, este servicio permite que muchos tipos de recursos de Azure se comuniquen de forma segura entre sí, con Internet y con las redes locales. Estos recursos de Azure incluyen máquinas virtuales (VM).
  (https://learn.microsoft.com/en-us/azure/virtual-network/virtual-networks-overview)
- Se implementaran dentro de Azure Virtual Network reglas de conectividad y NSG de manera eficiente para minimizar costos y garantizar la seguridad.
- Se buscará en Azure Virtual Network implementar el principio de least privilege para asignar los permisos de red de manera precisa y limitada, reduciendo superficies de ataque y minimizando riesgos de amenazas internas o externas.
- Azure NSG actuará como un tipo de cortafuegos que filtrará el tráfico de red hacia y desde los recursos de Azure. Al definir reglas específicas que permiten o niegan el tráfico según protocolos, puertos y direcciones IP, las NSG ofrecen un control preciso sobre la conectividad, permitiendo controles de acceso, segmentación de redes, monitoreo y adaptabilidad para ajustar dinamicamente los requisitos de seguridad o actualizaciones de la infraestructura.

App Services

- Utilizaremos Azure App Service con el plan de servicio de consumo para las aplicaciones web, ya que este plan se ajusta automáticamente a la demanda y puede resultar más económico en comparación con otros planes.
- Configuraremos certificados SSL mediante Azure Key Vault, aprovechando la gestión centralizada de certificados a un costo competitivo.

Sugerencias de seguridad

- Se implementará el principio de least privilege para asignar permisos y minimizar costos asociados con el consumo de servicios.

- Segmentación de redes para cumplir con requisitos de seguridad sin incurrir en costos adicionales.
- Se utilizarán los servicios de cifrado nativos de la nube para garantizar la seguridad de los datos.

Estimación de costos mensuales
La estimación de costos para este caso puede ser complejo debido a las variaciones en el consumo de recursos, elecciones de configuración y tarifas cambiantes por zonas de disponibilidad, dicho esto se proporcionara un escenario ficticio basado en ciertas suposiciones de DataTech, se considerará que es una startup y cantidad flexibles de datos a manejar.

Escenario

1. Máquinas Virtuales (VM):

Se implementan 5 instancias de máquinas virtuales de la serie B, utilizando instancias B2 para un equilibrio de rendimiento y costo.
Se asume que estas instancias son suficientes para manejar las cargas de trabajo previstas de DataTech.

2. Bases de Datos:

Se utiliza Azure SQL Database para las bases de datos PaaS, con capacidad de almacenamiento suficiente para manejar 1 TB de datos mensuales.
Se implementan 2 instancias de máquinas virtuales serie B para bases de datos IaaS, con almacenamiento adicional.

3. Storage Account - Fileshare:

Se configura Azure Blob Storage para File Share, con capacidad suficiente para almacenar archivos y datos no estructurados.

4. Configuración entre VNets:

Se establecen dos VNets para segmentar la arquitectura y se aplican reglas de NSG y conectividad según las necesidades de seguridad.

5. App Service:

Se implementa Azure App Service para alojar las aplicaciones web de DataTech, utilizando el plan de servicio de consumo para adaptarse a la demanda variable.

6. Estimación de Rango de Costos Mensuales:

- Máquinas Virtuales: Entre $500 y $1,000 por mes.

- Bases de Datos: Entre $300 y $600 por mes.
- Storage Account - Fileshare: Entre $50 y $100 por mes.
- Configuración entre VNets: Entre $20 y $40 por mes.
- App Service: Entre $100 y $200 por mes.

7. Justificación del Rango de Costos:

Los precios exactos dependerán de las configuraciones específicas, como el tamaño y tipo de instancias de máquinas virtuales, la cantidad de almacenamiento utilizado, y otros detalles de configuración.

Conclusiones

- Francisco: La decisión de probar con Microsoft Azure en lugar de quedarnos solo con AWS muestra que estamos abiertos a nuevas experiencias en la nube. Esto no solo amplía nuestros conocimientos, sino que también nos brinda opciones más allá de lo habitual. También es importante no solo conocer los servicios que brindan estas compañias de Cloud sino también tener conocimientos y capacidades de como conectar entre si todos estos para poder crear un entorno que permita al usuario de principio a fin tener una experiencia positiva sobre los servicios que brinda una empresa como DataTech.
- Juan Pablo: Hemos puesto mucho énfasis en la seguridad. Desde restringir permisos con el principio de least privilege hasta usar reglas de seguridad de red, queremos asegurarnos de que se cumplan las normas para disminuir vectores de ataque. La seguridad no es solo una palabra de moda, es una prioridad. Este enfoque equilibrado entre seguridad y eficiencia demuestra la prioridad de garantizar un entorno confiable y protegido para la infraestructura en la nube de DataTech.
- Aldo: La selección de instancias de máquinas virtuales de la serie B en Azure, junto con el uso eficiente de Azure Load Balancer y zonas de disponibilidad, resalta un la necesidad e importancia de optimizar la relación precio y rendimiento. Esta estrategia apunta a mantener costos bajo control mientras se garantiza una alta disponibilidad y capacidad de respuesta.
- Jose Alfredo: La gestión de bases de datos abarca tanto modelos PaaS como IaaS, evidenciando una estrategia integral para la escalabilidad y la seguridad de los datos. La flexibilidad para adaptarse a distintos escenarios, como la elección entre Azure SQL Database y máquinas virtuales con software de base de datos instalado una adaptabilidad clave y necesaria ante los entornos que pueden ser muy cambiantes de una empresa pequeña o mediana.